Proceedings of the 24th Amsterdam Colloquium



Edited by Fausto Carcassi, Tamar Johnson, Søren Brinck Knudstorp, Sabina Domínguez Parrado, Pablo Rivas Robledo & Giorgio Sbardolini

Foreword

The papers contained in this collection were presented at the 24th Amsterdam Colloquium, organised by the Institute for Logic, Language and Computation (ILLC) at the University of Amsterdam, December 18–20, 2024. The Amsterdam Colloquia aim at bringing together linguists, philosophers, logicians, cognitive scientists, and computer scientists who share an interest in the formal study of the semantics and pragmatics of natural and formal languages.

Besides the general programme, the 2024 edition featured two workshops. First, "Dynamic and Inquisitive: Themes from the Work of Jeroen Groenendijk", dedicated to the memory of Jeroen Groenendijk. Second, "Identifying and Explaining Universals of Semantics". The 2024 edition also included the Beth Lecture, jointly organised with the E.W. Beth Foundation. The programme included 7 invited talks, 34 contributed talks, and three poster sessions with 24 contributed posters.

We would like to thank all programme committee members, listed below, and all reviewers for their efforts in selecting the contributed talks and posters. We would also like to thank Alexandra Zieglerová, Peter van Ormondt, and Ewout Arends for their practical help in organising the conference, Hadas Kotek for the mentoring program, and the MoL and philosophy students for assistance.

Lastly, we would like to thank the ILLC, the E.W. Beth Foundation, the Netherlands Organization for Scientific Research (NWO), and the European Research Council (ERC) for financial support.

Fausto Carcassi
Tamar Johnson
Søren Brinck Knudstorp
Sabina Domínguez Parrado
Pablo Rivas Robledo
Giorgio Sbardolini

Programme Committee

Maria Aloni (chair)	ILLC, University of Amsterdam
Paul Dekker	ILLC, University of Amsterdam
Luca Incurvati	ILLC, University of Amsterdam
Karolina Krzyzanowska	ILLC, University of Amsterdam
Floris Roelofsen	ILLC, University of Amsterdam
Marieke Schouwstra	ILLC, University of Amsterdam

Table of Contents

Non-maximality effects in gestural plural predication 1
Stavroula Alexandropoulou, Kurt Erbach, Richard Breheny, Clemens Mayr, Jacopo Romoli, Yasutada Sudo

To compose a number is to compose a cardinality: evidence from intensional contexts 8
Luis Alonso-Ovalle & Bernhard Schwarz

Facts, intentions, questions: English 'come-to-know' predicates in deliberative environments 15
Pranav Anand & Natasha Korotkova

How is presupposition accommodation constrained, again? 22
Ido Benbaji-Elhadad & Omri Doron

Evaluating Unalternative Semantics on a large corpus of naturalistic speech 29
Fenna Berger & Matthijs Westera

More than one way to free choice: A view from child Romanian 36
Adina Camelia Bleotu, Mara Panaitescu, Alexandre Cremers, Andreea Nicolae, Anton Benz, Gabriela Bîlbîie, Lyn Tieu

Thinking Statively and Dynamically: A View from Georgian 44
Tatiana Bondarenko, Richard Luo, Vincent Rouillard

How Conditionals Restrict 51
David Boylan

Actuality Entailments and the Conditional Analysis 58
David Boylan & Ginger Schultheis

Priming NPI Acceptability Judgments and the Bagel Problem 65
Lisa Bylinina, Stavroula Alexandropoulou, Yasutada Sudo

Default Premise Semantics for Anankastic Conditionals 73
Fabrizio Cariani

Evaluativity in *even*-hosting comparatives: Information structure and salient scales 80
Zhuang Chen

Arbitrariness and Frege Arithmetic 88
Ludovica Conti

Vagueness and pragmatic reasoning in quantified sentences 95
Alexandre Cremers

Reanalysis in discourse comprehension: Evidence from reading times 102
John Duff & Daniel Altshuler

Ignorance under attitudes 111
Patrick Elliott & Filipe Hisao Kobayashi

Putting summative predicates into context 118
Kurt Erbach, Stavroula Alexandropoulou, Richard Breheny, Clemens Mayr, Jacopo Romoli, Yasutada Sudo

Ignorance and attenuation 125
Anamaria Fălăuş & Andreea Nicolae

Crossing different ontological domains: the case of *appena* 132
Valeria Gradimondo, Lucia M. Tovena, Timothée Bernard

What is special about *special*? 138
Yael Greenberg & Hedde Zeijlstra

Scalarity, Information Structure and Relevance in varieties of Hurford Conditionals 145
Adèle Hénot-Mortier

Additive Presuppositions and Logical Strength 152
Elena Herburger & Paul Portner

The Case for an Anchored CAUSE 159
Angelica Hill

A Pragma-Semantic Account for Negative Island Obviation by *Wa* in Japanese 167
Hitomi Hirayama

Same Type of Comparatives, But with Different Syntax and Semantics 174
Yaqing Hu

DUAL as a core concept and the pronounceability of alternatives 181
Paloma Jeretič, Aurore Gonzalez, Itai Bassi, Kazuko Yatsushiro, Uli Sauerland

A Semantics for Weak, Question-Sensitive Belief 188
Andrej Jovićević

Presuppositions project asymmetrically, *unless* they don't 196
Alexandros Kalomoiros, Matthew Mandelkern, Jacopo Romoli, Florian Schwarz

Relative quantification and equative scope-taking 203
Andrew Kato

Measuring event participants 211
Boram Kim

Counterpossibles, Functional Decision Theory, and Artificial Agents 218
Alexander W. Kocurek

Hurford's Constraint and Disjunctions over Speech Acts 226
Manfred Krifka

Circumstantial Indirectness: Future Orientation of Modals Revised 234
Jakob Majdič

Northern Khanty demonstratives: a markedness-based approach 241
Lera Marinina

Neg-raising interacts with implicatures: The case of *doubt* 248
Andrea Matticchio

Dutch Books, Indicative Conditionals, and Rational Updating 256
Calum McNamara & Xeuyin (Snow) Zhang

A localist approach to the semantics and distribution of X-marking 263
Teruyuki Mizuno

What Is Carnap's Problem and How Can We Solve It? 271
Julien Murzi & Brett Topey

Explaining universal semantic properties in the adjectival/adverbial domain 278
Irina Nemtcova & Jakub Szymanik

On the anti-exhaustive inference of *ya* 284
Andreea Nicolae, Yasutada Sudo, Muyi Yang

Generic Explanations for Generics 291
Tom Ralston

Dynamic Effects of Modalized Questions 298
Valentin D. Richard

Just-asking questions 308
Tom Roberts

Bounds for metaphysical modality 316
Paolo Santorio

Age stratification in collective readings of Present Day English *every*-DPs 323
Fabian Schlotterbeck, Polina Berezovskaya, Sigrid Beck

What appositives can tell us about names and definite descriptions 330
Eli Sharf

Generic Modals in Mandarin 337
Chun-Hung Shih

The interpretation of logophoric and ordinary pronouns in Ewe: an experimental study ... 344
Silvia Silleresi, Itai Bassi, Abigail Anne Bimpeh, Imke Driemel, Anastasia Nuworsu, Maria Teresa Guasti

Causal abductive reasoning in discourse processing: Evidence from eye tracking 351
Torgrim Solstad & Oliver Bott

Expressive inferences & mighty obstacles 358
Andres Soria-Ruiz

Loose nominals 365
Sreekar Raghotham

Making a statement: eventuality denoting nominals 371
Peter R. Sutton & Hana Filip

On Being Certain that Presuppositions don't Project Universally 378
Maik Thalmann & Andrea Matticchio

Distributivity facilitates ACD resolution 386
Jad Wehbe, Kate Kinnaird, Martin Hackl

Believe is not a propositional attitude verb 393
Simon Wimmer

On the availability of inverse scope reading in Dutch doubly-quantified thetic sentences ... 401
Hongchen Wu & Maxime Alexandra Tulling

Toward a Semantic Theory of Universal NCIs 408
Yusuke Yagi, Takanobu Nakamura, Kenta Mizutani

Testing theories of SDA via acquisition 415
Eleonora Zani, Ivano Ciardelli, Emanuela Sanfelici

An experimental assessment of the nall lexical gap 422
Ruizhe Zhou, Benjamin Spector, Mora Maldonado

Non-maximality effects in gestural plural predication

Stavroula Alexandropoulou
University College London
London, UK
stavroulaale@gmail.com

Kurt Erbach
Goethe-University Frankfurt am Main, Germany
& Saarland University, Saarbrücken, Germany
erbach@hhu.de

Richard Breheny
University College London
London, UK
r.breheny@ucl.ac.uk

Clemens Mayr
Georg-Austin-Universität Göttingen
Göttingen, Germany
clemens.steiner-mayr1@uni-goettingen.de

Jacopo Romoli
Heinrich-Heine-Universität
Düsseldorf, Germany
jacopo.romoli@gmail.com

Yasutada Sudo
University College London
London, UK
y.sudo@ucl.ac.uk

Abstract

The present paper investigates homogeneity and non-maximality of novel pro-speech gestures. Our experiment adopts and refines Tieu, Schlenker, and Chemla's (2019) experimental design and also incorporate Romoli et al.'s (2024) idea of context manipulation. The results indicate not only that homogeneity inferences can arise from unfamiliar gestures but also that non-maximality triggered by gestural means is contextually modulated and is more robustly so in positive than in negative sentences. In light of Romoli et al.'s (2024) observation for plural definites, these findings point to the possibility that the mechanism responsible for homogeneity and non-maximality extends beyond conventionalized linguistic expressions to include unfamiliar gestures, aligning with the view that homogeneity and non-maximality emerge from productive semantic and pragmatic processes.

1 Introduction

There has recently been growing interest in the meaning contributions of gestures (Schlenker 2019; Hinterwimmer, Patil, and Ebert 2021; Ebert 2024, *inter alia*). One of the central questions in this literature is whether or not certain types of meaning commonly observed with linguistic expressions can arise from utterances containing unfamiliar gestures. Tieu, Schlenker, and Chemla 2019, among others, provide experimental evidence that scalar implicatures, presuppositions, supplements, and homogeneity can indeed be triggered by novel gestures used in place of linguistic expressions, or 'pro-speech gestures'. This paper presents further evidence that homogeneity inferences can be triggered by unfamiliar pro-speech gestures and furthermore that they give rise to non-maximality effects similarly to definite plurals.

1.1 Definite plurals

Both homogeneity and non-maximality are typically discussed with respect to plural definites, although a number of other linguistic expressions also exhibit these two properties (Krifka 1996; Löbner 2000; Križ 2015; Križ 2019, *inter alia*). By way of illustration, consider the positive sentence in (1) and the negative sentence in (2).

(1) Sam took the marbles.

(2) Sam did not take the marbles.

Homogeneity has to do with a truth-value gap. Intuitively, (1) is true iff $\lambda x.$*Sam took* x applies to each relevant marble ('Sam took all of the marbles'), while (2) is true iff $\lambda x.$*Sam didn't take* x applies to every individual marble ('Sam didn't take any of the marbles'). Note that (2) is stronger than the logical negation of (1) ('It is not true that Sam took every marble'). Therefore, there are situations where neither of these sentences are true, namely, when Sam took some but not all of the marbles. In other words, homogeneity has to do with the inference that the atomic parts of the plural definite behave homogeneously—all true or all false—with respect to the predicate (e.g. $\lambda x.$*Sam didn't take* x).

Non-maximality, on the other hand, has to do with looseness in meaning. In fact, the above characterizations of the truth-conditions of (1) and (2) are arguably too strong, since, as a matter of fact, these sentences may be judged to be true, relative to certain nonhomogeneous scenarios. Concretely, (1) can well be judged to be true when Sam took nine out of ten marbles, at least if a small number of exceptions are somehow deemed insignificant and are tolerated.

Several competing theories have been proposed to capture these two interpretive properties of plural definites. One dimension along which the theories diverge in predictions is whether the context-sensitivity of non-maximality is asymmetric with respect to polarity such that negative sentences like (2) are less context sensitive with respect to exception tolerance (e.g., Bar-Lev 2018; Bar-Lev 2021), or symmetric (Križ 2015; Križ 2016). Recently, Romoli et al. 2024 experimentally tested these predictions with a picture-sentence verification task with context. They compared an existential context ("Opening the presents is prohibited") and a universal context ("Opening the presents is required"), and found an asymmetric interpretation pattern between definite plurals in positive (e.g., *Every girl opened her presents*) and negative sentences (e.g., *No girl opened her presents*) such that context had a more robust effect on positive sentences than on negative sentences. We incorporated their design in the experiment we will report below.

1.2 Gestures

Turning to gestural semantics, Tieu, Schlenker, and Chemla 2019 provide experimental evidence suggesting that unfamiliar pro-speech gestures can be interpreted with homogeneity. However, we think that their conclusion is not entirely convincing due to a certain feature of their experimental design, which we will discuss immediately below. Certain design considerations of our own experiment are related to this issue, so let us delve into the relevant aspect of their study.

Tieu, Schlenker, and Chemla 2019 tested two utterances with the same pro-speech gesture. In both cases, a narrator in a video first introduced a background scenario with two sets of objects—three crosses introduced on the left and three coins introduced on the right—using gestural plurals in the form of repeated pro-speech gestures at different adjacent spatial points, as in (3). "CROSS-REP3" and "COIN-REP3" are the pro-speech gestures and everything else was linguistically uttered.

(3) Sam is participating in a treasure hunt in the forest, and she is looking for crosses and coins. Very quickly, Sam will find $[\text{CROSS-REP3}]_{\text{left}}$ and $[\text{COIN-REP3}]_{\text{right}}$.

Next, the narrator made an utterance beginning either with "Sam will" or "Sam will not" followed by a pro-speech two-handed grabbing gesture, TAKE-2-HANDED, performed on the right-hand side, where the coins had been introduced. The question is whether this gesture is interpreted homogeneously similarly to 'take the coins' or not.

After seeing the target utterance with a pro-speech gesture, the participant did an inference task where they indicated on a slider how strongly they thought a written inference followed from the utterance. The positive utterance was paired with the inference *Sam will take all of*

Approach	Prediction	Context	
		Existential	Universal
Asymmetric	Sam will TAKE-2-HANDED.	✓	×
	Sam will not TAKE-2-HANDED.	×	×
Symmetric	Sam will TAKE-2-HANDED.	✓	×
	Sam will not TAKE-2-HANDED.	×	✓

Table 1: Predictions of two approaches for target sentences with respect to non-maximality.

the coins and the negative utterance was paired with the inference *Sam will not take any coins.*

The results indicate that for both positive and negative sentences, participants tended to judge these inferences to follow (relative to a baseline condition where the target inference does not follow). These results led Tieu, Schlenker, and Chemla to conclude that homogeneity can be triggered by unfamiliar gestures.

While their results are suggestive, we think they do not provide a very convincing argument for their conclusion. This has to do with the nature of the inference task that they employed. The inference task certainly has a number of practical advantages, but it also has some potential pitfalls. Above all, it does not make it clear whether the inference drawn is due to the semantic or pragmatic meanings of the expressions used, or alternatively, simply due to mere plausibility arising from the context. This latter possibility is highly relevant in interpreting Tieu, Schlenker, and Chemla's (2019) experimental results. That is, we think it is not too far-fetched to suppose that participants tended to assume, even without the target utterance, that Sam would take all of the coins, if he took any of them, because, after all, coins are valuable, and explicitly described as such, and Sam was looking for them. Our experimental design is free from such a contextual inference.

2 Experiment

Building upon Romoli et al.'s (2024) and Tieu, Schlenker, and Chemla's (2019) studies, we set out to investigate homogeneity and non-maximality with pro-speech gestures. Our experiment also aims to evaluate the predictions made by the symmetric and asymmetric approaches to the context-sensitivity of non-maximality. Table 1 summarizes the predictions of the two approaches.

2.1 Methods

We employed a picture-video judgment task. At the beginning of the experiment, a written background context introduced the main character, Sam, who cleans rooms and dusts shelves.

Following Romoli et al. 2024, we manipulated contexts between participants in terms of families with different cleaning instructions. In the Existential condition, the family instructions included the statement "Taking anything off the shelves is prohibited", and in the Universal condition, "Taking everything off the shelves is required". These instructions were mentioned as part of the instructions and also displayed at the top of the screen in each trial (see Figure 1).

In each trial, a narrator named Kurt in the video (played by the second author) verbally introduces the shelves Sam is going to dust along with the objects on them, using deictic co-speech gestures directed at the shelves and objects. The narrator then makes a prediction about what Sam will do while dusting the shelves. The narrator's prediction consists of a mixed utterance similar to those in Tieu, Schlenker, and Chemla 2019 such as (4).

(4) Sam will/will not TAKE-2-HANDED_  .

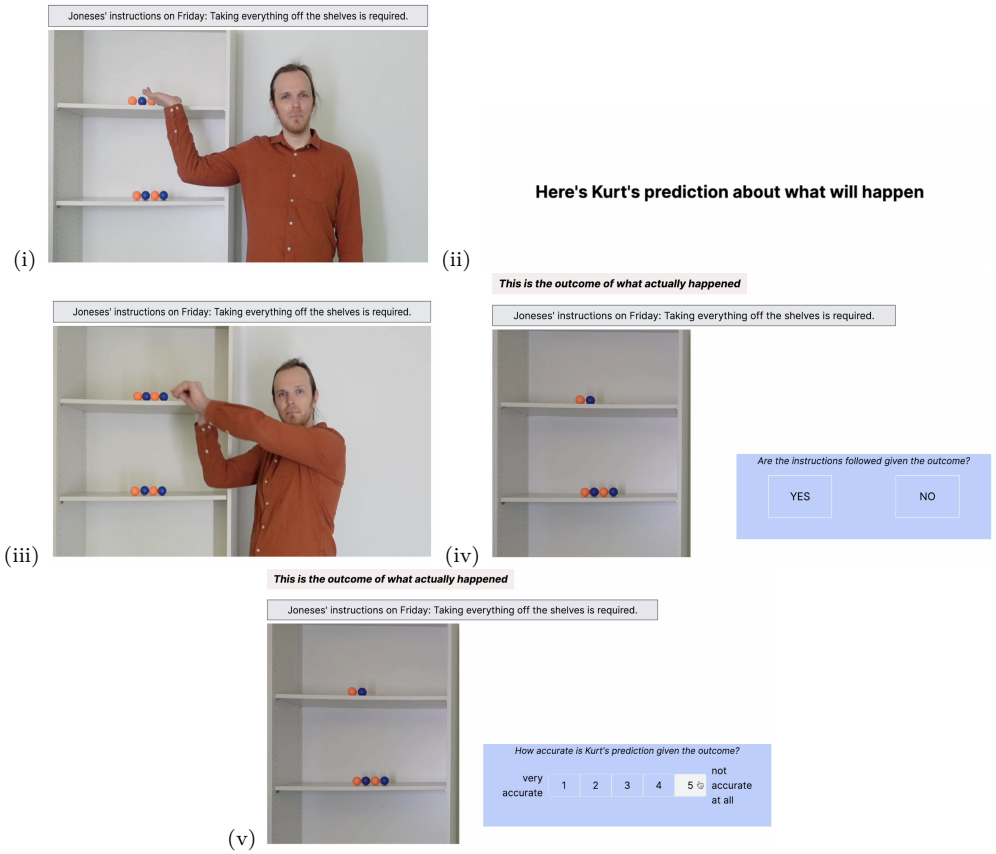


Figure 1: Example images of the displays in a trial: (i) The first display shows the video of the narrator presenting the shelves Sam will dust and the objects on them; (ii) the second display prepares participants for the narrator's prediction, containing the critical utterance; (iii) the third display shows the video with the narrator's prediction along with the critical gesture; (iv) the fourth display includes the secondary task; and (v) the fifth display contains the main task.

Unlike Tieu, Schlenker, and Chemla's (2019) items, which were about taking coins, we used objects of no particular value in all our items, in order to circumvent the potential issue of the contextual inference we pointed out above.

After the narrator's prediction, an image is presented depicting the outcome of what Sam actually did. Participants had to judge on a 5-point Likert scale how accurate the narrator's prediction was given the outcome of what Sam actually did. Before making this judgment, participants also answered a secondary question about whether Sam had followed the cleaning instructions (appearing at the top of the screen throughout the trial) given the actual outcome by *Yes* or *No*. This ensured that participants understood the context they were in and remembered it throughout the experiment. Each trial consisted of 5 displays (see Figure 1).

We manipulated the polarity of the narrator's prediction (positive: *will* vs. negative: *will not*), the type of outcome (True Control, False Control, Target), and the type of cleaning instructions (Context). In the Target outcome condition (see examples on OSF), the outcome picture showed two out of four blocks/balls, creating a non-homogeneous scenario that would only be true with non-maximality. The outcome pictures in the Control conditions depicted homogeneous scenarios, showing either all objects removed from the affected shelf or all objects intact (see OSF). The Control conditions tested homogeneity. Polarity and Type of outcome were within-participant factors and Context was a between-participant factor.

We created two experimental items: one featuring 4 balls on each shelf and the other featuring 4 blocks on each shelf. The position of the balls/blocks affected by the narrator's

gesture (top/bottom shelf), the direction of the gesture (left/right) and the narrator's outfit (3 different outfits) were counterbalanced across 12 lists. In half of them the order of labels of the response buttons (see Figure 1) was presented in one sequence, while in the other half, the order was reversed. In total, there were 24 lists—12 for each Context. Each item appeared in 6 conditions (2 Polarity × 3 Outcome type) per list, resulting in a total of 12 experimental trials per list. We also created 24 filler items, each featuring different types of objects (candles, mugs, toys) on both shelves and various gestures (e.g., a one- or two-handed pick-up gesture affecting different shelves, a one- or two-handed grabbing gesture affecting different shelves, a one-handed moving up/down gesture, and a one-handed knocking-down gesture). The same 24 filler items appeared in every list and were designed to balance the yes/no responses to the secondary question in experimental items regarding whether the cleaning instructions were followed by Sam, achieving a near-optimal balance.

Participants were randomly assigned to one of the two Context conditions. After reading the instructions, participants were presented with two example items designed to familiarize them with the tasks. For each example item, they received explicit feedback on the correct responses to the secondary task—whether Sam followed the cleaning instructions—and the critical task, if they initially gave an incorrect response. Unambiguous items resembling the filler items were used as example items. Participants could only proceed to the main experimental session after providing the correct responses to these example items.

The experiment was implemented in Gorilla Experiment Builder. Native speakers of English were recruited online via Prolific. To achieve a target of 40 participants per list, we recruited a total of 960 participants across two experimental batches. Each participant was compensated £2.61 (1st batch) or £2.55 (2nd batch) for their participation. Informed consent was obtained from each participant prior to participation. We excluded participants who completed the study on a smartphone, or on the Safari browser, which is not optimal for Gorilla experiments involving images and videos. Additionally, participants who reported not hearing any sound or who had an accuracy rate below 75% on the critical task on the Control items (with ratings higher than 3 for False Controls and ratings lower than 3 for True Controls considered errors) were excluded, resulting in a final dataset of 862 participants. This set of participants had a mean age of 39.94 years (age range: 18–77). Of these, 490 identified as female, 368 as male, and 4 preferred not to say.

2.2 Results

Following Romoli et al. 2024, the Context (Type of Instructions) factor was recoded into Lax (positive predictions in existential context, negative predictions in universal context) and Strict (positive predictions in universal context, negative predictions in existential context). Lax invites non-maximal interpretations while Strict prevents them. Figure 2 shows mean ratings per condition. Notice the high mean ratings obtained by the True control conditions overall, suggesting the availability of homogeneity effects for both positive and negative sentences. A similar, but inverse, pattern is observed for the False control conditions.

We fitted a cumulative link mixed-effects model to analyze prediction ratings for Target items, with sum-coding for Context and treatment-coding for Polarity of Prediction (with negative as the reference category).[1] The analysis was conducted using the ordinal package (Christensen 2023) in R (R Core Team 2022). The model also included random intercepts for Participants and random by-Participant slopes for Polarity. Due to the item factor having only two levels, the variance attributed to it was unlikely to be meaningful, and we excluded it as a random effect from our model.

[1]The model including the True and False Control conditions failed to converge, potentially due to insufficient variability in the corresponding ratings.

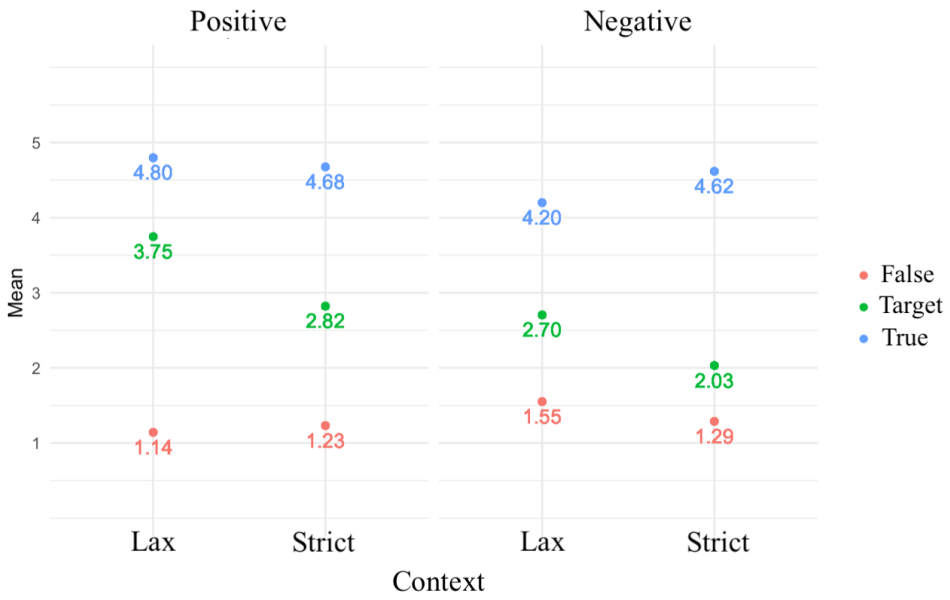


Figure 2: Mean prediction ratings per condition.

Our analysis revealed (see whole output in the Appendix on OSF): (i) a strong positive effect of Polarity of Prediction ($\beta = 2.25$, $SE = 0.13$, $z = 17.11$, $p < 0.0001$), with positive predictions eliciting higher ratings overall compared to negative predictions, (ii) a strong positive effect of Context ($\beta = 1.75$, $SE = 0.16$, $z = 10.64$, $p < 0.0001$), with Lax conditions receiving higher ratings than Strict ones overall, and importantly (iii) a significant positive interaction effect ($\beta = 0.46$, $SE = 0.19$, $z = 2.41, p < 0.05$), indicating greater context-sensitivity for non-maximal interpretations in positive predictions compared to negative ones.

3 Discussion and conclusion

The intermediate judgments in the Target conditions, which involved non-homogeneous scenarios, suggest that pro-speech gestures do trigger homogeneity.

Our results furthermore indicate that non-maximality of pro-speech gestures is more context-sensitive in positive sentences than in negative sentences, similarly to what Romoli et al. 2024 observed for plural definites. This points to the possibility that the mechanism responsible for homogeneity and non-maximality inferences may extend beyond linguistic expressions to include gestures. Following Schlenker 2019, the potential for gradient iconic effects of the pro-speech gesture tested in the present study makes it unlikely that the gesture was directly translated into the corresponding linguistic item.

Overall, our findings involving unfamiliar gestures further support the idea that types of inferences typically observed with linguistic expressions emerge through productive semantic and pragmatic processes rather than being exclusively tied to language.

Acknowledgements. The authors wish to thank Cornelia Ebert and Markus Steinbach for their valuable feedback on the experiment. The research in this paper was supported by an AHRC-DFG grant (AHRC: AH/V003526/1; DFG: EB 523/2-1, STE 958/12-1, STE 2555/3-1).

References

Bar-Lev, Moshe (2018). "Free choice, homogeneity and innocent inclusion". PhD thesis. The Hebrew University of Jerusalem.

Bar-Lev, Moshe E. (2021). "An Implicature Account of Homogeneity and Non-Maximality". In: *Linguistics and Philosophy* 44.5, pp. 1045–1097. DOI: 10.1007/s10988-020-09308-5.

Christensen, Rune H. B. (2023). *ordinal—Regression Models for Ordinal Data*. R package version 2023.12-4.1. URL: https://CRAN.R-project.org/package=ordinal.

Ebert, Cornelia (2024). "Semantics of gesture". In: *Annual Review of Linguistics* 10.1, pp. 169–189.

Hinterwimmer, Stefan, Umesh Patil, and Cornelia Ebert (2021). "On the interaction of gestural and linguistic perspective taking". In: *Frontiers in communication* 6, p. 625757.

Krifka, Manfred (1996). "Pragmatic Strengthening in Plural Predications and Donkey Sentences". In: *Proceedings of SALT VI*. Ed. by Teresa Galloway and Justin Spence. Ithaca, NY: CLC Publications, pp. 136–153.

Križ, Manuel (2015). "Aspects of homogeneity in the semantics of natural language". PhD thesis. University of Vienna.

— (2016). "Homogeneity, maximality, and *all*". In: *Journal of Semantics* 33, pp. 493–539.

— (2019). "Homogeneity effects in natural language semantics". In: *Language and Linguistics Compass* 13.11, e12350. DOI: 10.1111/lnc3.12350.

Löbner, Sebastian (2000). "Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences". In: *Linguistics and Philosophy* 23.3, pp. 213–308.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Romoli, Jacopo et al. (2024). *Putting plural definites into context*. submitted.

Schlenker, Philippe (2019). "Gestural semantics: Replicating the typology of linguistic inferences with pro-and post-speech gestures". In: *Natural Language & Linguistic Theory* 37, pp. 735–784.

Tieu, Lyn, Philippe Schlenker, and Emmanuel Chemla (2019). "Linguistic inferences without words". In: *Proceedings of the National Academy of Sciences* 116.20, pp. 9796–9801.

To compose a number is to compose a cardinality: evidence from intensional contexts

Luis Alonso-Ovalle
McGill University
Montréal, Québec, Canada
luis.alonso-ovalle@mcgill.ca

Bernhard Schwarz
McGill University
Montréal, Québec, Canada
bernhard.schwarz@mcgill.ca

Abstract

DPs of the form *the number of NP* can combine with predicates of numbers or with predicates of ordinary entities. The interpretation of such DPs in intensional contexts favours an account of this flexibility based on syntactic ambiguity (Alonso-Ovalle and Schwarz 2024) over an account based on constructing numbers from ordinary entities (cf. Scontras 2017).

1 Introduction

DPs whose NP is headed by the quantity noun *number* (NDPs), like (1), can saturate predicates of numbers, like *is below ten* in (2-a), and also predicates of ordinary entities, like *hire* in (2-b). This flexibility is surprising. Number predication is expected if NDPs are headed by the quantity noun, as in the parse in (3): if *number* is a predicate of numbers, NDPs should denote a number. But then, if NDPs denote numbers, how does entity predication arise?

(1) the number of cooks Al hired

(2) a. [The number of cooks Al hired] is below ten.
b. Bo hired [the number of cooks Al hired].

(3) $[_{DP}$ the $[_{NP}$ $[_{N}$ **number**] of cooks Al hired]]

In what we call the "rich number" approach, pioneered for closely related data in Scontras 2017, numbers are not primitives like cardinalities, but are constructed from ordinary entities. The simplest implementation of this idea takes numbers to be *sets* of pluralities of ordinary entities that have the same cardinality. In a world where Al hired three cooks, for example, the number denoted by (3) is then the set containing all cook pluralities of cardinality 3. These are the pluralities accessed in entity predication: (2-b) says that Bo hired at least one of them.[1]

In the "rich syntax" approach entertained in Alonso-Ovalle and Schwarz 2024, syntax enables entity predication. NDPs are taken to be structurally ambiguous. The parse in (3) is just one parse of (1). A second parse, in (4), contains two occurrences of *cooks*—one unpronounced. The higher occurrence of *cooks* is the head of the main NP, and is modified by the DP in (3).

(4) $[_{DP}$ $[_{NP}$ $\boxed{[_{DP}$ the $[_{NP}$ $[_{N}$ number] of cooks Al hired]] $}$ $[_{N}$ **cooks**]]]

The DP structure (3) is taken to denote a cardinality, such as 3 in a world where Al hired three cooks. Hence (3) supports number predication. The parse in (4) yields entity predication. At the level of the main NP, (4) parallels cases of modification by a numeral, as in *three cooks*.

[1] Reconstructing numbers in terms of entities is an old idea. Frege and Russell treat cardinal numbers as sets of sets with the same cardinality (see e.g. Hatcher 1990; Dummett 1991). Semanticists proposed, more generally, reconstructing degrees as equivalence classes of ordinary entities (see e.g. Cresswell 1976, Klein 1980 or Bale 2006.) Scontras 2017 does not construe rich numbers (more generally, amounts) as sets of entities, but as *nominalized properties of entities*. We stick to sets for simplicity of exposition, without affecting the argument.

Just like composing the numeral modifier with the head *cooks* in *three cooks* yields the set of cook pluralities of cardinality 3, composing the DP modifier with *cooks* in (4) yields the set of cook pluralities whose cardinality is the number of cooks Al hired. In a world where Al hired three cooks, this set is again the set of cook pluralities of cardinality 3. And again, (2-b) says that Bo hired at least one of them. Notably, this view makes it possible to maintain that numbers are cardinalities, understood as primitives, and that *number* is a predicate of cardinalities.

This paper presents a new argument for the rich syntax approach. The argument references NDPs in intensional contexts, where entity predication allows for "number transparency", truth conditions that reference a cardinality determined in the actual world. For instance, (5) can convey that Bo wants to hire a set of cooks whose cardinality is the number of cooks that Al actually hired. Since this cardinality can be given by the denotation of the modifier DP in (4), the rich syntax approach renders number transparency unsurprising. We argue that, in contrast, the rich number approach fails to offer a compositional path to number transparency.

(5) Bo wants [PRO to hire [the number of cooks Al hired]].

2 Entity predication

The sentence in (2-b) has truth conditions that make an existential statement (cf. Scontras 2017): they convey that Bo hired *some* plurality of cooks whose cardinality is the cardinality of the set of cooks that Al hired. We can state those truth conditions as in (6). This statement makes reference to the set in (7), whose description, in turn, refers to the cardinality in (8). The cardinality in (8) is the (maximal) number of cooks that Al hired. The set in (7) is the set of cook pluralities of that cardinality. The truth conditions in (6) quantify existentially over that set: they convey that Bo hired one of the pluralities in (7). In a world where Al hired three cooks, the cardinality in (8) is 3, the set in (7) is the set that contains all pluralities of three cooks, and the condition for the truth of (2-b) is that Bo hired at least one of them.

(6) $\exists x \in (7)[\text{hired}(\text{Bo})(x)]$

(7) *Domain component*: $\{y|\ \text{cooks}(y) \wedge |y| = (8)\}$

(8) *Cardinality component*: $\max\{n|\ \exists x \in \{z|\ \text{cooks}(z) \wedge |z| = n\}[\text{hired}(\text{Al})(x)]\}$

We will see next that both the rich syntax approach and the rich number approach derive these truth conditions, but that the approaches differ in terms of which syntactic constituents, if any, contribute the domain component in (7) and the cardinality component in (8). This difference will become consequential once we attend to NDPs in intensional contexts.

2.1 Composition: rich syntax

One of the parses that the rich syntax approach assumes for the NDP in (1) is repeated in (9). Here the noun *number* is assumed to be a predicate of cardinalities, and so the whole definite DP is assumed to denote one: the cardinality of the set of cooks that Al hired, in (8). Since (9) denotes a cardinality, it can saturate number predicates like *is below ten*.

(9) [$_{DP}$ the [$_{NP}$ [$_{N}$ **number**] of cooks Al hired]]

The second parse is repeated in (10). This is the parse involved in cases of entity predication like (2-b), repeated in (11), with the truth conditions given in (6) to (8).

(10) [$_{DP}$ [$_{NP}$ $\boxed{\text{[}_{DP}\text{ the [}_{NP}\text{ [}_{N}\text{ number] of cooks Al hired]]}}$ [$_{N}$ **cooks**]]]

(11) Bo hired [the number of cooks Al hired].

In (10), the main NP contributes the domain component in (7), and the modifying DP the cardinality component in (8). Modification restricts the denotation of the head noun, the set of all cook pluralities, to (7), the set of all cook pluralities with the cardinality in (8). Note the parallel to the NP *three cooks*, where *three* can be taken to restrict the denotation of *cooks* to the set of cook pluralities with cardinality 3. This restricting function could be enabled by a covert predicate MANY, as in *[$_{AP}$ three MANY] cooks* and in (12) (cf. Hackl 2000), which we will nevertheless omit from LFs. As for the source of the existential quantification in (6), we again merely assume that it is parallel to cases with numerals, like *Bo hired three cooks*. The quantification could be sourced to a covert existential head of the main DP. In an alternative composition, which we will adopt for ease of exposition, existence quantification is attached to the entity predicate, given a predication postulate such as (13).[2,3]

(12) [$_{DP}$ [$_{NP}$ [$_{AP}$ [$_{DP}$ the [$_{NP}$ [$_{N}$ number] of cooks Al hired]]] MANY] [$_{N}$ **cooks**]]]

(13) $[\![\,\text{hired}\,]\!](A)(y) \Leftrightarrow \exists x[x \in A \wedge [\![\,\text{hired}\,]\!](x)(y)]$

2.2 Composition: rich numbers

In an alternative based on Scontras 2017 (cf. Moltmann 2013), NDPs are unambiguously parsed as definite DPs, as in (14), which again repeats (3). Crucially, this definite DP is *not* taken to denote the cardinality (8), repeated in (15), but a "rich number": the set of equinumerous cook pluralities in (7), the domain component, repeated in (16). Under this analysis, it is the main DP, not the NP, that contributes the domain component.

(14) [$_{DP}$ the [$_{NP}$ [$_{N}$ **number**] of cooks Al hired]]

(15) $\max\{n|\ \exists x \in \{z|\ \text{cooks}(z) \wedge |z| = n\}[\text{hired}(Al)(x)]\}$

(16) $\{y|\ \text{cooks}(y) \wedge |y| = (15)\}$

Rich numbers convey information about both entities and cardinalities—(16) is a set of entities that are cooks and that share a certain cardinality. We then expect both quantity and entity predication to be possible. Entity predication simply conveys that the predicate is true of at least one member of the rich number, as in (17-b) for (17-a)—as before, the source of the existential quantification could be the predicate postulate in (13). Number predication can be taken to make a claim about the cardinality of the pluralities in a rich number, as in (18).

(17) a. Bo hired [the number of cooks Al hired].
b. $\exists x \in (16)[\text{hired}(Bo)(x)]$

(18) $[\![\,\text{is below ten}\,]\!](A) \Leftrightarrow \exists x[x \in A \wedge |x| < 10]$

The proposal in Scontras 2017 also suggests an internal composition for the DP. Within the

[2]In Scontras 2017, equivalences like (13) are the result of "derived kind predication", a composition principle adapted from Chierchia 1998, who builds on Carlson 1977. See Alonso-Ovalle and Schwarz 2024 for a critique.

[3]For the internal composition of (9), we assume partial reconstruction of the NP head of the relative clause, as in (i). Within the relative clause, the object DP may be headed by a covert existential determiner, or composition may proceed via (13). Either way, the relative clause denotes a set of numbers (ii-a), like *number* does (ii-b). The head noun and the relative clause can then compose intersectively. *The* picks up the largest number in the denotation of the main NP. The LF (i) is reminiscent of analyses of *how many NP* questions that decompose the wh-phrase into a wide scope quantifier ranging over cardinalities and a piece equivalent to the relative clause internal object DP in (i), which can be reconstructed. See Cresti 1995 and Rullmann 1995 and references therein.

(i) [$_{DP}$ the [$_{NP}$ [$_{NP}$ number] [$_{CP}$ wh$_n$ Al hired [$_{DP}$ n MANY (of) cooks]]]

(ii) a. $[\![[_{CP}\ \text{wh}_n\ [\text{Al hired}\ [_{DP}\ \text{n MANY}\ \underline{\text{(of) cooks}}]]]]\!] = \{n \mid \exists y[\text{cooks}(y) \wedge |y| = n \wedge \text{hired}(Al)(y)]\}$
b. $[\![\ \text{number}\ ]\!] = \{0,1,2,3\ldots\}$

main NP, the relative clause *Al hired* can be parsed as modifying the NP *number of cooks*, as detailed in (19). The NP *number of cooks* will denote the set of all rich numbers made of cooks, in (20-a). Each such rich number is a set that contains all cook pluralities of a given cardinality. The relative clause can likewise denote a set of rich numbers: those that contain a plurality that Al hired, (20-b). The NP *number of cooks* and the relative clause combine intersectively. That yields the set that contains all rich numbers made of cooks that contain at least one plurality that Al hired. If Al hired exactly three cooks, the highest NP then denotes a set that contains the set of all atomic cooks, the set of all two cook pluralities, and the set of all three cook pluralities. *The* picks up the last set, on the basis that it contains the largest pluralities.

(19) $[_{NP}\ [_{NP}\ [_{N}$ **number**] of cooks] $[_{CP}$ wh Al hired t]]

(20) a. $\{N \mid N \text{ is a rich number } \wedge\ \forall x \in N\ [\text{cooks}(x)]\}$
b. $\{N \mid N \text{ is a rich number } \wedge\ \exists x \in N\ [\text{hired}(\text{Al})(x)]\}$

A central feature of the rich number analysis of NDPs is that, unlike the rich syntax analysis, it does not posit any constituent that denotes a cardinality. In fact, we can understand the rich number analysis as the claim that there is no such constituent in NDPs, and perhaps more broadly in language.

3 Number transparency

We now consider cases of entity predication with NDPs in an intensional context, as in (5), repeated in (21). This sentence is ambiguous between readings that we call "opaque" and "number transparent". In the opaque reading, (21) reports on the relation desired by Bo between the numbers of cooks hired by Al and Bo. It conveys that Bo wants to match Al in terms of the number of cooks hired, that Bo wishes to do no worse than Al on that scale.

(21) Bo wants [PRO to hire [the number of cooks Al hired]].

We can also state this reading as in (22), with reference to the domain component in (23) and the cardinality component in (24), where, crucially, the meta-language predicates *cooks* and *hire* are applied to desire worlds, represented by *w*, as opposed to the actual world. Given this, (22) states that for each world w in the set of Bo's desire worlds, Acc, Bo hires at least m cooks in w, where m is the (maximal) number of entities that are cooks in w and that Al hires in w.

(22) $\text{Acc} \subseteq \{w \mid \exists x \in (23)[\text{hired}(w)(\text{Bo})(x)]\}$

(23) *Domain component:* $\{y \mid \text{cooks}(w)(y) \wedge |y| = (24)\}$

(24) *Cardinality component:* $\max\ \{n \mid \exists x \in \{z \mid \text{cooks}(w)(z) \wedge |z| = n\}[\text{hired}(w)(\text{Al})(x)]\}$

The "number transparent" reading of (21), in contrast, describes a desire of Bo that imposes a lower bound on the number of cooks Bo hires in absolute terms, not in comparison to Al. It states that Bo wants to hire some plurality of cooks whose cardinality is the number of cooks that Al happens to actually have hired. So under the assumption that the number of cooks that Al hired is three, (21) on that reading is equivalent to (25).

(25) Bo wants to hire three cooks.

The number transparent reading can also be stated as in (26), with reference to the domain component in (27) and the cardinality component in (28). In the cardinality term in (28), the meta-language predicates *cooks* and *hire* are now both evaluated in the actual world @. The truth conditions in (26) state that for every desire world w, Bo hires at least m cooks in w,

where m is the (maximal) number of cooks in the actual world that Al hired in the actual world.

(26) Acc $\subseteq$ {w| $\exists$x $\in$ (27)[hired(w)(Bo)(x)]}

(27) *Domain component:* {y| cooks(w)(y) $\wedge$ |y| = (28)}

(28) *Cardinality component:* max {n| $\exists$x $\in$ {z| cooks(@)(z) $\wedge$ |z| = n}[hired(@)(Al)(x)]}

The opaque and number transparent interpretation only differ in how the cardinality component is arrived at, in particular, in what worlds the metalanguage predicates *cooks* and *hired* take as arguments: either the desire worlds or the actual world. With this in mind, let us now discuss how the rich syntax analysis derives the opaque and number transparency readings.

3.1 Number transparency: rich syntax

On the rich syntax account, number transparent reading are unsurprising and compatible with various views of how transparent readings arise in general. For concreteness only, we will focus here on one familiar option. In this approach, the world arguments of predicates are syntactically represented, and a predicate's transparent interpretation can result from the assumption that its world argument is filled by a covert world pronoun @ that picks out the actual world (Percus 2000). Under this view, the embedded clause in (21) can be assigned the LF in (29).

(29) λw PRO to hire$_{w}$ [$_{DP}$ [$_{NP}$ [$_{DP}$ the number of cooks$_{@}$ Al hired$_{@}$]] [$_{N}$ cooks$_{w}$]]]

We have seen that the modifying DP in (29) contributes to the semantic composition the cardinality component of the truth conditions that arise from entity predication. In (29), the occurrences of *cook* and *hired* in this DP are indexed with the world pronoun *@* denoting the actual world. This yields the cardinality component in (28) and, in consequence, the number transparent reading of (21): with its embedded clause parsed as in (29), (21) conveys that Bo wants to hire some plurality of cooks that has the cardinality of the plurality of actual cooks that Al actually hired. So in a world where that cardinality is 3, (21) in the parse (29) conveys that Al wants to hire (at least) three cooks.

In contrast, in the parse in (30) all the word pronouns within the modifying DP are bound within the embedded clause. This yields the target domain cardinality in (27), and so (30) characterizes the set of worlds w such that Bo hires a cook plurality in w whose size is the (maximal) number of entities that are cooks in w and that Al hired in w. The intended opaque reading of (21) then falls out compositionally as the statement that all of Bo's desire worlds have this property.

(30) λw PRO to hire$_{w}$ [$_{DP}$ [$_{NP}$ [$_{DP}$ the number of cooks$_{w}$ Al hired$_{w}$]] [$_{N}$ cooks$_{w}$]]]

As long as predicates in intensional contexts can be evaluated in the actual world, then, the rich syntax approach captures the number transparent reading with no further stipulations. The modifier DP contributes the cardinality component, and number transparency arises if that component is evaluated in the actual world. We will see next that, in contrast, number transparency poses a challenge to the rich number approach.

3.2 Number transparency: rich numbers?

Under the rich number analysis, the opaque reading arises from the parse of the embedded clause in (31). With both *cooks* and *hired* evaluated in the desire worlds, the NDP in (31) contributes the domain component in (32), which references the cardinality component (33). The proposition expressed by (31) is true in w just in case in w Bo hires at least one member of

the set (32). (31) is accordingly equivalent to (30) as interpreted in the rich syntax analysis. And so the opaque reading is again derived as intended, as the statement that the set of worlds denoted by (31) includes all of Bo's desire worlds.

(31) λw PRO to hire$_{\text{w}}$ [$_{\text{DP}}$ the [$_{\text{NP}}$ number of cooks$_{\text{w}}$ Al hired$_{\text{w}}$]]

(32) *Domain component:* $\{\text{y}|\ \text{cooks(w)(y)} \wedge |\text{y}| = (33)\}$

(33) *Cardinality component:* $\max\ \{\text{n}|\ \exists \text{x} \in \{\text{z}|\ \text{cooks(w)(z)} \wedge |\text{z}| = \text{n}\}[\text{hired(w)(Al)(x)}]\}$

But the number transparent interpretation poses a challenge. Crucially, there is no available constituent that contributes the cardinality component by itself. To confirm that this does not leave a compositional path to the opaque reading, we observe that no choice of world arguments for the available predicates delivers the target meaning. With the lower *hire* evaluated in @ and the higher *hire* in w, just as in (29), only two options remain: the sole occurrence of *cooks* can be evaluated in w or in @, resulting in the structure in (34-a) and (34-b), respectively.

(34) a. λw PRO to hire$_{\text{w}}$ [$_{\text{DP}}$ the [$_{\text{NP}}$ number of cooks$_{\text{w}}$ Al hired$_{@}$]]
b. λw PRO to hire$_{\text{w}}$ [$_{\text{DP}}$ the [$_{\text{NP}}$ number of cooks$_{@}$ Al hired$_{@}$]]

The parse in (34-a) derives the meaning given in (35) to (37). This reading differs from the number transparent reading only in the cardinality component. The cardinality component of the number transparent reading is given in (38). When compared to it, the predicted cardinality reading in (37) is "under-transparent", as it applies *cooks* to the desire world w, and, therefore, the lower bound on the number of cooks that Bo wants to hire is not determined by a set of actual cooks, as in the target reading.

(35) $\text{Acc} \subseteq \{\text{w}|\ \exists \text{x} \in (36)\ [\text{hire(w)(Bo)(x)}]\}$

(36) *Domain component:* $\{\text{y}|\ \text{cooks(w)(y)} \wedge |\text{y}| = (37)\}$

(37) *Cardinality component:* $\max\ \{\text{n}|\ \exists \text{x} \in \{\text{z}|\ \mathbf{\text{cooks(w)}}(\text{z}) \wedge |\text{z}| = \text{n}\}[\text{hired}(\mathbf{@})(\text{Al})(\text{x})]\}$

(38) *Cardinality component of number transparent reading:*
$\max\ \{\text{n}|\ \exists \text{x} \in \{\text{z}|\ \mathbf{\text{cooks}(@)}(\text{z}) \wedge |\text{z}| = \text{n}\}[\text{hired}(\mathbf{@})(\text{Al})(\text{x})]\}$

On the other hand, the parse in (34-b) derives the meaning given in (39) to (41). This time, the choice of world arguments yields the cardinality component of the number transparent reading, but a different domain component. As we can see by comparing the predicted domain component in (40) to the target domain component of the number transparent interpretation in (42), the predicted domain component applies *cooks* to the actual world instead of w. So the overall meaning is "over-transparent": it requires Bo's desires to be about actual cooks only.

(39) $\text{Acc} \subseteq \{\text{w}|\ \exists \text{x} \in (40)[\text{hire(w)(Bo)(x)}]\}$

(40) *Domain component:* $\{\text{y}|\ \mathbf{\text{cooks}(@)}(\text{y}) \wedge |\text{y}| = (41)\}$

(41) *Cardinality component:* $\max\ \{\text{n}|\ \exists \text{x} \in \{\text{z}|\ \text{cooks}(@)(\text{z}) \wedge |\text{z}| = \text{n}\}[\text{hired}(@)(\text{Al})(\text{x})]\}$

(42) *Domain component of the number transparent reading:* $\{\text{y}|\ \mathbf{\text{cooks(w)}}(\text{y}) \wedge |\text{y}| = (41)\}$

4 Conclusion

Under the rich syntax analysis, entity predication with a NDP invokes an interpreted structure with a cardinality-denoting DP that modifies the NP. Number transparency diagnoses the presence of this DP in the interpreted structure. The next step is to clarify the syntactic principles that can yield this structure, including the covert duplication of the entity noun.

Acknowledgements. We gratefully acknowledge financial support from the Social Sciences and Humanities Research Council of Canada (Insight Grants 435-2018- 0524, and 435-2023-0146, PI: Alonso-Ovalle; and 435-2019-0143, PI: Schwarz). Our names are in alphabetical order.

References

Alonso-Ovalle, Luis and Bernhard Schwarz (2024). "Structural Ambiguity in Pseudo-Partitives: the Case of Quantity Nouns". In: *Proceedings of the Thirthieth-Third Semantics and Linguistic Theory Conference*. Ed. by Juhyae Kim et al. Cornell University: CLC Publications, pp. 603–621. DOI: https://doi.org/10.3765/dxn8d863.

Bale, Alan (2006). "The Universal Scale and the Semantics of Comparison". PhD thesis. McGill University.

Carlson, Gregory (1977). "A Unified Analysis of the English Bare Plural". In: *Linguistics and Philosophy* 1, pp. 413–457. DOI: https://doi.org/10.1007/BF00353456.

Chierchia, Gennaro (1998). "Reference to Kinds across Languages". In: *Natural Language Semantics* 6.4, pp. 339–405. DOI: https://doi.org/10.1023/A:1008324218506.

Cresswell, Max J (1976). "The Semantics of Degree". In: *Montague Grammar*. Ed. by Barbara H Partee. New York: Academic Press, pp. 261–292.

Cresti, Diana (1995). "Extraction and Reconstruction". In: *Natural Language Semantics* 3, pp. 79–122. DOI: https://doi.org/10.1007/bf01252885.

Dummett, Michael (1991). *Frege. Philosophy of Mathematics*. NY: Duckworth.

Hackl, Martin (2000). "Comparative Quantifiers". PhD thesis. Massachusetts Institute of Technology.

Hatcher, William S. (1990). *The Logical Foundations of Mathematics*. NY: Pergamin Press.

Klein, Ewan (1980). "A Semantics for Positive and Comparative Adjectives." In: *Linguistics and Philosophy* 4, pp. 1–45. DOI: https://doi.org/10.1007/BF00351812.

Moltmann, Friederike (2013). "Reference to Numbers in Natural Language". In: *Philosophical Studies* 162, pp. 499–536. DOI: https://doi.org/10.1007/s11098-011-9779-1.

Percus, Orin (2000). "Constraints on Some Other Variables in Syntax". In: *Natural Language Semantics* 8, pp. 173–229. DOI: https://doi.org/10.1023/A:1011298526791.

Rullmann, Hotze (1995). "Maximality in the Semantics of Wh-Questions". PhD thesis. Amherst, MA: University of Massachusetts, Amherst.

Scontras, Gregory (2017). "A New Kind of Degree". In: *Linguistics and Philosophy* 40, pp. 165–205. DOI: 10.1007/s10988-016-9200-9.

Facts, intentions, questions: English 'come-to-know' predicates in deliberative environments

Pranav Anand
UC Santa Cruz
Santa Cruz, CA, U.S.
panand@ucsc.edu

Natasha Korotkova
Utrecht University
Utrecht, The Netherlands
n.korotkova@ucla.edu

Abstract

This paper explores a distributional interaction in English between (i) deliberative environments, e.g., *intend*; (ii) 'come-to-know' verbs (CtKs): factive change-of-belief verbs, e.g., *discover* and *find out*; and (iii) complement selection. We show that deliberative environments permit CtKs only when they embed questions. We derive this constraint from a clash between two presuppositions of CtKs, their factivity and change-of-state requirements, that emerges due to the projection properties of deliberative environments. We show that when the factivity presupposition is either absent (as with non-factive change-of-belief predicates) or accommodated locally, CtKs in deliberative environments permit declarative complements.

1 Introduction

This paper investigates patterns of complement selection for 'come-to-know' verbs (CtKs), factive change-of-belief verbs such as *discover*, *find out*, *learn*, *notice*, and *realize* (see Korotkova and Anand 2024 for detailed discussion). All CtKs are responsive predicates (Rawlins and White 2019; Spector and Egré 2015): they take declarative and interrogative complements (1).

(1) Kira **discovered** / **found out** / **learned** {✓that Saturn has 95 moons, ✓which planet has 95 moons}.

However, CtKs are unable to embed declaratives in what we call 'deliberative environments' (following Cariani, M. Kaufmann, and S. Kaufmann 2013; Condoravdi and Lauer 2017). These are enviroments where a rational agent engages in selection of an action plan from a set of options, such as: vanilla attitude verbs of intention (*intend*, *try*), rationale clauses, volitional adverbs (*intentionally)*, desiderative predicates in anakastic conditionals, directive and commissive predicates (*command*, *order*, *promise*, *threaten*), and imperatives. Across these constructions, CtKs are felicitous when they embed an interrogative, but not a declarative:

(2) In his Stanford experiment, Zimbardo **intended** / **tried to discover** / **find out** / **learn** / **verify** {#that/✓how quickly} people conformed to the roles of guard and prisoner.

(3) You call 25 of your friends at random **in order to find out** / **learn** / **discover** / **verify** {#that/✓if} Psarras is the best restaurant in town.

(4) The detective **intentionally found out** / **learned** / **verified** {#that, ✓whether} the actor was sick.

(5) **If you want to find out** / **learn** / **discover** / **verify** {#that, ✓whether} it's humid, you should get a digital barometer.

(6) The doctor commanded their subordinate to **find out** / **learn** / **discover** / **verify** {#that, ✓whether} the patient was well.

This gives rise to the generalization in (7), which we will call DiD for short:

(7) **Declarative-in-Deliberative Prohibition (DiD)**
CtKs in a deliberative environment do not allow a declarative embedding.

Within responsive attitudes, this restriction is true only for CtKs. It does not hold for responsives like *say* (8), stative factives (9), or non-factive change-of-belief predicates such as non-veridical *conclude* or veridical *prove* (10). It is thus difficult to localize DiD to the change-of-state, doxastic, or factive components of CtKs.

(8) Kira **intended** to **say** {✓that Saturn has 95 moons, ✓which planet has 95 moons}.

(9) You need to test this hypothesis **in order to know** {✓that/✓if} the project is viable.

(10) Three criteria need to be met **in order to conclude** / **prove** {✓that/✓if} α causes β.

Furthermore, DiD is restricted to deliberative environments. Other future-oriented environments, including bouletic predicates (when not used deliberatively), admit CTKs with declaratives (11).

(11) Odo **hopes** / **wishes** / **is afraid** / **expects** to **learn** {✓that/✓if} there is life on Mars.

To sum up, declarative embedding is disallowed precisely when the embedder is a coming-to-know predicate in a deliberative environment. We argue that this restriction arises because of a conflict of two presuppositions that CtKs trigger: (i) the factivity of their declarative complements and (ii) a lack of belief in the prejacent by CtK experiencer prior to the change-of-doxastic-state. When these project across deliberative environments, they produce a clash. For reasons of space, we focus on intention verbs below, and assume that other deliberative environments behave the same way in aspects relevant to us.

2 Proposal

Coming-to-know eventualities involve belief change and knowledge. A deliberative agent does not have complete control over whether someone's beliefs will change or whether those beliefs are accurate. It is thus tempting to pin DiD on this lack of complete control, since deliberative predicates like *try* are known to be infelicitous with complements denoting eventualities the agent cannot control (e.g., the oddity of *?I tried to accidentally stumble*). But the data above shows things are not straightforward: we cannot categorically ban CtKs in deliberative environments, since they are acceptable when they embed interrogatives. Nor can we say containing doxastic change renders a predicate insufficiently controllable, since coming-to-believe predicates with declaratives are acceptable in deliberative environments.

It is instructive here to consider Jenkins (2019), who aims at providing a new theory of *try* and along the way puts forth the only discussion of something like DiD we are aware of. Capitalizing on previous work, Jenkins uses two cases of infelicity with doxastic change-of-state predicates to motivate his proposal:

(12) a. #Mary tried to realize that the oven was on. (Jenkins 2019)
b. #Mary tried to notice a picture. (Sharvit 2003)

Jenkins' semantics for *try* is couched in terms of expected utility (EU) of the prejacent ϕ. It assumes (i) a decision model that provides a subjective probability measure Pr_α over the set of doxastic alternatives A for deliberative agent α; and (ii) a set of possible actions, or choices $\mathcal{C} = \{\Delta_1, \ldots \Delta_n\}$ for α. Relative to these, α *try* ϕ is true only if α selects some Δ from $\mathcal{C}$ that increases the chances of bringing the prejacent about $Pr_\alpha(\phi|\Delta) > Pr_\alpha(\phi|\neg\Delta)$. In his discussion of the cases in (12), Jenkins assumes that *notice an x* is equivalent to *come to believe there is an x*, treating DP complements on a par with declaratives. He proposes that the oddity of these

cases arises because both Pr_{Mary} and the prejacents in (12) are defined over Mary's doxastic alternatives. His argument is then that no Δ could increase the probability of coming-to-believe the prejacent because A would have to already validate the prejacent.

Jenkins' discussion is phrased entirely in terms of belief, and so it directly predicts that come-to-believe predicates embedding declaratives should likewise be infelicitous in deliberative environments, contrary to the felicity we observed in (10) for *conclude* and *prove*. Nor does his account explain the felicity of CtKs with interrogatives (which he does not discuss).[1] However, we believe that Jenkins' underlying intuition is correct, once we see it as a product of factivity: the reason one cannot *try to figure out that p* is because, as a factive, *figure out* will project into the deliberative agent's doxastic state the requirement that p holds. And this, in turn, will clash with a change-of-doxastic-state requirement that *figure out* possesses. We turn to each of these requirements next.

2.1 CtK requirements: Factivity and lack of prior belief

CtKs are poster child cognitive factives (Degen and Tonhauser 2022; Djärv 2019; Karttunen 2016). Here, we will assume that they presuppose that their declarative complements are true in the evaluation world (13), a condition we gloss as Factivity:

(13) **Factivity**:
$CtK(x, p, t, w)$ is defined only if $p(w) = 1$.

Unlike stative factives, CtKs encode belief formation, a transition from not having a belief that p (unsettledness about p or belief in $\neg p$) to having it (Korotkova and Anand 2024; Rawlins and White 2019), akin to conditions for change-of-state verbs at large (Karttunen 1973).

(14) **Lack of Belief**:
$CtK(x, p, t, w)$ is defined iff $\forall t' < t[\neg \textsc{believe}(x, p, t', w) = 1]$,
where $\textsc{believe}(x, p, t, w) = 1$ iff $\forall w'' \in DOX(x, t, w)[p(w'') = 1]$.

CtKs are thus incompatible with a prior belief that p (15):

(15) # The children were sure that I was asleep, and then learned from their mother that I was asleep.

2.2 Presupposition projection in deliberative environments

Deliberative environments, like bouletics, invoke preferences. But they also share properties with doxastics. First, they require **realism**, the condition that the prejacent be epistemically possible for the deliberative agent, preventing an intention to do something impossible. Second, they demand **consistency** among the commitments of the deliberative agent, preventing commitment to incompatible propositions (Condoravdi and Lauer 2016; Grano 2017; Grano 2023; Grosz 2014; Hansson 2001):

(16) a. I {#intend, ✓want} to be a robot. REALISM
b. J. intends to move in w/ his girlfriend, #but also intends to keep living alone. (Condoravdi and Lauer 2016) CONSISTENCY

[1] In fact, unlike other CTKs, *notice* and *realize* are never acceptable in deliberative environments regardless of the complement shape (Korotkova and Anand 2024), cf. *#Mary tried to notice/realize whether the oven was on.* These verbs (but not CTKs at large) describe non-controllable actions and therefore are banned in deliberative environments altogether independetly of DiD.

Realism in turn yields a presupposition projection pattern for deliberative environments distinct from that of preference-oriented attitudes like bouletics (cf. Grano 2017; Sharvit 2003). In particular, like doxastics, it requires that the deliberative agent's doxastic state universally satisfies the presuppositions of the complement clause, as can be observed with the existence presupposition triggered by a possessive definite:

(17) Neo, unknowingly a brain in a vat, **planned** to raise **his arm**, #but believed he had none. (adapted from Grano 2017)

For our purposes, the descriptive characterization of this projection pattern in (18) is sufficient:

(18) **Deliberation Projection**:
If p presupposes q, $\textsc{intend}(x, p, t, w)$ is defined only if $\textsc{believe}(x, q, t, w)$.

2.3 Deriving DiD

The crux of our analysis is that when a declarative-embedding CtK is inside a deliberative environment, Deliberation Projection pits Factivity against Lack of Belief. Ordinarily, the two do not interact: Factivity references the evaluation world w and Lack of Belief references the deliberative agent's doxastic state in w at t. But when a CtK is within a deliberative environment, Deliberation Projection means that Factivity becomes a constraint of the deliberative agent's doxastic state, and this yields a temporal contradiction: (a) Factivity requires the deliberative agent to believe the CtK complement at deliberation time, (b) Lack of Belief requires no belief at deliberation time. Thus, like for Jenkins, this proposal involves a clash between belief in p and a lack of belief in p, but it is crucially mediated by factivity. Correspondingly, non-factive attitudes, including come-to-believe ones, correctly allow declarative complements in deliberative environments. So, too, do CtKs when they embed interrogatives: as we show in Section 3, no particular proposition projects into the deliberative agent's doxastic state, and hence nothing clashes with Lack of Belief. In what follows, we make several choices in order to provide a concrete proposal, but our account is compatible with many existing theories of deliberative environments and various approaches to responsive predicates and question embedding.

Deliberative attitudes involve selecting a course of action. This can be cashed out through filtering by a decision problem ordering (Cariani, M. Kaufmann, and S. Kaufmann 2013), an effective preference pre-order (Condoravdi and Lauer 2016), ordering by continuation branches of an initial event-stage (Sharvit 2003), continuation of an event-stage in intention worlds (Grano 2017), and computation of expected utility (Jenkins 2019). For simplicity, we will assume that *intend* quantifies over $INT(x, t, w)$, optimal worlds drawn from $DOX(x, t, w)$ and ordered by propositions representing the intentions linked to the deliberative event, predicating of those worlds that the complement property is true at some interval t'' after intention time t.

(19) $\textsc{intend}(x, P, t, w) = 1$ iff $\forall w' \in INT(x, t, w)\ \exists t'' > t\ [P(t'')(w') = 1]$

Consider now (20), where *intend* embeds the subject-controlled CtK *learn*; each $w'' \in INT$ thus has a future t'' where Kira learns p.

(20) #[Kira intends [to learn [that Saturn has 95 moons]]].
$\textsc{intend}(\text{Kira},\ \lambda t' \lambda w' \textsc{learn}(\text{Kira}, p, t', w'),\ t, w) = 1$
iff $\forall w'' \in INT(\text{Kira}, t, w)\ \exists t'' > t\ [\textsc{learn}(\text{Kira}, p, t'', w'') = 1]$

a. projected Lack of Belief:
$\forall w'' \in DOX(\text{Kira}, t, w)\ \exists t'' > t\ \forall t''' < t''\ [\neg \textsc{believe}(\text{Kira}, p, t''', w'') = 1]$

b. projected Factivity:
$\forall w'' \in DOX(\text{Kira}, t, w)\ p(w'') = 1$, or $\textsc{believe}(\text{Kira}, p, t, w) = 1$

Lack of Belief requires that prior to t'', Kira does not believe p, and by Deliberation Projection, this projects as a general belief by Kira about her own lack of belief stretching to variable points in her future (20a). Factivity projects to the requirement that Kira believes p at intention time t (20b). This is a temporal contradiction: it cannot be the case that Kira believes p at t and that, relative to a t'' after t, Kira never believed p.

The contradiction crucially only arises because of Deliberation Projection, which forces Factivity to project universally over all of $DOX(x, t, w)$, ensuring that x believes p at t. Were the factivity requirement to project to the INT worlds (or accommodate to them), we would merely require that the most ideal worlds are ones where the learned proposition were true. As this would only be a requirement for the ideal worlds, it would not require Kira's doxastic state at t to be settled on p, allowing Lack of Belief to be satisfied. We suggest this is precisely what occurs when CtKs occur in bouletic environments like those in (11), where Odo's optimal bouletic worlds are ones where there is life on Mars and he comes to know that fact. Thus, while desiderative predicates are known to disprefer local accommodation (Heim 1992), they allow it, enabling the embedding of a CtK without running afoul of Lack of Belief.

Our account relies on the identity between the deliberative agent and the CtK believer (so that Lack of Belief and Factivity target the same individual's doxastic state). We thus correctly predict that distinct subjects will not be subject to DiD (21):

(21) The teacher intended for the students to learn that Saturn has 95 moons.

We also predict felicity in cases where agent is the believer but (i) read non-*de se* or (ii) where there is a lack of continuity of the doxastic state. Thus, for example, if the deliberative agent is knowingly forgetful, and is planning to overcome their own later forgetting of a fact, DiD should not hold (we are grateful to an anonymous referee for pointing out these cases). Such scenarios may be quite unlikely, but we are here inspired by the Christopher Nolan film *Memento*, in which an investigator with anterograde amnesia uses messages written to his later self to overcome the inevitable forgetting. In such a situation, declarative embedding CtKs are acceptable:

(22) *A detective suffering from anterograde amnesia permanently forgets new information about a case upon waking the next morning.*
The detective wrote "don't trust him" on a Polaroid of a man in order to later (re)discover that the man is working against him.

3 Intermediate accommodation and question embedding

Our account for DiD relies on a clash of projected Factivity and Lack of Belief, and predicts that infelicity will not arise in at least two cases. The first case is when the predicate downstairs is not factive (hence nothing contradicts Lack of Belief). This is precisely what happens with come-to-believe predicates as in (10). The felicity of veridical predicates like *prove* indicates that it is not the truth of p *per se* that matters, but truth throughout the doxastic state, which we argues arises due to the interaction of factivity and Deliberation Projection.

Our second case is when the predicate is factive but the factive presupposition is accommodated before it can project via Deliberation Projection. We argue now that this is what occurs when CtKs embed interrogatives. While CtKs are factive when they embed declaratives, like other factive responsives, when they embed interrogatives, the resulting structure does not project a factive presupposition. In what follows, we will appeal to the system in Spector and Egré (2015) For concreteness (though our account can be couched in inquisitive approaches to responsive predicates). This system assumes that declaratives denote propositions and interrogatives denote sets of propositional answers. Responsive predicates take propositional complements, which means to that they cannot combine directly with interrogatives. Instead, an existential

quantifier ranging over elements in the interrogative is introduced, and the responsive takes this existentially bound element as its propositional argument, as schematized in (23):

(23) $\text{CtK}(x, Q, t, w) \leftrightarrow \exists p \in Q\ \text{CtK}(x, p, t, w)$

CtKs trigger Factivity, but the factive presupposition projects into the existential's domain:

(24) ✓[Kira intends [to learn [if Saturn has 95 moons]]].
$\textsc{intend}(\text{Kira},\ \lambda t'\lambda w'\textsc{learn}(\text{Kira}, Q, t', w'),\ t, w) = 1$ iff $\forall w'' \in INT(\text{Kira}, t, w)$
$\exists t'' > t\ \exists p \in Q$ [**p(w")=1** $\wedge\ \textsc{learn}(\text{Kira}, p, t'', w'') = 1$].

The factivity thus does not project further, and hence is not subject to Deliberation Projection, which means there is no presupposition that Kira believes any p. We thus have that relative to the ideal worlds in INT: (i) some answer $p \in Q$ is the case; (ii) at some future t'', Kira believes p is true; and (iii) that prior to t'', Kira does not believe p. Because we are restricting ourselves to Kira's ideal worlds here, there is no contradiction, so we correctly predict that CtKs with interrogative complements are felicitous in deliberative environments.

Nothing in this explanation is unique to question-embedding, and we thus predict that CtKs with declarative complements should be felicitous if other operators intervene between the CtK and the intention predicate, preventing Deliberation Projection from interacting with factivity. This is indeed the case—epistemic modals (25), negation (26) and coordination (27) all serve to rescue declarative-embedding CtKs:

(25) The student **intended** to **possibly** learn from the video that there were other ways to solve the problem.

(26) By deleting his email, the patron **intended** to **never** discover that his library books were overdue.

(27) The gambler **intended** to win the lottery **and** learn that he had won.

Under our account, this is expected, since such operators, like interrogative embedding, create a space for accommodation that blocks the application of Deliberation Projection, and the generation of a presupposition of belief by the deliberative agent that clashes with Lack of Belief.

4 Conclusion

Deliberative environments have received heavy scrutiny over the past two decades, with much attention to how controllability, agentivity, and action-guidance can inform the restrictions on the prejacents of such environments. Focusing on the behavior of factive change-of-state predicates, we have argued that presuppositional content is also a factor in explaining these restrictions, in line with Theiler, Reolofsen, and Aloni (2019) on the role of presuppositions in complement selection. We have shown that deliberative environments interact with presupposed content in a manner distinct from bouletics, but we have not here rigorously derived the source of that distinct behavior. All deliberative environments require choice from a set of actions, and we believe it is this focus on circumstantially accessible actions that scrupulously enforces that all presuppositions are satisfied by the doxastic state. We hope—and indeed intend—to more fully develop this tentative proposal in the future.

Acknowledgements. This material arose from our work on CtK predicates. For feedback and discussion, we thank audiences in Bochum, Konstanz, Stanford and Stuttgart, as well as Cleo Condoravdi, Paul Kiparsky, Judith Tonhauser and Igor Yanovich.

References

Cariani, Fabrizio, Magdalena Kaufmann, and Stefan Kaufmann (2013). "Deliberative modality under epistemic uncertainty". In: *Linguistics and Philosophy* 36.3, pp. 225–259. DOI: 10.1007/s10988-013-9134-4.

Condoravdi, Cleo and Sven Lauer (2016). "Anakastic conditionals are just conditionals". In: *Semantics and Pragmatics* 9.8, pp. 1–69.

— (2017). "Conditional imperatives and endorsement". In: *Proceedings of the Forty-Seventh Annual Meeting of the North East Linguistic Society*. Ed. by Andrew Lamont and Katerina Tetzloff.

Degen, Judith and Judith Tonhauser (2022). "Are there factive predicates? An empirical investigation". In: *Language* 98.3, pp. 552–591. DOI: https://doi.org/10.1353/lan.0.0271.

Djärv, Kajsa (2019). "Factive and assertive attitude reports". PhD thesis. University of Pennsylvania. URL: https://repository.upenn.edu/edissertations/3645.

Grano, Thomas (2017). "Logic of intention reports". In: *Journal of Semantics* 34, pp. 587–632. DOI: 10.1093/jos/ffx010.

— (2023). "A unified analysis of rationale clauses". In: *NELS 53: Proceedings of the Fifty-Third Annual Meeting of the North East Linguistic Society*. Ed. by Suet-Ying Lam and Satoru Ozaki. Amherst, MA: GLSA, pp. 277–289.

Grosz, Patrick (2014). "Modal particles in rationale clauses and related constructions". In: *Modes of Modality*. Ed. by Elisabeth Leiss and Werner Abraham. Amsterdam / Philadelphia: John Benjamins, pp. 263–290.

Hansson, Sven Ove (2001). *The Structure of Values and Norms*. Oxford University Press.

Heim, Irene (1992). "Presupposition projection and the semantics of attitude verbs". In: *Journal of Semantics* 9, pp. 183–221.

Jenkins, Robin (2019). "Expected utility and intentional action: The semantics of try". In: *Proceedings of Sinn und Bedeutung 23*. Ed. by M.Teresa Espinal et al.

Karttunen, Lauri (1973). "Presuppositions of compound sentences". In: *Liguistic Inquiry* 4.2, pp. 169–193.

— (2016). "Presupposition: What went wrong?" In: *Semantics and Linguistic Theory 26*. Ed. by Mary Moroney et al., pp. 705–731.

Korotkova, Natasha and Pranav Anand (2024). "Today we learned: English 'come-to-know' attitudes". Ms., Utrecht University and UC Santa Cruz.

Rawlins, Kyle and Aaron White (2019). "Question agnosticism and change of state". In: *Proceedings of Sinn und Bedeutung 21*. Ed. by Robert Truswell et al., pp. 1325–1342.

Sharvit, Yael (2003). "Trying to be progressive: The extensionality of try". In: *Journal of Semantics* 20.4, pp. 403–445. DOI: 10.1093/jos/20.4.403.

Spector, Benjamin and Paul Egré (2015). "A uniform semantics for embedded interrogatives: an answer, not necessarily the answer". In: *Synthese* 192.6, pp. 1729–1784. DOI: https://doi.org/10.1007/s11229-015-0722-4.

Theiler, Nadine, Floris Reolofsen, and Maria Aloni (2019). "Picky predicates: why *believe* doesn't like interrogative complements, and other puzzles". In: *Natural Language Semantics* 27, pp. 95–134. DOI: https://doi.org/10.1007/s11050-019-09152-9.

How is presupposition accommodation constrained, again?

Ido Benbaji-Elhadad
Massachusetts Institute of Technology
Cambridge, Massachusetts
ibenbaji@mit.edu

Omri Doron
Massachusetts Institute of Technology
Cambridge, Massachusetts
omrid@mit.edu

Abstract

According to the *satisfaction view*, presuppositions are conditions on the discourse context which must be satisfied for the utterances carrying them to be felicitous. These conditions can either hold in the context *prior* to the utterance of a presuppositional sentence, or they may be added to the context *post-hoc*, via a process known as *presupposition accommodation*, after a presuppositional utterance has been introduced. Since Lewis (1979), it is known that controversial or surprising information resists accommodation. However, this apparent constraint does not seem to follow from the theory of discourse underpinning the satisfaction view. Building on a proposal by Doron and Wehbe (2022), we argue that the facts can be derived within that theory without stipulating any primitive constraints on accommodation, once we take into account independent considerations regarding the structure of discourse.

1 Introduction

What is the difference between presupposed and asserted content? According to an influential view originating with Stalnaker 1970, discourse is held relative to a context set c (defined in (1)), and while assertions are used to *update* this context set, presuppositions impose a condition that must be *satisfied* for the update to go through successfully.[1] This condition, termed *Stalnaker's Bridge* (henceforth, SB) in (2), demands that the presupposition be true in every world in the context set. We will call this approach the *satisfaction view* of presuppositions.

(1) $c = \{w : w$ is compatible with what discourse participants consider common knowledge$\}$.

(2) **Stalnaker's Bridge (SB):** Updating a context c with an utterance ϕ (i.e., $c + [\![\phi]\!]$) is only felicitous if c entails any presupposition of ϕ at the time the update is preformed.

This simple view is complicated by the availability of *presupposition accommodation*, a process by which an addressee revises their assumptions about the context c *after* a presuppositional sentence has been uttered, to avoid violations of SB, as demonstrated in (3). In principle, then, the constraint imposed by a presupposition p on c can be satisfied in two ways: p can either hold in c *before* c is updated by an utterance that presupposes p, or c can be adjusted "quietly and without a fuss" after the fact, to ensure that context update does not fail (von Fintel 2008: 137).

(3) **Context:** A and B have just met, and know nothing about each other.
A: Do you want to join us for lunch?
B: I have to pick up my sister from the airport.

The availability of accommodation obscures the empirical difference between presupposition and assertion, given that both accommodation and context update by the asserted content of an

[1] There is no consensus in the literature as to the nature of *assertion*. On the so-called *two-dimensional* view, meanings of sentences are composed of a pair of propositions – a presupposition and an assertion. On the other hand, on the so-called *trivalent* view, the notion of assertion is only well-defined for sentences whose presupposition is already satisfied by the context. We remain neutral with respect to this debate, and avoid relying on a notion of assertion as an independent component. We use the term *assertion* here to refer to the proposition that is used to update the context set after presupposition accommodation has taken place.

utterance (henceforth, *context update*) result in a modified context set. However, certain cases of accommodation seem to be blocked, setting accommodation apart from context update. First, Lewis 1979 observes that information which is controversial or surprising in some intuitive sense seems to resist accommodation. We call this the *controversiality observation*, and while the notion of controversiality involved is often left vague in the literature, the canonical examples in (4) serve to illustrate the point. Second, as first observed in class notes by Irene Heim (2015), an accommodated presupposition cannot be used to address the question under discussion (*QUD*), as demonstrated in (5), where a sentence presupposing that Mary has a sister can be used to address a question involving Mary's lunch plans, but not one inquiring whether she has siblings. We call this the *QUD observation*.

(4) a. Daughter to her father, who does not know she is engaged: O dad, guess what? #My fiancé is subleasing my apartment this summer.(adapted from von Fintel 2008)
b. Sorry I'm late, #I was parking my camel. (adapted from Kissine and Pantazi 2020)

(5) $\mathcal{Q}_1$: Can Mary join us for lunch? $\mathcal{Q}_2$: Does Mary have any siblings?
A$_1$: ✓She is meeting her sister. A$_2$: #She is meeting her sister.

These observations raise a challenge for the satisfaction view, articulated in the following passage from Szabó 2006, which refers to the controversiality observation in particular:[2]

> It is a fundamental intuition that surprising things are better said then presupposed. How can we account for this? I don't see anything in the framework that can do the job. At one point Stalnaker suggests otherwise – he points out that the hearer may not be willing to accommodate a presupposition simply on the grounds of recognizing that this is what the speaker expects. True enough. But the hearer may also not be willing to accept an assertion simply on the grounds of recognizing that this is what the speaker expects. So I don't think there is an asymmetry here.

The goal of this paper is to sketch a reductive account of the controversiality observation. In particular, Doron and Wehbe (2022) show that the QUD observation is just a byproduct of the order of evaluation of SB relative to an independent *informativity* constraint on utterances. We show that once we pair *that* position with some auxiliary assumptions about the structure of discourse, we can explain the controversiality observation as just a special case of the QUD observation. The approach is appealing, as it allows us to avoid stipulating any primitive constraints on accommodation *qua* accommodation in our theory of natural language pragmatics.

The paper is organized as follows. In section 2, we show how the QUD observation follows from independent considerations regarding the informativity of utterances. In section 3, we propose that the controversiality observation follows from the interaction of the same considerations and an independent property of discourse structure. The approach makes testable predictions, which are shown to be borne out in section 4. Section 5 concludes.

2 Deriving the QUD observation

Doron and Wehbe (2022), henceforth D&W, show that resistance to accommodation in (5) is a byproduct of the architecture of context update. They assume that (i) discourse is relativized to a model $\langle c, Q\rangle$, where Q is the salient QUD partitioning c, as defined in (6) (Lewis 1988; Roberts 1996/2012 a.m.o.); and further that (ii) the *informativity* constraint in (7) constrains the set of admissible assertions by requiring that any utterance proposed as an update on c rule out at least one answer to Q (following, e.g., Groenendijk and Stokhof 1984; Roberts 1996/2012).

[2] As far as we know, this kind of challenge was first leveled against the satisfaction view in Gauker 1998.

(6) In a model $\langle c, Q\rangle$, Q is a partition over c; i.e. a set $\{q_1, ..., q_n\}$ s.t. (a) $Q = \{q_1\} \cup \{q_2\} \cup ... \cup \{q_n\}$; (b) $c = q_1 \cup q_2 \cup ... \cup q_n$; (c) $\forall i \leq n\ [q_i \neq \emptyset]$; and (d) $\forall i, j \leq n\ [i \neq j \rightarrow q_i \cap q_j = \emptyset]$.

(7) **Informativity:** ϕ is felicitous given a model $\langle c, Q\rangle$ only if $\exists Q' \subsetneq Q : [\![\phi]\!] \cap c \subseteq \bigcup Q'$.

D&W show that ordering the evaluation of informativity (7) *after* the evaluation of SB (2), which triggers accommodation of presuppositions not already entailed by the context c, reduces Heim's observation in (5) to a violation of informativity; i.e., that accommodation *bleeds* informativity. Assuming that to be asked felicitously, Q_2 in (5) must partition c in a non-trivial way, c cannot entail that Mary has a sister prior to the utterance of A_2. Hence, accommodation is required. But accommodating that Mary has a sister resolves Q_2 by ruling out its negative answer. The result is a partition consisting of the single cell compatible with the presupposition. And with respect to a single-celled partition, informativity (7), which requires an assertion to entail a subset of cells in the QUD while eliminating at least one, can never be satisfied.

A novel argument for D&W's account of (5) comes from the distinction between *full* and *partial* resolution of a question Q. A proposition fully resolves Q if it is only compatible with one of its cells, and partially resolves it if it is compatible with more than one but not with all of them. Given that polar questions like those in (5), when resolved, are always resolved in full, when such questions are resolved by an accommodated presupposition, the result always involves a violation of informativity. Yet when proposition p *partially* resolves Q, utterance ϕ presupposing p (henceforth, ϕ_p) can still in principle comply with informativity after p is accommodated. In such cases, accommodation does not rule out all but one of Q's cells, so if ϕ_p entails a proper subset of the cells remaining after accommodation, it is nonetheless informative. Accommodation should be allowed then, when the answer's presupposition *partially* resolves Q, as long as the answer *asserts* a proposition that satisfies informativity relative to Q.

The discourses in (8)-(9) show that this is indeed the case. Take (9), for instance, where the presupposition of B's utterance with *only*, namely, that John came, *partially* resolves A's question; i.e., it rules out those cells of A's question in which John did not come to the party. *Full* resolution is achieved only via B's assertion, which further conveys that no one other than John came to the party, and the resulting discourse is felicitous. The felicity of (8)-(9) is predicted on D&W's theory, further supporting their attempt to derive the QUD observation from a bleeding relation between accommodation and informativity, rather than from constraints on accommodation *qua* accommodation.[3]

(8) A: How many siblings do you have? B: Apart from my brother, I have two sisters.

(9) A: Who came to the party? B: Only John came.

3 Deriving the controversiality observation

The controversiality observation is described in the literature as a prohibition on accommodating controversial or surprising information. We propose that like the QUD observation, it is a by-product of independent constraints, and does not require any ad-hoc assumptions about accommodation *per se*. Moreover, we will show that the empirical picture has been mischaracterized; crucially, the differentiating factor which makes examples like (4) infelicitous does not pertain to the presupposition alone, but to the relation between the presupposition and the assertion, or more precisely, between the QUDs that they can be used to address.

The example of the controversiality observation in (4-a) conveys two things. That the speaker got engaged, as conveyed by the presupposition introduced by the DP "*my fiancé*", and that the speaker is subleasing her apartment this summer, as conveyed by the asserted content of

[3]Thank you to Mitya Privoznov for pointing this out to us.

(4-a). Notice that D&W's account of the QUD observation independently predicts (4-a) to be infelicitous if uttered relative to a QUD that inquires whether the speaker is engaged. Relative to such a question, (4-a) necessarily involves a violation of informativity (7), as discussed in Section 2 above. In principle, then, we can extend D&W's account of the QUD observation to the controversiality observation, as long as our model of discourse is designed in a way that *forces* (4-a) to be uttered relative to the question *Is the speaker engaged?*, which is fully resolved by its presupposition.

To achieve this desideratum, the following assumptions suffice. First, with (e.g.,) Farkas and Bruce (2010), we assume that discourse is relativized not just to a particular QUD, but to a *set* of questions that are *on the table*, i.e., that interlocutors aim to resolve, and that at each point in discourse the current QUD is chosen from this set and replaced with another once resolved. Second, we assume that the process of QUD-selection at any given stage of discourse is not free, but rather constrained in systematic way; among the QUDs to which the speaker can provide an answer, they must address the question which is most *urgent* by a certain criterion. Finally, we assume that of the questions in (10), (10-a) is generally ranked more urgent than (10-b) when both are part of the conversational question set. While the notion of *urgency* involved will remain somewhat of a black box throughout this paper,[4] these assumptions will provide us with testable predictions that extend beyond the realm of presupposition accommodation.

(10) a. *Is the speaker engaged?*
b. *Is the speaker subleasing her apartment this summer?*

Here is how these assumptions account for the infelicity of (4-a) in particular. We assume that the question in (10-a) is generally always on the table in a conversation between parent and child, and that the question in (10-b) is on the table in the context of (4-a), given that the daughter felt the need to convey her plans to sublease the apartment in her assertion. Since the speaker knows the answer to both (as indicated by her utterance), she is expected to share them with her father. The question now is in what order she should do so. Given that, by stipulation, (10-a) is ranked higher than (10-b), she should address the engagement issue first. The sentence in (4-a) therefore necessarily violates informativity, given that it is uttered mandatorily relative to a QUD which it fully resolves with its presupposition.

Note that the picture we have presented thus far seems to predict that any conversation between parent and child must begin by addressing the question of whether the child is engaged. This is obviously wrong. If the daughter in our example were not engaged, she could have gone on to talk about her apartment without any mention of engagement. In fact, it would have been quite odd to begin a conversation with the statement "I am not engaged!" The daughter only has to address the engagement question first if she did get engaged.

This demands that we be slightly more explicit about the ranking presumed to hold among questions on the discourse table. Specifically, we need to have the speaker's epistemic state inform how questions are ranked. An intuitive way to do so would be to require the speaker at any given point in discourse to first address those questions for which the answer she can provide would be most urgent to her addressee (with *urgent* remaining an intuitive primitive).

We can now be more explicit about how the infelicity comes about in (4-b), another well-known example of the controversiality observation. There, the utterance *I was parking my camel* again conveys two pieces of information; one, that the speaker owns a camel as a mode of transportation, is conveyed by the presupposition, and the other, that the speaker parked her mode of transportation, is conveyed by the assertion. Intuitively, the latter piece of information seems to address the question *Why is the speaker late?* and the former — *What is the speaker's mode of transportation?* Without the speaker's epistemic state contributing to the urgency

[4] In ongoing work, we attempt to cash it out with information-theoretic tools.

ranking of open questions, it would be unclear why a question about the speaker's mode of transport should be ranked higher than one about the reasons for her delay. In fact, had the speaker said *Sorry I'm late, I was parking my car*, no infelicity would have arisen, even when it is not taken for granted that the speaker owns a car. Thus, it is precisely the speaker's knowledge that her mode of transportation is a camel, which is an intuitively noteworthy piece of information, that elevates the latter question relative to the former, again resulting in a violation of informativity when this question is resolved by the presupposition as in (4-b).

4 Predictions

The notion of urgency is left vague, and indeed, it is perhaps reasonable to assume that like other contextual standards, it varies greatly among interlocutors. However, it is different from the intuitive notion of controversiality in an important respect: It is not a property of a single proposition, but a rather a relation between questions. Furthermore, despite its vagueness, the proposed analysis diverges quite clearly from previous literature in what it takes to be the source of infelicity in examples of the controversiality observation: It is not about illicit accommodation, but about the order in which information is introduced into the discourse.

This, in turn, has two empirical predictions. First, we predict that the same infelicity should arise even in discourses that contain no presuppositional utterances, or only presuppositional utterances that do not require accommodation, as long as asserted information is conveyed in an order that does not comply with what the urgency ranking on open questions dictates. Second, even when presuppositional utterances are involved, we predict that this infelicity should be dependent on the *relative urgency* of the asserted and presupposed information; a sentence should be acceptable whenever its assertion is judged more urgent than its presupposition, and unacceptable otherwise.

We show that the first prediction is borne out in (11)-(12). The original examples of the controversiality observation are repeated in the (a)-examples. Then, we repackage the information conveyed together by the presupposition and assertion in the (a)-cases into a sequence of presupposition-less utterances in the (b)-examples.[5] Crucially, however, the intuitively less urgent information is conveyed earlier in each sequence than the more urgent information.

(11) **Context:** A daughter to her father, who knows neither that she is engaged, nor that she is subleasing her apartment for the summer, and is interested in both:
   a. O dad, guess what? #My fiancé is subleasing my apartment this summer.
   b. O dad, guess what? I'm subleasing my apartment this summer... #I got engaged to the subleaser.

(12) a. Sorry I'm late, #I was parking my camel.
   b. Sorry I'm late, I was parking my vehicle. #It's a camel.

That the infelicity persists in the (b)-sentences seems to suggest that the role of accommodation in the (a)-sentences is purely accidental. The proposal we delineate above predicts the infelicity of, say, (12-b), where the second clause betrays that the speaker takes to be true a proposition that renders the question about her mode of transportation urgent, and thus her choice to first answer the less urgent question regarding the reasons for her delay flouts constraints on QUD-ordering.[6] The same explanation extends to (11).

[5] Or, rather, into a sequence of utterances such that each either lacks a presupposition altogether, or its presupposition is already taken for granted by the time it is uttered. This way, no accommodation is required.

[6] Note that there are also felicitous variants of the (b)-examples, formed by reordering the provided information in a way that complies with our urgency ranking. For instance, *Sorry I'm late. I use a camel to get around. I was parking it* seems to be an acceptable way to convey the information conveyed in (12-b). This is again predicted,

We turn to the second consequence of our analysis, namely that the governing factor in the controversiality observation is the relative urgency of the presupposition and the assertion. We expect to find two sentences which bear the same presupposition but contrast in acceptability in a given context in virtue of their assertion. An example of such a contrast is given in (13).

(13) **Context:** A daughter to her father, who knows neither that she is engaged, nor that she is subleasing her apartment for the summer, and is interested in both:

a. O dad, guess what? I got engaged to the person subleasing my apartment this summer.

b. O dad, guess what? ??I just had a really nice chat with the person subleasing my apartment this summer.

Both (13-a) and (13-b) presuppose that the daughter is subleasing her apartment for the summer. By stipulation, this is not common ground, and should thus be accommodated by the father. The difference lies in the asserted content of each sentence: In (13-a), the assertion addresses a QUD relating to the speaker getting engaged. As we have seen, given that the speaker got engaged, this kind of QUD is generally more urgent than the one relating to the subleasing of the speaker's apartment. For that reason, the sentence in (13-a) does not violate any ordering constraint, and is indeed judged felicitous. The sentence in (13-b), on the other hand, asserts information that corresponds to a QUD along the lines of *What were you doing before I called?* Intuitively, this QUD is less urgent than the one relating to the subleasing of the speaker's apartment. We thus correctly predict that (13-b) be infelicitous due to a violation of QUD-ordering.[7]

Our approach compares favorably to the alternative, usually at least implicitly assumed in the literature on accommodation, according to which controversiality restricts accommodation directly, via a primitive constraint posited in our theory of accommodation. This approach overgenerates the infelicitous examples in (11)-(12), as it views the infelicity of the (a)-examples as contingent on the presence of a presupposition that needs to be accommodated, thus failing to predict the infelicity of the (b)-examples which do not require accommodation. It also undergenerates the felicitous examples in (13), where the infelicity that arises from the presuppositional case disappears once we control for the level of urgency of what is conveyed by the assertion.

5 Conclusion

In this paper, we attempt to answer the question: Where do apparent constraints on accommodation come from? We promote the view that accommodation is not restricted directly, and is in fact as automatic and transparent as context update. Instead, when accommodation seems to be blocked, this reflects a violation of some independent constraint further down the line. We discuss a notable case of accommodation failure, the *controversiality observation*, and show that it can be explained as stemming from the interaction of constraints on informativity and on discourse structure. A central notion in our analysis, *urgency* – i.e., the relation between QUDs that requires a speaker to address one before the other – is left mostly unanalyzed throughout the paper. However, we show that regardless of the specifics, using such a notion allows us to better capture the empirical landscape. We thus conclude that the apparent failure to accommodate controversial information pertains directly neither to controversiality nor to accommodation.

Acknowledgements. Thanks to Kai von Fintel, Danny Fox, Benjamin Spector and Jad Wehbe for helpful comments and discussions.

as in this case, the more urgent question is addressed before the less urgent one.

[7] We use ?? rather than # to indicate that the infelicity of the (13-b) is not as stark as that of the canonical examples. This is expected given that the higher urgency of the presupposition here is not as intuitively clear.

References

Doron, Omri and Jad Wehbe (2022). "A constraint on presupposition accommodation". In: *Proceedings of the 23rd Amsterdam Colloquium*. Vol. 23.

Farkas, Donka F and Kim B Bruce (2010). "On reacting to assertions and polar questions". In: *Journal of semantics* 27.1, pp. 81–118.

von Fintel, Kai (2008). "What is presupposition accommodation, again?" In: *Philosophical perspectives* 22, pp. 137–170.

Gauker, Christopher (1998). "What is a context of utterance?" In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 91.2, pp. 149–172.

Groenendijk, Jeroen Antonius Gerardus and Martin Johan Bastiaan Stokhof (1984). "Studies on the Semantics of Questions and the Pragmatics of Answers". PhD thesis. Univ. Amsterdam.

Heim, Irene (2015). *Presupposition. Class notes, 24.954 (MIT, Fox/Heim)*.

Kissine, Mikhail and Myrto Pantazi (2020). "Pragmatic Accommodation". In: *The Wiley Blackwell Companion to Semantics*. John Wiley & Sons, Ltd, pp. 1–16. DOI: 10.1002/9781118788516.sem140.

Lewis, D (1979). "Scorekeeping in a Language Game". In: *Journal of Philosophical Logic* 8, pp. 339–359.

— (1988). "Relevant implication". In: *Theoria* 54, pp. 161–174.

Roberts, Craige (1996/2012). "Information structure: Towards an integrated formal theory of pragmatics". In: *Semantics and pragmatics* 5, pp. 6–1.

Stalnaker, Robert C. (1970). "Pragmatics". In: *Synthese* 22.1/2, pp. 272–289. ISSN: 00397857, 15730964.

Szabó, Zoltán Gendler (2006). "Comments on Kai von Fintel's paper. Unpublished manuscript, revised version of comments given at Ohio Workshop on Presupposition Accommodation." URL: https://bpb-us-w2.wpmucdn.com/campuspress.yale.edu/dist/d/1148/files/2015/10/On-Presupposition-Accommodation-2bdedla.pdf (visited on 12/02/2024).

Evaluating Unalternative Semantics on a large corpus of naturalistic speech

Fenna Berger
Leiden University / Vrije Universiteit Amsterdam
Leiden / Amsterdam, The Netherlands
f.berger@student.vu.nl

Matthijs Westera
LUCL, Leiden University
Leiden, The Netherlands
m.westera@hum.leidenuniv.nl

Abstract

While the experimental turn of the last decade has helped connect formal semantic/pragmatic theories to empirical reality, extending this empirical reach to naturally-occurring data has lagged behind. One of the challenges is the lack of large corpora in which the relevant technical notions have been annotated. Nevertheless, some insight may be gained by relying on automatically computed approximations of these notions. We take such an approach to focus realization – the relation between focus, grammar and pitch accents, taking inspiration in particular from Unalternative Semantics.

1 Introduction

Unalternative Semantics (Büring 2015) predicts pitch accents by combining a standard notion of focus (Rooth 1992), representing pragmatically 'new' information, with a set of rules that determine the *metric values* of nodes in a syntactic tree, representing the default stress patterns of a language. In this paper, we try to test whether this combination of focus and metric values indeed enables better prediction of prosodic stress, on a large corpus of spoken English. We do this by training standard machine learning models on predicting an automatically extracted measure of 'prosodic prominence' – a combination of pitch, duration and loudness – based on automatically extracted, approximate notions of pragmatic focus and syntactic structure (for calculating metric values). We find that, indeed, both sources of information contribute to predicting prosodic prominence, though with important caveats and open questions for future work.

Testing linguistic theories on naturalistic data is important (e.g., Chafe 1992) but remains challenging. In the case of focus semantics, no current corpora provide detailed information about the three dimensions involved: pitch accents, focus, and syntax. In this study, we rely on automatically extracted approximations of these notions.[1] The paper invites reflection on the use of such approximations for bridging the gap between theories and naturalistic data.

2 Background

2.1 Unalternative Semantics

As characterized in Büring 2015, standard focus theory (e.g., Rooth 1992) has three components. *Focus pragmatics* states that the focal target of a sentence (given the discourse context) must be among the set of focus alternatives. *Focus semantics* describes the relation between a sentence and its set of focus alternatives, standardly through the placement of F-markers on the syntactic tree. *Focus realization*, finally, concerns the relation between such F-markers and prosody, stating for instance that the main stress of a sentence must be placed within an F-marked constituent.

[1] Code for this paper can be found on https://github.com/fennaberger/prosodic-prominence.

Büring's Unalternative Semantics merges the latter two components – focus semantics and focus realization – by directly relating focus alternatives to prosody.

Unalternative Semantics relies on default rules about which words receive stress. One of these rules, called DEFAULT STRESS, is that, in a syntactic constituency tree, the right daughter of a node be prosodically stronger than the left daughter:

(1) DEFAULT STRESS
Label the left sister w(eak) and the right sister s(trong)

If this rule were the only rule in existence, the stress of a sentence would always be on the final word. This is called 'default prosody', and in such cases prosody and focus are related by the rule WEAK RESTRICTION:

(2) WEAK RESTRICTION (only applies in the case of default prosody)
If the weak daughter contains a constituent to be focused, the strong daughter does too.

This rule constrains which constituents may be the focus in case of default prosody. Consider example (3), which indeed shows default prosody:

(3) Mom made MUFFINS.

The rule WEAK RESTRICTION implies that 1) if "mom" is focused, the VP "made muffins" is focused too and 2) if "made" is focused, "muffins" is too. Hence, either the whole sentence is the focus, or the whole VP, or only the object DP – but not, for instance, only the subject DP.

If, contrary to the rule DEFAULT STRESS, a sentence contains a pitch accent somewhere on the left – so-called *prosodic reversal* – the rule STRONG RESTRICTION applies instead of WEAK RESTRICTION:

(4) STRONG RESTRICTION (only applies in the case of prosodic reversal)
The strong daughter contains a constituent to be focused, the weak daughter does not.

For example, (5) shows prosodic reversal at the S (sentence) node, since its left daughter is stronger than its right:

(5) DAD made pancakes.

Since DEFAULT STRESS cannot apply here, the rule STRONG RESTRICTION must be used, which implies that only "dad" can be the focus, nothing else.

At least two other rules (Büring 2016) play a role in determining default prosody (and, hence, co-define what counts as 'prosodic reversal'): "predicates (head) are weaker than their arguments (complements)" and "functional sisters (pronouns, auxiliars and similar) are weaker than lexical, open-class ones". These two rules will also be considered in the present study, along with DEFAULT STRESS, STRONG RESTRICTION and WEAK RESTRICTION.

2.2 The Helsinki Prosody Corpus

We seek to test the theory of Unalternative Semantics on naturally occurring spoken English. We will use the Helsinki Prosody Corpus (Talman et al. 2019), which consists of automatically generated annotations of 'prosodic prominence' on top of the LibriTTS corpus of recorded speech (Zen et al. 2019), using the Continuous Wavelet Transform (Suni et al. 2017) which measures fundamental frequency, energy, and duration in order to recognize prosodic events in speech files. They report an accuracy of 84% in detecting prosodic prominence. We use the 'train-360' partition of the Helsinki Prosody Corpus, containing around 2M words (see Table 1). In this corpus, each word has a *continuous* prominence score based on its prosodic prominence and a

Speakers	Sentences	Words	Prominence 0	Prominence 1	Prominence 2
904	116,262	2,076,289	1,003,454	569,769	503,069

Table 1: The train-360 partition of the Helsinki Prosody Corpus.

	Prosodic Prominence: 1 or 2	Prosodic Prominence: 0
Pitch Accent	737	99
No Pitch Accent	122	639

Table 2: Pitch accents and prosodic prominence

discrete prominence score of 0, 1, or 2, obtained simply by discretizing the continuous score. The distribution of prominence is skewed towards non-prominent, with 1M words receiving score 0, 569K words score 1 and 503K words score 2.

3 Method

One way to to test a theory of focus on naturalistic data, is to assess the ability of standard machine-learning models to predict pitch accent placement (the dependent variable), on the basis of features which the given focus theory deems necessary (the independent variables). For Unalternative Semantics, the features would include the focus of an utterance and the metric values for each syntactic node. As no corpus exists that contains all the necessary notions, we must rely on approximations: using prosodic prominence in place of pitch accents, using unpredictability in place of focus, and using automatic syntactic parsing for implementing Büring's rules for assigning metrical values (relative prosodic strength). These three approximations will be explained and discussed next, followed by the details of model fitting.

3.1 Prosodic Prominence as a Proxy for Pitch Accents

Whereas Büring's theory pertains to pitch accents, the Helsinki Prosody Corpus contains only automatic annotations of prosodic prominence – and not every word with high prosodic prominence will necessarily be perceived by human listeners as having a pitch accent. To assess the quality of this approximation, we annotate the pitch accents of 100 random sentences in the corpus, for a total of 1597 words. This was done by one rater (first author) who listened to the audio while using Praat (**broersma2019praat**) to visualize pitch contours. The results are shown in Table 2 (we group prominence scores 1 and 2 together, since both received a pitch accent in most cases).

According to these results, prominence scores are an accurate predictor of pitch accents 86% of the time. Recall that Suni et al. 2017 reported 84% accuracy for their method of automatically annotating prominence, although they did not look specifically at pitch accents. It is possible that the 86% accuracy we found is a compound of 1. inaccuracies in Suni et al.'s prosodic prominence scores to begin with, and 2. the difference between prosodic prominence as a graded notion and pitch accents as a discrete linguistic category.

3.2 Unpredictability as a Proxy for Focus

We will use unpredictability as an approximation of focus: words that are focused tend to be less predictable, more informative. Indeed, the notion of focus has often been characterized as the 'new' information, compared to what is 'given' (e.g., Schwarzschild 1999; Birner 1994), where 'given' does not simply mean to have been literally mentioned before, but, rather, to be unexpected given the context. We leave a quantitative assessment of the correlation between

unpredictability and focus for another occasion, and future work should try to tell the two apart (in particular because predictability is already known to correlate with prosodic prominence, Aylett and Turk 2004). (For the results reported here, we did not filter out the many tokens that were split up by BERT's word-piece tokenization, and used the first word-piece's probability.)

We implement this idea with the BERT language model (Kenton and Toutanova 2019), using the `transformers` library (Wolf et al. 2020). For each sentence, and for each word in that sentence, we created a version of the sentence where that word was masked (i.e., replaced by the special token `[MASK]`). We added one sentence before and one sentence after the sentence to provide context, and obtained the probability, according to BERT, of the original word being in the masked position. We used this probability, as a value between 0 and 1, directly as a measure of predictability, hence, as an approximation for *non*-focus. In order not to fool ourselves, we will henceforth use the term predictability instead of the notion of (non-)focus for which it is meant to be an approximation.

3.3 Automatically assigning metric values

We implement three rules that play a role in determining default prosody according to Büring 2016: "right sister is stronger than left sister" (DEFAULT STRESS), "complements are stronger than heads", and "content words are stronger than function words".

We implement the first rule ("right sister is stronger than left sister") using Stanza's constituency parser (Qi et al, 2020), by computing how often one has to turn 'right' to reach a given leaf node. However, whereas Büring's rule assumes binary trees, nodes in the constituency trees from Stanza can have more than two children (and do so in a variety of syntactic contexts). To deal with such cases, we assigned weights to each child corresponding to 'how far right' among the children they were. More precisely, we the score S_c of each child node, given its parent's score S_p can be calculated as follows, where n is the number of siblings the node has to its left, m is the total number of siblings (include the node itself), and $S_{root} = 0$:

$$S_c = S_p + \frac{n}{m-1} \tag{1}$$

We subsequently scaled all word-scores for a given sentence to the interval [0, 1] to control for sentence length (otherwise scores in longer sentences would be systematically higher).

To implement the second rule, "complements are stronger than heads", because Stanza's constituency parser does not provide information about heads and complements, we used SpaCy's (Honnibal et al. 2020) dependency parser instead (the model `en_core_web_md`; Unlabelled Attachment Score = 0.92, Labelled Attachment Score = 0.90). In general, the head of the non-head child of node X in a constituency tree depends in the dependency tree on the head child of node X (Xia and Palmer 2001), hence complements tend to be below their heads in the dependency tree. We thus implement the rule "complements are stronger than heads" by obtaining the depth of a word in the dependency tree, again scaled to control for sentence length.

Lastly, we implemented the rule "content words are stronger than function words" using SpaCy's part-of-speech tagger (`en_core_web_md`, POS accuracy 0.97), treating adjective, adverb, interjection, noun, proper noun, and verb as content words, and the rest as function words.

3.4 Models

We seek to evaluate Unalternative Semantics by assessing whether combining focus and metric values enables models to better predict pitch accents – or rather, given our approximations, whether combining predictability with metric values enables models to better predict prosodic prominence. We fit models in the 'predictability only' condition, the 'metric values only'

Total tokens	Total words (no punctuation)	Train set	Test set
100,000	70,667	53,000	17,667

Table 3: The train set and test set.

Type	Features	Accuracy	Precision	Recall	F1-score
Decision tree	Predictability	0.533	0.496	0.533	0.500
	Metric values	0.557	0.537	0.557	0.543
	Both	0.571	0.549	0.571	0.555
Random forest	Predictability	0.534	0.494	0.534	0.498
	Metric values	0.564	0.548	0.564	0.554
	Both	**0.579**	**0.557**	**0.579**	**0.563**
SVM	Predictability	0.521	0.409	0.522	0.457
	Metric values	0.566	0.551	0.566	0.556
	Both	0.568	0.547	0.568	0.553
Baseline	random	0.361	0.361	0.361	0.361

Table 4: Results of the classifiers (weighted average over classes), and weighted random baseline.

condition, and the combined 'predictability and metric values' condition, where the latter most closely resembles Büring's theory of Unalternative Semantics.

For each condition, we use `scikit-learn` to fit three types of classification models, to predict the discrete prominence scores (decision tree, random forest, and support vector machine (SVM)), and three types of regression models to predict the continuous prominence scores (ridge regression, random forest, and SVM).

Input features (predictability, metric values) were computed for an arbitrary total of 100K words, which after removing punctuation left 71K words for training and testing. We used the same train/test split for all six model types in all three conditions, using 25% as test set, and 5-fold cross-validation on the training set (see Table 3).

4 Results

See Table 4 for the results of the classification models and a weighted random baseline. For the decision tree and the random forest, the "both" (predictability and metric values) condition outperformed the other two conditions in all four metrics (all differences reported here are significant with $p < 0.0001$, according to Bowker-McNemar test with 3 dof). Not so for the SVM, where the "metric values only" condition in fact performed *very* slightly better than "both" in terms of F1-score (though not on accuracy and recall). As one possible explanation, we checked whether the SVM model in the "both" condition was perhaps overfitting on the training data, but this was not the case; the model did not perform better on the training data.

We inspect the best classification model more closely, the random forest in the "both" condition, by looking at the feature importances. The rule "content words are stronger than function words" carries the most weight (0.48), followed by predictability (0.37), "right sisters stronger than left sisters" (0.1) and "complements stronger than heads" (0.05).

The confusion table for our best classifier revealed substantial confusion between levels 1 and 2 (i.e., medium and strong prominence). Our annotations earlier revealed that both levels tend to reflect pitch accents, suggesting we could merge these levels, turning the task into a binary classification task. Moreover, this meant we could take our 86% accuracy of the prominence-as-pitch-accents approximation into account, by randomly reassigning labels in the test set according to this probability. Average F1-scores over 100 thusly resampled test sets revealed

Type	Features	Mean	Standard deviation
Ridge	Predictability	0.556	0.461
	Metric values	0.541	0.455
	Both	0.514	0.452
Random forest	Predictability	0.550	0.459
	Metric values	0.533	0.451
	Both	0.503	0.448
SVM	Predictability	0.519	0.539
	Metric values	0.499	0.531
	Both	**0.480**	**0.512**
Baseline	Mean	0.634	0.462

Table 5: Results (mean error and standard deviation) for the regression models (graded prominence).

that the best model was still the random forest in the "both" condition (mean F1-score of 0.690), marginally better than the "metric values only" condition (0.684) and more substantially so for the "predictability only" condition (0.640), and in each case significantly so. By contrast the decision tree proded significantly superior in the "metric values only" condition (0.681) compared to "both" (0.675) and "predictability only" (0.629), while no significant difference was found for the SVM classifier.

Table 5 shows the results of our regression models predicting graded prosodic prominence scores (between 0 and 2), compared to the average as a baseline. We report mean difference between predictions and targets. For all three types of regression models, the "both" condition had a better fit than the conditions with either only predictability or only metric values (all comparisons significantly so as per T-test, $p < 0.01$).

We inspect the best regressor more closely, i.e., the SVM model in the "both" condition, by estimating feature importances by permutation (Breiman 2001). These show the same pattern as the random forest classifier above: "content words are stronger than function words" carries most weight (0.244), followed by predictability (0.128), then "right sisters are stronger than left sisters" (0.04) and "complements are stronger than heads" (0.003).

The Ridge regressor is our only model which computes a likelihood function, so we can compute its Akaike information criterion (AIC) to assess whether the improved fit in the "both" condition is worth added model complexity due to it having more features. This appears to be the case: the AIC score in the "both" condition (111701) is significantly lower than the scores of the "predictability only" condition (119251) and the "metric values only" condition (116221).

5 Discussion

We conclude that having both predictability and metric values helps models predict prosodic prominence, as opposed to having only one of the two. Moreover, in the resulting models the content/function word distinction and predictability carry most of the weight. On the surface, none of the models seem particularly successful (top F1-score of 0.56, or 0.69 in the binary condition). Presumably this reflects, at least in part, inaccuracies in our approximations (unpredictability as focus, prominence as pitch accents, imperfect syntactic parsing). In addition, the default prosody rules identified by Büring may not be the only ones. However, ultimately, accent placement is determined not by context, but by a speaker's communicative intentions. For a more definitive assessment of the quality of model fit, it would be helpful to compare it to the ability of *humans* to predict prominence based only on the transcripts.

References

Aylett, Matthew and Alice Turk (2004). "The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech". In: *Language and speech* 47.1, pp. 31–56.

Birner, Betty J (1994). "Information status and word order: An analysis of English inversion". In: *Language*, pp. 233–259.

Breiman, Leo (2001). "Random forests". In: *Machine learning* 45, pp. 5–32.

Büring, Daniel (2015). "Unalternative semantics". In: *Semantics and linguistic theory*, pp. 550–575.

— (2016). "A beginner's guide to unalternative semantics". In: *Manuscript University of Vienna. http://homepage. univie. ac. at/daniel. buring/phpsite/index. php*.

Chafe, Wallace (1992). *The importance of corpus linguistics to understanding the nature of language*. De Gruyter.

Honnibal, Matthew et al. (2020). "spaCy: Industrial-strength natural language processing in Python". In.

Kenton, Jacob Devlin Ming-Wei Chang and Lee Kristina Toutanova (2019). "Bert: Pre-training of deep bidirectional transformers for language understanding". In: *Proceedings of naacL-HLT*. Vol. 1. Minneapolis, Minnesota, p. 2.

Rooth, Mats (1992). "A theory of focus interpretation". In: *Natural language semantics* 1.1, pp. 75–116.

Schwarzschild, Roger (1999). "GIVENness, AvoidF and other constraints on the placement of accent". In: *Natural language semantics* 7.2, pp. 141–177.

Suni, Antti et al. (2017). "Hierarchical representation and estimation of prosody using continuous wavelet transform". In: *Computer Speech & Language* 45, pp. 123–136.

Talman, Aarne et al. (2019). "Predicting prosodic prominence from text with pre-trained contextualized word representations". In: *arXiv preprint arXiv:1908.02262*.

Wolf, Thomas et al. (2020). "Transformers: State-of-the-art natural language processing". In: *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*, pp. 38–45.

Xia, Fei and Martha Palmer (2001). "Converting dependency structures to phrase structures". In: *Proceedings of the first international conference on Human language technology research*.

Zen, Heiga et al. (2019). "LibriTTS: A corpus derived from librispeech for text-to-speech". In: *arXiv preprint arXiv:1904.02882*.

More than one way to free choice: A view from child Romanian

Adina Camelia Bleotu *
University of Bucharest
Bucharest, Romania
adina.bleotu@lls.unibuc.ro

Mara Panaitescu *
University of Bucharest
Bucharest, Romania
mara.panaitescu@lls.unibuc.ro

Alexandre Cremers
University of Vilnius
Vilnius, Lithuania
alexandre.cremers@gmail.com

Andreea Nicolae
Leibniz-Centre General Linguistics (ZAS)
Berlin, Germany
andreea.nicolae@gmail.com

Anton Benz
Leibniz-Centre General Linguistics (ZAS)
Berlin, Germany
benz@leibniz-zas.de

Gabriela Bîlbîie
University of Bucharest
Bucharest, Romania
gabriela.bilbiie@lls.unibuc.ro

Lyn Tieu
University of Toronto / MARCS Institute for Brain, Behaviour & Development / Macquarie University
Toronto, Canada
lyn.tieu@utoronto.ca

Abstract

Studies show that children draw from modalized disjunctive statements of the structure *X is allowed to do P or Q* ($\Diamond(P \vee Q)$) a Free Choice (FC) inference, namely *X is allowed to do P and X is allowed to do Q* ($\Diamond P \wedge \Diamond Q$). Their ability to compute free choice inferences is surprising in light of their well-known difficulties with scalar implicatures involving non-modalized disjunction (Tieu, Romoli, et al. 2016), particularly on accounts that unify free choice inferences and scalar implicatures (e.g., Kratzer and Shimoyama 2002; Chierchia 2013). Recent work by Cochard, van Hout, and Demirdache (2024b), however, argues that some children only seemingly derive free choice: these children actually interpret $\Diamond(P \vee Q)$ as $\Diamond(P \wedge Q)$, which follows from their conjunctive understanding of non-modalized disjunction. In the present study, we extend the investigation by comparing the same children's performance on non-modalized and modalized utterances in Romanian, an understudied language. Specifically, we tested the same group of 5-year-old monolingual Romanian-speaking children and adult controls, balanced for order. Our findings provide partial evidence for Cochard, van Hout, and Demirdache (2024b)'s hypothesis: some children were inclusive with non-modalized disjunction, and appeared to derive genuine free choice on the free choice task, while some children indeed exhibited conjunctive interpretations in both tasks.

1 Introduction

Modalized disjunctive utterances of the structure *X is allowed to do P or Q* ($\Diamond(P \vee Q)$) are typically associated with Free Choice (FC) inferences, i.e. *X is allowed to do P and X is allowed to do Q* ($\Diamond P \wedge \Diamond Q$), as exemplified in (1-b) (see Kratzer and Shimoyama 2002; Chierchia 2013; Bar-Lev and Fox 2020; Aloni 2022). Such interpretations have been shown to arise both in adult language (Chemla and Bott 2014; Meyer and Feiman 2021) and child language (Zhou, Romoli,

*Joint first authors.

and Crain 2013; Tieu, Romoli, et al. 2016; Huang and Crain 2020). Interestingly, kindergarten-aged children seem to be adult-like in their computation of free choice inferences, while reportedly struggling with deriving exclusivity implicatures from non-modalized disjunction.

(1) a. Mary **is allowed** to eat the orange **or** the pear.
b. Mary **is allowed** to eat the orange, **and** Mary **is allowed** to eat the pear.

Recent work by Cochard, van Hout, and Demirdache (2024b), however, identifies multiple ways in which modalized disjunctive statements might be interpreted in child language. According to Cochard, Hout, and Demirdache (2024a), children may derive:

- a genuine FC reading, achieved through exhaustification above the modal:
  $Exh(Exh(\Diamond(P \vee Q))) = \Diamond P \wedge \Diamond Q$
- a MODAL AND reading, where disjunction is interpreted as conjunction via exhaustification under the modal:
  $\Diamond(Exh(Exh(P \vee Q))) \rightarrow \Diamond(P \wedge Q)$

Cochard, van Hout, and Demirdache (2024b) argue that, in previous studies (Zhou, Romoli, and Crain 2013; Tieu, Romoli, et al. 2016; Huang and Crain 2020), the contexts in which modalized disjunctive statements were presented were such that the statements were false under both the FC ($\Diamond P \wedge \Diamond Q$) and MODAL AND ($\Diamond(P \vee Q)$) readings. Consequently, these studies did not effectively distinguish between the two possible readings. To address this issue, Cochard, van Hout, and Demirdache (2024b) introduced a mutually exclusive disjunctive scenario in which, for example, the statement *With 1 coin, the bear can buy a croissant or a donut* was true under the FC reading ($\Diamond P \wedge \Diamond Q$) but false under the MODAL AND reading ($\Diamond(P \wedge Q)$), see Figure 1. Incorporating this trial type into their design allowed Cochard, van Hout, and Demirdache (2024b) to tease apart the two interpretations. They found that, while some children derived genuine FC inferences, others appeared to interpret the disjunction as conjunction. In the present study, we extend the investigation to Romanian, pursuing the possibility of a correlation between children's performance on modalized and non-modalized disjunctive utterances.



Figure 1: Mutually exclusive disjunctive scenario from Cochard, van Hout, and Demirdache (2024b) (image reproduced with permission).

2 Current experiments

To ensure a fair comparison of conjunctivity, inclusivity, and exclusivity in modalized and non-modalized disjunctive utterances, we designed two tasks: the Implicature Task and the Free Choice Task, which were administered to the same group of participants in a counterbalanced order. Additionally, we examined whether complex disjunctions would obligatorily trigger exhaustification (Spector 2014) compared to simple disjunctions. Romanian, with its rich array of disjunctions, provided the basis for our study. We focused on two specific forms: the simple disjunction *sau*, which prior research (Bleotu, Ivan, et al. 2023; Bleotu, Tieu, et al. 2024)

indicates is interpreted inclusively (verified when one or both disjuncts are true) or exclusively (verified when only one disjunct is true) by adults, but predominantly inclusively by children; and the complex disjunction *fie...fie*, which is interpreted exclusively by adults, but conjunctively (verified only when both disjuncts are true) or inclusively by children.

Building on Cochard, van Hout, and Demirdache (2024b), we hypothesized that some children would interpret both modalized and plain disjunctions conjunctively (i.e., as $\Diamond(P \wedge Q)$ and $P \wedge Q$, respectively). Conversely, children who were non-conjunctive in the Implicature Task were expected to derive genuine free choice inferences in the Free Choice Task. Additionally, based on prior studies showing that children interpret the non-modal disjunction *sau* inclusively and *fie...fie* either conjunctively or inclusively, we anticipated more conjunctive readings of both plain and modalized utterances with *fie...fie* than with *sau*.

We collected data from 69 monolingual Romanian-speaking 5-year-olds (ages 4;11–5;11, mean age: 5;05): 32 were tested on the simple disjunction *sau*, and 37 on the complex disjunction *fie...fie*. Of the 65 adult native speaker controls, 33 were tested on *sau* and 32 on *fie...fie*. The children were recruited from two kindergartens (Licurici and No. 248) in Bucharest, and the adults were undergraduate students at the University of Bucharest.

2.1 Implicature Task

The Implicature Task used a Truth Value Judgment Task in prediction mode, following Tieu, Yatsushiro, et al. (2017). Prediction mode licenses ignorance inferences regarding which disjunct will be true. As shown in (2) and Figure 2, participants were told a story about characters shopping, with the puppet Bibi guessing which items they would buy. They then saw what items were purchased, and had to evaluate Bibi's guesses by rewarding correct guesses with a smiley face and incorrect ones with a sad face. Each participant received 2 warm-up trials, 15 disjunctive test items, and 10 fillers. Disjunctive statements were presented in three kinds of contexts: (i) 1-disjunct-true (1DT), (ii) 2-disjunct-true (2DT), and (iii) a false 0-disjunct-true (0DT) control condition. Importantly, the visual context included three alternatives to avoid a potential task effect that could boost conjunctive interpretations (for discussion of this point, see Huang and Crain 2020; Skordos et al. 2020; Bleotu, Tieu, et al. 2024). Table 1 displays the predicted responses on each of the relevant interpretations, for each of the three contexts.

(2) *Scene 1*
Experimenter: It's Maria's turn to go to the store. Bibi, can you guess what's going to happen?
Bibi: *Maria o să cumpere portocala* ***sau*** *para.* 'Maria will buy the orange or the pear.'
Experimenter: Let's see if Bibi's right!
Scene 2 (outcome revealed)
Experimenter: Look, Maria bought this! So was Bibi right?
Scene 3
Experimenter: What smiley should we give Bibi?

2.2 Free Choice Task

In the Free Choice task, participants were introduced to a scenario where Mother set rules about what objects a character was allowed/not allowed to buy. They listened to Bibi describe a rule from memory, then saw a visual representation of the actual rule. The visual display included three images: Object 1, Object 2, and the combination of both, each surrounded by a circle. Green circles indicated permission to buy the items(s), while red circles indicated interdiction. Participants had to decide whether Bibi's statement matched the visual rule (see Tieu, Bill, and

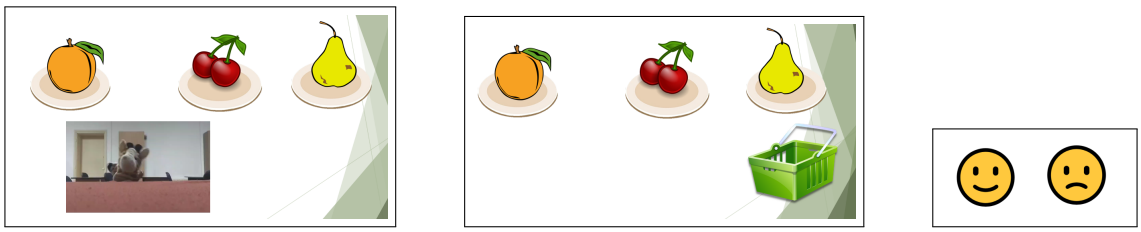


Figure 2: Implicature Task: The scenes of a trial in which *Mary will buy an orange or a pear* was uttered in what turned out to be a 1DT context.

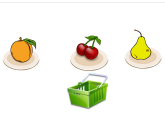
Interpretation of $(P \vee Q)$ *Maria o să cumpere P* ***sau*** *Q (e.g., portocala sau para)* 'Maria will buy P or Q' (e.g., the orange or the pear)		**1DT** 	**2DT** 	**0DT** 
$(P \vee Q) \wedge \neg(P \wedge Q)$	[+Excl]	YES	NO	NO
$P \vee Q$	[–Excl]	YES	YES	NO
$P \wedge Q$	[AND]	NO	YES	NO

Table 1: Conditions and expected responses in the Implicature Task.

Romoli 2024; Tieu, Godo, et al. 2024).

Participants saw 4 warm-up trials (two true, two false), followed by a pseudo-randomized sequence of 25 targets and 10 non-disjunctive modalized fillers (5 true, 5 false). Modalized disjunctive statements were presented across five different scenarios (see Table 2). In the OnlyOne scenario, only one disjunct was allowed. In the Each&NotBoth scenario, each disjunct was a possibility, but only one could be chosen, not both. In the Each&Both scenario, the character was allowed to buy each disjunct separately or both together. In the Both&NotEach scenario, the character could buy both disjuncts together, but not separately. Finally, in the False control condition, the character was not allowed to buy either disjunct, neither separately nor together. (3) shows an example of an OnlyOne trial, along with the pictured scenes in Figure 3.

(3) *Scene 1*
Experimenter: It's Maria's turn to go to the store! Bibi, can you remember the rule for Maria?
Bibi: *Maria* ***are voie*** *să cumpere portocala* ***sau*** *para.* 'Maria is allowed to buy the orange or the pear.'
Scene 2 (rule is visually depicted)
Experimenter: Let's see what Mother's rule is.

SCENARIO (Mother's rule) INTERPRETATION OF $\Diamond(P \vee Q)$ *Maria* ***are voie*** *să cumpere P* ***sau*** *Q.* 'Maria is allowed to buy P or Q.'		**Only one particular disjunct allowed (OnlyOne)** 	**Each disjunct, not both allowed (Each&NotBoth)** 	**Each disjunct and both disjuncts together allowed (Each&Both)** 	**Both disjuncts together allowed, not separately (Both&NotEach)** 	**False control** 
$(\Diamond P \wedge \Diamond Q) \wedge \neg\Diamond(P \wedge Q)$	[+FC +Excl]	NO	YES	NO	NO	NO
$\Diamond P \wedge \Diamond Q$	[+FC –Excl]	NO	YES	YES	YES/NO	NO
$\Diamond(P \wedge Q)$	[Modal AND]	NO	NO	YES	YES	NO
$\Diamond(P \vee Q)$	[–FC –Excl]	YES	YES	YES	YES	NO

Table 2: Conditions and expected responses in the FC Task.

Scene 3
Experimenter: Did Bibi remember correctly? What smiley would you give to Bibi?

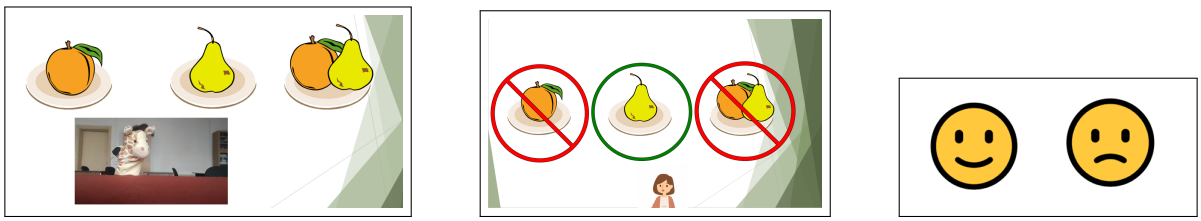


Figure 3: Free Choice Task: The scenes of a trial in which *Mary is allowed to buy an orange or a pear* was uttered in an OnlyOne context.

2.3 Results

Figure 4 shows the distribution of participants across different interpretive categories for both plain and modal disjunctive utterances.

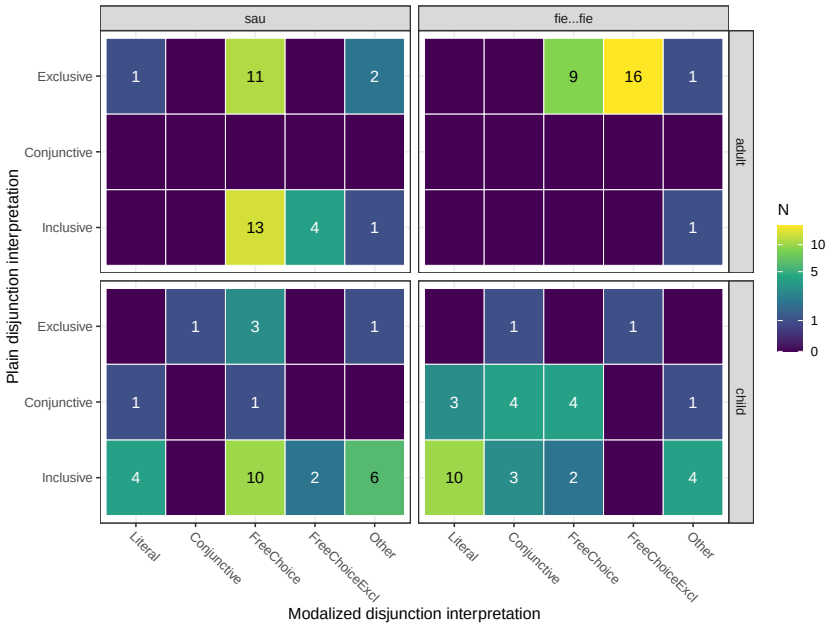


Figure 4: Distribution of participants per interpretive category in the Implicature and FC Conditions.

Following a strong version of Cochard, van Hout, and Demirdache (2024b), we expected the following types of behavior: (i) **Plain literal** responders should interpret both plain and modalized disjunctions inclusively; (ii) **Pragmatic children with access to the conjunctive alternative** should be adult-like; (iii) **Conjunctive** children should interpret plain disjunctions conjunctively and modalized disjunctions as MODAL AND. A weaker version of Cochard, van Hout, and Demirdache (2024b)'s model would add a fourth category, namely children who could be conjunctive in the Implicature Task but would nonetheless opt to derive inclusive FC. We fitted Poisson regressions to the count data with categorical parameters corresponding to each responder profile predicted by the model, and comparison was done using BIC and the approximation of Bayes' factors derived from it. In terms of BIC, the weak version of Cochard et al.'s model performed best (BIC = 70.32).

Regarding the difference in interpretation between *fie. . . fie* and *sau*, we had the expectation

that *fie... fie* would lead to more strengthened interpretations than *sau*, resulting in more exclusivity for adults, and more conjunctivity (or exclusivity) for children. We applied a binomial logistic regression to the categories assigned to each participant. The dependent variable indicated literal or strengthened interpretations, with predictors for Disjunction type, Group, and Condition, along with their interactions (excluding triple interactions). All factors were sum-coded. Results show that participants were more likely to access strengthened interpretations for *fie... fie* than *sau* ($p < .001$). The effect of disjunction type was stronger in plain contexts than in modalized ones. Moreover, adults and children showed different patterns, as indicated by a strong interaction between Group and Disjunction ($p < .001$).

3 Discussion

Our study provides valuable insights into how Romanian children interpret disjunction in plain and modalized utterances. While children's conjunctive interpretations are primarily associated with the complex disjunction *fie... fie* in both tasks, this conjunctive reading does not always transfer from non-modalized to modalized disjunctions. More specifically, in our Implicature Task, 12 children consistently interpreted *fie... fie* conjunctively, whereas in the FC Task, only 4 of these exhibited a MODAL AND interpretation; among the rest, 3 were literal, 4 derived FC, and 1 was mixed. Thus, while children can be conjunctive with non-modalized disjunctions, they are not necessarily so in modalized contexts. Conversely, in the FC Task, 8 children displayed a conjunctive interpretation of *fie... fie*. Of these, 4 were also conjunctive in the Implicature Task, while the remaining children interpreted the non-modalized *fie... fie* inclusively or exclusively.

The consistently conjunctive children align with Cochard, van Hout, and Demirdache (2024b)'s hypothesis that conjunctive readings of both plain and modal disjunctions are linked. These children consistently apply a conjunctive interpretation across modal and non-modal contexts. However, we also observed a subset of children who showed inconsistent conjunctive behavior. These children interpreted *fie... fie* conjunctively in one task but inclusively in the other. This inconsistency suggests that, for children, there may be multiple interpretations available for *fie... fie* (conjunctive/inclusive, see Bleotu, Tieu, et al. 2024). Overall, our results support a weak version of Cochard, van Hout, and Demirdache (2024b).

Interestingly, even our adult controls were not fully consistent and showed some differences in how they interpreted *fie... fie* across tasks: while most were exclusive with both plain and modal disjunctive utterances, they tended to be more exclusive in plain contexts and less exclusive and more inclusive in the modal environment. Tieu, Godo, et al. (2024) report the same finding for French-speaking adults, arguing that some participants opt to only exhaustify to FC. If FC inferences, like exclusivity implicatures, are derived through exhaustification (strengthening to $\Diamond P$ and $\Diamond Q$), then additional exhaustification to exclusivity might be optional for inclusive participants, perhaps for reasons of economy. The mechanism at play in the case of *fie... fie* then seems to be obligatory exhaustification rather than obligatory exclusivity.

4 Conclusion

Our findings, in addition to Cochard, van Hout, and Demirdache (2024b), provide further evidence that children's interpretations of modalized disjunctive statements can be driven by at least two distinct mechanisms: a genuine FC Inference ($\Diamond P$ and $\Diamond Q$), according to which each disjunct is independently possible, and a MODAL AND reading ($\Diamond(P \wedge Q)$), i.e., a conjunctive interpretation where both possibilities are allowed simultaneously. Additionally, our study also highlights that some participants may choose to resort to different exhaustification strategies in plain and modalized disjunctive utterances.

Acknowledgements. The current research was supported by the project "The Acquisition of Disjunction in Romanian" PN-III-P1-1.1-TE-2021-0547 (TE 140 din 30/05/ 2022), led by A. Bleotu. A. Nicolae was supported by the DFG grant NI-1850/2-1. A. Benz's work was partly funded by the "Linguistic Meaning and Bayesian Modelling" project within the Leibniz Collaborative Excellence Programme (PI Anton Benz, Application number: K535/2023). L. Tieu was supported in part by funding from the Social Sciences and Humanities Research Council of Canada and the Connaught Fund. We thank the students at the University of Bucharest for taking part in the experiments, our research assistants for helping with data collection, and the children from No. 248 Kindergarten and Licurici Kindergarten in Bucharest for their participation.

References

Aloni, Maria (2022). "Logic and conversation: The case of free choice". In: *Semantics and Pragmatics* 15, 5–EA.

Bar-Lev, Moshe and Danny Fox (Sept. 2020). "Free choice, simplification, and Innocent Inclusion". In: *Natural Language Semantics* 28, pp. 175–223. DOI: 10.1007/s11050-020-09162-y.

Bleotu, Adina Camelia, Rodica Ivan, et al. (2023). "Not all complex disjunctions are alike: On inclusive and conjunctive interpretations in child Romanian". In: *Proceedings of the Annual Conference of the Cognitive Science Society 45*, pp. 3062–3069.

Bleotu, Adina Camelia, Lyn Tieu, et al. (Jan. 2024). "Children Interpret Some Disjunctions Conjunctively: Evidence from Child Romanian". In: *PsyArXiv*. DOI: 10.31234/osf.io/bywj2.

Chemla, Emmanuel and Lewis Bott (2014). "Processing inferences at the semantics/pragmatics frontier: Disjunctions and free choice". In: *Cognition* 130.3, pp. 380–396. DOI: 10.1016/j.cognition.2013.11.013. URL: https://doi.org/10.1016/j.cognition.2013.11.013.

Chierchia, Gennaro (2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press.

Cochard, Antoine, Angeliek van Hout, and Hamida Demirdache (2024a). *When Modal Disjunction is interpreted as Modal Conjunction in Child French: "Liz can buy a croissant or a donut... that means both, right?"* Talk at *Free Choice Workshop*, 15 May 2024.

Cochard, Antoine, Angeliek van Hout, and Hamida Demirdache (2024b). ""Liz can buy a croissant or a donut... Both together, right?" Distinguishing target Free Choice from non-target Modal AND in Child French". In: *Proceedings of 3rd edition of Experiments in Linguistic Meaning (ELM3), University of Pennsylvania, Philadelphia*.

Huang, Haiquan and Stephen Crain (2020). "When *OR* is assigned a conjunctive inference in child language". In: *Language Acquisition* 27.1, pp. 74–97. DOI: 10.1080/10489223.2019.1659273.

Kratzer, Angelika and Junko Shimoyama (2002). "Indeterminate pronouns: The view from Japanese". In: *Proceedings of the Tokyo Conference on Psycholinguistics*. Ed. by Yukio Otsu. Vol. 3. Tokyo: Hituzi Syobo, pp. 1–25.

Meyer, M.-C. and R. Feiman (2021). "Priming reveals similarities and differences between three purported cases of implicature: Some, number and free choice disjunctions". In: *Journal of Memory and Language* 120, Article 104206. DOI: 10.1016/j.jml.2020.104206. URL: https://doi.org/10.1016/j.jml.2020.104206.

Skordos, Dimitrios et al. (2020). "Do Children Interpret 'or' Conjunctively?" In: *Journal of Semantics* 37.2, pp. 247–267.

Spector, Benjamin (2014). "Global positive polarity items and obligatory exhaustivity". In: *Semantics and Pragmatics* 7, pp. 1–61.

Tieu, Lyn, Cory Bill, and Jacopo Romoli (2024). "An experimental investigation of implicature and homogeneity approaches to free choice". In: *Natural Language Semantics*. DOI: 10.1007/s11050-024-09223-6. URL: https://doi.org/10.1007/s11050-024-09223-6.

Tieu, Lyn, Yawovi Godo, et al. (2024). *Experimentally Investigating the Strengthening Properties of Disjunction in French: When Exclusivity Meets Free Choice and Ad Hoc Implicatures*. Poster presentation at the 3rd Annual Conference on Experimental and Theoretical Linguistics (ELM 3).

Tieu, Lyn, Jacopo Romoli, et al. (May 2016). "Children's Knowledge of Free Choice Inferences and Scalar Implicatures". English. In: *Journal of Semantics* 33.2, pp. 269–298. ISSN: 1477-4593. DOI: 10.1093/jos/ffv001.

Tieu, Lyn, Kazuko Yatsushiro, et al. (Aug. 2017). "On the Role of Alternatives in the Acquisition of Simple and Complex Disjunctions in French and Japanese". In: *Journal of Semantics* 34.1, pp. 127–152.

Zhou, Peng, Jacopo Romoli, and Stephen Crain (2013). "Children's knowledge of alternatives". In: *Semantics and Linguistic Theory (SALT) 23*. Ed. by Todd Snider. Cornell University Ithaca, NY: CLC Publications, pp. 632–651.

Thinking Statively and Dynamically: A View from Georgian

Tatiana Bondarenko
Harvard University
Cambridge, MA, USA
tbondarenko@fas.harvard.edu

Richard Luo
Yale University
New Haven, CT, USA
richard.luo@yale.edu

Vincent Rouillard
Harvard University
Cambridge, MA, USA
vrouillard@fas.harvard.edu

Abstract

Whether verbs like *think* can embed questions depends on lexical aspect (Özyıldız 2021; Özyıldız 2024): stative 'think' is compatible only with declarative complements, whereas process 'think' can combine with both declaratives and interrogatives. In this paper we provide a refinement to this generalization based on the data from Georgian, and propose an account of why such dependence between lexical aspect and embeddability of questions holds. We propose that both stative and process 'think' are derived from a more abstract common core—the former via a genericity operator (Özyıldız 2024), the latter via a cumulativity operator—and argue that the genericity operator leads to interpreting 'think' as 'believe' with declaratives, but gives rise to a semantically deviant meaning with interrogatives.

1 The Puzzle

Although verbs like *think* have traditionally been assumed to be *anti-rogative* (Grimshaw 1979; Mayr 2019; Theiler, Roelofsen, and Aloni 2019), recent literature reveals a more complex empirical picture (Dayal 2016; White and Rawlins 2018; Özyıldız 2021; Özyıldız 2024; White 2021): the ability of these verbs to embed questions is contingent on their lexical aspect (see also Grohne 2016; Grohne 2017). Özyıldız (2021) proposed the following generalization for English *think*: it can combine with questions only if it is dynamic. In (2a), the interrogative CP is embedded under the verb *think* in the past simple form, which commonly receives a stative interpretation, and the resulting sentence is ungrammatical. In (2b), the present progressive form is used, giving rise to a process interpretation and a grammatical sentence with an embedded question.

(1) **Özyıldız's (2021) Generalization about 'think'**
When 'think' composes with a question,
the resulting eventuality description must be dynamic. (Özyıldız 2021, p. 34)

(2) a. *I thought whether to invite Bill to the party.
b. I'm thinking whether to invite Bill to the party. (Dayal 2016)

Interrogative clauses may receive different interpretations depending on the embedding verb. While the embedded CPs in (3a) and (3b) appear identical, the CP under *ask* must be interpreted as a question (Q), and the CP under *tell* is understood as the answer to the question ($\textit{Ans}$-Q).

(3) a. Nene **asked** [whether Shota came]. *Q-interrogative*
⇝ Nene uttered the question "Did Shota come?"
b. Nene **told** me [whether Shota came]. *Ans-Q-interrogative*
⇝ Nene uttered the answer to the question "Did Shota come?"

This distinction raises the question of how to interpret the generalization in (1): are only *Q*-interrogatives subject to the stativity restriction, or are *Ans*-*Q* interrogatives, which resemble declaratives in having a single proposition as their content, subject to it too?

Georgian (South Caucasian) is another language that obeys Özyıldız's (2021) generalization: to embed questions, *pikr-ob-s* '(s)he thinks' must be dynamic. In (4)-(5) we see that at least in this language, the stativity ban extends to *Ans*-*Q* interrogatives. (5) cannot be understood as Keti having a belief whose content is the answer to the question *"Who is the best candidate?"*. But the sentence does allow two process readings: one where Keti is wondering about what the answer is (*Q*), and one where Keti is having a recurring thought that is the answer to the question (*Ans*-*Q*).

(4) keti-m **i-pikr-a** [**rom** šota sauk'eteso k'andidat'i-a].
Keti-ERG **pv-think-AOR.3SG** **that** Shota best candidate-BE.PRS.3SG
Stative: Keti had the belief that Shota is the best candidate.
Process: Keti kept having the thought that Shota is the best candidate.

(5) keti-m **i-pikr-a** [**tu vin** aris sauk'eteso k'andidat'i].
Keti-ERG **pv-think-AOR.3SG** **Q** **who** BE.PRS.3SG best candidate
**Stative Q:* Keti had the question "Who is the best candidate?".
**Stative Ans-Q:* Keti had a belief which was an answer to the embedded question.
Process Q: Keti entertained the question "Who is the best candidate?".
Process Ans-Q: Keti had a recurring thought that is an answer to the question "Who is the best candidate?".

The **main goal** of this paper is to derive the refined generalization in (6).

(6) **Our generalization about 'think'**
Interrogative clauses with 'think' (both *Q* and *Ans*-*Q*) are compatible with the process interpretation of the verb but not with the stative one.

Any account of this generalization needs to make an assumption about how the process 'think' and the stative 'think' are related to one another. We will assume that neither of the readings is the basic one[1]: both are derived from a more abstract meaning of the verbal root.

2 Content and Subject Matter of Thinking Situations

We propose that the verbal root $\sqrt{\text{THINK}}$ denotes a predicate of episodic mental states where the cognitive attention is directed to some propositional content:

(7) $[\![\sqrt{\text{THINK}}]\!] = \lambda s.\ \mathsf{mental\text{-}state}(s)$

For these states, there are two properties which will be relevant to us—content and subject matter:

(8) a. **Content of** s ($\mathsf{cont}(s)$) is a set of propositions restricted to the doxastic set of the attitude holder that represents *the content of thinking*.
b. **Subject Matter of** s ($\mathsf{SM}(s)$) is a partition on the doxastic set of the attitude holder that represents *the topic of thinking*.

Under this approach, *to think* is to entertain propositional content compatible with one's beliefs. Assuming a neo-Davidsonian event semantics, we will introduce arguments of the $\sqrt{\text{THINK}}$ predicate by separate theta projections; in particular, the content of the thinking situation is introduced by the functional head Θ_{CONT} (cf. Θ_{EXP} for the experiencer):

[1] Cf. for example (Krifka et al. 1995; Özyıldız 2024), where stative 'think' is derived from dynamic 'think' via genericity and (Moens and Steedman 1988), where dynamic verbs are derived from stative ones via coercion.

(9) $[\![\Theta_{\text{CONT}}]\!] = \lambda Q_{(st)t}.\ \lambda s.\ \mathsf{cont}(s) = \{p \cap \mathsf{dox}_{\mathsf{exp}(s)} \mid p \in Q\}$ *abbr* $\mathsf{cont}(s) = Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}$

The Θ_{CONT} head takes a set of propositions as its first argument. We assume the *uniformity approach* to clausal embedding, according to which both declarative and interrogative clauses denote sets of propositions (Ciardelli and Roelofsen 2015; Theiler 2014; Theiler, Roelofsen, and Aloni 2018, a.o.), and so the head in (9) will be able to compose with both types of clauses; declarative CPs in this case will denote singleton sets. The content of the thinking situation will be a set of propositions obtained by intersecting every proposition in Q with the set of worlds compatible with the beliefs of the experiencer in s; we abbreviate this as $Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}$. Additionally, we propose that the propositional content of thinking events cannot be empty (10):

(10) NON-EMPTINESS OF CONTENT: $\forall s(\mathsf{mental\text{-}state}(s) \rightarrow \mathsf{cont}(s) \neq \varnothing)$

The full denotation of *x thinks Q* is shown in (11): it's true if and only if there is a mental state of x whose content is a set of propositions from Q intersected with the doxastic state of x.

(11) $[\![\Theta_{\text{EXP}}\ x\ \text{think}\ \Theta_{\text{CONT}}\ Q]\!] = 1$ iff $\exists s[\mathsf{exp}(s) = x \wedge \mathsf{mental\text{-}state}(s) \wedge \mathsf{cont}(s) = Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}]$

Now let us discuss subject matters. We propose the constraint in (12) that relates subject matters to the content of thinking situations in terms of *aboutness* (Lewis 1988b; Lewis 1988a). Given any pair of propositions from $\mathsf{cont}(s)$ and $\mathsf{SM}(s)$, they either have an empty intersection, or the proposition in $\mathsf{SM}(s)$ is a subset of the proposition in $\mathsf{cont}(s)$.

(12) CONSTRAINT ON SUBJECT MATTERS: For any thinking situation s, $\mathsf{cont}(s) \propto \mathsf{SM}(s)$ where "$\propto$" ("about") is defined as follows:
$\forall S_1, S_2 \in D_{st,t} : S_1 \propto S_2$ iff $\forall q \in S_2, \forall p \in S_1[p \cap q = \emptyset \vee q \subseteq p]$

Let us illustrate this constraint at work with the following example. Imagine that there is a thinking situation s whose subject matter is *Who came to the party?* This subject matter is a partition of the doxastic state of the attitude holder—for our purposes, we will consider a toy model with only two individuals, Ann and Bill, and thus a 4-cell partition. This $SM(s)$ is compatible with different propositional contents, but they all must stand in the aboutness relation to the subject matter. We illustrate this in Figures 1-3. The contents in (13a) and (13b) are possible propositional contents for the subject matter at hand, but (13c) is not.

(13) a. $\mathsf{cont}(s) = \{\{s\text{'}: \text{Ann came in } s\text{'}\} \cap \mathsf{dox}_{\mathsf{exp}(s)}\}$
b. $\mathsf{cont}(s) = \{\{s\text{'}: \text{Ann came in } s\text{'}\} \cap \mathsf{dox}_{\mathsf{exp}(s)}, \{s\text{'}: \text{Ann didn't come in } s\text{'}\} \cap \mathsf{dox}_{\mathsf{exp}(s)}\}$
c. $\mathsf{cont}(s) = \{\{s\text{'}: \text{It's raining in } s\text{'}\} \cap \mathsf{dox}_{\mathsf{exp}(s)}\}$

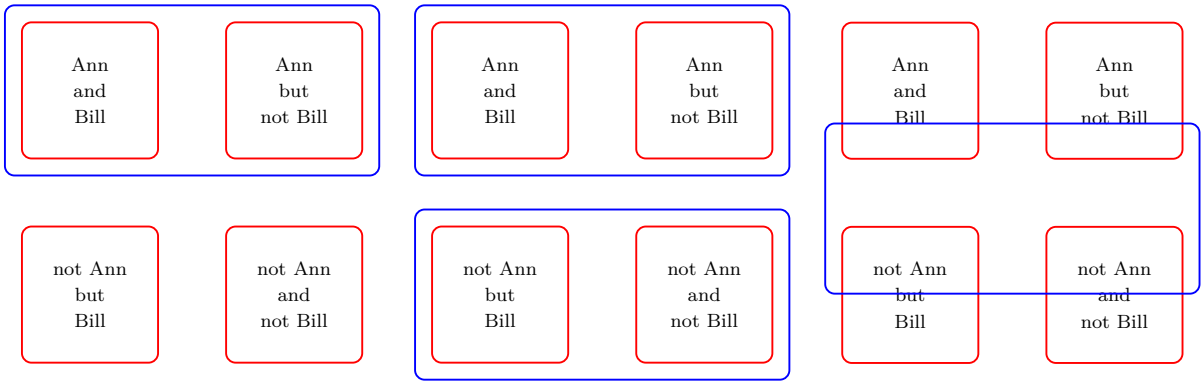


Figure 1: Ann came to the party is about Who came to the party?

Figure 2: Did A. come to the party? is about Who came to the party?

Figure 3: It's raining is not about Who came to the party?

Taking the content *Ann came to the party* in (13a), shown in Fig. 1, we see that $\mathsf{cont}(s) \propto \mathsf{SM}(s)$. This is because the single proposition in $\mathsf{cont}(s)$ makes no further distinctions between equivalence classes in the partition: it is either a superset of the propositions in $\mathsf{SM}(s)$ or has an empty intersection with them, satisfying the definition in (12). The content in (13b), which is the question *Did Ann come?*, is also about the subject matter *Who came?*, as is shown in Fig. 2. Propositions *Ann came* and *Ann didn't come* again make no further distinctions between worlds in the partition of the subject matter. Finally, the content in (13c), which is the proposition *It is raining*, is **not** about the subject matter *Who came?*, as illustrated in Figure 3. The proposition in cont(s) has a non-empty intersection with all of the propositions in SM(s), but the latter are not its subsets. In fact, cont(s) and SM(s) are **orthogonal**, that is, any two propositions where one is from cont(s) and the other is from SM(s) share a non-empty intersection:

(14) Orthogonality: $S \perp S' :\leftrightarrow \forall p \in S\ \forall q \in S' : p \cap q \neq \varnothing$

3 Stative 'Think'

We propose that the stative 'think' is derived from the basic meaning in (11) by application of the genericity operator (Chierchia 1995, a.m.o.), which results in the following denotation:

(15) $\mathsf{Gen}\ s[C(x)(s)][\exists s' \sqsubseteq s(\mathsf{exp}(s') = x \wedge \mathsf{mental\text{-}state}(s') \wedge \mathsf{cont}(s') = Q|_{\mathsf{dox}_x})]$

In all contextually restricted situations $C(x)$, there is a situation of the experiencer x having a mental state with propositional content Q. What kinds of situations might we find in the restrictor of Gen? We suggest that $C(x)$ consists of those situations in worlds maximally similar to ours where the felicity conditions for x having a thought are met and there are no inhibiting factors. We can think of these situations as pluralities of mental states experienced by x at a moment in time. (15) states that in each such plurality, we will be able to find a sub-situation whose propositional content is Q. Once we start considering pluralities of mental states and their parts, we need to make assumptions about how their subject matters are related to one another. Here we will assume the principle in (16): that the subject matter of sub-situations is the same as the subject matter of situations that they are part of.

(16) Subject Matter Inheritance:
$\forall s(\mathsf{exp}(s) = x \wedge \mathsf{mental\text{-}state}(s) \wedge (\exists s' \in C(x) : s \sqsubseteq s') \rightarrow \mathsf{SM}(s) = \mathsf{SM}(s'))$

Our proposal is that the situations in $C(x)$ are not restricted any further than what we have suggested above: the restrictor will contain absolutely any situation in which the pre-conditions for thinking are met and there are no inhibiting factors. $C(x)$ will thus, we hypothesize, include situations with all kinds of possible subject matters, and we will be able to find at least two (but probably many more) situations with orthogonal subject matters among them. We formulate this hypothesis as *diversity of subject matters* in (17): we will be able to find two situations in $C(x)$ such that each of their cells share a non-empty intersection.

(17) Diversity of Subject Matters:
$\exists s, s' \in C(x) : \mathsf{SM}(s) \perp \mathsf{SM}(s')$

Now let us see how our semantics for stative 'think' explains its behavior with declarative and interrogative CPs. For the declarative statement *x thinks p*, we will get the meaning in (18).

(18) $\mathsf{Gen}\ s[C(x)(s)][\exists s' \sqsubseteq s : \mathsf{exp}(s') = x \wedge \mathsf{mental\text{-}state}(s') \wedge \mathsf{cont}(s') = \{p \cap \mathsf{dox}_x\}]$

Suppose that the logical formula in (18) is true. Then by Diversity of Subject Matters (17), it follows that we will be able to find at least two situations in $C(x)$ with orthogonal subject

matters, say s_1 and s_2. Both of these situations have a part, let's call them s_1' and s_2' respectively, whose shared content is $\{p \cap \mathsf{dox}_x\}$. According to Subject Matter Inheritance (16), the subject matters of s_1' and s_2' will be the same as their corresponding bigger situations, and since $\mathsf{SM}(s_1) \perp \mathsf{SM}(s_2)$, they will be orthogonal to each other by extension: $\mathsf{SM}(s_1') \perp \mathsf{SM}(s_2')$. Thus, we arrive at the conclusion that if the statement is true, the attitude holder will hold mental states with the same content but of orthogonal subject matters. Given the Constraint on Subject Matters in (12), this means that the three conditions in (19) must hold simultaneously, which turns out to be possible only in two cases: either $p \cap \mathsf{dox}_x = \mathsf{dox}_x$ or $p \cap \mathsf{dox}_x = \varnothing$.

(19) a. $\{p \cap \mathsf{dox}_x\} \propto \mathsf{SM}(s_1')$
b. $\{p \cap \mathsf{dox}_x\} \propto \mathsf{SM}(s_2')$
c. $\mathsf{SM}(s_1') \perp \mathsf{SM}(s_2')$

Theorem:
the conditions in (19) are all satisfied iff either $p \cap \mathsf{dox}_x = \mathsf{dox}_x$ or $p \cap \mathsf{dox}_x = \varnothing$.

To see why this is the case, let us instead suppose that neither of these two options is true. Then $p \cap \mathsf{dox}_x$ will be a non-empty proper subset of dox_x. From $\{p \cap \mathsf{dox}_x\} \propto SM(s_1')$, it follows that $p \cap \mathsf{dox}_x$ is a superset for some of the propositions in $\mathsf{SM}(s_1')$ and has no overlap with others. Now because $\mathsf{SM}(s_1') \perp \mathsf{SM}(s_2')$, it follows that all the propositions in $\mathsf{SM}(s_1')$ will have non-empty intersections with all the propositions in $\mathsf{SM}(s_2')$, and thus all the propositions from $\mathsf{SM}(s_1')$ that $\{p \cap \mathsf{dox}_x\}$ is a superset of will partially overlap with propositions from $\mathsf{SM}(s_2')$. But this means that $\{p \cap \mathsf{dox}_x\}$ itself will partially overlap with propositions in $\mathsf{SM}(s_2')$, and so it will not be *about* $\mathsf{SM}(s_2')$—the condition in (19b) cannot be satisfied.

Hence we are left with two options: either $p \cap \mathsf{dox}_x = \mathsf{dox}_x$ or $p \cap \mathsf{dox}_x = \varnothing$. The latter, however, cannot be true due to Non-Emptiness of Content (10), so we conclude that $p \cap \mathsf{dox}_x = \mathsf{dox}_x$. But this implies that $\mathsf{dox}_x \subseteq p$; thus, we have shown that for the stative 'think' to be true with a declarative clause p, the attitude holder must believe p.

Now let us consider interrogatives with stative 'think', which will have the meaning in (20).

(20) $\mathsf{Gen}\ s[C(x)(s)][\exists s' \sqsubseteq s : \mathsf{exp}(s') = x \wedge \mathsf{mental\text{-}state}(s') \wedge \mathsf{cont}(s') = Q|_{\mathsf{dox}_x}]$,
where $Q|_{\mathsf{dox}_x}$ is a non-trivial partition of dox_x (i.e., has at least two equivalence classes)

Once again, Diversity of Subject Matters, (17), will ensure that at least two situations in $C(x)$ have orthogonal subject matters, and thus by Subject Matter Inheritance, (16), there will be at least two sub-situations with orthogonal subject matters but identical content: the partition of the doxastic state with respect to the question Q. This state of affairs turns out to be impossible—in all such cases, Constraint on Subject Matters in (12) will be violated. This is because if $Q|_{\mathsf{dox}_{exp}} \propto \mathsf{SM}(s_1')$ and $\mathsf{SM}(s_1') \perp \mathsf{SM}(s_2')$, then the propositions in $\mathsf{SM}(s_2')$ will necessarily make additional distinctions between the worlds within the cells of the $Q|_{\mathsf{dox}_{exp}}$ partition, and thus it will be necessarily false that $Q|_{\mathsf{dox}_{exp}} \propto \mathsf{SM}(s_2')$.

The remaining question that we have to explain is why *Ans-Q* interrogatives are incompatible with stative 'think' despite the fact that their propositional content is a singleton set. We would like to suggest that in sentences like (21), where the interrogative clause is interpreted as an/the answer to the embedded question, the subject matter of the thinking situation is always equal to the embedded question. Thus, (21) can be paraphrased as *"on the topic of whether Shota came, Nene has an opinion."* Given that $\mathsf{cont}(s) \propto \mathsf{SM}(s)$, the content of Nene's thoughts must be one of the two propositions—either *Shota came* or *Shota didn't come.*

(21) Nene **is certain** [whether Shota came]. *Ans-Q interrogative*
$\rightsquigarrow$ *presupposes* SM(s) = {Shota came, Shota didn't come}

This presupposition is simply incompatible with the generic operator. Consider the meaning that we will get (22): in all of the situations where pre-conditions for thinking are met, there is a sub-situation whose content is an answer to Q, and whose subject matter is Q.

(22) $\mathsf{Gen}\ s[C(x)(s)][\exists q \in Q,\ \exists s' \sqsubseteq s :$
$\mathsf{exp}(s') = x \wedge \mathsf{mental\text{-}state}(s') \wedge \mathsf{cont}(s') = \{q \cap \mathsf{dox}_x\} \wedge\ \mathsf{SM}(s') = Q|_{\mathsf{dox}_x}]$

Subject Matter Inheritance, (16), and Diversity of Subject Matters, (17), will together require that there be two distinct sub-situations whose content is an answer to Q intersected with dox_x, whose subject matter is $Q|_{\mathsf{dox}_x}$, and whose subject matters are orthogonal. The last two conditions cannot be met at the same time, as no subject matter is orthogonal to itself.

4 Process 'Think'

For process 'think,' a different operator than Gen is embedded within the basic meaning from (11). We propose that this dynamic interpretation of *think* involves a sum of mental state events; specifically, they are realized by the DO operator in (23).

(23) $[\![\text{DO}]\!] = \lambda P_{st}.\, \lambda s : \exists s', s''[s' \neq s'' \wedge s = s' \oplus s'' \wedge *P(s') \wedge *P(s'')].\, *P(s)$

This operator captures the **strictly cumulative** meaning of *dynamic eventualities*, taking a predicate of situations/events and returning only those which are strictly non-atomic. When applied to the root predicate from (7), we obtain the following denotation:

(24) $[\![\Theta_{\text{Exp}}\ x\ [\,\text{DO think}\,]\,\Theta_{\text{Cont}}\ Q]\!] = 1$ iff
$\exists s[\mathsf{exp}(s) = x \wedge [\![\text{DO}]\!](\lambda s.\mathsf{mental\text{-}state}(s'))(s) \wedge \mathsf{cont}(s) = Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}]$

To relate the content of these complex thinking situations in (24) to the contents of their subparts, we assume a **homomorphism** from the part structure of events to that of event contents (cf. Bondarenko and Elliott 2024). In this manner, the union of the sub-event contents is equated with the content of the maximal event sum in its mereology.

(25) $\forall s, s' \in \mathsf{Dom(cont)} : s \sqsubseteq s' \longrightarrow \mathsf{cont}(s) \subseteq \mathsf{cont}(s')$

This homomorphism between thinking events and their propositional contents in (25) yields the following predictions: given any situation s whose content is a singleton set, corresponding to a declarative complement p, we have $\mathsf{cont}(s) = \{p \cap \mathsf{dox}_{\mathsf{exp}(s)}\}$ and for any sub-situation $s' \sqsubseteq s$, it must also be the case that $\mathsf{cont}(s') = \{p \cap \mathsf{dox}_{\mathsf{exp}(s)}\}$ by Non-Emptiness of Content, (10). The resultant meaning is thus an **iterative** event of thinking p, which matches the 'recurring thought' interpretation of the example from Georgian with process 'think' in (4).

For interrogative complements Q, the content of a complex thinking situation s in (24) is a *plural* set of propositions which form a partition of $\mathsf{dox}_{\mathsf{exp}(s)}$, i.e. $\mathsf{cont}(s) = Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}$. By the homomorphism in (25) and Non-Emptiness of Content (10), any sub-situation $s' \sqsubseteq s$ must have as its content some non-empty subset of $Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}$, which may include more than one proposition from the partition. The resultant meaning is thus an event of entertaining different possible answers to Q, which matches the interpretation with process 'think' + Q in (5). As for process 'think' with *Ans*-Q, the content of s is also a singleton proposition corresponding to an answer to Q, which will therefore reduce to the declarative case and yield an iterative meaning, with the added presupposition that $\mathsf{SM}(s) = Q|_{\mathsf{dox}_{\mathsf{exp}(s)}}$.

5 Conclusion

Consistent with Özyıldız's (2021) generalization in (1), we confirm that Georgian 'think' embeds questions only if it receives a dynamic interpretation — this fact extends both to Q and *Ans*-Q readings for interrogative complements. We propose that 'think' starts from a basic meaning, whose stative and process interpretations result from applying different operators, Gen and DO.

References

Bondarenko, Tanya and Patrick D. Elliott (2024). *Monotonicity via mereology in the semantics of attitude reports*.

Chierchia (1995). "Individual level predicates as inherent generics". In: *The Generic Book*. Edited by Carlson & Pelletier. Backup Publisher: Chicago University Press. Chicago University Press, pp. 176–223.

Ciardelli, Ivano A. and Floris Roelofsen (2015). "Inquisitive dynamic epistemic logic". In: *Synthese* 192.6, pp. 1643–1687.

Dayal, Veneeta (2016). *Questions*. Oxford University Press. ISBN: 978-0-19-928126-8.

Grimshaw, Jane (1979). "Complement Selection and the Lexicon". In: *Linguistic Inquiry* 10.2, pp. 279–326.

Grohne, Gisela (2016). "Embedding Questions". PhD thesis. Georg-August-Universität Göttingen, Master's Thesis.

— (2017). *Aspect and Embedded Questions*. Talk at the University of Konstanz.

Krifka, Manfred et al. (1995). "Genericity: An Introduction". In: *The Generic Book*. Edited by Carlson & Pelletier. Backup Publisher: Chicago University Press. Chicago University Press, pp. 1–124.

Lewis, David (1988a). "Relevant implication". In: *Theoria* 54.3, pp. 161–174.

— (1988b). "Statements partly about observation". In: *Philosophical papers* 17.1, pp. 1–31.

Mayr, Clemens (2019). "Triviality and Interrogative Embedding: Context sensitivity, factivity, and neg-raising". In: *Natural Language Semantics* 27.3, pp. 227–278.

Moens, Marc and Mark Steedman (1988). "Temporal ontology and temporal reference". In: *Computational linguistics* 14, pp. 15–28.

Özyıldız, Deniz (2021). "The Event Structure of Attitudes". PhD thesis. University of Massachusetts, Amherst.

— (2024). "Aspect shapes the meaning of attitude reports". DGfS meeting, Bochum.

Theiler, Nadine (2014). "A multitude of answers: Embedded questions in typed inquisitive semantics". PhD thesis. (MSc Thesis), ILLC, University of Amsterdam.

Theiler, Nadine, Floris Roelofsen, and Maria Aloni (2018). "A uniform semantics for declarative and interrogative complements". In: *Journal of Semantics* 35.3, pp. 409–466.

— (2019). "Picky predicates: why believe doesn't like interrogative complements, and other puzzles". In: *Natural Language Semantics* 27.2, pp. 95–134.

White, Aaron Steven (2021). "On believing and hoping whether". In: *Semantics and Pragmatics* 14.6, pp. 1–21.

White, Aaron Steven and Kyle Rawlins (2018). "Question agnosticism and change of state". In: *Sinn und Bedeutung*. Vol. 21. 2, pp. 1325–1342.

How Conditionals Restrict

David Boylan
University of Glasgow
Glasgow, United Kingdom
david.boylan@glasgow.ac.uk

Conditionals restrict modal domains: in sentences of the form $A > MODAL\ B$, the modal quantifies only over worlds where the antecedent is true. This paper has two aims. The first is to derive these data axiomatically for basic modals, giving a logic for the conditional which has as theorems the axiomatic regimentations of the restrictor data. The second is to show that this logic does not require the conditional to be a *pure* restrictor, as in Kratzer (1981) and Kratzer (1991); it is sound on a *restricting operator* semantics, where conditionals shift worlds and accessibility relations. The restricting operator semantics offers a picture whereby conditionals can restrict without that being their only semantic function, thus sidestepping a notorious difficulty for Kratzer's semantics, the existence of bare conditionals.

1 Restrictor Data for Basic Modals

This paper focuses on restrictor data for *basic modals*, possibility and necessity modals of epistemic or historical flavour.

For basic modals, the following object language principle expresses the restrictor data for necessity modals:

Boxiness. $(A > \Box C) \leftrightarrow \Box(A \supset C)$

This is easy to see with epistemic *must*. Take a conditional like:

(1) If the butler didn't commit the crime, the gardener must have done it.

The intuitive truth-conditions here, reported by Kratzer and others, are that in all the epistemic possible worlds where the butler didn't commit crime, the gardener committed the crime. But this is just to say that in all epistemically possible worlds it is true that $\neg butler \supset gardener$.

To motivate Boxiness for historical modals, we focus on cases where it is known that what is historically necessary goes beyond what is epistemically necessary:

> *The bet.* You have a coin which is either double-heads or double-tails. You may or may not toss the coin at 12pm. I make two bets. First I bet either that the coin won't land heads: that is, I bet it won't be flipped or it will land tails. I also bet it won't land tails, that it won't be flipped or will land heads. Right now it is not settled whether the coin will be flipped.

I could describe my situation as follows:

(2) Either if the coin is flipped, I can't win the first bet; or if the coin is flipped, I can't win the second bet. (I just don't know which.)

Here the modal appears to be restricted; and it does not appear to be epistemic.

The following principle captures the restrictor data for possibility modals:[1]

[1] See Gillies (2010) for a similar principle.

Conjunctivitis. $\Diamond A \supset ((A > \Diamond\ B) \leftrightarrow \Diamond(A \wedge B))$

Why the rider that A is possible? Because when A is not possible, $A > \Diamond B$ will be trivial true but $\Diamond(A \wedge B)$ will be false.

Again, take an epistemic example first. The following seem interchangeable:

(3) a. If the butler is the murderer, the gardener might have helped him.
b. Maybe the butler is the murderer and the gardener helped him.

(3-a) says that among the epistemically possible worlds where the butler is the murderer there is some world where the gardener helped him; which is just to say there is an epistemically possible world where the butler is the murderer and the gardener helped him (given the rider that it is possible the butler is the murderer).

With historical modality, we need to again focus on cases where epistemic and historical possibility come apart. One such case is that of action. Deciding I will leave work early to go to a party this evening enables me to now know that I will not be there; but it does not yet mean that it is historically impossible for me to go. Now consider:

(4) a. If I leave work early, I could go to the party. (But I know I won't.)
b. I could leave work early and go to the party. (But I know won't.)

Here we see the same Conjunctivitis pattern; but the modal cannot be epistemic.

2 The R Logic

I show how to derive both principles in a minimal logic which I take to be a regimentation of the restrictor theory, a logic I call R.

The first principle of R is Boolean Persistence:[2]

Boolean Persistence. $A > (B > A)$, when A is Boolean.

In update terms, Persistence says that successive updates do not undo earlier ones (for Booleans at least). The plausibility of Persistence is also witnessed by the status of claims like:

(5) If it rained, then if it was cold, then it rained and it was cold.

The second principle is Flattening:[3]

Boolean Flattening. $((A \wedge B) > C)) \leftrightarrow (A > ((A \wedge B) > C))$ when A, B, C are Boolean

Given the rest of R, this principle boils down to a simpler one: when A is logically stronger than B then $A > (B > C)$ is equivalent to $B > C$. This seems correct:

(6) a. If the first coin is flipped, then if it lands head, the second coin will also land heads.
b. If the first coin lands heads, the second coin will also land heads.

The final principle is Modal CEM:

Modal CEM. $(A > (\neg C > \bot)) \vee (A > \neg(C > \bot))$

As we will shortly see, $\neg C > \bot$ is a normal modal operator and $\neg(C > \bot)$ its dual. This principle thus says, in this proprietary sense, either if A, C is necessary or if A, then C is possible.

[2] See Boylan (2024b) for further discussion of this principle.
[3] This first appears in Mandelkern (forthcoming) and attributed to Cian Dorr.

The complete R logic is obtained by adding these three principles to a restricted version of the B conditional logic.[4] Here is the entire R logic:

PC. All theorems of PC.

Identity. $A > A$

Boolean Cautious Mon. $((A > B) \wedge (A > C)) \supset ((A \wedge B) > C)$, when A, B, C are Boolean.

Boolean OR. $((A > C) \wedge (B > C)) \supset ((A \vee B) > C)$, when A, B, C are Boolean.

Boolean Flattening. $((A \wedge B) > C) \leftrightarrow (A > ((A \wedge B) > C))$, when A, B, C are Boolean.

Boolean Persistence. $A > (B > A)$, when A is Boolean.

Modal CEM. $(A > (\neg C > \bot)) \vee (A > \neg(C > \bot))$

LLE. If $\vdash A \leftrightarrow B$ then $\vdash (A > C) \leftrightarrow (B > C)$

Agglomeration. If $\vdash (B_1 \wedge ... \wedge B_n) \supset C$ then $\vdash ((A > B_1) \wedge ... \wedge (A > B_n)) \supset (A > C)$

Detachment. $A, A \supset C \vdash C$

An important theorem of this logic is a restricted version of the Import-Export Principle:

Boolean Import-Export. $((A \wedge B) > C) \leftrightarrow (A > (B > C))$, when A, B, C are Boolean.

The various restrictions throughout are to avoid triviality: as Boylan (2024b) shows, triviality results can be proved from unrestricted Persistence plus Identity and Agglomeration; and from Boolean Persistence with unrestricted Cautious Monotonicity and OR. Triviality results can also be obtained from unrestricted Flattening, given the rest of R. And of course, famously, various triviality results can be proved from unrestricted Import-Export.

3 Deriving Boxiness and Conjunctivitis

We now show how to derive both Boxiness and Conjunctivitis for each flavour of modality.

First of all, it's helpful to introduce by stipulation a conditional necessity operator: $\Box_> A$ says that $\neg A > \bot$; its dual, $\Diamond A$, says that $\neg(A > \bot)$. The B conditional logic suffices to show that $\Box_>$ is a normal modal operator. We can then prove versions of Boxiness and Conjunctivitis stated in terms of this operator:

Boxiness$_{\Box_>}$. $(A > \Box_> C) \leftrightarrow \Box_>(A \supset C)$

Conjunctivitis$_{\Diamond_>}$. $\Diamond_> A \supset ((A > \Diamond_> B) \leftrightarrow \Diamond_>(A \wedge B))$

I sketch the proofs here. For Boxiness, suppose first that $A > \Box_> B$. Given the definition of $\Box_>$, this is equivalent to a purely conditional claim, namely $A > (\neg B > \bot)$. Applying Import-Export, we derive $(A \wedge \neg B) > \bot$ and then, applying LLE, $\neg(A \supset B) > \bot$. But this is just equivalent to $\Box_>(A \supset B)$. In fact, the whole argument is a chain of equivalences: simply run it in reverse to obtain the other direction of Boxiness.

For Conjunctivis, first suppose that we have $A > \Diamond_> B$ and $\Diamond_> A$. By definition of $\Diamond_>$, the former is equivalent to $A > \neg(B > \bot)$. When $\Diamond_> A$, then $A > \neg C$ holds only if $\neg(A > C)$ holds. Thus from $A > \neg(B > \bot)$ and $\Diamond_> A$ it follows that $\neg(A > (B > \bot))$. From Import-Export it then follows that $\neg((A \wedge B) > \bot)$. This is just equivalent to $\Diamond_>(A \wedge B)$. To obtain the

[4] The B logic is due to Burgess (1981).

other direction, we run the argument in reverse except Modal CEM is used to transition from $\neg(A > (B > \bot))$ to $A > \neg(B > \bot)$

To extend these arguments to the epistemic case, note that the default reading of the indicative is usually taken to be epistemic. For instance, indicatives usually presuppose the epistemic possibility of their antecedents:

(7) a. If the butler didn't do it, the gardener did.
b. $\rightsquigarrow$ The butler might have done it.

In general, such patterns make sense if the conditional has an epistemic accessibility relation; which amounts to assuming that the modality defined from the conditional as above is epistemic necessity, as suggested by Dorr and Hawthorne (ms.):

Identity of Accessibility, Epistemic. $\Box_E A \leftrightarrow (\neg A >_E \bot)$.

Given this principle we can move between claims like $\neg A >_E \bot$ and $\neg(A >_E \bot)$ on the one hand and $\Box_E$ and $\Diamond_E$ on the other. This allows us to run the same arguments as above to obtain Boxiness and Conjunctivitis for the epistemic modal operators $\Box_E$ and $\Diamond_E$ and $>_E$.

To obtain Boxiness and Conjunctivitis for historical modal operators, I postulate a second, historical reading of the indicative, $>_H$. I assume that, on this reading, the conditional has a historical accessibility relation, giving us:

Identity of Accessibility, Epistemic. $\Box_H A \leftrightarrow (\neg A >_H \bot)$.

This allows us to move between $\neg A >_H \bot$ and $\neg(A >_H \bot)$ on the one hand and $\Box_H$ and $\Diamond_H$. We can then once more run the arguments above to obtain Boxiness and Conjunctivitis for the historical modal operators.

This historical reading of the indicative is in fact independently plausible. Consider:

> *Don't Jump!* Jones, the office worker standing on the roof of her 20 story office building, looks down over the edge towards the ground. When we looked half an hour ago, there was no net anywhere near the building; that being said, we cannot be completely certain that there isn't one hanging now. Jones is a known thrill-seeker, but she has no death wish. (Adapted from Moss (2016), §4.3)

Here we can hear true (or highly likely) readings of both of the following:

(8) If Jones jumps, she will die.

(9) If Jones jumps, she will live.

As Ciardelli and Ommundsen (2024) note, it is plausible that (8) is a historical and (9) is an epistemic reading of the conditional.

4 A Restricting Operator Semantics

I now outline a *restricting operator* semantics, on which the R logic is sound.

This semantics has three core ideas. First, sentences are evaluated relative to a pair of a world and an accessibility relation. Second, conditionals have a selection function semantics: $A > C$ is true just in case C is true at the "closest" indices where A is true. Finally, conditionals shift the accessibility relation, as well as the world: the closest accessibility relation is one that *accepts* the antecedent.[5]

[5]For important precedents, see Yalcin (2007) and Goldstein and Santorio (2021).

Here is the semantic entry for the conditional presented more precisely. The selection function f is a function from a set of world-accessibility relation pairs (i.e. a set of indices) and a world-accessibility relation pair (i.e. an index) to a set of world-accessibility relation pairs. Then we have:

Def of $>$. $\langle w, R\rangle \models A > C$ iff for all $\langle w', R'\rangle \in f([\![A]\!], \langle w, R\rangle) : \langle w', R'\rangle \models C$

In order to validate the axioms of R, various constraints are imposed on f.

A particularly important constraint is the *Update Constraint*. Say that a pair $\langle w, R\rangle$ accepts A iff for all w' such that wRw' $w', R \models A$; that is, holding the original accessibility relation fixed, A is true at all worlds accessible from w. The Update Constraint ensures that, for all elements of $f(\mathbf{A}, \langle w, R\rangle)$, the accessibility relation parameter accepts $\mathbf{A}$. Say that $R^{\mathbf{A}}$ is the set of maximal subrelations R' of R which see only points where A is accepted, relative to R' itself.

Def. of $R^{\boldsymbol{A}}$. $R^{\mathbf{A}} = max\{R' \subseteq R : \text{if } wR'w' \text{ then } \langle w', R'\rangle \in \mathbf{A}\}$

The Update Constraint says that at all the "closest" pairs where A is true, the accessibility parameter is always one of the maximal subrelations of the original which accepts A:

Update Constraint. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ then $R' \in R^{\mathbf{A}}$.

Note that usually the relations in $R^{\mathbf{A}}$ are non-reflexive: for instance, if p is false at w, we have counterexamples to reflexivity at w. To ensure that no such points are selected by f we impose:

Propriety. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ then $w' \in R'(w')$.

We can now see that both Boolean Persistence and Identity will be valid, for essentially the same reason. Identity is valid because if $\langle w', R'\rangle \in f([\![A]\!], \langle w, R\rangle)$ then, by the Update Constraint, R' is such that if $w''R'w'''$ then $w''', R' \models A$; and by Propriety $w'R'w'$. Boolean Persistence is valid because Booleans persist through multiple updates: if A is Boolean, then $(R^{[\![A]\!]})^{[\![C]\!]}$ still accepts A; and so for all $\langle w', R'\rangle \in f([\![A]\!], \langle w, R\rangle)$ if $\langle w'', R''\rangle \in f([\![C]\!], \langle w', R'\rangle)$ then $R'' \in (R^{[\![A]\!]})^{[\![C]\!]}$ and, by Propriety, $w''R''w''$.

Many of the remaining constraints are simply analogues of the usual variably strict ones. The complete list is as follows, where $f_w(\mathbf{A}, \langle w, R\rangle) = \{w' : \exists R' : \langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)\}$:[6]

Update Constraint. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ then $R' \in R^{\mathbf{A}}$.
Propriety. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ then $w' \in R'(w')$.
Visibility. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ then $w' \in R(w)$.
Non-Vacuity. If for some $R' \in R^{\mathbf{A}}$ $R'(w) \neq \emptyset$, then $f(\mathbf{A}, \langle w, R\rangle) \neq \emptyset$.
CMon. If $f(\mathbf{A}, \langle w, R\rangle) \subseteq \mathbf{B}$, then $f_w(\mathbf{A} \cap \mathbf{B}, \langle w, R\rangle) \subseteq f_w(\mathbf{A}, \langle w, R\rangle)$.
Union. $f_w(\mathbf{A} \cup \mathbf{B}, \langle w, R\rangle) \subseteq f_w(\mathbf{A}, \langle w, R\rangle) \cup f_w(\mathbf{B}, \langle w, R\rangle)$.
Flat. $f_w(\mathbf{A} \cap \mathbf{B}, f(\mathbf{A}, \langle w, R\rangle)) = f_w(\mathbf{A} \cap \mathbf{B}, \langle w, R\rangle)$.
Relational Uniqueness. If $\langle w', R'\rangle \in f(\mathbf{A}, \langle w, R\rangle)$ and $\langle w'', R''\rangle \in f(\mathbf{A}, \langle w, R\rangle)$, then $R' = R''$ and $R'(w') = R'(w'')$.

[6] Soundness for R is proved in Boylan (2024a).

5 Comparison to the Pure Restrictor Semantics

I close by comparing this semantics to its famous cousin, Kratzer's pure restrictor semantics:

Kratzer Semantics. $w, R \models A > C$ iff $w, R + [\![A]\!] \models C$, where $R + \mathbf{A}(w) = \{w : \langle w, R\rangle \in \mathbf{A}\}$.

The main difference in these views lies in how many parameters are shifted:[7] the Kratzer semantics shifts *only* the accessibility relation; the restricting operator semantics shifts both the world and the accessibility parameter. Does this make a difference?

It does. Consider bare conditionals like:

(10) If it rained, there was no party.

As Kratzer immediately noted, such conditionals pose a problem, prima facie at least, for a pure restrictor semantics: there is no modal to restrict and so the if-clause is vacuous. Kratzer posits a covert modal in sentences like (10), suggesting their true structure is something like:

(11) If it rained, MODAL there was no party.

The restricting operator semantics needs no covert modals here. When there is no modal, although the shift in accessibility relation is vacuous, the shift in world is not: (10) simply receives the standard variably strict truth-conditions for indicatives.

I think there are at least two things to be said in favour of the restricting operator approach. The first is that its explanation reflects better the structure of the data. Our earlier examples, like (1), (2), (3-a) and (4-a), show that conditionals can restrict modals; bare conditionals like (10) suggest this is not the *only* thing conditionals do. This is exactly the account given by the restricting operator semantics.

Second, the assumption of covert modals raises its own problems. A variety of authors, including Schulz (2010) and Ciardelli (2021), note that, without heavily restricting the places where this modal can appear, the theory massively overgenerates modal readings of non-conditionals. For instance, take claims like:

(12) Either it's raining or it's snowing.

(13) I doubt that it's raining.

Neither can be read respectively as:

(14) Either it must be raining or it must be snowing

(15) I doubt that it must be raining.

I thus conclude that its simple account of bare conditionals is a genuine advantage of the restricting operator theory. Further debate will examine how extensively the view can be developed. For the pure restrictor view has been applied to a large range of data, in particular, to both q-adverbs and deontic conditionals. In Boylan (2024a) I extend the semantics in this paper to capture those cases too.

Acknowledgements. For very helpful discussion, thanks to audiences at the Dubrovnik Philosophy of Language and Linguistics Course and the UMD Meaning Group and to Matt Mandelkern, Paolo Santorio and the referees for the Amsterdam Colloquium. Special thanks to Ginger Schultheis for extensive feedback.

[7] Not the only difference though: as Mandelkern (2021) shows, Kratzer's semantics fails to validate Identity. This however could be avoided by shifting instead to an element of $R^{\mathbf{A}}$, rather than simply intersecting.

References

Boylan, David (2024a). "Conditional Restriction". ms.
— (2024b). "Persistence is Trivial". ms.
Burgess, John P. (1981). "Quick Completeness Proofs for Some Logics of Conditionals". In: *Notre Dame Journal of Formal Logic* 22.1, pp. 76–84. DOI: 10.1305/ndjfl/1093883341.
Ciardelli, Ivano (2021). "The Restrictor View, Without Covert Modals". In: *Linguistics and Philosophy* 45.2, pp. 293–320. DOI: 10.1007/s10988-021-09332-z.
Ciardelli, Ivano and Adrian Ommundsen (2024). "On the ambiguity of will-conditionals". ms.
Dorr, Cian and John Hawthorne (ms.). "If...: A theory of conditionals". book manuscript. New York University and University of Southern California.
Gillies, Anthony S (2010). "Iffiness". In: *Semantics and Pragmatics* 3.
Goldstein, Simon and Paolo Santorio (2021). "Probability for Epistemic Modalities". In: *Philosophers' Imprint* 21.33.
Kratzer, Angelika (1981). "The Notional Category of Modality". In: *Words, Worlds and Contexts*. Ed. by Hans-Jürgen Eikmeyer and Hannes Reiser. Walter de Gruyter and Co., pp. 38–74.
— (1991). "Modality". In: *Semantics: An International Handbook of Contemporary Research*. Ed. by Arnim von Stechow and Dieter Wunderlich. de Gruyter, pp. 639–650.
Mandelkern, Matthew (2021). "If P, Then P!" In: *Journal of Philosophy* 118.12, pp. 645–679.
— (forthcoming). "*Bounds: the Dynamics of Interpretation*". Book ms. New York University.
Moss, Sarah (2016). *Probabilistic Knowledge*. Oxford, United Kingdom: Oxford University Press.
Schulz, Katrin (2010). "Troubles at the Semantics/Syntax Interface: Some thoughts about the modal approach to conditionals". In: *Proceedings of Sinn und Bedeutung*. Vol. 14, pp. 388–404.
Yalcin, Seth (2007). "Epistemic Modals". In: *Mind* 116.

Actuality Entailments and the Conditional Analysis*

David Boylan
University of Glasgow
Glasgow, UK
David.Boylan@glasgow.ac.uk

Ginger Schultheis
The University of Chicago
Chicago, USA
vks@uchicago.edu

1 Introduction

Consider the following ability ascription.

(1) Luc was able to lift the suitcase last night.

On one prominent interpretation, (1) entails that Luc lifted the suitcase. This implication of (1) is known as an *actuality entailment*. Actuality entailments are associated with *perfective aspect*: past perfective ability claims have actuality entailments; past imperfective ability claims do not.[1]

We give a new theory of this phenomenon. In the first part of the paper, we present a new observation: that ⌜S was able$_{\text{perfective}}$ to V⌝ presupposes that S tried to V. And we show that, given a conditional analysis of ability, we can derive the actuality entailment—that S Ved—from the presupposition—that S tried to V.

The second part of the paper concerns the source of the trying presupposition. We say that perfective aspect carries a *settledness presupposition*: ⌜Perfective P⌝ presupposes that P is settled true or settled false. The conditional analysis tells us that it is either settled true or false that S is able to V at t only if (2) is settled true or false at t.

(2) If S tried to V at t, S would V at t.

We argue that, in any ordinary circumstance, a conditional like (2) is settled true or settled false at t only if its antecedent is true. And so, in any ordinary circumstance, speakers will accommodate the settledness presupposition by presupposing that S tried to V at t.

2 Actuality Entailments and Perfective Aspect

We said past perfective ability ascriptions carry actuality entailments and that past imperfective ones do not. We can see this in a language like French in which there is an overt morphological distinction between past perfective ability ascriptions and past imperfective ones. (English does not overtly mark perfective and imperfective aspect on modals.) Consider the following sentence.

(3) #Luc a pu soulever la valise, mais il ne l'a pas soulevée parce qu'il avait peur de se faire mal au dos.
Luc was able$_{\text{perfective}}$ to lift the suitcase, but he didn't lift it because he was afraid of hurting his back.

(3) is unacceptable: you cannot say that Luc was perfectively able to lift the suitcase and, at the same time, say that he did not lift the suitcase. Contrast (3) with (4), which replaces the perfective ability claim with an imperfective one.

*Both authors contributed equally and are listed in alphabetical order.
[1]For this observation, see Bhatt (2006). For some existing theories, see Hacquard (2009) and Homer (2021).

(4) Luc pouvait soulever la valise, mais il ne l'a pas soulevée parce qu'il avait peur de se faire mal au dos.
Luc was able$_{\text{imperfective}}$ to lift the suitcase, but he didn't lift it because he was afraid of hurting his back.

(4) is acceptable: you can say that Luc was imperfectively able to lift the suitcase and, at the same time, say that he did not lift it.[2]

3 A New Observation

We observe that ⌜S was able$_{\text{perfective}}$ to V⌝ presupposes that S tried to V.

To test whether a proposition P is a presupposition of a sentence S, we check whether it projects through various embeddings: in particular, through S's negation ⌜not S⌝, the question ⌜S?⌝, and the conditional ⌜if S, . . .⌝? Our trying presupposition passes these tests.

(5) Luc n'a pas pu soulever la valise.
Luc was not able$_{\text{perfective}}$ to lift the suitcase.

(6) Est-ce que Luc a pu soulever la valise?
Was Luc able$_{\text{perfective}}$ to lift the suitcase?

(7) Si Luc a pu soulever la valise, ses parents doivent être heureux.
If Luc was able$_{\text{perfective}}$ to lift the suitcase, his parents must be thrilled.

Each of these suggests that Luc tried to lift the suitcase. Moreover, you cannot say that he was not perfectively able and, at the same time, say that he did not try: (8) is unacceptable.

(8) #Simplement en le regardant, Luc savait qu'il n'a pas pu soulever la valise. Donc il n'a pas essayé.
Just looking at it, Luc knew he wasn't able$_{\text{perfective}}$ to lift the suitcase. So he didn't try.

Only perfective ability ascriptions generate the trying presupposition: you can say that Luc was not imperfectively able and, at the same time, say that he did not try.

(9) Simplement en le regardant, Luc savait qu'il ne pouvait pas soulever la valise. Donc il n'a pas essayé.
Just looking at it, Luc knew he wasn't able$_{\text{imperfective}}$ to lift the suitcase. So he didn't try.

4 Deriving the Actuality Entailment

We show that, given a conditional analysis of ability claims, we can derive the actuality entailment of ⌜S was able$_{\text{perfective}}$ to V⌝ from its trying presupposition.

Here is the Conditional Analysis.

Conditional Analysis
⌜S is able to V⌝ is true if and only if: if S tried to V, S would V.

To see how to derive the entailment, suppose that I truly and felicitously assert:

(1) Luc was able$_{\text{perfective}}$ to lift the suitcase last night.

[2] It is standardly assumed that aspect is covert in English ability claims and that (1) is ambiguous. On one disambiguation, it contains a covert perfective operator; on another, it has a covert imperfective one. This is why the actuality entailment of (1) appears cancellable: the imperfective disambiguation has no actuality entailment.

Since my assertion of (1) is felicitous, its presupposition is satisfied: (10) is true.

(10) Luc tried to lift the suitcase last night.

Since my assertion of (1) is true, it follows from the Conditional Analysis that:

(11) If Luc tried to lift the suitcase last night, he did lift the suitcase last night.

It follows from (10), (11), and Modus Ponens that:

(12) Luc lifted the suitcase last night.

In what follows, we sharpen this account and give a story of how the presupposition arises. We state precisely the Conditional Analysis (§6) and we give our semantics for perfective aspect (§7). We then explore our theory's predictions: for perfective eventive sentences (§8) and for perfective ability sentences (§9). We begin, in §5, with the syntax of tense and aspect.

5 Syntax of Tense and Aspect

Following Cable (2021), we assume the syntactic tree of an atomic finite complementiser phrase — that is, a finite CP that does not contain any other CP — consists of a verb phrase VP dominated by an aspectual phrase AspP, which is in turn dominated by a tense phrase TP:

(13) $[_{CP}\ \emptyset_C\ [_{\text{TP}}\ \text{Tense}\ [_{\text{AspP}}\ \text{Aspect}\ [_{\text{VP}}\ \text{VP}\ ]\ ]\ ]\ ]$

The VP denotes a *temporal proposition*: a function from times to a function from worlds to truth values. For example, 'Matt walk to the store' denotes a function that takes a time t and returns a function that takes a world w and returns true if and only if an event of Matt walking to the store occurs at t in w. We assume the VP does not contain any expressions referring to times. Those occur higher in the tree: they modify tense. The VP combines with aspect to produce an aspectual phrase. This aspectual phrase denotes another temporal proposition—one that situates the time at which VP is true with respect to a *reference time*. In the absence of temporal modifiers like 'last night', we assume that the reference time is the time supplied by tense. Finally, the aspectual phrase combines with a tense morpheme to produce a tense phrase, which denotes an *atemporal proposition*—a function from worlds to truth values.

The VP may be *eventive* or *stative*. For example, 'Matt walk to the store' is eventive, whereas 'Matt know the answer' is stative. An *eventive sentence* is a sentence with the LF in (13) that has an eventive VP. An *eventive proposition* is the semantic value of an eventive sentence. A *stative sentence* is a sentence with the same LF with a stative VP. A *stative proposition* is the semantic value of a stative sentence.

6 Semantics for Ability Verb Phrases

We assume that an ability verb phrase has a syntactic structure that looks like this:

(14) $[_{\text{VP}}\ [_{\text{DP}}\ \text{Luc}]\ \text{able to}\ [_{\text{VP}}\ \text{Luc lift the suitcase}\ ]\ ]\ ]$

We say that 'able' denotes a function that takes a temporal proposition P to a function that takes a subject S and returns another temporal proposition—one that is true at a time t and world w if and only if, if S tried to bring about P at t in w, S would bring about P at t in w.

We make this more precise using a conditional operator $>$. Following Stalnaker (1968), the semantics for $>$ is given in terms of a *selection function* s that takes a world w and an atemporal proposition A and returns another world— the way things would be if A were true.

Semantics for $>$.
$> =_{df} \lambda A_{\langle s,t\rangle}\lambda C_{\langle s,t\rangle}[\lambda w.C(s(A,w)) = 1]$

We assume that the selection function obeys *Minimality*: if $w \in A$, then $s(A,w) = w$. Minimality ensures that conditionals obey *Strong Centering*: ⌜A and B⌝ is true if and only if ⌜A and if A, then B⌝ is true.

Following Thomason (2005), we use this semantics for $>$ to state our semantics for 'able'.

Conditional Analysis.
$[\![\text{able}]\!] = \lambda P_{\langle i,st\rangle}\lambda S_e\lambda t. > (\lambda w.S \text{ tries to bring about } P \text{ at } t \text{ in } w)(\lambda w.S \text{ brings about } P \text{ at } t \text{ in } w)$.

Let Try_t = λw.Luc tries to lift the suitcase at t in w. Let $Lift_t$ = λw.Luc lifts the suitcase at t in w. Then the truth conditions for (14) are:

$[\![(14)]\!] = [\![\text{able}]\!]([\![\text{Luc lift the suitcase}]\!])([\![\text{Luc}]\!]) = \lambda t\lambda w.Lift_t(s(Try_t, w)) = 1$

7 Semantics for Tense and Aspect

We assume a referential theory of tense. Tense morphemes are indexed to free variables ranging over times, whose values are supplied by a variable assignment g. Following Heim (1994), we assume that the past tense morpheme Past_i carries a presupposition of temporal precedence: $g(i)$ precedes the time of utterance t_c. This gives us the following entry for Past_i.

Semantics for Past.
$[\![\text{Past}_i]\!]^{g,c}$ is defined only if $g(i) < t_c$. If defined, $[\![\text{Past}_i]\!]^{g,c} = g(i)$.

Aspect in eventive sentences concerns the relationship between two times: the *event time* and *reference time*. Perfective aspect says the event time is contained in the reference time. Imperfective aspect says the reference time is contained in the event time. Compare:

(15) Matt walked to the store last night.

(16) Matt was walking to the store last night.

(15) has perfective aspect: it says that the time of Matt's walk is contained in the reference time—in this case, last night. (16) has imperfective aspect: it says the time of Matt's walk contains the reference time.

We offer a referential theory of aspect that closely resembles referential theories of tense. Aspectual operators are indexed to free variables ranging over times, whose values are given by g. ⌜Perf_j P⌝ says that P is true at $g(j)$. Perf_j carries two presuppositions. First is the *temporal presupposition*: $g(j)$ must be contained within the reference time. Second is the *settledness presupposition*. Where P be any temporal proposition, and $P_{g(j)}$ is the atemporal proposition that results from saturating P's temporal input with $g(j)$, ⌜Perf_j P⌝ presupposes that $P_{g(j)}$ is settled true or settled false by the end of $g(j)$.

To state the settledness presupposition, we introduce a class of *settling operators*: for any time t, $\Box_t A$ says that A is settled true by the end of t. Following Thomason (1984), we state the semantics of $\Box_t$ in terms of a historical modal base H_t that takes a world w and returns the set of worlds that are exactly like w with respect to matters of particular fact up to the end of t.

Settling Operators.
$\Box_t A = \lambda w.\text{for all } w' \in H_t(w) : A(w') = 1$

Note that very few, if any, propositions about the future are settled: if A_{t^+} is a proposition about times after t, then, almost always, there will be worlds that are exactly like the actual world in matters of particular fact up to t in which A_{t^+} is true, and worlds that are exactly like the actual world in matters of particular fact up to t in which A_{t^+} is false.

We can now state our semantic entry for Perf_j.

Semantics for Perfective.
$[\![\text{Perf}_j]\!]^{g,c} = \lambda P_{\langle i,st\rangle}\lambda t : g(j) \subseteq t \lambda w : \Box_{g(j)} P_{g(j)}(w) \vee \Box_{g(j)} \neg P_{g(j)}(w).P_{g(j)}(w)$

8 Predictions: Eventive Sentences

The eventive sentence (15), repeated below, has the logical form in (17).

(15) Matt walked to the store.

(17) Past_i [Perf_j [Matt walk to the store]]

Let $W = \lambda t \lambda w$. Matt walks to the store at t in w. Then we have the following entry for (17).

$$\begin{aligned}\text{If defined, } [\![(17)]\!]^{g,c} &= [\![\text{Perf}_j]\!]^{g,c}(W)([\![\text{Past}_i]\!]^{g,c}) \\ &= \lambda w : \Box_{g(j)} W_{g(j)}(w) \vee \Box_{g(j)} \neg W_{g(j)}(w).\, W_{g(j)}(w);\ \text{and } g(j) \subseteq g(i) \text{ and } g(i) < t_c\end{aligned}$$

The truth conditions and the temporal presupposition tell us that (17) is true only if Matt walked to the store at $g(j)$—a time that is contained in the time supplied by tense.

The settledness presupposition says that (17) is defined only if whether Matt walked to the store by the end of $g(j)$ is settled by the end of $g(j)$. We will show that this presupposition is trivially satisfied: it is satisfied no matter what world we're in. Thus the settledness presupposition is not felt to contribute anything to the meaning of (17), and so (17) is felt to be true if and only if Matt walked to the store.

The VP 'Matt walk to the store' does not contain any expressions referring to times. This has an important consequence: $W_{g(j)}$ is *entirely about* times up to $g(j)$. If $H_{g(j)}(w) = H_{g(j)}(w')$, then $W_{g(j)}(w) = W_{g(j)}(w')$. With this in mind, it is easy to see that (17)'s settledness presupposition is trivially satisfied: satisfied no matter what world we're in. Consider any world w. Either $W_{g(j)}$ is true in w or $\neg W_{g(j)}$ is true in w. Suppose $W_{g(j)}$ is true. Then, since $W_{g(j)}$ is entirely about times up to $g(j)$, it's true in all $w' \in H_{g(j)}(w)$. That is to say, $W_{g(j)}$ is settled true by the end of $g(j)$. Suppose $\neg W_{g(j)}$ is true in w. Since $W_{g(j)}$ is entirely about times up to $g(j)$, $\neg W_{g(j)}$ is also entirely about times up to $g(j)$. It follows that $\neg W_{g(j)}$ is true in all $w' \in H_{g(j)}(w)$. That is to say, $\neg W_{g(j)}$ is settled true by the end of $g(j)$. This reasoning generalizes. It is easy to show:[3]

Eventive Settling.
For any eventive proposition P, time t, and world w: either $\Box_t P_t(w) = 1$ or $\Box_t \neg P_t(w) = 1$.

9 Predictions: Ability Sentences

The following perfective ability ascription, (18), has the logical form in (19). .

(18) Luc a pu soulever la valise hier soir.

(19) Past_i [Perf_j [Luc be able to lift the suitcase]]

[3] *Proof.* Let P be an eventive proposition, t a time and w a world. Either $P_t(w) = 1$ or $\neg P_t(w) = 1$. Suppose $P_t(w) = 1$. Then since P_t is entirely about times up to t, it follows that $P_t(w') = 1$ for all $w' \in H_t(w)$, and so $\Box_t P_t(w) = 1$. Suppose $\neg P_t(w) = 1$. Since P_t is entirely about times up to t, $\neg P$ is also entirely about times up to t. It follows that $\neg P_t(w') = 1$ for all $w' \in H_t(w)$, and so $\Box_t \neg P_t(w) = 1$.

Let $Able = \lambda t \lambda w.Lift_t(s(Try_t, w))$. Then we have the following semantic entry for (19).

$$\text{If defined, } [\![(19)]\!]^{g,c} = [\![\text{Perf}_j]\!]^{g,c}(Able)([\![\text{Past}_i]\!]^{g,c})$$
$$= \lambda w : \Box_{g(j)} Able_{g(j)}(w) \vee \Box_{g(j)} \neg Able_{g(j)}(w).Able(g(j))(w);\ \text{and } g(j) \subseteq g(i) \text{ and } g(i) < t_c$$

We said that (19) presupposes that Luc tried to lift the suitcase. And we showed that, given the Conditional Analysis, we can derive the actuality entailment of (19)—that Luc lifted the suitcase—from its presupposition—that Luc tried to lift the suitcase. In the remainder of the paper, we will show how the settledness presupposition of (19) combines with the semantics of 'able' to generate the trying presupposition. We begin by showing that, given the Conditional Analysis, *one* way for hearers to accommodate the settledness presupposition of (19) is to presuppose that Luc tried to lift the suitcase. We then show that, in any ordinary circumstance, this is the only way to accommodate the settledness presupposition.

Why presupposing trying is one way to accommodate the settledness presupposition.
We show that if $Try_{g(j)}$ is true in w, then the settledness presupposition of (19) is satisfied in w: $\Box_{g(j)} Able_{g(j)}(w) = 1$ or $\Box_{g(j)} \neg Able_{g(j)}(w) = 1$. Suppose $Try_{g(j)}$ is true in w. It follows from Eventive Settling that $\Box_{g(j)} Try_{g(j)}$ is true in w. There are two cases: either $Lift_{g(j)}$ is true in w or $\neg Lift_{g(j)}$ is true in w. First suppose $Lift_{g(j)}$ is true. By Eventive Settling, $\Box_{g(j)} Lift_{g(j)}$ is true, and so $\Box(Try_{g(j)} \wedge Lift_{g(j)})$ is true in w. Strong Centering tells us that the conjunction $Try_{g(j)} \wedge Lift_{g(j)}$ entails the conditional $Try_{g(j)} > Lift_{g(j)}$. So, by the semantics for $\Box_{g(j)}$, it follows that $\Box_{g(j)}(Try_{g(j)} > Lift_{g(j)})$ is true in w: $\Box_{g(j)} Able_{g(j)}(w) = 1$. Now suppose $\neg Lift_{g(j)}$ is true in w. It follows from Eventive Settling that $\Box_{g(j)} \neg Lift_{g(j)}$ is true in w, and so $\Box(Try_{g(j)} \wedge \neg Lift_{g(j)})$ is true in w. Strong Centering tells us that the conjunction $Try_g(j) \wedge \neg Lift_{g(j)}$ entails the conditional $Try_{g(j)} > \neg Lift_{g(j)}$. So, by the semantics for $\Box_{g(j)}$, it follows that $\Box_{g(j)}(Try_{g(j)} > \neg Lift_{g(j)})$ is true in w. And finally, by the semantics for $>$, it follows that $\Box\neg(Try_{g(j)} > Lift_{g(j)})$ is true: $\Box_{g(j)} \neg Able_{g(j)}(w) = 1$.[4]

Why presupposing trying is the only way to accommodate the presupposition.
We now show that, in any ordinary circumstance, $Able_{g(j)}$ is settled true or settled false only if if $Try_{g(j)}$ is true. We begin with an assumption about when counterfactuals are settled. A counterfactual antecedent that concerns a particular time t shifts the world of evaluation to worlds that match the actual world up to some time t^- shortly before t. This time is known as the *time of the fork*. We assume that whenever the antecedent A_t of a counterfactual is false, the counterfactual $A_t > C_t$ is settled true at t only if the material conditional $A_t \supset C_t$ is settled true at the time of the fork. We call this principle *Conditional Settling*.[5]

Let $g(j)^-$ be the time of the fork for the counterfactual $Try_{g(j)} > Lift_{g(j)}$. Given Conditional Settling, this counterfactual is settled true or settled false in w only if at least one of the following conditions obtains: (a) $Try_{g(j)}$ is true in w; (b) $\Box_{g(j)^-}(Try_{g(j)} \supset Lift_{g(j)})$ is true in w; or (c) $\Box_{g(j)^-}(Try_{g(j)} \supset \neg Lift_{g(j)})$. But, in any ordinary circumstance, both necessitated material conditionals are false. That is, in any ordinary circumstance, it is possible, at the time of the fork, that Luc will try and succeed in lifting the suitcase, and it is possible that Luc will try and fail to lift the suitcase. It follows that, in any ordinary circumstance $Able_{g(j)}$ is settled true or settled false only if $Try_{g(j)}$ is true.

Acknowledgements. Thanks to audiences at the Center for Human Abilities, PhLiP 9 and NYU and to Nathaniel Baron-Schmitt, Cian Dorr, Matt Mandelkern, Barbara Vetter and the referees for the Amsterdam Colloquium.

[4] Here we assume: $\Box_{g(j)} \Diamond_{g(j)} Try_{g(j)}$, guaranteeing that $\Box_{g(j)}(Try_{g(j)} > \neg Lift_{g(j)})$ entails $\Box_{g(j)} \neg Able_{g(j)}$.
[5] See Thomason and Gupta (1980) for a similar constraint.

References

Bhatt, Rajesh (2006). *Covert modality in non-finite contexts*. Mouton de Gruyter.

Cable, Seth (2021). "A basic introduction to the semantics of aspect". Class handout for Semantics of Tense and Aspect Across Languages, University of Massachusetts, Amherst.

Hacquard, Valentine (2009). "On the interaction of aspect and modal auxiliaries". In: *Linguistics and philosophy* 32, pp. 279–315.

Heim, Irene (1994). "Comments on Abusch's theory of tense". In: *Ellipsis, Tense and Questions*. Ed. by Hans Kamp.

Homer, Vincent (2021). "Actualistic interpretations in French". In: *Semantics and Pragmatics* 14, pp. 12–1.

Stalnaker, Robert C (1968). "A theory of conditionals". In: *Ifs: Conditionals, belief, decision, chance and time*. Springer, pp. 41–55.

Thomason, Richmond (1984). "Combinations of tense and modality". In: *Handbook of Philosophical Logic: Volume II: Extensions of Classical Logic*. Springer, pp. 135–165.

— (2005). "Ability, action and context". Ms. University of Michigan.

Thomason, Richmond and Anil Gupta (1980). "A theory of conditionals in the context of branching time". In: *Ifs: Conditionals, Belief, Decision, Chance and Time*. Springer, pp. 299–322.

Priming NPI Acceptability Judgments and the Bagel Problem

Lisa Bylinina
Utrecht University
Utrecht, The Netherlands
e.g.bylinina@uu.nl

Stavroula Alexandropoulou
University College London
London, United Kingdom
stavroulaale@gmail.com

Yasutada Sudo
University College London
London, United Kingdom
y.sudo@ucl.ac.uk

Abstract

We report on a series of priming experiments on the acceptability of two types of Negative Polarity Items (NPIs) in two languages, namely, *any* in English and *libo*-NPIs in Russian. Our results show that (i) acceptability judgments of NPIs can be primed both in English and in Russian, and (ii) only unacceptablity of the same type, i.e., unacceptability caused by an unlincensed NPI, triggers priming effects on the acceptability of NPIs. The results also indicate that (iii) priming effects for English *any* and Russian *libo*-NPIs are counter-directed: For *any*, an unlicensed instance has a positive priming effect, while for *libo*-NPIs an unlicensed instance has a negative priming effect. We attribute this contrast to the so-called 'bagel' distribution of *libo*-NPIs in Russian: Although licensed in all other contexts where weak NPIs are commonly licensed, *libo*-NPIs are hardly ever licensed by negation. We claim that this makes the crucial difference for priming effects triggered by unlicensed NPIs in English and Russian. That is, unacceptable primes trigger priming effects by evoking possible ways to 'fix' them, and different NPIs evoke different such ways, depending on their typical contexts of use, whereby exhibiting qualitatively different priming effects.

1 Introduction

The general idea behind *priming* is that if a preceding trial (a *prime*) affects one's behavior in a following trial (a *target*), the two trials must have something in common. Typically priming studies involve comparing two or more types of primes (often including trivial primes, i.e., primes with no priming effects) with respect to the same target. The priming paradigm has been extensively employed to investigate various levels of linguistic representations from lexical semantics (e.g., D. E. Meyer and Schvaneveldt 1971; Neely 1976; Neely, Keefe, and Ross 1989; McNamara 2005; Neely 2012) to syntactic structure (e.g., Bock 1986; Ferreira and Bock 2006; Branigan 2007; Pickering and Ferreira 2008; Branigan and Pickering 2017) to quantifier scope (e.g., Raffray and Pickering 2010; Chemla and Bott 2015) to scalar implicatures (e.g., Bott and Chemla 2016; Rees and Bott 2018; M.-C. Meyer and Feiman 2021; Marty et al. 2024). However, it has rarely, if ever, been used in conjunction with an acceptability judgment task.

In this study, we report on a series of experiments that applied the priming paradigm to the acceptability judgments of NPIs. We tested two specific types of NPIs in two languages, *any* in English and *libo*-NPIs in Russian. The results show that acceptability judgments of NPIs can be primed both in English and in Russian, but only unacceptable primes trigger priming effects. We claim that this observation is harmonious with previous findings that large priming effects are only triggered by primes with 'unexpected' properties (e.g., Jaeger and Snider 2013; Marty et al. 2024). That is, since speakers normally do not encounter unlicensed, unacceptable NPIs in their daily lives, they trigger large priming effects. Importantly, we also observe that not all unacceptable primes trigger priming effects on acceptability judgments of NPIs, but only unacceptable primes that are unacceptable due to unlicensed NPIs do. This indicates that the priming paradigm can selectively tap into a representation or mechanism pertaining to NPI licensing. Moreover, we observe a curious crosslinguistic difference between the priming effects in English and Russian such that unlicensed *any* in English leads to a positive priming effect, but

unlicensed *libo*-NPIs in Russian leads to a negative priming effect. We attribute this contrast to the so-called 'bagel' distribution of *libo*-NPIs in Russian.

2 NPIs in English and Russian

2.1 English *any*

English *any* is seen as a prototypical NPI, more precisely a weak NPI, that is licensed by a variety of Downward Entailing (DE) operators (Fauconnier 1975; Ladusaw 1979 and much subsequent work), including clausemate nagation, as well as by certain other non-monotonic operators (see below).

One complication that will be relevant for the present study is the fact that *any* has a Free Choice (FC) use (unlike other weak NPIs like *ever*), although the exact relationship between the NPI and FC uses of *any* is immaterial for our purposes (see, e.g., Giannakidou 2001; Dayal 2004; Menéndez-Benito 2010).

2.2 Russian *libo*-NPIs

Turning now to Russian, *libo*-NPIs, which are formed by suffixing *libo* onto a *wh*-word, exhibit an interesting distribution. *Wh-libo* is an NPI and as such needs a licenser, and it is in fact licensed in most typical contexts that license canonical weak NPIs including *any* in English such as the scope of *without* and the restrictor of a universal quantifier. This is illustrated by the contrast between (1-a-b) and (1-c), where the former, but not the latter, contain an NPI licenser (*každij* 'every' and *bez* 'without').

(1) a. Každij, kto gde-**libo** byval, znaet èto čuvstvo.
every who where-LIBO was knows this feeling
'Everybody who has been anywhere knows this feeling.'
b. Ivan opozdal na urok bez kakoj-**libo** pričiny.
Ivan came.late on class without which-LIBO reason
'Ivan came to class late without any reason.' (Denić 2015)
c. *Ivan opozdal na urok po kakoj-**libo** pričine.
Ivan came.late on class on which-LIBO reason
Intended: 'Ivan came to class late for some reason.'

In light of this, it is striking that *libo* NPIs are not licensed by clausemate negation, the most prototypical NPI licenser, as in (3).

(2) *Ja ne videl kogo-**libo**.
I NEG saw who-LIBO
Intended: 'I didn't see anybody.'

Explanations of this distributional 'hole' usually involve competition with another series of pronouns, *ni*-NPIs, which are strong NPIs that are canonically licensed by clausemate negation. In fact, the distributions of *libo* NPIs and *ni* NPIs are nearly complementary, as illustrated by (3) (although with some overlap; see Padučeva 2015; Padučeva 2016; Pisarenko 2023).

(3) a. Ja ne videl {**ni-**kogo / *kogo-**libo**}.
I NEG saw who-LIBO
'I didn't see anybody.'
b. Každij, kto {*ni-gde / gde-**libo**} byval, znaet èto čuvstvo.
every who NI-where / where-LIBO was knows this feeling

'Everybody who has been anywhere knows this feeling.'

For our purposes, details of theories of *libo* NPI's 'bagel' distribution are immaterial. Importantly, *libo* is also licensed in non-negative contexts—such as deontic modals—where they exhibit non-NPI readings (for the discussion of these readings and the range of non-negative licensers, see Padučeva 2015 and Pisarenko 2023).

2.3 NPIs in non-monotonic environments

Apart from DE environments, weak NPIs can also appear in certain non-monotonic environments (Linebarger 1980; Linebarger 1987). Among such non-monotonic environments that license weak NPIs, we will focus on the scope of *exactly n* (Heim 1984; Rothschild 2006; Crnič 2014). Both English *any* and Russian *libo* NPIs are licensed in the scope of *exactly n*, as illustrated by (4) and (5).

(4) Exactly two restaurants served any vegan dishes.

(5) Rovno dva restorana dobavili v men^{j}u kakie-**libo** vegetarianskie bl^{j}uda.
exactly two restaurants added in menu which-LIBO vegetarian dishes
'Exactly two restaurants added any vegetarian dishes to their menu'

The acceptability judgments of NPIs in the scope of *exactly n* are not as crisp as those of NPIs in canonical DE environments, as previously experimentally shown by Alexandropoulou, Bylinina, and Nouwen 2020 for *any*. These intermediate judgments are advantageous for our purposes, as they are more likely to be affected by priming by more extreme judgments and may be affected in either direction. For this reason, our experiments use such sentences in target items.

3 Experiment 1: English

3.1 Methods

We collected acceptability judgments of 16 target sentences that contain *exactly n NP* in subject position and *any* in object position, as in (4). As weak NPIs are considered to be judged as more acceptable for smaller *n*'s (Heim 1984; Rothschild 2006; Crnič 2014), we used numerals between *two* and *eight*, each of which was used in two separate target items.

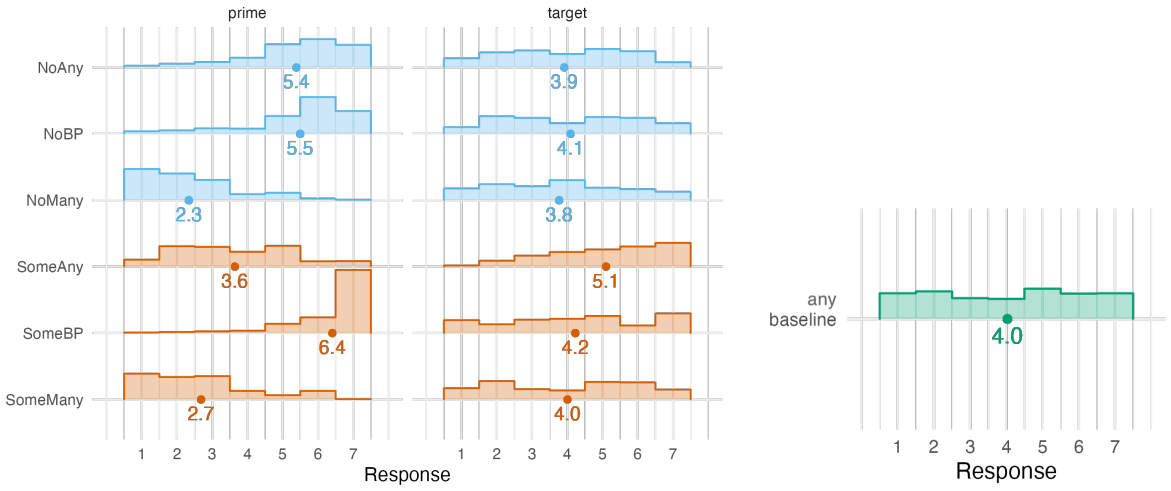
Each target sentence was preceded by two primes. There were six types of primes altogether, with *no* or *some* as the subject quantifier and one of the following as the object quantifier: (a) *any*, (b) bare plural, or (c) *many* + singular NP. The six types of primes are illustrated in (6) and (7).

(6) a. No artists sold any paintings.
b. No artists sold paintings.
c. No artists sold many painting.

(7) a. Some artists sold any paintings.
b. Some artists sold paintings.
c. Some artists sold many painting.

Regardless of the subject quantifier, (b) is expected to be grammatical, and (c) is expected to be ungrammatical due to an irrelevant grammatical factor, namely, number mismatch between *many* and the noun. On the other hand, (a) should be sensitive to the subject quantifier such that it is acceptable with *no* and unacceptable with *some*.

We had a 2 (SUBJECT: Some vs. No) × 3 (OBJECT: Any vs. BarePlural vs. Many) factorial design, with both factors being between participants. This between-participant design was chosen to avoid potential spillover effects across primes, as a target trial could be influenced by priming trials occurring earlier in the experiment, rather than only by the two preceding prime trials. In addition to the six priming conditions, furthermore, we also prepared a baseline



(a) Primed *any* (b) Unprimed *any* (baseline)

Figure 1: **Acceptability of English** *any*. (a) Ratings of primes and target items by condition in the priming experiment. (b) Ratings of target items outside of the priming context (baseline ratings).

condition, which was created by removing all the primes. The purpose of this condition was to collect baseline ratings for the target trials.

For the main six priming conditions, the 16 prime-prime-target triplets were interspersed by 72 filler items with varying acceptability. The order of stimulus presentation was randomized for each participant. Participants were randomly assigned to one of the six priming conditions. Each of them provided acceptability ratings of 120 sentences (16 prime-prime-target triplets = 48 items + 72 filler items) on a 7-point Likert scale (labelled 'Completely ungrammatical' and 'Completely grammatical' at endpoints), after reading instructions that were modelled after those used by Sprouse, Schütze, and Almeida 2013 and completing one practice item. The baseline condition was carried out in essentially the same way except that the participants saw no primes and half the filler items, so judged 48 sentences in total (16 targets + 36 filler items).

The experiment was designed on Gorilla. 90 participants were recruited on Prolific Academic (`www.prolific.com`). Each participant was compensated £3.00 for their participation and provided informed consent prior to participating. Two participants were excluded for low accuracy ($\leq 75\%$) on filler items, where accuracy was measured *post hoc*: Ratings outside 1.5 times the inter-quantile range from the median were considered wrong.

3.2 Results

The results are summarized in Figure 1. The acceptability judgments of primes (Figure 1a) are overall as expected. We fitted an ordinal mixed-effects regression model to the target data using the `ordinal` package (Christensen 2023) for R (R Core Team 2024). The model included two fixed effect variables, Subject (Some vs. No) and Object (Any vs. BarePlural vs. Many), each treatment-coded with Some and Any as reference levels, respectively. The model also had by-item variance on the intercept as the sole random effect (including any other random effect resulted in estimation error). The model reveals that target items following ungrammatical Some+Any primes were judged as more acceptable than those following grammatical No+Any primes ($\beta = -1.14, z = -7.24, p < 0.001$), and that target items following Some+Any primes were judged as more acceptable than those following the other two kinds of Some primes (BarePlural: $\beta = -0.83, z = 5.04, p < 0.001$; Many: $\beta = -1.04, z = -6.21, p < 0.001$). Baseline non-primed judgments (Figure 1b) confirm the direction of priming: Unlicensed *any* as a

prime drives amelioration in target acceptability (BarePlural: $\beta = -0.83, p < 0.001$; Many: $\beta = -1.04, p < 0.001$). In addition, significant positive interaction effects were observed (BarePlural: $\beta = 0.99, z = 4.24, p < 0.001$; Many: $\beta = 0.88, z = 3.89, p < 0.001$) that counteract the negative effect of SUBJECT and suggest that the acceptability of target items following the three types of No primes did not differ substantially from each other.

4 Experiment 2

4.1 Methods

Methods and procedure were the same as in Experiment 1. The main experiment is a priming study, followed by a baseline study where non-primed acceptability judgments were collected for the target items. The basis for the materials was translations of the English items used in Experiment 1 with *libo*-NPIs as objects in target items. Importantly, we changed the design of primes. There were 4 types of primes – each either had sentential negation (*ne* 'not') or did not have it (SUBJECT factor), and either contained *libo* or *ni* in object position (OBJECT factor). Example primes of each type are given below (items were shown to participants in Cyrillic):

(8) Xudožniki **ne** prodali kakix-**libo** kartin.
artists not sold which-LIBO paintings
'No artists sold any paintings.'

(9) Xudožniki **ne** prodali **ni**kakix kartin.
artists not sold NI-which paintings
'No artists sold any paintings.'

(10) Xudožniki prodali kakix-**libo** kartin.
artists sold which-LIBO paintings
'Artists sold any paintings.'

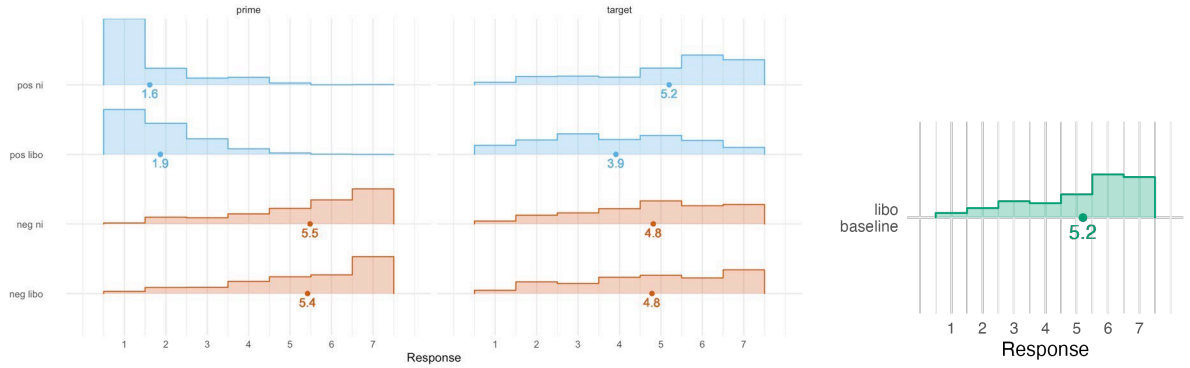
(11) Xudožniki prodali **ni**kakix kartin.
artists sold NI-which paintings
'Artists sold any paintings.'

As in Experiment 1, SUBJECT and OBJECT were between-participant factors. Each participant was randomly assigned to one of the priming conditions in the main experiment and rated 120 sentences (16 prime-prime-target triplets + 72 fillers) on a 7-point Likert scale (labeled *Užasno zvučit po-russki, tak nikto ne govorit* 'It sounds bad in Russian; no one speaks like that' and *Xorošo zvučit po-russki, ja by sam(a) mog(la) tak skazat*^{j} 'It sounds good in Russian; I could say that myself' at endpoints). In addition, there was a baseline condition that had no primes and half the fillers but was otherwise identical to the six priming conditions.

The experiment was designed using Gorilla. 68 native speakers of Russian were recruited via the Toloka platform. Each participant was compensated £3.00 for their participation and provided informed consent prior to participation. Five participants were excluded for low accuracy ($\leq 75\%$) on the filler items (same participant exclusion criterion as in Experiment 1).

4.2 Results

The obtained ratings are presented in Figure 2. As per Experiment 1, we fitted an ordinal mixed-effects regression model, with treatment-coded fixed effects: SUBJECT (absence of sentential negation (Pos) as reference level) and OBJECT (*libo* as reference level). The model revealed: (i) a difference in target acceptability between Pos+Libo and Pos+Ni prime conditions ($\beta = 2.10$, $z = 2.80$, $p < 0.01$), with the former receiving lower ratings overall; (ii) a difference in target acceptability between Pos+Libo and Neg+Libo primes ($\beta = 1.45$, $z = 1.99$, $p < 0.05$), with the former generally receiving lower ratings, and (iii) this difference being larger than the difference



(a) Primed *libo* (b) Unprimed *libo* (baseline)

Figure 2: **Acceptability of Russian** *libo*. (a) Ratings of primes and target items by condition in the priming experiment. (b) Ratings of target items outside of the priming context (baseline ratings).

between Pos+Ni and Neg+Ni (significant interaction: $\beta = -2.06$, $z = -2.03$, $p < 0.05$). Given the non-primed baseline (Figure 2b), we conclude that the Pos+Libo prime (unlicensed *libo*) degrades the *libo* target.

5 General discussion

The results indicate that acceptability judgments of NPIs under *exactly n* can be primed, but only by unacceptable primes containing an unlicensed NPI. This selective nature of NPI priming gives credence to the existence of a mental representation dedicated to NPI licensing. Licensed NPIs don't show priming effects—we attribute this to the more general property of priming that only 'unexpected events' trigger priming effects (see Jaeger and Snider 2013; Waldon and Degen 2020; Marty et al. 2024).

Our results also indicate a contrast between English and Russian: unlicensed *any* triggers a positive priming effect, while unlicensed *libo* triggers a negative priming effect. We suggest that this difference arises from an interplay between the difference in their distributions and how acceptability judgments are primed. Specifically, for Experiment 1, we hypothesize that positive sentences containing unlicensed *any* evoked in participants' mind possible ways to make them grammatical, most plausibly, if not exclusively, by adding negation. Then upon encountering the following target item with *exactly n*, they are more likely to notice that it is another way of licensing the NPI *any*. Turning now to Experiment 2, we assume that priming effects arose in a similar manner, but crucially, *libo* does not have prototypical NPI uses under direct negation and as a result, salient ways of fixing ungrammatical primes involved other types of licensors. We hypothesize that one class of licensors that participants came up with was modals, in whose scope *wh-libo* receives a non-NPI reading. Then when judging acceptability of the following target item containing *exactly n*, licensing *wh-libo* as a non-NPI was still on the participant's mind and they were less likely to notice that *exaclty n* is a possible licensor of its NPI use.

Note that English *any* also has a non-NPI use, namely a Free Choice use, under modals, but our explanation assumes that the most salient way that was evoked in Experiment 1 involved negation and other ways of licensing its NPI use. This assumption needs to be given independent evidence, and if true, a theoretical explanation, but we leave these questions for future research.

Acknowledgements. The research was supported by an AHRC-DFG grant (AHRC: AH/V003526/1; DFG: EB 523/2-1, STE 958/12-1, STE 2555/3-1) and by a grant from Toloka.

References

Alexandropoulou, Stavroula, Lisa Bylinina, and Rick Nouwen (2020). "Is there any licensing in non-DE contexts? An experimental study". In: *Proceedings of Sinn und Bedeutung*. Vol. 24. 1, pp. 35–47.

Bock, Kathryn (1986). "Syntactic persistence in language production". In: *Cognitive psychology* 18.3, pp. 355–387.

Bott, Lewis and Emmanuel Chemla (2016). "Shared and distinct mechanisms in deriving linguistic enrichment". In: *Journal of Memory and Language* 91, pp. 117–140.

Branigan, Holly (2007). "Syntactic priming". In: *Language and Linguistics Compass* 1.1-2, pp. 1–16.

Branigan, Holly and Martin J Pickering (2017). "An experimental approach to linguistic representation". In: *Behavioral and Brain Sciences* 40, e282.

Chemla, Emmanuel and Lewis Bott (2015). "Using structural priming to study scopal representations and operations". In: *Linguistic Inquiry* 46.1, pp. 157–172.

Christensen, Rune H. B. (2023). *ordinal—Regression Models for Ordinal Data*. R package version 2023.12-4.1. URL: `https://CRAN.R-project.org/package=ordinal`.

Crnič, Luka (2014). "Against a dogma on NPI licensing". In: *The art and craft of semantics: A festschrift for Irene Heim* 1, pp. 117–145.

Dayal, Veneeta (2004). "The universal force of free choice". In: *Linguistic variation yearbook* 4.1, pp. 5–40.

Denić, Milica (2015). "Minimizing Scope Ambiguity Hypothesis". Master's Dissertation. Université Paris-Diderot – Paris 7.

Fauconnier, Gilles (1975). "Polarity and the scale principle". In: *Proceedings of the Chicago Linguistic Society*. Vol. 11, pp. 188–199.

Ferreira, Victor S and Kathryn Bock (2006). "The functions of structural priming". In: *Language and cognitive processes* 21.7-8, pp. 1011–1029.

Giannakidou, Anastasia (2001). "The meaning of free choice". In: *Linguistics and philosophy* 24.6, pp. 659–735.

Heim, Irene (1984). "A Note on Negative Polarity and Downward Entailingness". In: *Proceedings of NELS*.

Jaeger, T Florian and Neal E Snider (2013). "Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience". In: *Cognition* 127.1, pp. 57–83.

Ladusaw, William A. (1979). "Polarity Sensitivity as Inherent Scope Relations". Ph.D. Dissertation. University of Texas, Austin.

Linebarger, Marcia Christine (1980). "The grammar of negative polarity". PhD thesis. Massachusetts Institute of Technology.

— (1987). "Negative polarity and grammatical representation". In: *Linguistics and philosophy*, pp. 325–387.

Marty, Paul et al. (2024). "Implicature priming, salience, and context adaptation". In: *Cognition* 244, p. 105667.

McNamara, Timothy P (2005). *Semantic priming: Perspectives from memory and word recognition*. Psychology Press.

Menéndez-Benito, Paula (2010). "On universal free choice items". In: *Natural Language Semantics* 18, pp. 33–64.

Meyer, David E and Roger W Schvaneveldt (1971). "Facilitation in recognizing pairs of words: evidence of a dependence between retrieval operations." In: *Journal of experimental psychology* 90.2, p. 227.

Meyer, Marie-Christine and Roman Feiman (2021). "Priming reveals similarities and differences between three purported cases of implicature: Some, number and free choice disjunctions". In: *Journal of Memory and Language* 120, p. 104206.

Neely, James H (1976). "Semantic priming and retrieval from lexical memory: Evidence for facilitatory and inhibitory processes". In: *Memory & cognition* 4.5, pp. 648–654.

— (2012). "Semantic priming effects in visual word recognition: A selective review of current findings and theories". In: *Basic processes in reading*, pp. 264–336.

Neely, James H, Dennis E Keefe, and Kent L Ross (1989). "Semantic priming in the lexical decision task: roles of prospective prime-generated expectancies and retrospective semantic matching." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 15.6, p. 1003.

Padučeva, E. V. (2015). "Snjataja utverditel'nost' i neveridiktal'nost' (na primere russkix mestoimenij otricatel'noj poljarnosti)/Suspended assertion and nonveridicality: the case of Russian Negative Polarity Items". In: *Russian Linguistics*, pp. 129–162.

— (2016). "Mestoimenija tipa čto-nibud' v otricatel'nom predloženii / Pronouns of the čto-nibud' type in negative sentences". In: *Voprosy jazykoznanija* 3, pp. 22–36.

Pickering, Martin J and Victor S Ferreira (2008). "Structural priming: a critical review." In: *Psychological bulletin* 134.3, p. 427.

Pisarenko, Denis (2023). ""Bagel problem" v russkom jzyke: semantika mestoimenij serij -libo i -nibud'". MA thesis. Higher School of Economics.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Raffray, Claudine N and Martin J Pickering (2010). "How do people construct logical form during language comprehension?" In: *Psychological science* 21.8, pp. 1090–1097.

Rees, Alice and Lewis Bott (2018). "The role of alternative salience in the derivation of scalar implicatures". In: *Cognition* 176, pp. 1–14.

Rothschild, Daniel (2006). *Definite descriptions and negative polarity*.

Sprouse, Jon, Carson T Schütze, and Diogo Almeida (2013). "A comparison of informal and formal acceptability judgments using a random sample from Linguistic Inquiry 2001–2010". In: *Lingua* 134, pp. 219–248.

Waldon, Brandon and Judith Degen (2020). "Symmetric alternatives and semantic uncertainty modulate scalar inference." In: *CogSci*.

Default Premise Semantics for Anankastic Conditionals

Fabrizio Cariani
University of Maryland, College Park
College Park, MD, USA
fabrizio@umd.edu

Abstract

I show that a generalization of premise semantics allows for a new set of solutions to the problems posed by anankastic conditionals.

Anankastic conditionals have been a steady source of puzzlement for the orthodox analysis of priority modals. I show that a generalization of premise semantics allows for a new set of solutions to the problems they pose. My proposal integrates standard premise semantics with established ideas from the literature on defeasible reasoning in logic and computer science — particularly Horty's theory of reasons as defaults (e.g. Horty 2012, but see also Stojnic forthcoming, Willer 2016).

1 The core problems of anankastic conditionals

Here is a small variant on the classic example from Sæbø 2001.

(1) **context:** Lila is located in Jersey city.
If Lila wants to go to Harlem, she has to cross a state line. [True]

The *compositional problem* is to characterize the update provided by the antecedent of (2). Intuitively, it seems to operate on the considerations that affect the domain (Sæbø 2001) and "ignore" the desire predicate, but the compositional mechanics that achieve this are opaque (von Fintel and Iatridou 2005, and §§4.2.4-4.2.5 of Huitink 2008).

The *conflict problem* arises if we additionally suppose that in the actual world, Lila has a desire that is incompatible with the desire introduced by the antecedent of (2) .

(2) **context:** Lila is located in Jersey city. She wants to go to Hoboken.
If Lila wants to go to Harlem, she has to cross a state line. [True]

An account of sentences like (2) must systematically predict that when actual goals and hypothetical goals conflict, the hypothetical ones win out (von Fintel and Iatridou 2005, Huitink 2008, (Condoravdi and Lauer 2016).

It is well established in the literature, that counting as an anankastic conditional is not just a matter of matching the surface form of (2). Indeed, prior to the development of the literature, Hare 1973 introduced the pair:

(3) a. If you want to put sugar in your soup, you should use a spoon.
b. If you want to put sugar in your soup, you should get checked by a doctor.

The same difference in readings occurs in variants of (3) that, like (2), involve *has to*.

(4) a. If you want to put sugar in your soup, you have to use a spoon.
b. If you want to put sugar in your soup, you have to get checked by a doctor.

The critical difference is that (4-a) is anankastic, while (4-b) is not. Anankastic conditionals are roughly characterized by the fact that they express that the content of the prejacent of the teleological modal in the consequent is a necessary condition of the proposition that is the object of the desire in the antecedent. A teleological modal is one whose domain is shaped by desires, goals, and other bouletic states of the relevant agent. (Unlike much of the literature, I do not draw a distinction between teleological and bouletic modality.)

2 Standard Solutions

It is standard (see von Fintel and Iatridou 2005, Huitink 2008, Condoravdi and Lauer 2016) to approach the compositional problem by positing a double modal, as opposed to single-modal, construal for (2) (see Figs 1 and 2). As a result, the *if-clause* shifts the world of evaluation, and not the modal base of *has to*.

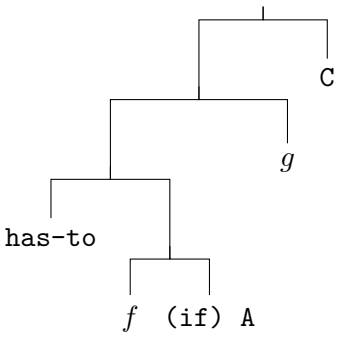


Figure 1: Single modal LF

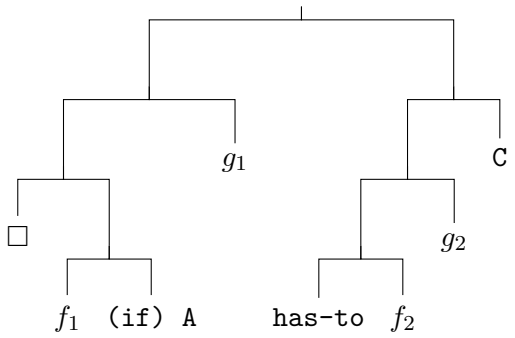


Figure 2: Double Modal LF

There is more disagreement about the conflict problem. For space reasons, I single out Condoravdi and Lauer's (2016, henceforth C&L) approach as a primary contrast to my proposal.[1] C&L distinguish between *weak* and *effective* interpretations of desire predicates. Effective desires capture the conclusions of an agent's deliberation (somewhat similar, but not identical, to the state expressed by *I intend to*). They suggest that (i) anankastics involve effective desires in the antecedent; (ii) one can't have effective desires whose satisfaction is ruled out by one's beliefs. Thus in the effective sense (5) is defective when uttered by someone who was not born in the United States.

(5) I want to become president of the United States.

The proposal encounters problems that, while perhaps not reaching the status of refutations, motivate the search for an alternative. Methodologically, it relies on substantial and under-motivated assumptions about the nature of desires. Specifically, it assumes that the perceived existence of two readings of desire predicates is to be accounted in terms of two kinds of desire-like attitudes. Empirically, I note two problems. First, there are anankastic-like sentences with weak desire predicates like (6-a), as well as other non-desire pro-attitudes, like (6-b).[2]

(6) a. If Lila would like to go to Harlem, she has to cross a state line.
b. If Lila hopes to go to Harlem, she has to cross a state line.

[1] A fuller discussion of the literature must address the well-known analyses by von Fintel and Iatridou 2005; von Stechow, Krasikova, and Penka 2006; Huitink 2008.

[2] Data of this kind was originally reported in von Fintel and Iatridou 2006 where it is attributed to Brian Weatherson (p.c.).

To account for these, the effective analysis would have to extend beyond the domain of simple desire ascriptions, and turn into a sweeping ambiguity claim.[3] Additionally, the analysis falters if in some worlds the agent doesn't believe that the desires are incompatible, as pointed out by Phillips-Brown (2019).

In contrast, the approach advocated here does not require specific assumptions about desire-talk or the metaphysics of desire, nor does it require a double-modal construal of anankastics.[4]

3 The idea of default premise semantics

Standard premise semantics for deontics assumes an ordering source that is a function from worlds to sets of propositions. Inspired by the formal theory of reasons in Horty 2012 this is replaced here by a *default ordering source* (a function from worlds to) sets of *pairs* of propositions $[\mathbf{A} \Rightarrow \mathbf{B}]$. A helpful heuristic is to think of this as expressing that **A** functions as a reason for **B**.[5] We refer to $[\mathbf{A} \Rightarrow \mathbf{B}]$ as a *consideration*; to **A** as the *premise* and to **B** as the *conclusion*

This enriched perspective on ordering sources is not just motivated by the application to anankastic conditionals, even though that is the only case discussed in this talk. Default premise semantics is a strict generalization of classical premise semantics that is capable of modeling a variety of phenomena including the puzzles of information-sensitive *oughts* (Cariani, M. Kaufmann, and S. Kaufmann 2013; Charlow 2013b; Kolodny and MacFarlane 2010; Willer 2016).

The formalism presented here is based on the idea of filtering the context-initialized default ordering source twice over. In the first stage, we only restrict attention to those considerations whose premise is entailed by the relevant modal base at the world of evaluation. In the second stage, we resolve those conflicts among triggered considerations that can be resolved by attending to the relevant priority of different considerations. (Not all conflicts are of this sort: see the discussion in Cariani ms. specifically §2.4). Fig. 3 displays the architecture of the resulting theory.

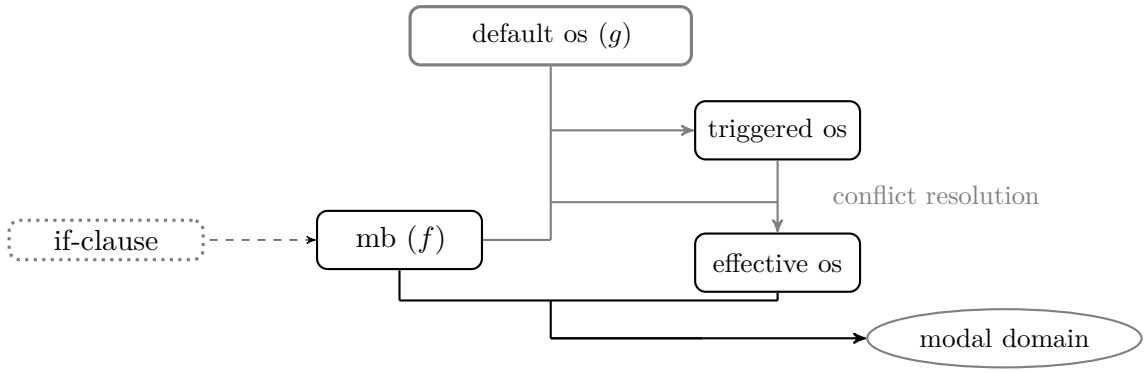


Figure 3: Figure 3: Architecture of Default Premise Semantics

[3]For the record, C&L (§7.2.3) disagree, and propose a sketch of an account of (6-a) on which these emerge as "near-anankastic readings" based on contextual reasoning. However, the proposed account is rather promissory and incomplete, so the difficulty for their view cannot be treated as resolved.

[4]One type of theory that agrees with the present one on not requiring the double modal construal is advanced by Charlow 2013a. Charlow states a constraint on a selection function for deontic conditional which covers, in more schematic form, similar ground as the domain construction recipe identify below.

[5]For empirical reasons too complex to detail here, the system cannot be replicated with single conditional propositions in the ordering source. See the full paper Cariani ms.

4 Formalities of default premise semantics

Recall that a member $[\mathbf{A} \Rightarrow \mathbf{B}]$ of a default ordering source is a **consideration**. To extract the two elements of a consideration, rely on the functions $prem([\mathbf{A} \Rightarrow \mathbf{B}]) = \mathbf{A}$ and $con([\mathbf{A} \Rightarrow \mathbf{B}]) = \mathbf{B}$. A **default ordering source** g is a function from worlds to sets of considerations. When Σ is a set of propositions $\Sigma \sqsubseteq \mathbf{A} =_{def} \bigcap \Sigma \subseteq \mathbf{A}$, (read: " Σ entails $\mathbf{A}$"). The considerations in $g(w)$ are ordered by a strict partial order $\succ$ tracking their relevant importance.

In (7) and (8), we identify the two stages of filtering:

(7) **triggered ordering source**: $\mathbf{tr}(f, g, w) = \{\pi \in g(w) \mid f(w) \sqsubseteq prem(\pi)\}$.

(8) **effective ordering source**: $\mathbf{eff}(f, g, w, \succ) = \{\pi \in g(w) \mid \neg\exists \pi' \in tr(f, g, w)$
$(\pi' \succ \pi \wedge f(w) + con(\pi') \neq \emptyset \wedge f(w) + con(\{\pi, \pi'\}) = \emptyset\}$

(7) filters out the considerations whose premise is not entailed by the modal base; (8) eliminates considerations that are defeated by a higher ranking incompatible consideration.

Given an effective ordering source as in (8), stipulate the limit assumption, and follow the Kratzerian (e.g. 2012) recipe for determining the domain of `has-to`:

(9) a. $w \preceq_{f,g,z} u$ iff $\{\mathbf{P} \in con(\mathbf{eff}(f, g, z)) \mid w \in \mathbf{P}\} \subseteq \{\mathbf{P} \in con(\mathbf{eff}(f, g, z)) \mid u \in \mathbf{P}\}$
b. $\mathbf{dom}(f, g, z) = \{u \in f(z) \mid \neg\exists w, w \geq_{f,g,z} u\}$

To complete our account of deontic conditionals, let us capture Kratzer's 1991 account of conditionals as restrictors of modals as a constraint on the semantic value of certain complex LF's, and specifically:

(10) $[\![\texttt{has-to}_{f,g}(\texttt{if}\ \cdot)[\cdot]]\!] = \lambda\mathbf{A}\lambda\mathbf{C}\lambda w.\forall u \in \mathbf{dom}(\lambda z.f(z) \cap \mathbf{A}, g, w), u \in \mathbf{C}$

This analysis is in need of further refinement to deal with chaining of considerations, but it will do for present purposes.[6]

5 New solutions

.

While default premise semantics is a fully general framework, solving the two core problems of anankastic conditionals requires novel specific assumptions. With regards to the compositional problem, the key assumption is that in teleological flavors all instances of the schema [**want** $\mathbf{A} \Rightarrow \mathbf{A}$] that are relevant in w belong to the teleological ordering source at w. Informally, wanting a proposition to be true is a defeasible reason to bring about that proposition.

With regards to the *conflict problem*, having represented the priority relation $\succ$, we need to enforce the constraint that the *local goal* (e.g. going Harlem) is to take priority over the *global goal* (e.g. going to Hoboken). To this end, we first represent the distinction between local and global considerations. Consider evaluation parameters $\langle f, g, w\rangle$, and a modal conditional `if A, have-to C`, and a consideration γ. Stipulate:

> γ is globally triggered (relative to $\langle f, r\rangle$) iff $f(w)$ entails $prem(\gamma)$.
> γ is locally triggered (relative to $\langle f, r\rangle$) iff $f(w) + \mathbf{A}$ entails $prem(\gamma)$, but $f(w)$ does not.

[6] Suppose for instance that the ordering source at w contains $[\mathbf{A} \Rightarrow \mathbf{B}]$ and $[\mathbf{B} \Rightarrow \mathbf{C}]$, but that the modal base at w only entails $\mathbf{A}$. Once $\mathbf{B}$ is triggered by $\mathbf{A}$, we want it in turn to trigger $\mathbf{C}$. This is not provided by the present scheme. There are established tools to capture this kind of chaining, see Horty 2012; Horty 2014, as well as the coarse solution in Cariani ms.

As a first step in our account of the conflict problem, we consider the constraint:

LOCALITY: locally triggered considerations take priority over globally triggered ones.

LOCALITY is intended as a pragmatic constraint on the priority relation $\succ$. The basic idea behind locality is that highlighting a goal, e.g. with a conditional antecedent, results in that goal taking priority over others. In this respect, despite the substantial difference in implementation, we are after similar intuitions to proposals that treat the goal introduced by the antcedent as special or designated (e.g. von Fintel and Iatridou 2005). Thus understood, LOCALITY entails the desired result – i.e., that *local goal* (e.g. Harlem) $\succ$ *global goal* (e.g. Hoboken).

Despite this success, there are challenging cases in which LOCALITY seems to overgenerate. As we noted, it is a plausible lesson of the data on anankastic conditionals that, when it comes to teleological modals, desires in the antecedent take priority over incompatible desires emerging from the global context. However, LOCALITY makes the additional prediction that this should happen also when the antecedent triggers some piece of information that is not itself a desire. That this prediction is wrong can be seen with fairly simple examples:

(11) **context:** Your fancy restaurant meal is approaching the end, and a gruesome surprise is announced: there will only be one dessert, cake, and it is poisoned.

a. If the cake is tasty, you have to eat the cake. [False]
b. If you want to eat dessert, you have to eat the cake. [True]
c. If the cake is tasty and you want dessert, you have to eat the cake. [True]

The analysis that combines default premise semantics and locality hypothesis incorrectly predicts that (11-a) should be true and more generally that (11-a), (11-b), and (11-c) should be on a par. To see why, note that it is plausible that in the provided context, there is a consideration roughly representable as

(12) [**the cake is tasty** $\Rightarrow$ **eat the cake**].

This consideration is not triggered by the global context (which leaves it open that the cake is not tasty), but it is triggered by the local context of the modal in (11-a) after the update with the antecedent. A seeming consequence of the locality hypothesis is that this consideration should take priority over ones that emerge from the global context—including the unconditionally triggered consideration to not get sick. But if it does take priority, the theory predicts that (11-a) is true. The worry, then, is that by relying on something as general as LOCALITY, we have missed out on what's distinctive about anankastics. A symptom of this is that the LOCALITY does not distinguish between (11-a), (11-b), and (11-c).

A natural next idea is to restrict the locality hypothesis to cases that can properly be characterized as instances of *conditioning on desires or goals*. A first hunch would be to think that the hypothesis only applies to considerations whose enabling condition is a pro-attitude (e.g., a desire, an intention, a hope).

However, it is not clear that this is the right restriction. There are cases like (13-a) below where the antecedent picks out a state that is not even a state of mind.

(13) **context:** Hansel and Gretel are driving on a freeway, they need to get to their destination as fast as possible. Hansel says:

a. If you are hungry for lunch, we have to stop at this gas station. [True]

Here is a way out of this impasse: although being hungry is not itself a desire, it contextually entails one. Someone who is hungry typically wants to eat food. Similarly, the antecedent of (11-c) — *the cake is tasty and you want to eat dessert* — is not itself a desire, but it entails one.

Accordingly the locality hypothesis might be restricted to antecedents that contextually entail constraints on one's structure of norms, desires, and goals.

Say that a proposition **A** is **teleologically relevant** iff **A** entails a proposition **B** that itself expresses a norm, desire or goal (or lack thereof).

> REVISED LOCALITY: When a global modal base f for a teleological modal gets updated by a teleologically relevant proposition **A** to a new modal base f^+, considerations that are triggered by f^+ but not by f have higher priority (in the sense of $\succ$) over considerations that are triggered by f.

This revised form of the locality hypothesis appears to strike the right middle ground between prioritizing the considerations that matter most directly to a teleological modal and not prioritizing ones that do not. Of course, deriving this sort of specific pragmatic constraints from more general pragmatic principles requires additional work, which is not carried out here. The intuition that carries the constraint is fairly clear nonetheless: mentioning a desire – or indeed anything that matters to desire-like states – has the effect of giving the considerations associated with that desire extra significance. As noted, this idea is convergent with accounts that maintain that the content of the desire in the antecedent gets to play a special role (e.g. von Fintel and Iatridou 2005; von Stechow, Krasikova, and Penka 2006). But it builds in flexibility so that the account does not face some of the same difficulties. Thus, C&L attack these alternate proposal for their inability to deal with certain "near-anankastics", like:

(14) If you want to go to Disneyland, you have to spend at least five days there.

As they point out, (14) is not well understood by paraphrasing it as *To go to Disneyland, you have to spend at least five days there.* By contrast, it is entirely unproblematic for the default analysis once the details about chaining discussed in footnote 6 are addressed. Summing up, the present analysis allows us to separate the idea of anankastics as involving "goal promotion" from its implementation in terms of rationale clauses.

6 Conclusion

I have developed default premise semantics, and within it new proposals for a semantic treatment of anankastic conditionals. From a motivational point of view, default premise semantics is inspired by the idea that deontic claims are made true by reasons. The technical reflection of these motivational considerations is the assumption that considerations should have the binary structure that is attributed to reasons in default models of reasons. This innovation can be imported into the classical semantics for modals in a modular and conservative way. It is modular because it only involves how considerations shape the domain of a priority modal. It is conservative because it preserves the structure of the original framework, limiting its intervention to the local structure of reasons. Under the weak (though not uncontroversial) assumption that having a desire with a proposition as content is a defeasible reason for one to bring about that proposition, the framework solves the compositional problem without invoking a dual modal analysis of deontic conditionals. Furthermore, the framework vindicates Saebo's (2001) idea that anankastics involve a kind of update to the ordering source. They indeed do, but this update is best viewed from the perspective of a filtered ordering source, and it is mediate by the standard update of the modal base. For the conflict problem, I explored the prospects for a pragmatic constraint, to the effect that the locally triggered considerations should be given higher priority. While the unrestricted version of the constraint overgenerates, the constraint can be restricted appropriately to those updates that match the flavor of the modal. This too matches the motivating ideas of important contributions in the literature (von Fintel and Iatridou 2005; von Stechow, Krasikova, and Penka 2006) while offering a new technical implementation.

References

Cariani, Fabrizio (ms.). "Anankastic conditionals and the default theory of reasons". University of Maryland, College Park. Available at `https://philpapers.org/rec/CARACA-22`.

Cariani, Fabrizio, Magdalena Kaufmann, and Stefan Kaufmann (2013). "Deliberative Modality under Epistemic Uncertainty". In: *Linguistics and Philosophy* 36, pp. 225–259.

Charlow, Nate (2013a). "Conditional preferences and practical conditionals". In: *Linguistics and philosophy* 36, pp. 463–511.

— (2013b). "What We Know and What To Do". In: *Synthese* 190, pp. 2291–2323.

Condoravdi, Cleo and Sven Lauer (2016). "Anankastic conditionals are just conditionals". In: *Semantics & Pragmatics* 9.

von Fintel, Kai and Sabine Iatridou (2005). "What to do if you want to go to Harlem?" MIT.

— (2006). "Anankastic Conditionals and Related Matters". slides from a talk given in Tuebingen, MIT. URL: `https://web.mit.edu/fintel/fintel-2006-harlem-tubingen.pdf`.

Hare, R. M. (1973). "Wanting: Some Pitfalls". In: *Agent, Action, and Reason*. Ed. by Roger Trigg. Wiley-Blackwell, pp. 81–127.

Horty, John F. (2012). *Reasons as Defaults*. New York & Oxford: Oxford University Press.

— (2014). "Deontic modals: Why abandon the classical semantics?" In: *Pacific Philosophical Quarterly* 95.4, pp. 424–460.

Huitink, Janneke (2008). "Modals, conditionals and compositionality". Radboud Universiteit Nijmegen. PhD thesis.

Kolodny, Niko and John MacFarlane (2010). "Ifs and oughts". In: *Journal of Philosophy* 107.3, pp. 115–143. ISSN: 0022-362X.

Kratzer, Angelika (1991). "Conditionals". In: *Semantics: An International Handbook of Contemporary Research*. Ed. by A. von Stechow & D. Wunderlich. from the *Semantics Archive*. De Gruyter.

— (2012). *Modals and Conditionals*. New York Oxford: Oxford University Press.

Phillips-Brown, Milo (2019). "Anankastic conditionals are still a mystery". In: *Semantics & Pragmatics* 12.13, pp. 1–19.

Sæbø, Kjell Johan (2001). "Necessary conditions in a natural language". In: *Audiatur vox sapientiae: A Festschrift for Arnim von Stechow*, pp. 427–449.

Stojnic, Una (forthcoming). "The inferential constraint and the *if* ϕ, *Ought* ϕ problem". In: *Philosophical Studies*.

von Stechow, Arnim, Sveta Krasikova, and Doris Penka (2006). "Anankastic conditionals again". In: *A Festschrift for Kjell Johan Sæbø: In partial fulfilment of the requirements for the celebration of his 50th birthday*, pp. 151–171.

Willer, Malte (2016). "Dynamic Foundations for Deontic Logic". In: *Deontic Modals*. Ed. by N. Charlow and M. Chrisman. Oxford University Press, pp. 324–354.

Evaluativity in *even*-hosting comparatives: Information structure and salient scales

Zhuang Chen
Bar-Ilan University
Ramat Gan, Israel
chenzhuangcq@gmail.com

Abstract

Even-hosting comparatives trigger an evaluative inference that is absent in *even*-less counterparts. Greenberg 2018 offers a degree-based account to formally explain this phenomenon, but this account is criticized by Bi 2022 for (i) failure to capture cases where *even* associates with the comparison target and, more fatally, (ii) running the risk of over-generation. Arguing for the degree-based account, this work shows how it can systematically explain this phenomenon. More crucially, by offering a novel, information structure-based argument, we claim that this account is free of any over-generation risk, and further demonstrate how this novel perspective presents a simple account for cases involving negation. Thus, apart from addressing this phenomenon per se, this work sheds light on how information structure helps to constrain the choice of the salient scales with scalar, alternative-sensitive expressions.

1 Introduction

English comparatives per se do not trigger any evaluative inference in the sense of Rett 2014, but intriguingly, it is observed that with the particle *even*, an evaluative inference arises. For instance, (1) triggers no evaluative inference that Alex and Bill are both tall, nor the inference that they are both short. In contrast, (2-a), where *even* associates with Bill, implies that they are both tall (see e.g., Ippolito 2007; Greenberg 2015; Greenberg 2018); and (2-b), where *even* associates with Alex, implies that they are both short (see e.g., Bi 2022). Why does such an evaluative inference arise in *even*-hosting comparatives? There must be some underlying systematic mechanism since similar observations are also made for Russian *daže*, Hebrew *afilu* and German *sogar* (Miashkur and Greenberg 2019), and Mandarin *even*-like *dōu* (Guo 2022).

(1) Alex is taller than Bill. ⇝ no **Both are tall / short** inference

(2) a. Alex is **even** taller than $[\text{Bill}]_F$ ⇝ implying **Both are tall**
b. $[\text{Alex}]_F$ is **even** taller than Bill. ⇝ implying **Both are short**

In the formal linguistics literature, Greenberg 2018 offers an account that assumes a degree-based semantics of English *even* (the degree-based account below), partly addressing the evaluativity issue in *even*-hosting comparatives. This account, however, is criticized by Bi 2022 for, among other issues, (**i**) failure to capture cases like (2-b) where *even* associates with the comparison target and, more fatally, (**ii**) running the risk of over-generation regarding the choice of the salient scale (see below). Instead, Bi 2022 offers an alternative account that assumes the canonical likelihood-based semantics of *even*.[1] This paper argues in defense of the degree-based account. Specifically, we will illustrate how it CAN systematically capture the evaluative inference with both (2-a) and (2-b) (Sect. 2), and more crucially, we offer a novel, information structure-based argument which displays that the choice of the salient scale can be independently made, thus ruling out the over-generation risk (Sect. 3). Then, this novel argument, with the degree-based account, is extended to cases involving negation (Sect. 4) before we conclude (Sect. 5).

[1] For space limits, we refer readers to Bi 2022 for details.

2 Systematically capturing (2) via the degree-based account

2.1 Greenberg's (2018) degree-based account for English *even*

Canonically, English *even* is argued to presuppose that the prejacent p is less likely than all its distinct alternatives q in the context C and assert that the prejacent is true (Karttunen and Peters 1979; Rooth 1985; Chierchia 2013), roughly formulated as (3) [2]:

(3) $\|even\|^{g,c} = \lambda C \lambda p. : \forall q \in C\ p \neq q \rightarrow p <_{likely} q.p(w)$

This line of analysis has been challenged in terms of, among other issues, the nature of the scale *even* operates on (see e.g., Kay 1990; Rullmann 1997; Zhang 2022). In particular, Greenberg (2016, 2018) shows that p being less likely than q is neither necessary nor sufficient to license *even*, as respectively illustrated in (4), where the prejacent is actually more likely than its alternative but *even* is felicitous, and in (5), where giving birth to a boy and to a girl are equally likely and are both less likely than giving birth, but *even* is felicitous only with giving birth to a boy. Instead, as argued by Greenberg, the scale in (5) is contextually decided, e.g., happiness of Princess Jane.

(4) (Seller to client:) Both tools are strong. The one on the right is made of strong aluminum, and the one on the left is **even** made of $[\text{steel}]_F$. (Greenberg 2016)

(5) Context: Any princess who gives birth can stay in the palace. If she gives birth to a boy, she becomes a queen.
Princess Jane gave birth. She (**even**) gave birth to $[\text{a boy}]_F$/#$[\text{a girl}]_F$. (Greenberg 2018)

In addition, Greenberg 2018 observes that *even* requires both the prejacent and its contextually salient alternative to indicate a degree above the standard on some scale. Consider (6). *Even* is licensed only in Seller A's reply, where both tools are above the standard on the scale of strength.

(6) Context: W.r.t. strength, plastic < aluminum < NORM < iron < steel. (Greenberg 2018)
Client:I need a strong tool. What about the red and blue tools over there?
Seller (a):The red one is made of iron and the blue one is (**even**) made of $[\text{steel}]_F$.
Seller (b):The red one is made of plastic and the blue one is (#**even**) made of $[\text{aluminum}]_F$.
Seller (c):The red one is made of plastic and the blue one is (??**even**) made of $[\text{steel}]_F$.

Given such observations, Greenberg 2018 proposes a degree-based semantics of *even*, which carries a twofold presupposition (7). (see Zhang 2022 for a similar view) Specifically, as in (7), *even* contributes (**i**) a comparative presupposition (7-a) that some non-focused item x in the prejacent p has a higher degree on a scale associated with a contextually supplied gradable property G in accessible w_1 worlds, worlds where the prejacent p holds, than in accessible w_2 worlds, worlds where the alternative q holds but the prejacent p is false, and (**ii**) an evaluative presupposition (7-b) that this x is above the standard on the G scale in accessible w_2 worlds (; given that x ranks higher on the G scale in accessible w_1 worlds, x is above the standard on the G scale in accessible w_1 worlds as well).

(7) $\|even\|^{g,c} = \lambda C.\lambda p.\lambda w. : \forall q \in C[q \neq p \rightarrow \forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_1 \in p \wedge w_2 \in [q \wedge \neg p] \rightarrow$
a. $[\max(\lambda d_1.G(d_1)(x)(w_1)) > \max(\lambda d_2.G(d_2)(x)(w_2)) \wedge$ **Comparative Ps.**
b. $\max(\lambda d_2.G(d_2)(x)(w_2)) > \text{Stand}_G]]]$ **Evaluative Ps.**
c. $p(w)$ **Assertion**

[2] Debates exist regarding whether *even* also has an additive presupposition that some alternative to p is true in the world of evaluation, which is orthogonal to us. See e.g., Rullmann 1997; Wagner 2013; Greenberg 2016.

2.2 How the degree-based account systematically captures (2-a) and (2-b)

First, let's illustrate how entry (7) accounts for (2-a) (repeated as (8-a)), where *even* associates with the comparison standard Bill.[3] For simplicity, let's assume that the focused item Bill has only one relevant alternative in the context, say, Chris, thus an alternative set as in (8-b).

(8) a. Alex is **even** taller than $[\text{Bill}]_F$. $\rightsquigarrow$ implying **Both are tall**
b. Alt. Set ={Alex is taller than Bill, Alex is taller than Chris.}(prejacent underlined)

Assuming for (8-a) that the non-focused item is Alex and that the contextually salient gradable property G is `tallness` (see more on this below), we obtain interpretation (9) via entry (7):

(9) $\|(8\text{-a})\|^{g,c} = . : \forall q \in C[q \neq p \rightarrow \forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_1 \in p \wedge w_2 \in [q \wedge \neg p] \rightarrow$
a. $[\max(\lambda d_1.\text{tallness}(d_1)(\text{Alex})(w_1)) > \max(\lambda d_2.\text{tallness}(d_2)(\text{Alex})(w_2)) \wedge$ **Com. Ps.**
b. $\max(\lambda d_2.\text{tallness}(d_2)(\text{Alex})(w_2)) > \text{Stand}_{\text{tallness}}]]]$ **Eval. Ps.**
c. Alex is taller than Bill. **Assertion**

where p = Alex is taller than **Bill**, q = Alex is taller than **Chris**.

In prose, the comparative presupposition says that Alex ranks higher on the `tallness` scale in accessible w_1 worlds, worlds where p (Alex exceeds Bill in height) holds, than in accessible w_2 worlds, worlds where $[q \wedge \neg p]$ (Alex exceeds Chris but not Bill in height) holds. This is trivially satisfied. But crucially, the evaluative presupposition says that Alex is above the standard on the `tallness` scale in w_2 worlds; given that in w_2 worlds Alex does not exceed Bill in height, Bill must be also above the standard on the `tallness` scale.[4] Note that the focused item, i.e., Bill in this case, remains invariant across w_1 and w_2 worlds w.r.t. G (i.e., `tallness`), and is thus above the standard on the `tallness` scale in accessible w_1 worlds, too. Now the prejacent asserts that in height Alex exceeds Bill, whose `tallness` is invariant and above the standard on the `tallness` scale, Alex is therefore also inferred to be above the standard on the `tallness` scale, thus the evaluative inference that they are both tall. (see Fig. 1)

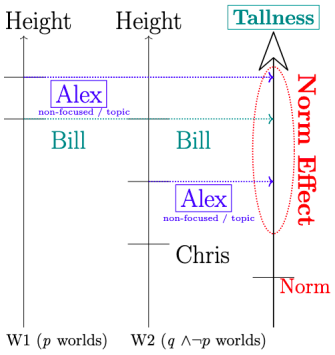


Figure 1: Interpretation of (8-a) via (7)

Crucially, pace Bi 2022, we argue that entry (7) naturally accounts for (2-b)(repeated as (10-a)) as well, where *even* associates with the comparison target Alex, and evokes the inference that Alex and Bill are both short. Likewise, for simplicity, we assume Chris as the only relevant alternative to the focused item Alex in the context, thus the alternative set (10-b).

(10) a. $[\text{Alex}]_F$ is **even** taller than Bill. $\rightsquigarrow$ implying **Both are short**
b. Alt. Set = {Alex is taller than Bill, Chris is taller than Bill.}(prejacent underlined)

Let's assume for (10-a) that the non-focused item x is Bill, and that the contextually supplied gradable property G is `shortness` (see more on this below). Then, applying entry (7) to (10-a) returns interpretation (11):

(11) $\|(10\text{-a})\|^{g,c} = . : \forall q \in C[q \neq p \rightarrow \forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_1 \in p \wedge w_2 \in [q \wedge \neg p] \rightarrow$
a. $[\max(\lambda d_1.\text{shortness}(d_1)(\text{Bill})(w_1)) > \max(\lambda d_2.\text{shortness}(d_2)(\text{Bill})(w_2)) \wedge$ **C. Ps.**
b. $\max(\lambda d_2.\text{shortness}(d_2)(\text{Bill})(w_2)) > \text{Stand}_{\text{shortness}}]]]$ **Eval. Ps.**
c. Alex is taller than Bill. **Assertion**

where p = **Alex** is taller than Bill, q = **Chris** is taller than Bill.

[3] Note in the original example by Greenberg 2018, *even* is assumed to associate with comparative marker *-er*.
[4] Whether the alternative Chris is above the standard on the G (i.e., `tallness` for (8-a)/(9)) scale or not is unknown and makes no difference. This is also the case for the interpretation in (11) and (18) later.

The comparative presupposition says that Bill ranks higher on the `shortness` scale in accessible w_1 worlds, worlds where p (Alex exceeds Bill in height) holds, than in accessible w_2 worlds, worlds where $[q \wedge \neg p]$ (Chris but not Alex exceeds Bill in height) holds. This is trivially satisfied. Crucially, the evaluative presupposition says that Bill is above the standard on the `shortness` scale in accessible w_2 worlds; given that in w_2 worlds Alex does not exceed Bill in height, Alex must be also above the standard on the `shortness` scale. Recall that the focused item, i.e., Alex in this case, remains invariant across w_1 and w_2 worlds w.r.t. to G (i.e., `shortness` here), and is consequently above the standard on the `shortness` scale in w_1 worlds, too. Now the prejacent asserts that Alex (, who is above the standard on the `shortness` scale and remains invariant across worlds) exceeds Bill in height, Bill is thus also inferred to be above the standard on the `shortness` scale, evoking the inference that they are both short. (see Fig. 2)

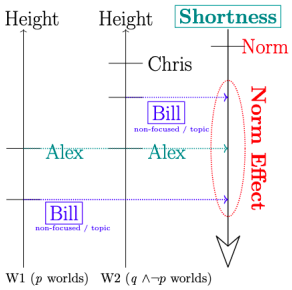


Figure 2: Interpretation of (10-a) via (7)

3 No over-generation: Constraining G via information structure

The successful derivation of the evaluative inference above crucially depends on the assumption that the contextually supplied gradable property G is taken to be `tallness` in (2-a) but `shortness` in (2-b). But is this not stipulated? Is the interpretation of G as `tallness` in (2-a) but as `shortness` in (2-b) too flexible, thus running the risk of over-generation? Remember that Bi 2022 argues the G should be `tallness` in both cases in (2), as in both cases we use the salient adjective `tall`. In this section, we show how the information structure helps to determine G independently, thus ruling out the over-generation risk.

First let's recall that following entry (7), *even* maps to a contextually salient scale based on G.[5] If we assume, reasonably, that (2), out of the blue, concerns the height of the non-focused item, and assume, quite standardly, that `tallness` and `shortness` are essentially associated with the same height scale with a reversed ordering, then the most plausible G scale *even* maps to in (2) is either `tallness` or `shortness`. Bearing this in mind, we make the novel observation that even without *even* (i.e., for the prejacent 'Alex is taller than Bill' itself), the difference in information structure between (2-a) and (2-b) – namely the difference in the focused material – independently leads to the `tallness` scale in (2-a) and the `shortness` scale in (2-b).

To see this point, let's look at the interaction of the prejacent with the connective *but* and the particle *also* respectively. As per Winterstein 2012; Winterstein 2018, two conjuncts conjoined by *but*, as in q *but* p, stand in opposed argumentative orientation; for *but* to be licensed, there must be some apparent goal such that the first conjunct q argues for this goal while the second conjunct p argues against this goal. In contrast, the prejacent of *also*, p, and its salient antecedent q, as in q *also* p, are in argumentative co-orientation. For *also* to be licensed, q and p should argue for the same goal. Now consider (12) (adapted from Winterstein 2012) where the second conjunct of *but* is the prejacent of (2-a).

(12) a. Alex is tall, (#but) he is taller than $[\text{Bill}]_F$.
b. Alex is short, but he is taller $[\text{Bill}]_F$.

Winterstein 2012 observes and experimentally establishes that in data like (12-a) *but*, with no other contextual information, is not licensed. Following Winterstein 2012, this is because out of the blue, there lacks an apparent goal which the first conjunct in (12-a) would argue for but the second conjunct would argue against; instead, in Winterstein's (2012) words, the two conjuncts are 'co-oriented', so *but* is odd. In particular, if we assume, plausibly, that 'Alex is tall', without

[5] See also Zhang 2022 who argues that *even* necessarily addresses a contextually salient degree question.

any other contextual clue, argues for the `tallness` of Alex, this indicates that 'Alex is taller than $[\text{Bill}]_F$' (i.e., the prejacent of (2-a)), also argues for his `tallness` (following Winterstein's (2012) ideas – it raises the conditional probability of 'Alex is tall', although it does not entail that). We now predict then, that *but* would be licensed only if the first conjunct argues for the reversal of `tallness`, i.e., `shortness`. This is indeed borne out in (12-b). We take this to be independent evidence that out of the blue, in (2-a), where Bill is focused, the scale is based on `tallness` but not `shortness`. Now consider the interaction with *also* as in (13):

(13) a. Alex is tall; he is (also) taller than $[\text{Bill}]_F$.
b. Alex is short; he is (#also) taller than $[\text{Bill}]_F$.

Recall that *also* requires both q and p to argue for the same goal. In (13), *also* is licensed in (13-a) where the prejacent of (2-a) is conjoined with 'Alex is tall' that argues for `tallness` of Alex, but *also* is not licensed in (13-b), where the prejacent is conjoined with 'Alex is short' that argues for the `shortness` of Alex. The felicity contrast lends further support to our point.

Crucially, we can now show that this reasoning can also correctly predict that the scale in (2-b), where the subject Alex is focused, is that of `shortness`. In (14), *but* is licensed only when the other conjunct argues for `tallness` of Bill (14-a) and *also* is licensed only when the salient antecedent argues for the `shortness` of Bill (15-b), indicating that the prejacent of (2-b) (i.e., '$[\text{Alex}]_F$ is taller than Bill') can only argue for `shortness` without other contextual information.

(14) a. Bill is tall, (but) $[\text{Alex}]_F$ is taller than him.
b. Bill is short, (#but) $[\text{Alex}]_F$ is taller than him.

(15) a. Bill is tall; $[\text{Alex}]_F$ is (#also) taller than him.
b. Bill is short; $[\text{Alex}]_F$ is (also) taller than him.

4 Extending to cases involving negation

Now we extend the degree-based account to comparative cases involving negation, and show how the choice of the salient scale via information structure offers an alternative, straightforward mechanism to account for the evaluative inference in such cases. Consider (16) from Bi 2022.

(16) a. Alex is not **even** taller than $[\text{Bill}]_F$. ⇝ implying **both are SHORT**
b. $[\text{Alex}]_F$ is not **even** taller than Bill. ⇝ implying **both are TALL**

Intuitively, (16-a) implies that Alex and Bill are both *short*, instead of *tall* (cf.(2-a)) while (16-b) implies that they are both *tall*, instead of *short* (cf. (2-b)). This flip is apparently due to the interaction between the downward entailing (DE) operator *not* and the particle *even*. There are two well-known relevant lines of analysis of *even* that commonly assume *even*'s scale to be based on likelihood, i.e., the scope theory (e.g., Karttunen and Peters 1979; Wilkinson 1996; Guerzoni 2004; Nakanish 2012) and the lexical ambiguity theory (e.g., Rooth 1985; Rullmann 1997; Giannakidou 2007). Roughly, the scope theory argues that there is just one *even* (i.e., (3)) but it can somehow scope out of the DE operator at the logical form. In contrast, the lexical ambiguity theory argues that there exist two *even*s: the positive polarity item *even* ($even_{ppi}$) and the negative polarity item *even* ($even_{npi}$) licensed in the scope of DE operators. $Even_{ppi}$ and $even_{npi}$ trigger a reversed presupposition: $even_{ppi}$ presupposes that the prejacent p is the *least likely* alternative in C, while $even_{npi}$ presupposes that p is the *most likely* alternative in C.

To account for the flipped evaluative inference in (16), an attempt assuming a likelihood-based semantics of *even* needs to decide between the two theories. For instance, Bi's (2022) account assumes the likelihood-based semantics. To explain the flip in (16), Bi assumes the lexical ambiguity theory and adopts $even_{npi}$ for such cases (see Bi 2022 for details). In principle, Bi's

account can also adopt the scope theory and assumes entry (3) for *even*, but if so, *even* has to be assumed to scope out of *not* for (16-a) and (16-b) in order to obtain our intuitive readings.

We are totally neutral regarding the two theories and have no intention to step into the debate. Instead, we intend to show how the degree-based account, along with determining the salient scale via information structure, presents a simple alternative explanation for this flip. First, reconsider (16-a). Akin to what we did above, let's assume that (16-a), without any other contextual clue, has something to do with the height of the non-focused item, i.e., Alex here. Given that `tallness` and `shortness` are associated with the same height scale, the most salient, contextually supplied gradable property G should be either `tallness` or `shortness`. Now consider the interaction between the prejacent of (16-a) (i.e., 'Alex is not taller than $[\text{Bill}]_F$') with the connective *but* (17): The connective *but* is licensed in (17-a) where the first conjunct argues for the `tallness` of Alex, but unlicensed in (17-b) where the first conjunct argues for the `shortness` of Alex. We take this felicity contrast to indicate that the prejacent of (16-a), out of the blue, only leads to the scale of `shortness`. Analogously, this reasoning independently leads to the scale of `tallness` for (16-b), as indicated by the felicity contrast (17-c) vs. (17-d).

(17) a. Alex is tall, but he is not taller than $[\text{Bill}]_F$.
b. Alex is short, (#but) he is not taller than $[\text{Bill}]_F$
c. Bill is tall, (#but) $[\text{Alex}]_F$ is not taller than him.
d. Bill is short, but $[\text{Alex}]_F$ is not taller than him.

For illustration, let's apply entry (7) to (16-a). Assuming for (16-a) that (i) the non-focused item x is Alex, (ii) the contextually supplied gradable property G is `shortness`, and (iii) Chris is the only relevant alternative to the focused Bill, we obtain interpretation (18) via entry (7):

(18) $\|(16\text{-a})\|^{g,c} = . : \forall q \in C[q \neq p \rightarrow \forall w_1, w_2[w_1 R w \wedge w_2 R w \wedge w_1 \in p \wedge w_2 \in [q \wedge \neg p] \rightarrow$
a. $[\max(\lambda d_1.\text{shortness}(d_1)(\text{Alex})(w_1)) > \max(\lambda d_2.\text{shortness}(d_2)(\text{Alex})(w_2)) \wedge$ **C. Ps.**
b. $\max(\lambda d_2.\text{shortness}(d_2)(\text{Alex})(w_2)) > \text{Stand}_{\text{shortness}}]]]$ **Eval. Ps.**
c. Alex is NOT taller than Bill. **Assertion**

where p = Alex is NOT taller than **Bill**, q = Alex is NOT taller than **Chris**.

The comparative presupposition says that Alex ranks higher on the `shortness` scale in accessible w_1 worlds, worlds where p (Alex does not exceed Bill in height) holds, than in accessible w_2 worlds, worlds where $[q \wedge \neg p]$ (Alex does not exceed Chris but does exceed Bill in height) holds. This is trivially satisfied. Crucially, the evaluative presupposition says that Alex is above the standard on the `shortness` scale in accessible w_2 worlds; given that in w_2 worlds Alex exceeds Bill in height, Bill is inferred to be above the standard on the `shortness` scale in w_2 worlds as well. Recall that the focused item, i.e., Bill here, remains unaltered across accessible w_1 vs. w_2 worlds w.r.t. G (i.e., `shortness`); therefore, Bill is above the standard on the `shortness` scale in w_1 worlds, too. Now the prejacent asserts that Alex does not exceed Bill in height, Alex is also inferred to be above the standard on the `shortness` scale, thus the inference that they are both short.

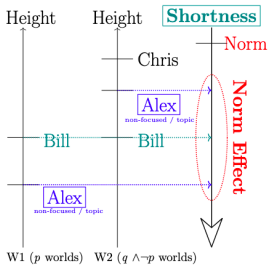


Figure 3: Interpretation of (16-a) via (7)

5 Concluding remarks

Recently, various linguistic constructions are argued to necessarily address a contextually salient degree question or involve a salient scale, e.g., incremental particles (Greenberg 2010; Grubic and Wierzba 2021), hyperbole (Nouwen 2024), cumulative readings (Zhang 2023), multi-head

comparatives (Zhang 2024). What factors play a role in the choice of this degree QUD or salient scale is under debate. If we are on the right track, information structure is a plausible one.

Acknowledgements. I am very grateful to the anonymous reviewers of Amsterdam Colloquium 2024 for the valuable feedback! Special thanks to Yael Greenberg for the long discussion and many helpful suggestions! Part of the idea further developed here was presented at the 35th European Summer School in Logic, Language and Information (ESSLLI 35); thanks to the reviewers there! The research reported in this work was funded by ISF Grant 682/22 to Yael Greenberg. Any error is, of course, mine.

References

Bi, Ruyue Agnes (2022). "Evaluativity in 'even'-comparatives via presupposition accommodation". In: *Proceedings of Sinn und Bedeutung*. Ed. by Daniel Gutzmann and Sophie Repp. Vol. 26, pp. 112–129. DOI: https://doi.org/10.18148/sub/2022.v26i0.991.

Chierchia, Gennaro (2013). *Logic in Grammar: Polarity, Free choice, and Intervention*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199697977.001.0001.

Giannakidou, Anastasia (2007). "The landscape of EVEN". In: *Natural Language & Linguistic Theory* 25.1, pp. 39–81. DOI: https://doi.org/10.1007/s11049-006-9006-5.

Greenberg, Yael (2010). "Additivity in the domain of eventualities (or: Oliver Twist's 'more')". In: *Proceedings of Sinn und Bedeutung*. Ed. by Martin Prinzhorn, Viola Schmitt, and Sarah Zobel. Vol. 14, pp. 151–167. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/465.

— (2015). "*E*ven, comparative likelihood and gradability". In: *Proceedings of the Amsterdam Colloquium*. Ed. by Thomas Brochhagen, Floris Roelofsen, and Nadine Theiler. Vol. 20, pp. 147–156. URL: https://semanticsarchive.net/Archive/mVkOTk2N/AC2015-proceedings.pdf.

— (2016). "A novel problem for the likelihood-based semantics of *even*". In: *Semantics and Pragmatics* 9, pp. 1–28. DOI: https://doi.org/10.3765/sp.9.2.

— (2018). "A revised, gradability-based semantics for *even*". In: *Natural Language Semantics* 26.1, pp. 51–83. DOI: https://doi.org/10.1007/s11050-017-9140-0.

Grubic, Mira and Marta Wierzba (2021). "The German additive particle noch: Testing the role of topic situations". In: *Glossa: a journal of general linguistics* 6.1. DOI: https://doi.org/10.5334/gjgl.1275.

Guerzoni, Elena (2004). "*Even*-NPIs in Yes/No questions". In: *Natural Language Semantics* 12.4, pp. 319–343. DOI: https://doi.org/10.1007/s11050-004-8739-0.

Guo, Rui (2022). "Bi zi ju zhong hái he dōu de yuyi zuoyong. (Chinese) [The semantic analysis of adverbs 'hái' and 'dōu' in Chinese comparative constructions]". In: *Jornal of Shanxi Univeristy (Philosophy & Social Scicences)* 45.5, pp. 62–74. DOI: 10.13451/j.cnki.shanxi.univ(phil.soc.).2022.05.007.

Ippolito, Michela (2007). "On the meaning of some focus-sensitive particles". In: *Natural Language Semantics* 15, pp. 1–34. DOI: https://doi.org/10.1007/s11050-007-9004-0.

Karttunen, Lauri and Stanley Peters (1979). "Conventional lmplicature". In: *Syntax and Semantics 11: Presupposition*. Ed. by Choon-Kyu Oh and David A. Dinneen. New York: Academic Press, pp. 1–56.

Kay, Paul (1990). "Even". In: *Linguistics and Philosophy* 13.1, pp. 59–111. DOI: https://doi.org/10.1007/BF00630517.

Miashkur, Elena and Yael Greenberg (2019). "Two types of scalar additives in Russian and other languages". In: *Syntax and Semantics Conference in Paris (CSSP) 2019*. URL: http://www.cssp.cnrs.fr/cssp2019/abstracts/CSSP_2019_paper_29.pdf.

Nakanish, Kimiko (2012). "The scope of *even* and quantifier raising". In: *Natural Language Semantics* 20, pp. 115–136. DOI: https://doi.org/10.1007/s11050-011-9077-7.

Nouwen, Rick (2024). "Meiosis and hyperbole as scalar phenomena". To appear in *Proceedings of Sinn und Bedeutung 28*, edited by Liefke et al. URL: https://ricknouwen.org/d/mhs.pdf.

Rett, Jessica (2014). *The semantics of evaluativity*. Oxford University Press. DOI: https://doi.org/10.1093/acprof:oso/9780199602476.001.0001.

Rooth, Mats (1985). "Association with focus". PhD thesis. University of Massachusetts, Amherst. URL: https://hdl.handle.net/1813/28568.

Rullmann, Hotze (1997). "*Even*, polarity, and scope". In: *Papers in Experimental and Theoretical Linguistics*. Ed. by Martha Gibson, Grace Wiebe, and Gary Libben. Vol. 4. 40-64. Edmonton, Alberta: University of Alberta. URL: https://semanticsarchive.net/Archive/WZhOWY5N/rullmann1997.pdf.

Wagner, Michael (2013). "Additivity and the syntax of *even*". Colloquium talk given at University of Chicago in December 2015. URL: https://prosodylab.org/~chael/papers/wagner15chicago.pdf.

Wilkinson, Karina (1996). "The scope of *even*". In: *Natural Language Semantics* 4.3, pp. 193–215. DOI: https://doi.org/10.1007/BF00372819.

Winterstein, Grégoire (2012). "What *but*-sentences argue for: An argumentative analysis of *but*". In: *Lingua* 122.15, pp. 1864–1885. DOI: https://doi.org/10.1016/j.lingua.2012.09.014.

— (2018). "A Bayesian approach to Argumentation within Language". Manuscript. Université du Québec à Montréal. URL: https://semanticsarchive.net/Archive/mFkZGM0N/Argumentation.pdf.

Zhang, Linmin (2022). "The presupposition of *even*". In: *Semantics and Linguistic Theory*. Ed. by John R. Starr, Juhyae Kim, and Burak Öney, pp. 249–269. DOI: https://doi.org/10.3765/salt.v1i0.5355.

— (2023). "Cumulative Reading, QUD, and Maximal Informativeness". In: *Logic and Engineering of Natural Language Semantics. LENLS 2022.* Ed. by Daisuke Bekki, Koji Mineshima, and Elin McCready. Lecture Notes in Computer Science. Springer, Cham, pp. 1–17. DOI: https://doi.org/10.1007/978-3-031-43977-3_1.

— (2024). "The semantics and pragmatics of multi-head comparatives". Manuscript. NYU Shanghai. URL: https://ling.auf.net/lingbuzz/007777.

Arbitrariness and Frege Arithmetic

Ludovica Conti
University of Vienna
Wien, Österreich
ludovica.conti@univie.ac.at

Abstract

This paper focuses on a less-known version of Abstractionism, that we'll call Arbitrary Frege Arithmetic. This system aims to restore Fregean Logicist project (of a logical derivation of second-order Peano Arithmetic) by adopting, as much as possible, its original tools, namely a logical second-order system augmented with a (consistent) version of Basic Law V. The major differences from Frege's Grundgesetze will be a recasted presentation of the abstractionist vocabulary and a weakening of first-order classical logic in a negative free logic one. Both these choices are supported by philosophical and formal motivations that allow us to consider them completely compatible with the abstractionist spirit.

Keywords: Abstraction principles, Logicism, Free logic, Arbitrariness, Frege Arithmetic.

1 Introduction

Abstractionist theories in philosophy of mathematics are systems composed by a logical theory augmented with an abstraction principle (AP), of form: $\forall X \forall Y(@X = @Y) \leftrightarrow E(X, Y)$[1] – that introduces, namely rules and implicitly defines, a term-forming operator @ by means of a formula satisfying the conditions of an equivalence relation E. As is well-known, the seminal abstractionist program, Frege's Logicism, failed[2]: Russell's Paradox proved its inconsistency and, *a fortiori*, its non-logicality. In the last century, both the issue of consistency and the issue of logicality have been resumed in the abstractionist debate (cf. Tarski 1956, Fine 2002, Antonelli 2010, Cook 2016, Boccuni and Woods 2020). More precisely, on the one side, different revisions of Frege's original system have been proposed in order to avoid Russell's Paradox and to obtain a consistent system that is strong enough to derive (at least, a relevant portion of) Peano Arithmetic. On the other side, given a semantical definition of logicality as permutation invariance, some abstraction principles have been proved to be logical (Antonelli 2010, Cook 2016).

Nevertheless, many concerns are still open. Particularly, regarding the preliminary condition of consistency, the ways out of Russell's Paradox proposed so far do not precisely mirror a corresponding explanation of the origin of the contradiction and often imply a weakening of the hoped strength of the theory (cf. Heck 1996, Wehmeier 1999, Ferreira and Wehmeier 2002)[3]; regarding the issue of logicality, an undesired dilemma overshadows the abovementioned results: precisely in case of logical (i.e. permutation invariant) abstraction principles, their implicit *definienda* turn out to be non logical (Antonelli 2010) – so preventing a full achievement of the Logicist goal.

[1] In the rest of the paper, I'll adopt this axiomatic version of AP. Given full Comprehension Axiom Schema (that will be assumed in the systems that we'll investigate), it is provably equivalent to the schematic form: $@x.\alpha(x) = @x.\beta(x) \leftrightarrow E(\alpha(x), \beta(x))$. Cf. Linnebo 2016

[2] It was proposed with the foundational purpose to derive arithmetical laws as logical theorems and to define arithmetical expressions by logical terms.

[3] In Ferreira 2018 and Boccuni 2022, second-order Peano Axioms are recovered but by appealing to stronger logical resources – i.e. double-sorted variables

My preliminary hypothesis is that these – apparently unrelated –problems have a common source in some unquestioned assumptions of Frege's project (inherited also by the following abstractionist programs). I argue that such assumptions are part of what we can call the traditional view of the abstraction, that includes the choice of classical logic as the base theory, with the related semantical consequence of full referentiality of the vocabulary, and the choice of a so-called canonical interpretation function for all the (both primitive and defined) expressions of the language.

In the rest of the paper, I show that by renouncing one or both of these problematic assumptions we can recover consistency and/or logicality. More precisely, I propose a double revision of Frege's Logicist program: on the one side, weakening the canonical interpretation function for the implicitly defined (abstract) expressions of the vocabulary (cf. Boccuni and Woods 2020), I prove that any consistent revision of BLV turns out to be logical (i.e. permutation invariant); on the other side, I show that such an arbitrary interpretation, on a (negative) free logic background, allows us to identify a restriction of BLV, able to precisely exclude the paradoxical concepts, namely to avoid Russell's Paradox, but, at the same time, to preserve the derivational strength necessary to derive second-order Peano axioms. This means that this system, precisely renouncing to the Traditional assumptions mentioned above, is able to recover both Frege's goals of logicality and consistency.

2 Non-canonical Abstraction

By a non-canonical account of the abstraction I mean the result of renouncing the meta-semantic assumption of a uniform interpretation function for all the expressions of a same syntactic category. Such an account emphasises the semantic difference between undefined expressions (such as non-logical constants) and defined ones, whose meaning should be fixed by the theory itself. Accordingly, in a non-canonical perspective, the semantic indeterminacy of the abstractionist vocabulary (as exhibited by mixed identity statements) is not a problem but a feature to be accounted for. A formulation of this attitude has been characterised by Antonelli (**Antonelli:2010**)[4] as a deflationary reading of the abstraction. We can summarise this approach as a negative thesis: abstractionist theories are categorical, and abstraction principles do not provide sufficient information to identify an intended model – understood, in the Neologicist framework, as the model in which abstract objects are identified as the singular and determined interpretation of the abstract terms and are provably distinct from all other non-abstract objects. On the contrary, abstraction principles are silent about the particular function that their operator selects – among all those that are able to map equivalent concepts in a same object and non-equivalent concepts in different objects; accordingly, abstraction functions only index classes of their arguments by objects of the codomain – they impose a lower bound on the cardinality of the domain, but are neutral with respect to the identity of their values; thus, abstract terms should only denote possible indexes of equivalence classes – indifferent to their specific nature.

As a negative thesis, deflationism is still compatible with different notions of indeterminacy and different implementations of it. In what follows, I suggest that a fruitful way to spell out such an indeterminacy is through the notion of arbitrary reference, which in turn can be further spelled out through "quantificational" and "parametrical" meanings (according to their prototypes of arbitrary expressions). In the "quantificational" approach (cf. Woods 2014), classical interpretation function from sub-sentential expressions to their singular denotations must be substituted by a generalised interpretation, (i.e. a function which sends some domains

[4] "We characterize this view of abstraction as deflationary in that the main role it ascribes to abstraction is to provide such a lower bound, while denying the objects delivered by abstraction any special status.

D to a set of objects of the same type in the type-hierarchy over D) that assigns a whole range of candidate ones (made available in the different models). In this perspective, abstraction principles are able to single out a whole class of (compatible) knowable and standard functions as a range of candidate denotations of the abstractionist operator. Accordingly, the abstract terms exhibit a "quantificational behaviour", similar to that of the variables, i.e. are plurally referring expressions, "ranging over" a class of admissible elements (functions/objects) that could serve as denotation. Correspondingly, the notions of truth and logical consequence need to be rephrased by renouncing the classical valuation function (bivalence) and split in two branches: the sub-truth evaluated on each singular model (namely, relative to any possible precisification of the interpretation function) and the super-truth evaluated by quantifying over all the candidate models and relative to all the possible precisifications of the interpretation function.

On the contrary, in the "parametrical" approach (cf. Andreas and Schiemer 2016), the classical interpretation function from sub-sentential expressions to their singular denotations must be substituted by a choice interpretation, that picks up an indeterminate item into the whole range of candidate ones (made available in the different models). In this perspective, arbitrary terms are singularly referring expressions, denoting particular, though indeterminate, elements (functions/objects). Correspondingly, the notions of truth and logical consequence need to be rephrased by renouncing the classical valuation function (bivalence) and split in two branches: the local truth evaluated respect to the existence of at least a possible choice interpretation function (namely, relative to any possible value) and the general truth evaluated by quantifying over all the possible choice interpretation functions, namely relative to all the candidate models in which they insist[5].

Besides successfully handling semantical indeterminacy (thereby dissolving the cross-sortal identity problem) and inheriting the above mentioned advantages of deflationism in general (i.e. full compatibility with abstractionist results and enhancement of the language priority thesis), an arbitrary interpretation of the abstractionist vocabulary exhibits further pleasant aspects (at least in a Fregean and Neologicist perspective), such as a more pronounced analyticity, a new kind of non-arrogance and an epistemically safe impredicativity. In the next sections, I aim to emphasise that, given an arbitrary interpretation – or, more precisely, a corresponding paraphrase – of the abstractionist vocabulary, the criterion of permutation invariance (and then the desirable feature of logicality) is satisfied by any abstraction operator defined by consistent revisions of BLV; on the contrary, the extensional operator defined by the original version of BLV fails precisely for the arguments that determine its inconsistency. Accordingly, an alternative, consistent and logical, extensional operator will be (partially) defined by means of a corresponding restriction of the axiom.

3 Logicality

Arbitrary interpretation of the abstractionist vocabulary allows us to provide an arbitrary rephrase of the criterion of Weak Objectual Invariance: an abstraction operator $@_R$ – understood as denoting any of the range of its candidate denotation – is weakly objectually invariant if and only if, for any model $M =< \Delta, I >$ and permutation π of Δ, $\pi(@) = \{< \pi(X), \pi(y) > | < X, y > \in @\} = @$, i.e. $\forall X, \forall y\ @(X) = y$ if and only if $@(\pi[X]) = \pi[y]$[6].

[5]To be more precise, there are at least two (equivalent) ways to model this choice-functional meaning of arbitrariness. On the one side, we can build choice functions on each (fully determined model) – cf. Andreas and Schiemer 2016; on the other side, we could build choice functions as part of the different models themselves.

[6]This criterion has been further strengthened (cf. Boccuni and Woods 2020) as Weak Objectual Invariance under isomorphism: an abstraction operator $@_R$ is weakly objectually invariant under isomorphism if and only if, for any model $M =< \Delta, I >$ and bijection i of $\Delta \to \Delta'$, $i(@^{\Delta}) = \{< i(X), i(y) > | < X, y > \in @\} = @^{\Delta'}$, i.e. $\forall X, \forall y\ @(X) = y$ if and only if $@(i(X)) = i(y)$.

By extending this result about the cardinal operator, we can prove that many consistent revisions of Fregean BLV implicitly define extensional operators that are weakly objectually invariant. The most obvious example is constituted by what we can call Boolosean restrictions of BLVb, cf. Boolos 1986, Shapiro 2003, (BLV^B): $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow (\phi(X) \wedge / \vee \phi(Y) \rightarrow \forall x(Xx \leftrightarrow Yx)))$. As is well known, principles instantiating this schema rule abstraction functions that perform as an injective function when they take arguments satisfying the condition ϕ but are free to map the other concepts in the same object. No matter the specific condition ϕ and the unruled behaviour of the function for the arguments that do not satisfy it, the function (arbitrarily interpreted) is weakly objectual invariant (under permutation and also under isomorphism).

Theorem 3.1. *The extensional operator ϵ^* defined by BLV^B is weakly objectually invariant.*[7]

A fortiori ϵ^*, arbitrarily interpreted, is weakly objectually invariant under permutation. Intuitively, given any permutation π of Δ, every concept, both satisfying and not satisfying ϕ, is co-extensional only with itself. Then, the partition of the second-order domain – both of the portion that constitutes the domain of the injection and of its complement – is invariant under permutation (and isomorphism). Given an arbitrary interpretation of the abstractionist vocabulary, permutation does not affect the set of candidate denotations of the abstract terms. Thus $\pi(\epsilon^*) = \epsilon^*$.

A similar reasoning can be rephrased for any consistent version of BLV, obtained both by weakening the logical background (e.g., in case of predicative subsystems of *Grundgesetze*, cf. Heck 1996, Wehmeier 1999, Ferreira and Wehmeier 2002) or by different restrictions of the principle itself (e.g. restrictions of BLVa on a free logic background (BLV^F): $\forall X \forall Y(\epsilon X = \epsilon Y \leftrightarrow (\phi(X) \wedge / \vee \phi(Y) \wedge \forall x(Xx \leftrightarrow Yx)))$).

This kind of results extends the solution of the logicality dilemma from Neologicism to all the consistent revisions of Frege's program. Not only the cardinal operator, but also the extensional operator, implicitly defined by means of the logical (weakly invariant) relation of co-extensionality, is logical (weakly objectually invariant) as well.

4 Consistency

The logical part of the language of Arbitrary Logicism, L_F, includes denumerably many first-order variables $(x, y, z, ...)$, denumerably many second-order variables $(X, Y, Z, ...)$, logical connectives $(\neg, \rightarrow)$ and a first-order existential quantifier $(\exists)$[8]. We can also usefully define a predicative monadic constant (E!), whose extension is equal to the range of identity: E!a $=_{def} \exists x(x = a)$. The only non-logical primitive symbol is the term-forming operator ϵ which applies to monadic second-order variables to produce complex singular terms $(\epsilon(X))$[9].

[7] Given D, D' and an isomorphism $\zeta : D \rightarrow D'$, we have to prove that $\zeta(\epsilon^{*D}) = \epsilon^{*D'}$. (Given an arbitrary interpretation, ϵ^{*D} and $\epsilon^{*D'}$ respectively denote the set of candidate canonical denotations, namely the set of abstraction functions, respectively, $f : \wp(D) \rightarrow D$ and $f' : \wp(D') \rightarrow D'$, that satisfy BLV^B).
$\forall f \in \epsilon^{*D}, \zeta(f) = \zeta \circ f \circ \breve{\zeta}^{1} = f' : \wp(D') \rightarrow D'$ satisfying BLV^B, i.e. $f' \in \epsilon^{*D'}$. Then, $\zeta(\epsilon^{*D}) \subseteq \epsilon^{*D'}$.
$\forall f' \in \epsilon^{*D'}, \zeta^{-1} \circ f' \circ \zeta = f : \wp(D) \rightarrow D$ satisfying BLV^B, i.e. $f \in \epsilon^{*D}$ such that $\zeta(f) = f' \in \zeta(\epsilon^{*D'})$. Then, $\epsilon^{*D'} \subseteq \zeta(\epsilon^{*D})$.

[8] We can also define the other connectives and the universal quantifier $\forall x Ax =_{def} \neg \exists x \neg Ax$.

[9] Let D be the full first-order domain (then, the second-order domain is constituted by its power-set $\wp(D)$). The satisfaction clauses for the formulas of L_F are defined in terms of an evaluation function V and an assignment function I that ascribes elements of D to the first-order terms and elements of $\wp(D)$ to the second-order terms:

- $V(Pt_1, ..., t_n) = 1 \leftrightarrow I(t_1), ..., I(t_n) \in D \wedge < I(t_1), ..., I(t_n) > \in I(P)$; 0 otherwise;
- $V((s) = (t)) = 1 \leftrightarrow I(s), I(t) \in D \wedge I(s) = I(t)$; 0 otherwise;
- $V(E!t) = 1 \leftrightarrow I(t) \in D$; 0 otherwise;
- $V(\neg\alpha) = 1 \leftrightarrow V(\alpha) = 0$; 0 otherwise;

The theory involves, as its logical part, the axioms and inference rules of non-inclusive negative free logic with identity ($NF^{=}$):

NF1) $\forall v\alpha \rightarrow (E!t \rightarrow \alpha(t/v))$;

NF2) $\exists v E!v$;

NF3) $s = t \rightarrow (\alpha \rightarrow \alpha(t//s))$[10];

NF4) $\forall v(v = v)$;

NF5) $P\tau_1, ..., \tau_n \rightarrow E!\tau_i$ (with $1 \leq i \leq n$);

$\forall$I): $E!a ... \phi(a/x) \vdash \forall x\phi$;

$\exists$E): $\phi(a/x), E!a ... \psi, \exists x\phi \vdash \psi$, where a is a new individual constant which does not occur in ϕ and ψ.

Additionally, the theory involves an axiom-schema of universal instantiation for second-order variables ($\forall X\phi(X) \rightarrow \phi(Y)$), a rule of universal generalisation (GEN), a second-order comprehension axiom schema (CA: $\exists X\forall x(Xx \leftrightarrow \alpha)$) and *modus ponens* (MP)[11].

The abstraction principle that characterizes this theory is obtained by weakening the right-to-left conditional of Basic Law V (BLV: $\forall F\forall G(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx))$, i.e. BLVa (arbitrarily interpreted), by means of the condition of Permutation Invariance (cf. Antonelli 2010, Boccuni and Woods 2020).

W-BLV: $\forall F\forall G(\epsilon F = \epsilon G \leftrightarrow \forall x(Fx \leftrightarrow Gx) \wedge \epsilon(\pi(F)) = \pi(\epsilon F))$[12]

As is well known, the ϵ operator (as defined by standard BLV), also arbitrarily interpreted, is not Permutation Invariant – because, roughly speaking, by being inconsistent it is unable to define or rule any function. We can emphasize that, given an arbitrary interpretation, Permutation Invariance fails precisely for the argument that determines its inconsistency. In other words, as can be pointed out for other consistent revisions of BLV, in any case in which it is safely restricted, ϵ turns out to satisfy Permutation Invariance, namely it is such that $\pi(\epsilon) = \epsilon$, i.e. $\forall X\forall y(\epsilon X = y \leftrightarrow \epsilon(\pi(X)) = \pi(y))$. Then, the second conjunct of the right-hand side of W-BLV requires that – no matter which object y is identical to ϵF – ϵ satisfies Permutation Invariance for the considered arguments[13].

- $V(\alpha \wedge \beta) = 1 \leftrightarrow \alpha = 1 \wedge \beta = 1$; 0 otherwise;
- $V(\alpha \vee \beta) = 1 \leftrightarrow \alpha = 1 \vee \beta = 1$; 0 otherwise;
- $V(\forall v\alpha) = 1 \leftrightarrow \forall s \in D, V_{(t,s)}(\alpha(t/v)) = 1$ – where t is not in α and $V_{(t,s)}$ is the valuation function on the model $< D, I^* >$ such that $I^* = I$, except that $I^*(t) = s$.
- $V(\forall V\alpha) = 1 \leftrightarrow \forall S \subseteq D, V_{(T,S)}(\alpha(T/V)) = 1$ – where T is not in α and $V_{(T,S)}$ is the valuation function on the model $< D, I^* >$ such that $I^* = I$, except that $I^*(T) = S$.

[10]Where $\alpha(t//s))$ is the result of replacing one or more occurrences of s in A by t.

[11]From these axioms we can also derive the following theorems: T1) $\forall x E!x$; T2) $t = t \leftrightarrow E!t$; T3) $(\neg E!s \wedge \neg E!t) \rightarrow (\alpha \rightarrow \alpha(t//s))$.

[12]This abstraction principle is clearly circular because the extensional operator occurs on both the sides of the biconditional. The idea that circularity defeats the definitional role of such principle (or, in general, of implicit definitions) is controversial. Anyway, in this framework, what we need is a principle that rules the behavior of a new symbol of the language and W-BLV carries out this task.

[13]This revision of BLV (particularly of BLVa) is featured by a restriction that, with respect to many other (syntactical ones), is expressible into the language. Indeed, the permutation π of the operator or of the concepts mentioned in the right-hand side of the bi-conditional can be defined as abbreviation of the effects of any first-order bi-jective function f: $D_1 \rightarrow D_1$ on the entities (sets, relations or functions) further up in the type hierarchy.

Accordingly, W-BLV, as a bi-conditional, turns out to be satisfied by any concept instantiating the universal quantifier. On the one side, given an arbitrary interpretation of the abstraction operator, for any concept different from Russellian concept (R), $\pi(\epsilon) = \epsilon$. On the other side, we can consider Russell's Paradox as a *reductio ad absurdum* of the alleged truth of both the sides of the bi-conditional for the concept R: the contradiction proves that ϵR – as legitimately admitted on a free logical background – does not exist, namely it is a term devoid of denotation; accordingly, it is not identical to itself (so, falsifying the left-hand side of W-BLV) and, even if R, as any other concept, is co-extensional with itself, it falsifies Permutation Invariance of the operator[14]. Accordingly, also the right-hand side of W-BLV is false and also the instance of the bi-conditional for the concept R is verified.

Such a restricted version of W-BLV allows us to derive a corresponding restricted version of Hume's Principle (W-HP). However, as already proved (Boccuni and Woods 2020), cardinal operator (in our choice-functional translation or in an arbitrary interpretation) satisfies permutation invariance without restrictions. For this reason, even if we are only able to prove W-HP, it is equivalent to full HP, so it allows us to derive the main arithmetical results, including Frege's Theorem.

References

Andreas, Holger and Georg Schiemer (2016). "A choice-semantical approach to theoretical truth". In: *Studies in History and Philosophy of Science Part A* 58, pp. 1–8.

Antonelli, G Aldo (2010). "Notions of invariance for abstraction principles". In: *Philosophia Mathematica* 18.3, pp. 276–292.

Boccuni, Francesca (2022). "Plural Frege Arithmetic". In: *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences* 26-1, pp. 189–206.

Boccuni, Francesca and Jack Woods (2020). "Structuralist neologicism". In: *Philosophia Mathematica* 28.3, pp. 296–316.

Boolos, George (1986). "Saving Frege from contradiction". In: *Proceedings of the Aristotelian Society*. Vol. 87. JSTOR, pp. 137–151.

Cook, Roy T (2016). "Abstraction and four kinds of invariance (or: What's so logical about counting)". In: *Philosophia Mathematica* 25.1, pp. 3–25.

Ferreira, Fernando (2018). "Zigzag and Fregean arithmetic". In: *The philosophers and mathematics: Festschrift for Roshdi Rashed*, pp. 81–100.

Ferreira, Fernando and Kai F Wehmeier (2002). "On the consistency of the Δ 1 1-CA fragment of Frege's Grundgesetze". In: *Journal of Philosophical Logic* 31, pp. 301–311.

Fine, Kit (2002). *The limits of abstraction*. Clarendon Press.

Heck, Richard G (1996). "The consistency of predicative fragments of Frege's Grundgesetze der Arithmetik". In: *History and Philosophy of Logic* 17.1-2, pp. 209–220.

Linnebo, Øystein (2016). "Impredicativity in the neo-Fregean program". In: *[Ebert and Rossberg, 2016]*, pp. 247–268.

Shapiro, Stewart (2003). "Prolegomenon To Any Future Neo-Logicist Set Theory: Abstraction And Indefinite Extensibility". In: *The British journal for the philosophy of science* 54.1, pp. 59–91.

Tarski, Alfred (1956). "The concept of truth in formalized languages". In: *[Tarski, 1983]*, pp. 152–278.

[14]This last claim follows from the definition of π and the result of non-existence of ϵR: on the one side, $\epsilon(\pi(R)) = \epsilon(\{\pi(x)|x \in R\}) = \epsilon(X)$ – where X is any other concept (based on π); on the other side, $\pi(\epsilon R)$, given that ϵR is not denoting, is another well-formed term without denotation; then, the identity between ϵX (for any X that is obtained by means of a permutation of R) and the empty term $\pi(\epsilon R)$ is false.

Wehmeier, Kai F (1999). "Consistent fragments of Grundgesetze and the existence of non-logical objects". In: *Synthese* 121, pp. 309–328.
Woods, Jack (2014). "Logical indefinites". In: *Logique et Analyse* 227, pp. 277–307.

Vagueness and pragmatic reasoning in quantified sentences

Alexandre Cremers
BKKI, Filologijos Fakultetas, Vilniaus Universitetas
alexandre.cremers@gmail.com

Abstract

Vagueness has been extensively studied as a standalone topic, but little is known of its interactions with other semantic and pragmatic phenomena. We present an experiment testing a Rational Speech-Act model from Cremers (2022) which aims to offer a general account of the pragmatic use of vague sentences and the implicatures they give rise to on a new class of sentences where a vague term restricts a universal quantifier. We found that an implicature predicted by the model is present, but a separate prediction on borderline cases isn't verified. We discuss consequences for the modeling of pragmatic uses of vague sentences.

1 Vagueness and implicatures

Standard theories of implicatures (be it Neo-Gricean or grammatical) predict that a sentence like (1) should compete with an alternative without the modifier, and this would lead to an implicature that Mary is (somewhat) tall (Simons, 2001). Yet, Leffel et al. (2019) show that as long as *very* is not stressed, this implicature is absent (if anything, participants conclude that Mary is rather short). They explain this observation as the result of an interaction between the vagueness of the adjective and the mechanism deriving implicatures.

(1) Mary is not very tall $\not\rightsquigarrow$ Mary is tall

Informally, Leffel et al. (2019) argue that no height can at the same time count as clearly 'not very tall' and clearly 'tall', and this would prevent the implicature from being derived. To capture this, they extend Fox's (2007) mechanism of innocent exclusion (which blocks some contradictory implicatures in a grammatical theory) to also rule out "borderline contradictions" (vague sentences such as 'Peter is tall and not tall'; Ripley, 2011). Cremers (2022) later recast this proposal in slightly more pragmatic terms within the Rational Speech-Act framework (Frank and Goodman, 2012), using a probabilistic version of supervaluationism from Spector (2017). We refer to Cremers (2022) for a general presentation of the model, but in short, the utility of a candidate vague utterance takes into account all possible (first-order) vague denotations, in proportion to their likelihood (second-order vagueness). This, combined with explicit reasoning on possible strengthened meanings (Franke and Bergen, 2020), assigns a probability close to zero to the exhaustified reading in (1).

Nevertheless, few other cases of potential interactions between vagueness and implicatures have been tested, and given the complexity and degrees of freedom in these RSA models, this raises concerns of possible overfit. We test quantified vague sentences which also have the potential to raise implicatures in order to test the model, both qualitatively and quantitatively.

2 Quantified sentences and predictions

(2) Every tall student laughed $\overset{?}{\rightsquigarrow}$ not every student laughed

Take (2) for instance. Superficially, this example shares a lot with (1): in both cases, a vague term occurs in a downward entailing environments and a more informative alternative

can be derived by deletion (~~*very*~~ in (1), ~~*tall*~~ in (2)). Unlike (1), the proposals discussed above predict that (2) should have an implicature that some students didn't laugh (the strengthened meaning is **not** a "borderline contradiction" in this case since a short student can validate the implicature). Cremers (2022) makes an additional prediction: for (2) to be felicitous, borderline tall students should laugh too (a student who counts as tall under at least *some* denotations but didn't laugh would invalidate the sentence). In probabilistic terms, this theory predicts a slight decrease in laughing probability for short students (due to the implicature) but a marked increase for borderline-tall students (due to the supervaluationist mechanism). As usual with vagueness, it is difficult to delineate between clearly-not-tall and borderline-tall (higher-order vagueness problem), and specific predictions from the model depend on the value of all model parameters. In §3.3 we present details specific to our experiment (worlds, messages, costs. . .)

3 Experiment

3.1 Design

We designed an experiment to test how participants interpret sentences like (2), with both relative (vague) and absolute (non-vague) gradable adjectives. Participants first saw a table with 8 items and their dimensions on a scale relevant to the gradable adjective, and were asked about three of them "Would you say that the following [items] are [adj]?" (Q1) on a slider from 'no' to 'yes'. The three items were chosen to represent a clear true, clear false, and borderline case (or a case very close to the boundary for absolute adjectives). We then probed participants' out-of-the-blue prior beliefs about an unrelated predicate (e.g., playing video games) for each of the three items tested on the previous question, using a slider from 'certainly the case' to 'certainly not the case' (Q2). Participants had no reason to make different a priori judgments for this predicate on the different items, unless they assumed a correlation with the vague predicate, so measuring them all was safer. Finally, the story introduced someone familiar with the matter, who uttered a sentence like (2). We then asked participants to convey their posterior belief about the predicate of Q2 given what they heard, again using continuous sliders (Q3).

We tested 12 gradable adjectives (6 relative, 3 min standard, 3 max standard) in the restrictor of either 'every' or 'no'. Each adjective came with a different scenario (different items, different predicate), all listed in Table 1. For each adjective, we tested two potential "borderline" cases (of which each participant saw only one, randomly selected). After rejecting participants who took the survey multiple times, 417 participants were recruited on Amazon's MTurk and paid $0.47 for their participation. We excluded 51 participants whose responses on Q1 suggested they were not focused on the task,[1] plus one participant whose data wasn't properly recorded.

3.2 Results

Data and scripts available at `https://github.com/Alex-Cremers/vague-restrictors-AC2024`.

Figure 1 shows participants' response to Q1 (i.e., how much they agree with an attribution of the adjective for each item that was tested). We see that even for absolute adjectives, it is

[1]The initial criterion was to reject all participants whose response at one or both scale ends fell on the wrong side of the slider, e.g., participants who assigned the shortest student a "tallness" of over 50%, or the tallest one, less than 50%. This led to the rejection of 89 participant (21%), which an anonymous reviewer found surprisingly high for this task. On further inspection, it turned out that many participant seemed to genuinely find the lowest value to fall to some extent in the positive extension of some predicates. This was particularly true of 'hot' (possibly because the experiment was run in winter?). I therefore updated the criteria to include all participants who saw a relative adjective as long as their responses increased with the underlying measure, even if it started too high or ended too low. I didn't apply such an inclusion criterion to absolute adjectives, because their scale ends had to be clear cases, even under a loose interpretation. I thank the reviewer for drawing my attention to this.

	Adjective	Measure	Item	Scope Predicate
relative	tall	ft, in	students	plays video games
	powerful	hp	car	has a touchscreen
	hot	°F	day	was a workday
	expensive	$	car	has the gas tank door on the left side
	large	sq ft	apartment	has West-facing windows
	young	yr	employee	is left-handed
min	late	time	employee	works in HR
	spicy	🌶	(Thai) dish	contains lemongrass
	profitable	$	company	is a construction company
max	complete	%	investigation	was about credit card fraud
	safe	crime rate	neighborhood	is situated North of city center
	full	free seats	flight	departed in the morning

Table 1: Adjectives by type, and associated items and predicates. 'late' was relative to 8:30am. For 'young', 'full' and 'safe', the measure is inversely correlated to measurement units.

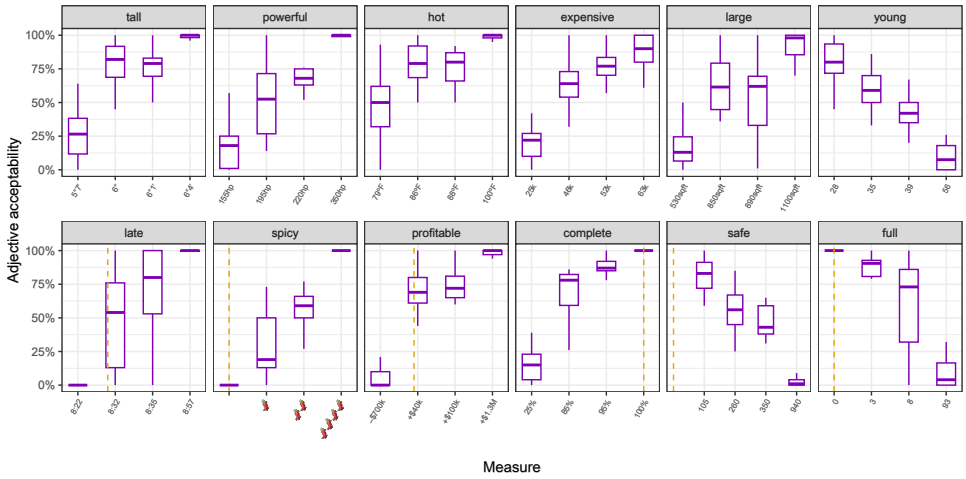


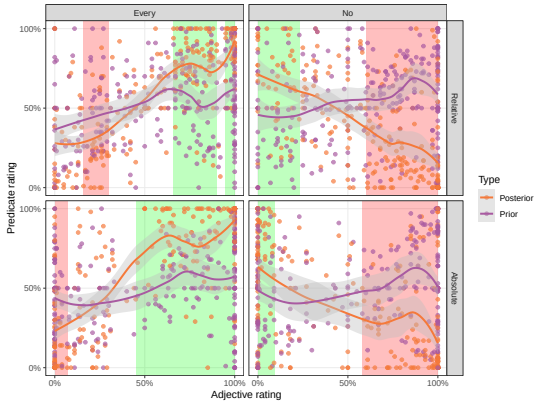
Figure 1: Responses to Q1 as a function of the measure given to participants. For closed scales, we indicate the approximate position of the relevant boundary (dashed orange line).

possible to find cases where ratings are mixed (be it genuine vagueness or a different phenomenon). Figure 2a shows responses to Q2 and Q3 as a function of Q1. A linear mixed-effects model on responses to Q2 with the interaction of scaled adjective rating (answers to Q1) and adjective as a predictor showed a significant relation between Q2 and Q1 ($\chi^2(12) = 122.7, p < .001$), with coefficients ranging from -0.31 for 'late' (late employees are less likely to work in HR) to 0.81 for 'powerful' (powerful cars are more likely to have a touchscreen), confirming that we cannot assume the target predicates to be independent from the vague restrictors, a fact which will complicate the Bayesian model. We then run a theory-neutral analysis aimed at identifying areas where the posterior differs significantly from the prior. The results are presented in Table 2b. In all 4 cases we found a large cluster corresponding to the literal contribution of the sentence, as well as a smaller cluster of opposite sign at the bottom of the scale, corresponding to the existential implicature. We therefore confirm the presence of this implicature with 'no' and 'every', for both relative and absolute adjectives, in line with the prediction of Leffel et al. (2019).

3.3 Model evaluation

Implementation: For simplicity, we restrict our attention to relative adjectives only and leave the analysis of absolute adjectives for future research.[2] We also simplify the model compared to

[2]The case of minimum standard adjectives is complicated by two facts: they can be ambiguous between a strict reading and a relative-like reading (Qing, 2021), and the threshold of their relative-like interpretation tends



(a) Prior (Q2) and posterior (Q3) ratings on the scope predicate as a function of applicability of the restrictor adjective (Q1). Significant positive/negative clusters are highlighted in green/red. Smoothing lines added for illustrative purposes.

Quant.	Adj. class	cluster range	C	p
every	relative	[16%, 30%]	−42	.002
		[66%, 90%]	85	< .001
		[95%, 100%]	35	.004
	absolute	[0%, 7%]	−29	.008
		[45%, 100%]	213	< .001
no	relative	[0%, 23%]	66	< .001
		[60%, 100%]	−220	< .001
	absolute	[0%, 9%]	28	.008
		[58%, 100%]	−170	< .001

(b) Significant clusters of differences between prior and posterior judgments. We adapted the non-parametric permutation test proposed in Maris and Oostenveld, 2007. The analysis was run independently for each Quantifier and Adjective class (relative vs. absolute).

Figure 2: Responses to Q2 and Q3 as as function of Q1, and cluster analysis results.

Cremers (2022) by fitting a unique set of costs and rationality parameters for all participants (instead of using by-subject random effects on these parameters).

In the context of our experiment, the question of the vague restrictor (e.g., tall) is settled, and we take the set of worlds to represent the $2^8 = 256$ possible ways to realize the binary scope predicate (e.g., laughing), which correspond to the cells of the unique QUD we will consider ("which students laughed?"). Table 2 presents the set of messages we considered. The quantifier 'no' is formally identical to 'every' if we flip the target predicate ("no tall student laughed" is equivalent to "every tall student didn't laugh").

	message	cost
u_0	"..."	c_0
u_{WE}	"A, B, and C"	$n \times c_1$
u_{SE}	"only A, B, and C"	$n \times c_1 + c_{\text{SE}}$
$u_{\forall\text{adj}}$	"every tall student"	$c_{\text{every}} + c_{\text{adj}}$
$u_{\forall}$	"every student"	c_{every}

Table 2: Messages considered in the model, and associated cost. List answers receive a cost dependent on the length of the list. We used a different cost for 'no', and each adjective had its own cost.

We assume that the threshold θ for the adjective follows a normal distribution parametrized by $\Theta = (\mu, \sigma)$ (first-order vagueness). Θ itself is assumed to follow a hybrid multivariate normal/lognormal distribution parametrized by $\Omega = (m_\mu, m_\sigma, s_\mu, s_\sigma, \rho)$ (Fletcher and Zupanski, 2006) (second-order vagueness plays a crucial role in the RSA-SvI). Following Cremers (2022), we obtained these parameters by fitting an independent Bayesian mixed-effects model for each adjective on responses to Q1, where each participant is assigned a Θ.

For $u_{\forall\text{adj}}$, we propose the literal interpretation in (3), where $d_{\text{last}}(w)$ is the height of the tallest student who didn't laugh in w (every tall student laughed if and only if the tallest student who didn't laugh, if any, doesn't exceed θ_{tall}), and d_{max} is the height of the tallest student (which unlike d_{last} is settled in this case).[3]

$$L_0(w|u_{\forall\text{adj}}, \Theta) \propto P(w)P(d_{\text{last}}(w) \leq \theta < d_{\text{max}}) \tag{3}$$

To compute (3), we need the prior on all 256 worlds, but Q2 only probed participants' priors on 3 items, so we first used their response to fit a prior on the whole 8 item scale, and then

not to follow a normal distribution (in Cremers (2022), they are modeled with exponential distributions).

[3] d_{max} is there to account for the presupposition that the restrictor isn't empty (at least one student must count as tall, otherwise the target sentence becomes a null message), which we assimilate with falsity.

derived priors for worlds.[4] This was again done independently for each adjective via mixed-effects models with random by-participant intercepts and slope, using the logit of priors squeezed onto $[0.05, 0.95]$ as the dependent variable and normalized degrees as the regressor.

Having obtained the parameters Ω and the prior $P(w)$, we were able to precompute the informativity component of U_1 for all the messages, in each world, and for each participant. We used it to fit a single Bayesian model on data from Q3 with the remaining parameters of the model: the costs variables $(c_0, c_1, c_{SE}, c_{\text{every}}, c_{\text{no}}, (c_{\text{adj}}))$ and the rationality parameter α. We obtain the L_1 posterior on proposition φ by summing posterior on world-parse pairs on both parses of the target sentence across φ-worlds.

We compared this model to two simpler "literal" models where instead of $L_1(\varphi(X)|u)$ we assumed that the participants' responses follow $\int P(\theta)L_0(\varphi(X)|u^i, \Theta)d\Theta$ for $i = \text{LIT}$ or EXH. The model adopts the grammatical theory of implicatures, so we can have a "literal" listener with an exhaustified parse (but it involves no reasoning on parses, second-order vagueness, or alternative messages).

parameter	prior	mean	SD	90% CrI
c_0	$\mathcal{N}_+(0, 5)$	0.17	0.17	$[0.01, 0.51]$
c_1	$\mathcal{N}_+(0, 0.5)$	2.28	0.35	$[1.74, 2.88]$
c_{SE}	$\mathcal{N}_+(0, 1)$	3.59	0.81	$[2.22, 4.89]$
c_{every}	$\mathcal{N}_+(0, 1)$	1.10	0.58	$[0.22, 2.11]$
c_{no}	$\mathcal{N}_+(0, 1.5)$	3.26	1.00	$[1.68, 4.97]$
$c_{\text{expensive}}$	$\mathcal{N}_+(0, 1)$	1.48	0.72	$[0.32, 2.71]$
c_{hot}		0.69	0.49	$[0.07, 1.62]$
c_{large}		0.58	0.51	$[0.03, 1.60]$
c_{powerful}		0.80	0.65	$[0.05, 2.09]$
c_{tall}		1.58	0.76	$[0.39, 2.91]$
c_{young}		1.67	0.76	$[0.43, 2.95]$
α	$\Gamma(4, 1)$	1.69	0.20	$[1.38, 2.03]$

(a) Prior and posterior distribution for all parameters of the RSA-SvI L_1 model. $\mathcal{N}_+$ indicates a truncated Gaussian. c_0 is the cost of the null message, c_1 the linear coefficient for list answers.

Model	elpd_{loo}	Δ_{elpd}	SE_Δ	p_{loo}
L_0^{EXH}	-96.4	0	0	1.9
L_0^{LIT}	-117.6	-21.2	4.4	1.8
L_1	-118.0	-21.5	5.0	9.2

(b) PSIS LOO-CV model comparison. The expected log predictive density (elpd) is a measure of how well the model predicts unseen data (Gelman, Hwang, and Vehtari, 2014), and naturally penalizes models with too many degrees of freedom. Δ_{elpd} is the difference with the best model, and SE_Δ the standard error on this difference. p_{loo} is the estimated effective number of parameters. Pareto influence values indicated that the approximation was valid (all $\hat{k} < 0.34$ across all three models, values under 0.7 are considered good).

Table 3: Posterior distributions of the parameters in the L_1 model (a) and comparison with the L_0 models (b).

Results: All models converged perfectly. Figure 3 shows the aggregated predictions of each of the three models and the measured posteriors. Table 3a gives the priors for each parameter in the L_1 model and the mean, SD, and quantiles of the posterior distribution. Looking at parses, we found that the L_1 model assigned a higher posterior probability to the exhaustified parse for all participants but one outlier (who had a posterior probability of 27%). The posterior probability for other participants ranged from 50% to 92%, with an average of 55%.

The models were compared using the PSIS Leave-one-out CV approximation (Vehtari, Gelman, and Gabry, 2017; Vehtari, Simpson, et al., 2024). The details are presented in Table 3b. In short, the L_0^{EXH} model outperforms the two other models, which are indistinguishable.

[4] We assumed that the predicate on each individual item is independent. For instance, a world in which only students A, B, and E laughed would receive prior probability $p_A p_B(1-p_C)(1-p_D)p_E(1-p_F)(1-p_G)(1-p_H)$, where $p_A \dots p_H$ are the value fitted from the participant's response to 3 of these items.

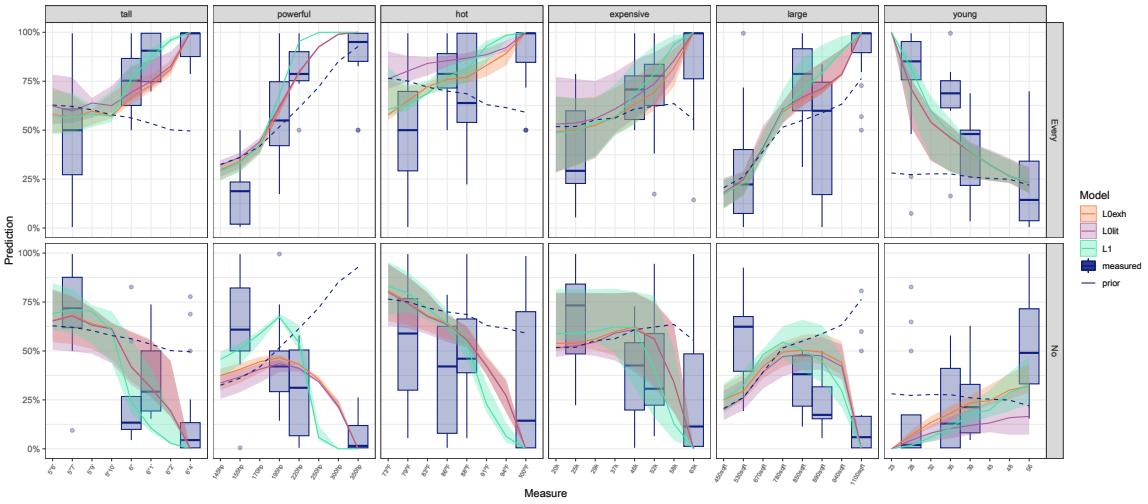


Figure 3: Predictions of the pragmatic L_1 model and the two L_0 models, compared to priors and measured posteriors, by degree. The ribbons indicate the median and quartiles of the predictions at each degree. The "supervaluationist" prediction of the L_1 model manifests itself as the green curve standing above the orange one in the upper-mid range with 'every , under for 'no' (most visible with 'powerful').

4 Discussion

Our results expand our empirical knowledge of the complex interactions between vagueness and pragmatic reasoning and confirm Leffel et al.'s (2019) prediction that an implicature should this time be present. Yet, they seem to invalidate an additional prediction about borderline cases from Cremers (2022), a concrete implementation of this theory.

As usual, falsifying a probabilistic model isn't straightforward. We showed that a simple L_0 model performed better on relative adjectives in our data, but we already know that it would not capture Leffel et al.'s (2019) data. Since the supervaluationist component that allowed Cremers (2022) to capture the contrast between absolute and relative adjectives (probabilistic supervaluationism on second-order vagueness) makes an incorrect prediction, we could imagine recasting Leffel et al.'s explanation without it, assuming the explanation for the puzzle (1) can be saved. Yet, it remains unclear whether supervaluationism is the main culprit for the poor fit here. In particular, the fact that both L_0 and L_1 models underestimate the impact of the "not every" implicature points to the well-known exacerbated effect of priors in the RSA framework (Degen, Tessler, and Goodman, 2015; Schreiber and Onea, 2021; Cremers, Wilcox, and Spector, 2023). Alternatively, one of our simplifying assumptions might be the issue (independence assumption to derive priors on worlds, shared adjective denotations across participants). Finally, we postponed an evaluation of the model on absolute adjectives, which may favor the L_1 model (after all, it was initially designed to account for a contrast between absolute and relative adjectives).

To conclude, our data suggests that giving a precise, explicit account of the interaction between vagueness and pragmatics remains elusive (and we haven't touched on speakers with partial knowledge Egré et al., 2023). Understanding the pragmatics of vague sentences is challenging on empirical grounds (requiring precise quantitative data) and theoretical grounds (models tend to be computationally more complex than for other phenomena), but it could eventually shed new light on long-standing questions in the literature on vagueness, in particular, why does human communication rely so heavily on vague terms?

Acknowledgements. This research was made possible by funding from the Research Council of Lithuania and the European Social Fund under Measure 09.3.3-LMT-K-712.

References

Cremers, Alexandre (2022). "A Rational Speech-Act model for the pragmatic use of vague terms in natural language". In: *Proceedings of CogSci 44*. Ed. by J. Culbertson et al., pp. 149–155.

Cremers, Alexandre, Ethan Wilcox, and Benjamin Spector (2023). "Exhaustivity and anti-exhaustivity in the RSA framework: Testing the effect of prior beliefs". In: *Cognitive Science* 47.5, e13286. DOI: 10.1111/cogs.13286. URL: https://arxiv.org/abs/2202.07023.

Degen, Judith, Michael Henry Tessler, and Noah D Goodman (2015). "Wonky worlds: Listeners revise world knowledge when utterances are odd." In: *CogSci*.

Egré, Paul et al. (2023). "On the Optimality of Vagueness: "Around", "Between", and the Gricean Maxims". In: *Linguistics and Philosophy* 46.5. DOI: 10.1007/s10988-022-09379-6.

Fletcher, Steven J and Milija Zupanski (2006). "A hybrid multivariate normal and lognormal distribution for data assimilation". In: *Atmospheric Science Letters* 7.2, pp. 43–46.

Fox, Danny (2007). "Free Choice Disjunction and the theory of Scalar Implicature". In: *Presupposition and implicature in compositional semantics*. Ed. by Uli Sauerland and Penka Stateva. New York, NY: Palgrave Macmillan, pp. 71–120.

Frank, Michael C and Noah D Goodman (2012). "Predicting pragmatic reasoning in language games". In: *Science* 336.6084, pp. 998–998.

Franke, Michael and Leon Bergen (2020). "Theory-driven statistical modeling for semantics and pragmatics: A case study on grammatically generated implicature readings". In: *Language* 96.2, pp. 77–96.

Gelman, Andrew, Jessica Hwang, and Aki Vehtari (2014). "Understanding predictive information criteria for Bayesian models". In: *Statistics and computing* 24, pp. 997–1016.

Leffel, Timothy et al. (2019). "Vagueness in Implicature: The Case of Modified Adjectives". In: *Journal of Semantics* 36.2, pp. 317–348. ISSN: 0167-5133. DOI: 10.1093/jos/ffy020.

Maris, Eric and Robert Oostenveld (2007). "Nonparametric statistical testing of EEG-and MEG-data". In: *Journal of neuroscience methods* 164.1, pp. 177–190.

Qing, Ciyang (2021). "Zero or minimum degree? Rethinking minimum gradable adjectives". In: *Proceedings of Sinn und Bedeutung* 25, pp. 733–750. DOI: 10.18148/sub/2021.v25i0.964.

Ripley, David (2011). "Contradictions at the borders". In: *Vagueness in communication*. Ed. by Rick Nouwen et al. Springer, pp. 169–188.

Schreiber, Alexander and Edgar Onea (2021). "Are Narrow Focus Exhaustivity Inferences Bayesian Inferences?" In: *Frontiers in Psychology* 12.

Simons, Mandy (2001). "On the conversational basis of some presuppositions". In: *Proceedings of SALT*. Ed. by R. Hastings, B. Jackson, and Z. Zvolensky. Vol. 11, pp. 431–448.

Spector, Benjamin (2017). "The pragmatics of plural predication: Homogeneity and Non-Maximality within the Rational Speech Act Model". In: *Proceedings of the 21st Amsterdam Colloquium*. Ed. by Alexandre Cremers, Thom van Gessel, and Floris Roelofsen, p. 435.

Vehtari, Aki, Andrew Gelman, and Jonah Gabry (2017). "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC". In: *Statistics and Computing* 27.5, pp. 1413–1432. DOI: 10.1007/s11222-016-9696-4.

Vehtari, Aki, Daniel Simpson, et al. (2024). "Pareto Smoothed Importance Sampling". In: *Journal of Machine Learning Research* 25, pp. 1–57.

Reanalysis in discourse comprehension: Evidence from reading times

John Duff
Saarland University
Saarbrücken, Germany
jduff@lst.uni-saarland.de

Daniel Altshuler
University of Oxford
Oxford, UK
daniel.altshuler@ling-phil.ox.ac.uk

Abstract

This paper explores the evidence for 'discourse reanalysis,' the hypothetical case where an initial interpretation of a sequence of discourse units must be revised due to further context. If discourse reanalysis does occur, our representations may need to permit selective non-monotonic update of discourse-pragmatic meaning like coreference and coherence (Lascarides and Asher 1993; Haug 2014). But a simpler alternative looms: fully-underspecified representations, which could maintain all grammatical discourse interpretations, without any intermediate selection or ranking to later be reanalyzed. To see which approach best models the representations used in actual incremental comprehension, we turn to psycholinguistic evidence. Discourse reanalysis should be associated with momentary difficulty akin to 'garden path' effects in syntactic processing. But previous studies looking for this difficulty in discourse processing have been equivocal. Here, we investigate a case missing from those previous experiments, a joint ambiguity of coreference and coherence, for which we show robust initial preferences. Then, in a self-paced reading experiment, we observe evidence for difficulty when a preferred interpretation must be abandoned. We take this to solidify the empirical basis for discourse reanalysis as a consequence of incremental composition at the discourse level, and discuss two ways this might be accomplished, depending on one's approach to the interface between logical representation and real-time comprehension.

1 Introduction

In some sequences of discourse units $[S_a \dots S_n, S_{n+1}]$, the natural interpretation that would be given for $[S_a \dots S_n]$ alone is not part of the natural interpretation of the completed discourse. For instance, given only S_1 and S_2 of (1), the discourse would seem to describe a series of parallel events, where *him* in S_2 picks out George. But given the continuation in S_3, *him* must instead pick out Bill, and S_1 and S_2 would now seem to describe an action and a reaction.[1]

(1) [Bill pinched George]$_{S_1}$ [and Ava elbowed him.]$_{S_2}$ [Bill didn't elbow her in return,]$_{S_3}$...

In dynamic logics where the first representation of S_1 through S_2 is a step along the construction of the final meaning of the discourse, that representation must somehow be revised in light of S_3. We will say that such representations would have to undergo 'discourse reanalysis.'

As noted by Haug (2014), discourse reanalysis in sequences like (1) poses problems for logics that don't permit selective non-monotonic update of coreference. For instance, in Discourse Representation Theory (DRT; Kamp and Reyle 1993), once a pronoun *him* has been represented as coreferent with its antecedent, the downstream semantics cannot distinguish propositions involving the uncertain referent of *him* from those involving the antecedent. This observation

[1] A reviewer points out that intonation associated with focus structure could play a role in marking this interpretation already at S_2, and wonders whether, as a result, this ambiguity may be somewhat artificial. We are less worried. Because broad focus-marking for S_1 and S_2 permits both interpretations, and given that intonational disambiguation is not especially common in natural production (Allbritton, McKoon, and Ratcliff 1996; Syrett, Simon, and Nisula 2012), we suspect that such ambiguities could arise commonly even in speech. However, we acknowledge that assigning focus structure in silent reading is its own process which certainly interacts with other aspects of interpretive selection (e.g. Carlson et al. 2009), and its role here should be further examined.

motivated Haug (2014) to develop a partial compositional variant of DRT (PCDRT), where dynamic meaning is separated into a monotonically-updated semantic representation, which underspecifies for coreference, and a flexibly-edited set of pragmatic enrichments, through which coreference is ultimately resolved.[2]

To be sure, discourse reanalysis in (1) may involve more than just updating coreference. If the initial interpretation of $[S_1,\ S_2]$ also includes the inference of a coherence relation between the two segments, this relation must also be revised, from Parallel to Result (in the sense of Asher and Lascarides 2003). We follow the standard view that coherence is intricately linked to coreference resolution (e.g. Hobbs 1979; Kehler et al. 2008). On this view, the initial resolution of *him* as George is a consequence of establishing a Parallel relation. But a different coherence relation is preferred once *him* is revised to refer to Bill.[3] Lascarides and Asher (1993) and Asher and Lascarides (2003) use similar examples to argue that coherence too must permit non-monotonic update; to accomplish this in their Segmented Discourse Representation Theory (SDRT), they adopt a glue logic that determines coherence relations through defeasible entailment.

One objection to theories capable of selective non-monotonic update is that a simpler alternative looms: persistent underspecification. Discourse reanalysis could be avoided wholesale if the dynamic representation of the discourse was limited to only an underspecified representation identifying all the interpretations consistent with the monotonic content (cf. common approaches to quantifier scope, see Egg 2010). This would leave open, e.g., uncertain coherence and coreference, only applying defeasible heuristics at the end of the discourse. In such an approach, there is no need for the logic to permit any step-wise non-monotonic update.

So, do we need non-monotonicity here, or would persistent underspecification suffice?[4] If our goal is to model the representations used in actual incremental comprehension, we can turn to behavioral observations from psycholinguistic experiments. True discourse reanalysis would require recognition of an inconsistency, followed by the identification and construction of a suitable revised interpretation; as a whole this should incur momentary difficulty akin to 'garden path' effects in the processing of lexical and syntactic ambiguity (Frazier and Rayner 1982; Duffy, Morris, and Rayner 1988).[5] Readers indeed exhibit slower and more effortful reading when there are inconsistencies in the monotonic content of a text (Albrecht and O'Brien 1993). Yet, to date, reading studies have generally failed to observe clear difficulty for cases of hypothetical discourse reanalysis, suggesting that readers may not experience inconsistency, or launch any subsequent process of costly revision for, e.g., ambiguous pronouns (Stewart, Holler, and Kidd 2007), temporal ordering (Dickey 2001; Sasaki 2021), and coherence-related causal inferences (Mak and Sanders 2013; Duff 2023). This could be explained if readers persistently underspecify interpretations at the level of discourse, as a rule.

In the present work, we push this body of evidence further, looking for the first time (we believe) at the incremental comprehension of discourses with joint ambiguities of coherence and

[2]For an application of this approach to fictional narrative, which often seems to provoke discourse reanalysis, see Altshuler and Haug (2017), Altshuler and Haug (forthcoming), and Altshuler and Kim (2024). See also Kameyama (1996) and Jaspars and Kameyama (1998) for another non-monotonic approach to coreference.

[3]As a reviewer points out, in SDRT it is possible to represent multiple relations, so that a final representation here would, in addition to Result, possibly maintain a weak Parallel relation (weak in that it relies on mere semantic similarity without argument identity). Such an analysis could be built without non-monotonic edits to coherence, depending on one's willingness to equate this weaker Parallel to its stronger predecessor. Nevertheless, this derivation would still require non-monotonic edits to coreference.

[4]Asher and Lascarides (2003) highlight that SDRT is flexible about when pragmatic heuristics are compiled to deliver a preferred interpretation. It is in principle possible in SDRT to provide and revise "best guesses" at every discourse unit, or to retain persistent underspecification and avoid discourse reanalysis. Given this, we could refine our question: Is there any evidence that our representations need to permit the former?

[5]Out of an abundance of caution, we will avoid calling cases of discourse reanalysis 'garden paths' *per se*, as we make no claim that they instantiate the same high degree of difficulty. See the papers in Fodor and Ferreira (1998) for careful discussion of the individual components of that difficulty and how they might vary.

Table 1: An example stimulus set from Experiment 1, following "The kids were misbehaving at a fancy dinner."

Pronoun	**Context**	**Text**
Ambiguous	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table.
	Result-biased V	Mia hit Winona with a pea and Ian scolded her.
	Result (*so*)	Mia hit Winona with a pea so Ian kicked her under the table.
Disambig (S)	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table. Mia didn't kick him in return, she simply...
Disambig (O)	Parallel-biased V	Mia hit Winona with a pea and Ian kicked her under the table. Winona didn't bother the other kids in return, she simply...
Unambig (S/O)	Parallel-biased V	Mia hit Harrison with a pea and Ian kicked (her/him) under the table.
Is it more likely that Ian kicked Mia or Winona? (Mia, Winona)		*How likely is the other option?* (1–4)

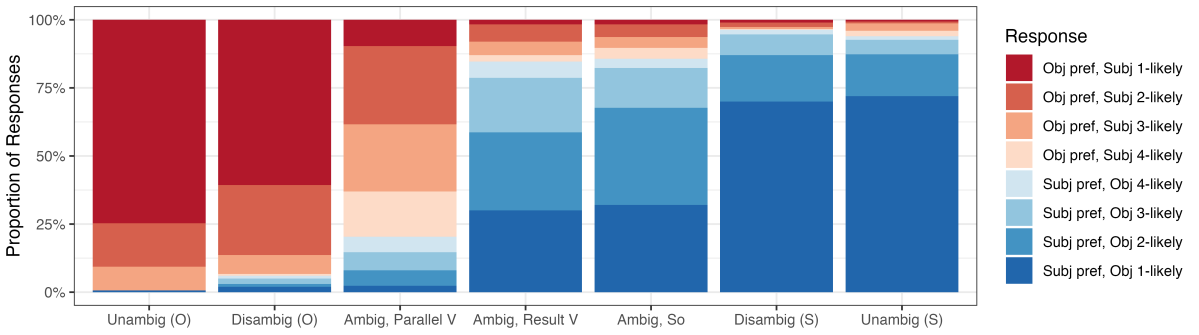


Figure 1: Results from Experiment 1.

coreference, like (1). Such discourses are uniquely interesting because of the complexity of the reanalysis they might engender, and because of the strength of their preferred interpretations. We present two experiments, first validating strong preferences for our set of narratives in a judgment study, and then examining their processing in self-paced reading. We find that when late information contradicts the preferred interpretation, readers experience a slowdown consistent with costly reanalysis. We take this to diagnose discourse reanalysis as a real phenomenon of incremental comprehension, with important consequence for formal theories.

2 Experiment 1: Interpretation preferences

In an initial experiment,[6] we collected preferred interpretations of potentially-ambiguous pronouns in 60 narratives like (1), with and without disambiguating continuations. Each narrative featured coordinated clauses with parallel syntactic structure $[S_1, S_2]$, where S_2 contained the critical pronoun in object position. The interpretation of such pronouns is biased towards an object antecedent due to parallelism (Smyth 1994), but this can be reversed, either by adding a preceding causal connective (Wolf, Gibson, and Desmet 2004), e.g. *so*, or by using a second verb biased more towards the causal relation Result than Parallel (Kehler et al. 2008). We elicited preferred antecedents for the object pronoun as a binary choice, followed by a 4-point Likert scale asking about the likelihood of the participant's dispreferred antecedent. These measures were compared across six key conditions (Table 1) to verify the contributions of the verb and connective, and further examine how preferred interpretations changed when we added an S_3 where the presupposition of *in return* disambiguated the previous pronoun.

[6]The materials, methods, and analytical choices for both experiments were pre-registered (Duff and Altshuler 2024a). Data and analysis scripts are available in our supplementary materials (Duff and Altshuler 2024b).

Table 2: Stimuli from Experiment 2. "/" indicates a chunk boundary. Critical regions in bold.

The kids/were misbehaving/at a fancy dinner.	
Ambig.	Mia hit Winona/with a pea/and/**Ian kicked her**/under the table.
Unambig.	Mia hit Harrison/with a pea/and/**Ian kicked her**/under the table.
so + Unambig.	Mia hit Harrison/with a pea/so/**Ian kicked her**/under the table.
Name	Winona hit Harrison/with a pea/and/**Ian kicked Mia**/under the table.
Mia didn't kick him/**in return**,/**she simply**/went to get more peas.	

We tested 30 native English speakers recruited on Prolific (Figure 1). Responses were combined to create an index of the direction and strength of interpretation preferences,[7] which was then analyzed using Bayesian ordinal mixed-effects regression.[8] Participants exhibited a strong preference (80%) for an object antecedent in the ambiguous, Parallel condition. This object antecedent preference was sharply reversed by causal-biased verbs, $\delta_{.95} = (-1.92, -1.24)$, by the causal connective *so*, $\delta_{.95} = (-2.05, -1.31)$, and, most strongly, by an S_3 disambiguating to a subject antecedent, $\delta_{.95} = (-3.34, -2.31)$.

These results replicate a strong object-antecedent bias in such narratives that is conditional on certain discourse coherence, and they provide a proof of concept that initially-preferred interpretations of ambiguous pronouns can be reversed by a subsequent sentence.

3 Experiment 2: Phrase-by-phrase reading times

From Experiment 1, we identified 32 narratives with the strongest parallelism bias in S_2 and the largest reversal given S_3, to investigate in a chunked, non-cumulative self-paced-reading experiment testing for costly reanalysis. We compared these 32 across four conditions (Table 2). In Ambiguous conditions, the critical object pronoun is expected to prefer an initial object antecedent due to establishing Parallel, but it is followed by an S_3 which requires a subject antecedent. In all other conditions, S_3 doesn't motivate any change in interpretation, because S_2 already makes a subject antecedent unambiguous, or lacks a pronoun. If comprehenders indeed commit to a firm interpretation of S_2 that they must then reanalyze, we should see particularly slow reading of S_3 in Ambiguous conditions. Materials were presented with 60 filler items of similar length, including 28 narratives where the discourse remained globally consistent with a Parallel relation and an object antecedent for the pronoun.

Reading times (Figure 2) from 96 participants (recruited as in Experiment 1) were analyzed using Bayesian log-normal mixed-effects regressions. We find that the final sentences were generally read slightly slower in the Ambiguous condition (red) than the other conditions, with our models estimating credible marginal slowdowns of about 15ms at *in return*, $\delta_{.95} = (1, 29)$, and *she simply*, $\delta_{.95} = (3, 25)$.[9] This result suggests that readers experienced some degree of

[7] An 8-point ordinal variable ranging from object-antecedent preferences where a subject antecedent was deemed unlikely, to subject-antecedent preferences where an object antecedent was deemed unlikely.

[8] All analyses for both experiments were conducted using `brms` (Bürkner 2017; Bürkner 2018) in R, using regularizing priors. In the text, we report 95% highest-density intervals over the posterior marginal differences between conditions extracted from our models, $\delta_{.95}$. When this interval excludes 0, we take the associated effect to be noteworthy. Note that δ in Experiment 1 is expressed in abstract model-internal units, not the units of the response scale; in Experiment 2, δ is the difference in reading times in milliseconds.

[9] Evidence for costly revision is strongest for the first trials of the experiment, and disappears by the last few, resembling adaptation effects well-known in work on syntactic garden paths (Prasad and Linzen 2021). We conjecture that participants may have learned to postpone discourse interpretation after experiencing that early interpretations often required reanalysis. An alternative possibility is that participants learned to anticipate the final subject-antecedent interpretation correctly, although we note that there is no evidence of a corresponding increase in discourse reanalysis for the fillers featuring late disambiguation to an object-antecedent interpretation.

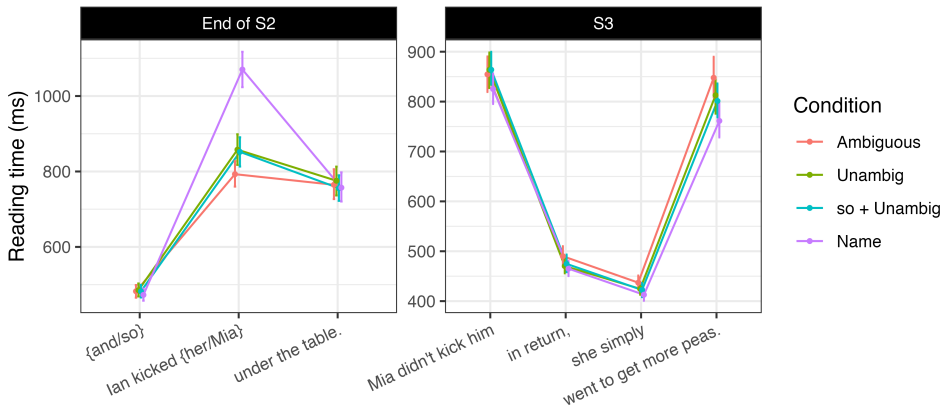


Figure 2: Results from Experiment 2.

costly reanalysis, revising an initial interpretation that yielded an object antecedent for *her*. This cost may include revision of the coherence relation between S_2 and S_3, from Parallel to Result; the cost of any such process cannot be untangled from pronominal revision.

Reading times on the core content of S_2 offer further insight. Besides costs for a new discourse referent (purple), we also see here an inversion of the later pattern; conditions with unambiguous subject antecedents (green, blue) were slower than Ambiguous (red), although this fell just below our standards for a notable contrast $\delta_{.95}$ = (-2, 57). This helps us rule out an alternative hypothesis for the later slow-down; ambiguous narratives were not globally slower as a function of some ambiguity-specific processing (cf. Stewart, Holler, and Kidd 2007; Creemers and Meyer 2022; see also Logačev and Vasishth 2016). This also cannot be a simple cost of processing an unambiguous pronoun (cf. Swets et al. 2008; Grant, Sloggett, and Dillon 2020), as unambiguous object-antecedent conditions in our filler items are not similarly slow. We take this instead as weak evidence that participants begin reading this region already expecting Parallel and thus parallelism in coreference, and comprehension is facilitated when the input matches this.[10]

4 Discussion

The reading data from Experiment 2 provides evidence for a small slowdown just when comprehenders encounter input that encourages a previously despised set of coreference and coherence choices. We take this to diagnose costly reanalysis of discourse interpretation, consistent with the idea that, at least in some cases, comprehenders make and revise firm pragmatic representations during incremental discourse comprehension, and do not somehow underspecify indefinitely.

Zooming out, the consequences of our results for theories of semantic representation depend on one's theoretical taste. We follow Brasoveanu and Dotlačil (2015) in considering two paths for how formal semantics may deal with the nature of incremental comprehension. On the first path, we could define a logic capable of representing and updating meaning incrementally. Among the theories available along this path, our results arbitrate for a dynamic semantics which permits selective non-monotonic update of coreference and other aspects of discourse meaning.

On the second path, we could maintain a simpler theory of representation, and shift the explanatory burden of incrementality onto our theory of processing. Along this path, we think our results help discriminate between two options for how a processing theory could treat discourse reanalysis. On the most extreme approach, we might imagine that discourse reanalysis requires the wholesale reconstruction of an alternative analysis permitted by the grammar, as carried out by a separate meta-system aware of the constraints from later input (a 'semantic

[10] As for implicit Results (Keenan, Baillet, and Brown 1984) and Explanations (Mak and Sanders 2013).

parser'). This would require no flexibility for discourse-pragmatic meaning in the semantic representations themselves; for instance, the conflation of pronouns and their antecedents that Haug (2014) objects to in DRT would not be a problem if the semantic parser re-parsed the entire relevant sequence with hindsight. On a more balanced approach, one could treat discourse reanalysis as an operation of a semantic parser, but using representations which do distinguish coreference (etc.) in a way which allows them to be selectively edited.[11] We take the specific reanalysis costs we observe in Experiment 2 to be far more compatible with the latter option. If our (surprisingly small) 15ms estimate is generalizable, the reanalysis of coreference seems to be rather less difficult than syntactic garden paths observed in self-paced reading, e.g., 50ms in Ferreira and Henderson (1990). This would be easier to understand if discourse reanalysis is unique in that it does not require complete reconstruction of the input—and even in a system where incrementality is largely handled outside the grammar, this would require a representation of meaning which permits selective revisions to discourse-pragmatic content.

Another open question across any of these approaches is *how often* comprehenders make and revise firm pragmatic representations; that is, we have established a case where we observe incremental interpretation and reanalysis, but is incremental interpretation the rule? Given the literature on processing coherence and coreference, we think we can best explain the diversity of previous findings if incrementality is flexible, and temporary underspecification may persist without preference in some cases. After all, other reading experiments have failed to observe costly discourse reanalysis, in cases of ambiguous coreference in a short narrative without connections to coherence (Stewart, Holler, and Kidd 2007), and in cases of ambiguous coherence without connections to coreference (Mak and Sanders 2013; Duff 2023), or with consequences for temporal ordering (Dickey 2001; Sasaki 2021). In finding some cost here, we join a few other experiments, where costly reanalysis was observed for coreference when an initial interpretation was more strongly cued by prominence as a repeated subject (Gordon and Scearce 1995), consistency with an explicit coherence relation (Mak and Sanders 2010), or consistency with world knowledge (Jones and Bergen 2024).[12] We think this split is likely the result of adaptively-timed interpretation, as would be permitted in logics like Asher and Lascarides (2003), or processing models like Karimi and Ferreira (2016) or Duff (2023), which lays out a larger argument that comprehenders learn to postpone interpretation for aspects of meaning which are contextually uncertain or practically irrelevant.

If costly reanalysis is a general matter of discourse comprehension, we predict similar effects when a strong coherence bias drives resolution of e.g. temporal ordering. To this end, our next experiments will test cases like (2), where S_1 and S_2 first appear to relate via Explanation, such that the event described by S_2 would be understood to precede the event described by S_1. S_3 would then force reanalysis to Result, ensuring the opposite event ordering.

(2) [Bob sued Liz.]$_{S_1}$ [She defaced his shop,]$_{S_2}$ [because his lawsuit was cruel.]$_{S_3}$

This reanalysis would be from a subordinating coherence relation to a coordinating one, in the sense of Asher and Vieu (2005), whereas our present study featured only coordinating relations. If coherence revision is a source of difficulty, as we hypothesize, this discourse structural shift may drive especially large reanalysis costs.

Acknowledgements. Thanks to Kelsey Sasaki, Christina Kim, Julian J. Schlöder, Adrian Brasoveanu and three anonymous reviewers for insightful input. The experiments were financially supported by the John Fell Fund, awarded to the authors by Oxford University Press.

[11] For a starting point, see Brasoveanu and Dotlačil's (2020) incremental DRT parser in ACT-R.

[12] Experiments with other designs are likewise split, compare Gernsbacher (1989) and Greene, McKoon, and Ratcliff (1992) vs. Arnold et al. (2000), Pyykkönen and Järvikivi (2010), and Rohde, Levy, and Kehler (2011).

References

Albrecht, Jason E. and Edward J. O'Brien (1993). "Updating a mental model: Maintaining both local and global coherence". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19, pp. 1061–1070. DOI: 10.1037/0278-7393.19.5.1061.

Allbritton, David W., Gail McKoon, and Roger Ratcliff (1996). "Reliability of prosodic cues for resolving syntactic ambiguity". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 22.3, pp. 714–735. DOI: 10.1037/0278-7393.22.3.714.

Altshuler, Daniel and Dag T. T. Haug (2017). "The Semantics of provisional, temporal anaphors and cataphors". Talk presented at Deutsche Gesellschaft für Sprachwissenschaft, Saarland University, Saarbrücken.

— (forthcoming). *Literature as a formal language: The semantics and pragmatics of narrative garden path in 'Sylvie'*. Routledge.

Altshuler, Daniel and Christina S. Kim (2024). "Towards a typology of narrative frustration". In: *Topoi* 43, pp. 1193–1210. DOI: 10.1007/s11245-023-09942-1.

Arnold, Jennifer E. et al. (2000). "The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking". In: *Cognition* 76, B13–B26. DOI: 10.1016/s0010-0277(00)00073-1.

Asher, Nicholas and Alex Lascarides (2003). *Logics of conversation*. Cambridge.

Asher, Nicholas and Laure Vieu (2005). "Subordinating and coordinating discourse relations". In: *Lingua* 115, pp. 591–610. DOI: 10.1016/j.lingua.2003.09.017.

Brasoveanu, Adrian and Jakub Dotlačil (2015). "Incremental and predictive interpretation: Experimental evidence and possible accounts". In: *Proceedings of SALT 25*, pp. 57–81.

— (2020). *Computational cognitive modeling and linguistic theory*. Springer. DOI: 10.1007/978-3-030-31846-8.

Bürkner, Paul-Christian (2017). "brms: An R package for Bayesian multilevel models using Stan". In: *Journal of Statistical Software* 80.1, pp. 1–28. DOI: 10.18637/jss.v080.i01.

— (2018). "Advanced Bayesian multilevel modeling with the R package brms". In: *The R Journal* 10, pp. 395–411. DOI: 10.32614/rj-2018-017.

Carlson, Katy et al. (2009). "Information structure expectations in sentence comprehension". In: *Quarterly Journal of Experimental Psychology* 62.1, pp. 114–139. DOI: 10.1080/17470210701880171.

Creemers, Ava and Antje S. Meyer (2022). "The processing of ambiguous pronominal reference is sensitive to depth of processing". In: *Glossa Psycholinguistics* 1.1, p. 3. DOI: 10.5070/G601166.

Dickey, Michael Walsh (2001). *The processing of tense*. Kluwer. DOI: 10.1007/978-94-010-0568-5.

Duff, John (2023). *On the timing of decisions about meaning during incremental comprehension*. Dissertation, UC Santa Cruz.

Duff, John and Daniel Altshuler (2024a). *Probing Parallel-to-Result reanalysis*. Preregistration hosted by the OSF, filed June 24, 2024. URL: https://osf.io/4sz79?view_only=bc75f487b79e47da9fe5c7f3765805e3.

— (2024b). *Reanalysis in discourse comprehension: Supplementary materials*. Repository hosted by the OSF. URL: https://osf.io/mk38v/?view_only=e01b727827b1465c95848f467b0713b1.

Duffy, Susan A., Robin K. Morris, and Keith Rayner (1988). "Lexical ambiguity and fixation times in reading". In: *Journal of Memory and Language* 27, pp. 429–446. DOI: 10.1016/0749-596X(88)90066-6.

Egg, Markus (2010). "Semantic underspecification". In: *Language and Linguistics Compass* 4.3, pp. 166–181. DOI: 10.1111/j.1749-818X.2010.00188.x.

Ferreira, Fernanda and John M. Henderson (1990). "Use of verb expectation in syntactic parsing: Evidence from eye movements and word-by-word self-paced reading". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16.4, pp. 555–568. DOI: 10.1037/0278-7393.16.4.555.

Fodor, Janet Dean and Fernanda Ferreira (1998). *Reanalysis in sentence processing*. Springer. DOI: 10.1007/978-94-015-9070-9.

Frazier, Lyn and Keith Rayner (1982). "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences". In: *Cognitive Psychology* 14.2, pp. 178–210. DOI: 10.1016/0010-0285(82)90008-1.

Gernsbacher, Morton A. (1989). "Mechanisms that improve referential access". In: *Cognition* 32, pp. 99–156. DOI: 10.1016/0010-0277(89)90001-2.

Gordon, Peter C. and Kimberly A. Scearce (1995). "Pronominalization and discourse coherence, discourse structure and pronoun interpretation". In: *Memory & Cognition* 23.3, pp. 313–323. DOI: 10.3758/BF03197233.

Grant, Margaret, Shayne Sloggett, and Brian Dillon (2020). "Processing ambiguities in attachment and pronominal reference". In: *Glossa* 5.1, p. 77. DOI: 10.5334/gjgl.852.

Greene, Steven B., Gail McKoon, and Roger Ratcliff (1992). "Pronoun resolution and discourse models". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18.2, pp. 266–283. DOI: 10.1037/0278-7393.18.2.266.

Haug, Dag T. T. (2014). "Partial dynamic semantics for anaphora: Compositionality without syntactic coindexation". In: *Journal of Semantics* 31, pp. 457–511. DOI: 10.1093/jos/fft008.

Hobbs, Jerry R. (1979). "Coherence and coreference". In: *Cognitive Science* 3, pp. 67–90. DOI: 10.1207/s15516709cog0301_4.

Jaspars, Jan and Megumi Kameyama (1998). "Discourse preferences in dynamic logic". In: *Computing natural language*. Ed. by Atocha Aliseda, Rob van Glabbeek, and Dag Westertahl. CSLI, pp. 67–96.

Jones, Cameron R. and Benjamin Bergen (2024). "Does word knowledge account for the effect of world knowledge on pronoun interpretation?" In: *Language and Cognition*. Advance online publication. DOI: 10.1017/langcog.2024.2.

Kameyama, Megumi (1996). "Indefeasible semantics and defeasible pragmatics". In: *Quantifiers, deduction, and context*. Ed. by Makoto Kanazawa, Christopher Piñón, and Henriette de Swart. CSLI, pp. 111–138.

Kamp, Hans and Uwe Reyle (1993). *From discourse to logic: Introduction to model theoretic semantics of natural language, formal logic and Discourse Representation Theory*. Kluwer. DOI: 10.1007/978-94-017-1616-1.

Karimi, Hossein and Fernanda Ferreira (2016). "Good-enough linguistic representations and online cognitive equilibrium in language processing". In: *Quarterly Journal of Experimental Psychology* 69.5, pp. 1013–1040. DOI: 10.1080/17470218.2015.1053951.

Keenan, Janice M., Susan D. Baillet, and Polly Brown (1984). "The effects of causal cohesion on comprehension and memory". In: *Journal of Verbal Learning and Verbal Behavior* 23, pp. 115–126. DOI: 10.1016/S0022-5371(84)90082-3.

Kehler, Andrew et al. (2008). "Coherence and coreference revisited". In: *Journal of Semantics* 25.1, pp. 1–44. ISSN: 0167-5133. DOI: 10.1093/jos/ffm018.

Lascarides, Alex and Nicholas Asher (1993). "Temporal interpretation, discourse relations and commonsense entailment". In: *Linguistics and Philosophy* 16, pp. 437–493. DOI: 10.1007/BF00986208.

Logačev, Pavel and Shravan Vasishth (2016). "A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension". In: *Cognitive Science* 40, pp. 266–298. DOI: 10.1111/cogs.12228.

Mak, Willem M. and Ted J. M. Sanders (2010). "Incremental discourse processing: How coherence relations influence the resolution of pronouns". In: *The linguistics enterprise: From knowledge of language to knowledge in linguistics*. Ed. by Martin Everaert et al. John Benjamins, pp. 167–182.

— (2013). "The role of causality in discourse processing: Effects of expectation and coherence relations". In: *Language and Cognitive Processes* 28.9, pp. 1414–1437. ISSN: 0169-0965. DOI: 10.1080/01690965.2012.708423.

Prasad, Grusha and Tal Linzen (2021). "Rapid syntactic adaptation in self-paced reading: Detectable, but only with many participants". In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 47.7, pp. 1156–1172. DOI: 10.1037/xlm0001046.

Pyykkönen, Pirita and Juhani Järvikivi (2010). "Activation and Persistence of Implicit Causality Information in Spoken Language Comprehension". In: *Experimental Psychology* 57.1, pp. 5–16. DOI: 10.1027/1618-3169/a000002.

Rohde, Hannah, Roger Levy, and Andrew Kehler (2011). "Anticipating explanations in relative clause processing". In: *Cognition* 118.3, pp. 339–358. DOI: 10.1016/j.cognition.2010.10.016.

Sasaki, Kelsey (2021). *Components of coherence*. Dissertation, UC Santa Cruz.

Smyth, Ron (1994). "Grammatical determinants of ambiguous pronoun resolution". In: *Journal of Psycholinguistic Research* 23.3, pp. 197–229. DOI: 10.1007/BF02139085.

Stewart, Andrew J., Judith Holler, and Evan Kidd (2007). "Shallow processing of ambiguous pronouns: Evidence for delay". In: *Quarterly Journal of Experimental Psychology* 60.12, pp. 1680–1696. DOI: 10.1080/17470210601160807.

Swets, Benjamin et al. (2008). "Underspecification of syntactic ambiguity: Evidence from self-paced reading". In: *Memory & Cognition* 36.1, pp. 201–216. DOI: 10.3758/mc.36.1.201.

Syrett, Kristen, Georgia Simon, and Kirsten Nisula (2012). "Prosodic disambiguation of scopally ambiguous sentences". In: *Proceedings of NELS 43*, pp. 141–152.

Wolf, Florian, Edward Gibson, and Timothy Desmet (2004). "Discourse coherence and pronoun resolution". In: *Language and Cognitive Processes* 19.6, pp. 665–675. DOI: 10.1080/01690960444000034.

Ignorance under attitudes

Patrick D. Elliott
Heinrich-Heine University Düsseldorf
Düsseldorf, Germany
patrick.d.elliott@gmail.com

Filipe Hisao Kobayashi
Paris Lodron University of Salzburg
Salzburg, Austria
filipe.kobayashi@plus.ac.at

Abstract

Yablo (2014) observed that *Ann and Beth agree that p or q* conveys that p and q are both compatible with the beliefs of Ann and Beth. While Yablo argued for a radical departure from possible world semantics, we propose that this is due to local pragmatic reasoning. We frame our proposal within Update Semantics and develop a system where *p or q* requires the context to be compatible with both p and q. When disjunction is unembedded, our system yields ignorance inferences. Yablo's puzzle is accounted for because, under attitude verbs the requirement is imposed in the complement clause's local context.

1 Yablo's problem

Consider the following valid argument, based on a case discussed by Yablo (2014, p. 12). Even when x and y's individual beliefs differ, we can still truthfully say that x and y *agree* on certain propositions.

(1) a. Stephen believes [that snow is white and expensive].
b. Daniel believes [that snow is white, but not expensive].
c. (1-a) ∧ (1-b) ⇒ Stephen and Daniel agree [that snow is white].

A first stab at a semantics for *agree* that captures validity of (1) is given in (2) — it says that x and y *agree* on ϕ if each of x and y individually believe the proposition expressed by ϕ.[1] This is effective simply because a sentence of the form "ϕ and ψ" classically entails both ϕ, and ψ.

(2) Truth-condition for an *agree*-report (first attempt):
$[\![x \text{ and } y \text{ agree that } \phi]\!]^w = 1$ if $\forall w' \in Dox^x_w, [\![\phi]\!]^{w'} = 1 \wedge \forall w' \in Dox^y_w, [\![\phi]\!]^{w'} = 1$

As pointed out by Yablo (2014), this simple-minded assumption about the semantics of *agree* immediately leads to strange results. For example, it predicts the validity of (3), which does not seem to be intuitively correct. The prediction straightforwardly follows from the classical validity of disjunction introduction: both "snow is white" and "snow is black" classically entail that "snow is white or black".

(3) a. Stephen believes [that snow is white].
b. Daniel believes [that snow is black].
c. Stephen and Daniel agree [that snow is white or black].

It is difficult to see how to block this prediction with classical resources while maintaining an account of (1), but it is very natural to suspect that the use of disjunction in (3-c) leads to additional non-semantic inferences that contribute to the sense that (3) is not valid. To spell this out, it is useful to consider a very simple case of disjunction embedded in a *belief* report.

(4) a. Daniel believes [that snow is white].

[1] We use Dox^x_w to talk about the set of worlds that x considers a candidate for the actual world in w, i.e., x's *doxastic alternatives*.

b. Daniel believes [that snow is white or black].

(5) $[\![x \text{ believes } \phi]\!]^w = 1$ iff $\forall w' \in Dox^x_w, [\![\phi]\!]^{w'} = 1$

Intuitively, (4) is felt not to be valid, despite what is predicted by a relatively orthodox, Hintikkan semantics for *believe* Hintikka 1969, (5). There is however a pragmatic explanation for the feeling of non-validity, which follows from basic Gricean reasoning (Grice 1989). Let's make this concrete: the sentence in (4-b) evokes, among others, alternative *belief* reports involving the individual disjuncts, which are each presumably relevant to the current topic of conversation, (6-a). Given that the speaker has chosen to assert (4-b) rather than the logically stronger alternatives in (6-a), on the assumption that she is operating in accordance to the Maxim of Quantity, we can conclude that, for each $\psi \in$ (6-a), it's not the case that she believes ψ. Together with (4-b), the resulting implicature conveys speaker ignorance about the status of each alternative.

(6) a. $\{$ D. believes that snow is white, D. believes that snow is black $\} \subseteq Alt$(4-b)

We can now consider the disjunctive *belief*-report together with its implicatures (7-a) and (7-b). Note that the resulting *enriched* meaning of the disjunctive *belief* report is incompatible with the each simpler alternative. This is plausibly responsible for the feeling that (4) is not valid.

(7) Daniel believes that snow is white or black.
   a. $\rightsquigarrow$ *the speaker isn't certain that Daniel believes that snow is white.*
   b. $\rightsquigarrow$ *the speaker isn't certain that Daniel believes that snow is black.*

This reasoning however does not extend to Yablo's problem (see Rothschild 2017 for discussion). To see this, consider the simpler alternatives evoked by a disjunctive *agree* report. In a context of (3), each of Alt_1 and Alt_2 are already known to be false, since Daniel and Stephen have incompatible beliefs about the color of snow.

(8) Daniel and Stephen agree [that snow is white or black].
   a. Alt_1: *Daniel and Stephen agree that snow is white.*
   b. Alt_2: *Daniel and Stephen agree that snow is black.*

Yablo uses the problem of agreement, among many other cases, to motivate a theory of partial content, couched in truthmaker semantics (see also Fine 2017). Within the framework of truthmaker semantics, a non-classical notion of entailment can be stated according to which "ϕ and ψ" entails "ϕ", and "ψ", but "ϕ" doesn't entail "ϕ or ψ". In this paper however we'll suggest that it would be too hasty to dismiss a pragmatic explanation for the invalidity of (3), in light of the possibility of local pragmatic enrichment. We turn to this notion in the next section.

2 Local pragmatic enrichment

2.1 Background: minimal worlds exhaustification

There is a significant body of evidence indicating that class of inferences traditionally derived as Gricean conversational implicatures are instead amenable to an analysis in terms of an embeddedable, covert operator *Exh* (see Chierchia, Fox, and Spector 2012 for an overview). One such inference is the so-called *scalar implicature* associated with indefinites

(9) Alex ate some of the cookies.
   a. $\rightsquigarrow$ *Alex didn't eat all of the cookies*

There are various formulations of Exh (see especially Spector 2016 for discussion), and the one that we build upon in the present work is due to van Rooij and Schulz (2004): the so called *minimal worlds* exhaustivity operator (Exh_{mw}). As is standard, we relativize the interpretation of Exh_{mw} applied to a sentence ϕ to the set of *alternatives* to ϕ, $Alt(\phi)$. The definition of Exh_{mw} proceeds in two steps. First we define an ordering on worlds relative to $Alt(\phi)$: $w \leq_{Alt(\phi)} v$ when every element of $Alt(\phi)$ that is true in w is also true in v.

(10) $w \leq_{Alt(\phi)} v := \{ \psi \in Alt(\phi) \mid [\![\psi]\!]^w = 1 \} \subseteq \{ \psi \in Alt(\phi) \mid [\![\psi]\!]^v = 1 \}$

To see how this works, it will be useful to consider a schematic example. Consider a disjunctive sentence $a \vee b$. The standardly assumed alternatives (Sauerland 2004) are listed below:

(11) $Alt(a \vee b) = \{ a \vee b, a \wedge b, a, b \}$

Assume the following worlds, where subscripts indicate which disjuncts are true: $w_{ab}, w_a, w_b, w_\emptyset$. First, let's determine which alternatives are true in which worlds.

- $\{ \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_{ab}} = 1 \} = \{ a \vee b, a \wedge b, a, b \}$
- $\{ \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_a} = 1 \} = \{ a \vee b, a \}$
- $\{ \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_b} = 1 \} = \{ a \vee b, b \}$
- $\{ \psi \in Alt(a \vee b) \mid [\![\psi]\!]^{w_\emptyset} = 1 \} = \emptyset$

The result is that the worlds which make the *fewest alternatives true* are the most minimal:

(12) $w_\emptyset < w_a, w_b < w_{ab}$

Now, minimal worlds exhaustification can be derivatively defined: it's true in any world w such that (i) the prejacent is true (i.e., the assertion) in w, and (ii) there is no world v that is more minimal than w and the prejacent is true in v.

(13) **Minimal worlds exhaustification:**
$[\![Exh_{mw}\, \phi]\!]^w = 1$ iff $\underbrace{[\![\phi]\!]^w = 1}_{\text{assertion}}$ and $\underbrace{\neg\exists v[[\![\phi]\!]^v = 1 \wedge v <_{Alt(\phi)} w]}_{\text{implicature}}$

Ignoring $w_\emptyset$ (since there, the prejacent is false), the minimal worlds in which the prejacent is true are w_a and w_b. It follows that the proposition expressed by $Exh_{mw}(a \vee b)$ is $[\lambda w\,.\, w \in \{ w_a, w_b \}]$, i.e., either a or b is true, but not both — the scalar implicature is derived.

An interesting property of minimal worlds exhaustification is that it does not rely on the presence of the scalar alternative $a \wedge b$ — the so-called domain alternatives are sufficient in order to derive the scalar implicature. This is easy to see — even ignoring $a \wedge b$, both w_a and w_b are more minimal than w_{ab} with respect to alternatives $\{ a, b \}$, since w_{ab} makes both alternatives true, and each of w_a and w_b make just one alternative true.

As it stands, note that minimal worlds exhaustification does not predict so-called ignorance inferences associated with disjunction. If a speaker asserts $a \vee b$, it typically implicates not just that she is certain that $a \wedge b$ is false (the scalar implicature), but also that she is not certain that a, and is not certain that b (the ignorance inferences). The classical strategy is to derive ignorance resources via Gricean reasoning (see, e.g., Fox 2007 for an explicitly hetereogeneous approach to scalar implicatures and ignorance inferences). Our main innovation will be to extend minimal worlds exhaustification in order to capture both of these inferences.

2.2 Exhaustification and local contexts

Since ignorance inferences amount to requirements on information states, it is not possible to develop an embeddable account of ignorance inferences without reference to a *local context*. We implement local contexts in a straightforward way by adopting *update semantics*; sentences describe ways of updating information states (Heim 1982; Dekker 1993). Certain operators, such as the logical connectives, modals/attitudes, and (crucially for us) Exh, are defined directly in terms of updates.

(14) For all sentences ϕ not defined directly as updates,
$c[\phi] := \{ w \in c \mid [\![\phi]\!]^w = 1 \}$

(15) For all sentences ϕ directly defined as updates,
$[\![\phi]\!]^w = 1$ iff $w \in W[\phi]$

(16) Updates for complex sentences:
a. $c[\phi \wedge \psi] := c[\phi][\psi]$
b. $c[\phi \vee \psi] := c[\phi] \cup c[\psi]$
c. $c[\neg\phi] := c - c[\phi]$

With the additional expressive power of update semantics, we will define a new formulation of minimal worlds exhaustification, crucially stated with respect to a local context c. We follow Bassi, Del Pinal, and Sauerland 2021 in that, like other presupposition triggers, Exh places constraints on c.[2] First we'll provide the definition of Exh we assume, and then we'll go through it step by step:

(17) $$c[Exh(\phi)] := \begin{cases} c[\phi] & \text{if } \forall\psi \in Alt(\phi)[\exists w[[\![Exh_{mw}\,\phi]\!]^w = 1 \wedge [\![\psi]\!]^w = 1] \leftrightarrow c[\psi] \neq \emptyset] \\ \text{undefined} & \text{otherwise} \end{cases}$$

$Exh(\phi)$ places a definedness condition on c. Informally, the definedness condition placed on c by $Exh(\phi)$ is that the alternatives in $Alt(\phi)$ that are compatible with c are exactly those alternatives in $Alt(\phi)$ that are compatible with $Exh_{mw}(\phi)$. If this definedness condition is met, then the context is simply updated with the prejacent ϕ.

Let's see how this works for disjunction. We already know that $Exh_{mw}(a \vee b)$ is the proposition is true in worlds where only one of a and b is true. We now must compute the definedness condition on c triggered by $Exh(a \vee b)$:

- $a \wedge b$ is logically incompatible with $Exh_{mw}(a \vee b)$, hence $c[a \wedge b] = \emptyset$.
- $a \vee b$ is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[a \vee b] \neq \emptyset$.
- a is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[a] \neq \emptyset$.
- b is logically compatible with $Exh_{mw}(a \vee b)$, hence $c[b] \neq \emptyset$

Thus:

(18) $$c[Exh\,(a \vee b)] = \begin{cases} c[a \vee b] & \text{if } c[a \wedge b] = \emptyset \wedge c[a] \neq \emptyset \wedge c[b] \neq \emptyset \wedge c[a \vee b] \neq \emptyset \\ \text{undefined} & \text{otherwise} \end{cases}$$

Updating with $Exh\,(a \vee b)$ therefore requires that $a \wedge b$ is known to be false (the scalar implicature), that the domain alternatives a, b are considered possible, and that the prejacent is considered

[2] See Bassi, Del Pinal, and Sauerland 2021 for discussion of this treatment of scalar implicatures as presuppositions.

possible. N.b., that *contextual ignorance* only follows in conjunction with the scalar implicature (see Degano et al. 2024 for experimental evidence that *possibility* should be derived separately from uncertainty).

3 Accounting for Yablo's problem

Back to Yablo's problem — recall that a simple-minded understanding of what it means for two individuals x and y to *agree* on a proposition ϕ is that both of x's and y's beliefs individually entail ϕ. This works for some simple cases involving disjunctive beliefs, but incorrectly predicts that x and y may agree on disjunctive propositions that are weaker than their respective beliefs, due to properties of classical entailment. Our central claim will be that Yablo's problem can be resolved without shifting to a non-classical notion of entailment by invoking Exh, in order to disrupt the predicted entailment. Concretely, we'll assume that the Logical Form of an *agree* report is as follows:

(19) Stephen and Daniel agree [Exh that snow is white or black].

In order to figure out what our analysis predicts, we need a principled account of the local context of *agree*, based on how presuppositions project from the complement of *agree*. Our contention is that "x and y agree that ϕ_π" (where π is the presupposition of ϕ) presupposes that each of x and y believe that π, mirroring the projection properties of other attitude verbs (Heim 1992). We demonstrate this via standard filtering tests:

(20) Context: *Sally doesn't smoke.*

a. Both Stephen and Daniel believe that Sally used to smoke, and they agree that she stopped.

b. Both Stephen and Daniel believe that Sally used to smoke, but they don't agree that she stopped.

Following Heim's entries for attitude verbs, we build this behavior into the update contributed by the *agree* report, as in (21). It presupposes that the context entails that x and y both individually believe the presuppositions of ϕ, and asserts that both x and y's beliefs entail that ϕ. In other words, we maintain a simple entailment-based semantics for *agree*, while being a bit more precise about how presuppositions project.[3]

(21) $c[x \text{ and } y \text{ agree } \phi] \neq \#$ when $\forall w \in c, Dox^x_w[\phi]$ is defined $\wedge\, Dox^y_w[\phi]$ is defined.
When defined, $c[x \text{ and } y \text{ agree } \phi] = \{\, w \in C \mid Dox^x_w[\phi] = Dox^y_w \wedge Dox^y_w[\phi] = Dox^y_w \,\}$

We can now resolve Yablo's problem. Consider what is predicted as the update associated with "Exh that snow is white or black" ($Exh(Wh \vee Bl)$): it asserts that snow is white or black, presupposes that snow isn't both white and black, and that it's contextually possible that snow is white, and possible that snow is black.

(22) $c[Exh\,(Wh \vee Bl)] \neq \#$ only if $\neg\exists w \in c[Wh(w) \wedge Bl(w)], \exists w \in c[Wh(w)], \exists w \in c[Bl(w)]$

Now we can compute the result of updating the global context with "Stephen and Daniel agree that [Exh Snow is white or black]". What we predict is that this presupposes that Stephen and Daniel both consider it false that Snow is white and black, and that they each consider it possible that Snow is white, and possible that snow is black. This is demonstrated formally below:

[3] We set aside the presuppositions triggered by *agree* as they are not relevant for this particular set of data.

(23) $c[s \text{ and } d \text{ agree that } Exh\,(Wh \vee Bl)] \neq \#$ only if:
$\forall w \in c, \neg\exists w' \in Dox^s_w[Wh(w) \wedge Bl(w)], \exists w \in Dox^s_w[Wh(w)], \exists w \in Dox^s_w[Bl(w)],$
$\neg\exists w' \in Dox^d_w[Wh(w) \wedge Bl(w)], \exists w \in Dox^d_w[Wh(w)], \exists w \in Dox^d_w[Bl(w)]$
When defined, $c[s \text{ and } d \text{ agree that } Exh\,(Wh \vee Bl)]$
$= \{\, w \in c \mid \forall w \in Dox^s_w, Wh(w) \vee Bl(w) \text{ and } \forall w \in Dox^d_w, Wh(w) \vee Bl(w)\}$

Intuitively, the reasons that this doesn't jointly follow from "Stephen believes that snow is white", and "Daniel believes that snow is black", is because former doesn't entail that Stephen considers it possible that snow is black, and the latter doesn't entail that Daniel considers it possible that snow is white. To be concrete, we assume a simple Heimian semantics for *belief* reports as in (24).[4]

(24) $$c[x \text{ believes } \phi] = \begin{cases} \{\, w \in c \mid Dox^x_w[\phi] = Dox^x_w \,\} & \forall w \in c, Dox^x_w[\phi] \text{ is defined} \\ \text{undefined} & \text{otherwise} \end{cases}$$

4 Conclusion

In this paper, we have showed that Yablo's problem can be resolved once we take into consideration how implicatures are calculated within embedded clauses. We proposed a novel way of deriving the ignorance implicatures of disjunction via a silent *Exh* operator whose semantics imposes constraint on its local context.

A clear advantage of our proposal is that it can account for the observation of Rothschild (2017) that the following entailment goes through:

(25) Ann believes Beth ate the cookies
Cleo believes Deb ate the cookies ⇒ Ann and Cleo agree that a girl ate the cookies.

Our approach predicts that the validity of the inference in (25) should directly track the robustness of the corresponding *ignorance inference*. Here, we observe that there is a clear contrast between (26-a) and (26-b) — conveys that both *Cleo eating the cookies*, and *Beth eating the cookies* are open possibilities. (26-a) does not always lead to the same inference; for example, perhaps it is relevant to the topic of conversation that a *girl* rather than a *boy* ate the cookies.

(26) a. A girl ate the cookies.
b. Cleo or Beth ate the cookies.

Acknowledgements. We're grateful to Alexandros Kalomoiros, Jacopo Romoli, and Raven Zhang. This research was funded in whole or in part by the Austrian Science Fund (FWF) 10.55776/F1003.

References

Bassi, Itai, Guillermo Del Pinal, and Uli Sauerland (2021). "Presuppositional exhaustification". In: *Semantics and Pragmatics*. DOI: 10.3765/sp.14.11. URL: https://doi.org/10.3765/sp.14.11.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (2012). "Scalar implicature as a grammatical phenomenon". In: *Semantics: an international handbook of natural language meaning*. Ed. by Claudia Maienborn, Klaus von Heusinger, and Paul Portner. Vol. 3. De Gruyter Mouton.

[4] There is however a sense in which entailment goes through. Specifically, "S. believes snow is white and D. believes snow is black" *Strawson entails* "S. and D. agree that *Exh* snow is white or black" (von Fintel 1999).

Degano, Marco et al. (2024). "The ups and downs of ignorance". Unpublished ms., lingbuzz/007389. URL: https://ling.auf.net/lingbuzz/007389.

Dekker, Paul (1993). "Transentential meditations: ups and downs in dynamic semantics". PhD thesis. University of Amsterdam.

Fine, Kit (2017). "Truthmaker Semantics". In: *A Companion to the Philosophy of Language*. John Wiley & Sons, Ltd, pp. 556–577. ISBN: 978-1-118-97209-0. DOI: 10.1002/9781118972090.ch22.

Fox, Danny (2007). "Free Choice and the Theory of Scalar Implicatures". In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. London: Palgrave Macmillan UK, pp. 71–120.

Grice, Herbert Paul (1989). *Studies in the Way of Words*. Cambridge: Harvard University Press.

Heim, Irene (1982). "The semantics of definite and indefinite noun phrases". PhD thesis. University of Massachusetts, Amherst.

— (1992). "Presupposition Projection and the Semantics of Attitude Verbs". In: *Journal of Semantics* 9.3, pp. 183–221. ISSN: 1477-4593. DOI: 10.1093/jos/9.3.183. URL: http://dx.doi.org/10.1093/jos/9.3.183.

Hintikka, Jaakko (1969). "Semantics for Propositional Attitudes". In: *Philosophical Logic*. Ed. by J. W. Davis, D. J. Hockney, and W. K. Wilson. Synthese Library 20. Springer Netherlands, pp. 21–45.

Rothschild, Daniel (2017). "Yablo's semantic machinery". In: *Philosophical Studies* 174, pp. 787–796. DOI: 10.1007/s11098-016-0759-3. URL: https://doi.org/10.1007/s11098-016-0759-3.

Sauerland, Uli (2004). "Scalar implicatures in complex sentences". In: *Linguistics and Philosophy* 27, pp. 367–391. DOI: 10.1023/B:LING.0000023378.71748.db. URL: https://doi.org/10.1023/B:LING.0000023378.71748.db.

Spector, Benjamin (Apr. 27, 2016). "Comparing Exhaustivity Operators". In: *Semantics and Pragmatics* 9.0 (0), pp. 11-1–33. ISSN: 1937-8912. DOI: 10.3765/sp.9.11. URL: https://semprag.org/index.php/sp/article/view/sp.9.11 (visited on 01/15/2021).

van Rooij, Robert and Katrin Schulz (2004). "Exhaustive interpretation of complex sentences". In: *Journal of Logic, Language and Information* 13, pp. 491–519. DOI: 10.1007/s10849-004-2118-6. URL: https://doi.org/10.1007/s10849-004-2118-6.

Von Fintel, Kai (1999). "NPI Licensing, Strawson Entailment, and Context Dependency". In: *Journal of Semantics* 16, pp. 97–148.

Yablo, Stephen (2014). *Aboutness*. Princeton, New Jersey: Princeton University Press.

Putting summative predicates into context

Kurt Erbach
Goethe-University Frankfurt am Main, Germany
& Saarland University, Saarbrücken, Germany
erbach@hhu.de

Stavroula Alexandropoulou
University College London
London, UK
stavroulaale@gmail.com

Richard Breheny
University College London
London, UK
r.breheny@ucl.ac.uk

Clemens Mayr
Georg-Austin-Universität Göttingen
Göttingen, Germany
clemens.steiner-mayr1@uni-goettingen.de

Jacopo Romoli
Heinrich-Heine-Universität
Düsseldorf, Germany
jacopo.romoli@gmail.com

Yasutada Sudo
University College London
London, UK
y.sudo@ucl.ac.uk

Abstract

The present paper investigates homogeneity and non-maximality effects of summative predicates like color terms (e.g., *blue*). It specifically tests theoretical accounts that make diverging predictions as to whether summative predicates in positive and negative sentences exhibit non-maximality effects to the same extent. To adjudicate between these accounts, we conducted experiments that tested how the availability of non-maximal readings is modulated by the conversational context. Our results were largely symmetric: Non-maximality readings of summative predicates show similar context-sensitivity in simple positive ('The flag is blue') and negative sentences ('The flag is not blue'), and under the quantifiers 'every' ('Every flag is blue') and 'no' ('No flag is blue'). We discuss the theoretical consequences of these findings in light of Romoli et al.'s (2024) observation that the non-maximal readings of plural definites are asymmetric with greater context sensitivity in the scope of 'every' than in the scope of 'no'.

1 Introduction

Summative predicates like color terms give rise to homogeneity effects (Krifka 1990; Löbner 2000; Paillé 2022; 2023a; 2023b). For example the positive sentence in (1) is true iff ⟦blue⟧ is true of every part of the flag under discussion, and its negative counterpart in (2) is true iff ⟦blue⟧ is true of no part of the flag, which is a stronger meaning than the logical negation of (1) ('⟦blue⟧ is true of not every part of the flag').

(1) The flag is blue.

(2) The flag is not blue.

It is also known that summative predication sometimes allows for *non-maximality effects*. That is, for (1) to be judged as true, it is often not absolutely necessary that the entire flag be blue. For instance, in many actual conversational contexts, (3) is judged as true of the EU flag (🇪🇺), despite its yellow stars.

In recent years, homogeneity and non-maximality effects have primarily been discussed in connection with plural definites (e.g. Malamud 2012; Križ 2016; Križ and Spector 2021; Bar-Lev 2018; Bar-Lev 2021), but in a series of recent works (Paillé 2022; Paillé 2023a; Paillé 2023b) Mathieu Paillé has turned the spotlight on summative predicates, discussing multiple ways of accounting for their homogeneity and non-maximality effects in terms of exhaustification. The

different theories he considers make diverging predictions as to whether summative predicates in positive and negative sentences exhibit non-maximality effects to the same extent. To adjudicate between them, we ran experiments inspired by Romoli et al.'s (2024) study on plural definites to test how the availability of non-maximal readings is modulated by the conversational context. Our results were largely symmetric: Non-maximality readings of summative predicates show similar context-sensitivity in the simple sentences (1) and (2), and under the quantifiers *every* and *no*. We discuss the theoretical consequences of these findings in light of Romoli et al.'s (2024)'s observation that the non-maximal readings of plural definites are asymmetric with greater context sensitivity in the scope of *every* than in the scope of *no*.

2 Theoretical background

Paillé 2022 proposes to account for homogeneity effects of summative predicates in terms of exhaustification, and discusses versions of his account that make different predictions for non-maximality effects (see also Paillé 2023a; Paillé 2023b). They can be said to be either *asymmetric* in that non-maximality readings are predicted to be more robustly available in positive than negative sentences, or *symmetric* in that polarity should not affect the availability of non-maximal readings. Let us look into them in some detail.

Both versions of Paillé's theory posit an existential meaning for predicates like *blue*—e.g. $[\![\text{blue}]\!] = \lambda y.\ \exists x \sqsubseteq y[\mathsf{blue}(x)]$ (where $\sqsubseteq$ is the mereological part-of relation). The truth-conditions this analysis derives straightforwardly account for the negative sentence in (2), but are too weak for the positive sentence in (1). Paillé proposes that sentences like (1) are strengthened via exhaustification with respect to alternative sentences containing other color terms, and discusses different ways of cashing out this idea that make different predictions for non-maximality.

What we call the asymmetric version of Paillé's theory assumes the standard bivalent exhaustification mechanism Exh (Chierchia, Fox, and Spector 2012; Fox 2007 among others).

(3) $[\![\text{Exh } \phi]\!] = 1$ iff $[\![\phi]\!] = 1$ and for each innocently excludable alternative ψ to ϕ, $[\![\psi]\!] = 0$.

For (1), it delivers the following meaning by negating alternative sentences formed with other color terms: $\exists x \sqsubseteq f[\mathsf{blue}(x)] \wedge \neg\exists x \sqsubseteq f[\mathsf{yellow}(x)] \wedge \neg\exists x \sqsubseteq f[\mathsf{red}(x)] \wedge \cdots$ (where f is the relevant flag). This means that the flag only has blue parts.

This theory accounts for non-maximality effects in terms of pruning of alternatives (see Bar-Lev 2018; Bar-Lev 2021 for plural definites). If *yellow* is excluded from the set of alternatives of *blue*, for example, the exhaustified meaning will no longer entail that there are no yellow parts, allowing the flag to have some yellow parts, as in the case of the EU flag (although, as Paillé himself points out, the resulting non-maximal meaning is too weak, failing to require the flag to be mostly blue, rather than partly blue; cf. the Ukranian flag 🇺🇦). This theory is *asymmetric* because pruning of alternatives, and hence a non-maximal reading, should be a possibility only when exhaustification takes place, and by assumption, it does not take place in negative sentences like (2).

The symmetric version of Paillé's theory uses a different exhaustification mechanism that gives rise to trivalent truth-conditions, called Pexh and defined as in (4) (after Bassi, Del Pinal, and Sauerland 2021).

(4) $$[\![\text{Pexh } \phi]\!] = \begin{cases} 1 & \text{if } [\![\text{Exh } \phi]\!] = 1 \\ 0 & \text{if } [\![\phi]\!] = 0 \\ \# & \text{otherwise} \end{cases}$$

Crucially, this version of the theory always works with the full set of alternatives. With Pexh,

View	Statement		Context	
	Experiment 1	Experiment 2	Existential	Universal
Asymmetric	The flag is blue.	Every flag is blue.	✓	×
	The flag is not blue.	No flag is blue.	×	×
Symmetric	The flag is blue.	Every flag is blue.	✓	×
	The flag is not blue.	No flag is blue.	×	✓

Table 1: Predictions of two views for target sentences with respect to non-maximality.

(1) will entail that the flag only has blue parts, as before. Its non-maximality reading is explained, not in terms of pruning of alternatives, but in terms of the pragmatics of #: It is assumed a sentence denoting # may nonetheless be perceived as true depending on contextual considerations. For instance, with *the flag* referring to the EU flag, (1) semantically denotes #, but in a context where all that is relevant is whether or not the flag has any blue at all ('existential context'), it is judged as practically true, while in a context where all that is relevant is whether or not the flag is entirely blue ('universal context'), it is judged to be untrue. Turning now to the negative sentence in (2), negation flips the truth and falsity conditions and Pexh is assumed to be obligatorily present under the negation, giving rise to a trivalent meaning (we ignore the version of this account where Pexh is optional to save space). It is then predicted that a non-maximal reading should be available, except that the judgments should be reversed: (2) should be perceived as true in the universal context and untrue in the existential context.

3 Experimental background

Romoli et al. 2024 investigated definite plurals and experimentally evaluated the predictions of symmetric and asymmetric theories of non-maximality effects. Building on Križ and Chemla 2015, who investigated non-maximal readings of plural definites in positive and negative statements, they conducted a picture-sentence verification experiment manipulating both the conversational context and the polarity of the target statement (positive/negative). Their study included an existential context ("Opening the presents is prohibited") and a universal context ("Opening the presents is required"). Their findings showed that context differentially affected non-maximality interpretations of positive sentences with a plural definite in the scope of *every* (e.g., *Every girl opened her presents*) and of genuine negative sentences with a plural definite in the scope of *no* (e.g., *No girl opened her presents*). More specifically, positive conditions were more context-sensitive than negative conditions in non-homogeneous situations where, for two out of four depicted girls, some but not all of their presents were opened.

4 Experiments

We conducted experiments that are similar in design to the experiments of Romoli et al. 2024 in order to investigate how context influences the interpretation of the positive and negative sentences in (1) and (2) (Experiment 1) as well as the quantified sentences in (5) and (6) (Experiment 2), in light of the predictions of the symmetric and asymmetric views of non-maximality, as summarized in Table 1 (We assume strong Kleene quantifiers for the trivalent theory).

(5) Every flag is blue.

(6) No flag is blue.

Since the two experiments have similar materials, design, procedure, and methods, we provide a joint description in what follows.



Figure 1: Example stimuli of Experiment 1. The flag is (not) blue. (A) is True Control for positive sentences and False Control for negative sentences, (B) is False Control for positive sentences and True Control for negative sentences. (C) is the Target for both positive and negative sentences.

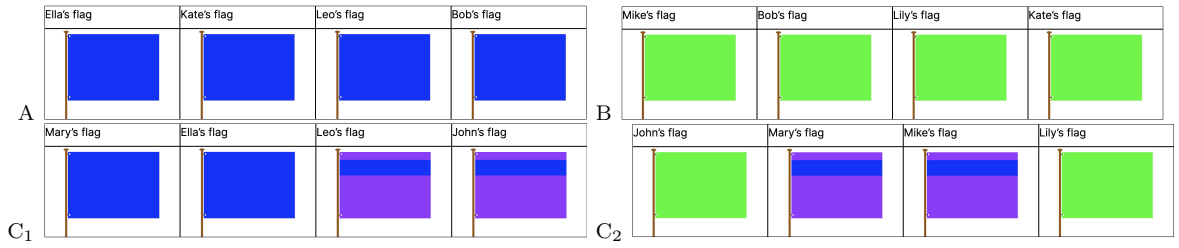


Figure 2: Example stimuli of Experiment 2. Every/No flag is blue. (A) is True Control for positive sentences and False Control for negative sentences, (B) is False Control for positive sentences and True Control for negative sentences. (C_1) is the Target for positive sentences and (C_2) is the Target for negative sentences.

4.1 Methods

Building on Romoli et al.'s (2024) study on non-maximality of plural definites, we conducted a picture-sentence verification task with a (7-point) Likert scale. In each experiment, participants judged how true the sentences were in three different pictorial situations: (i) True Control, (ii) False Control, and (iii) Target. See Figures 1 and 2 for example visual stimuli paired with the sentences in (1)/(2) and (5)/(6) in Experiments 1 and 2, respectively. True and False Control pictures depicted homogeneous situations, so the corresponding target sentences were expected to receive high and low ratings, respectively, on the response scale.

The Target pictures depicted non-homogeneous situations (see Pictures C in Figures 1 and 2). In Experiment 1, the Target picture shows a flag that is only partially blue (less than 50%), while in Experiment 2, the Target pictures (C_1/C_2) depict a situation where half of the flags are partially blue, as in Experiment 1, and the rest are uniformly one color (homogeneous flags).

It was explained in the instructions at the beginning of each experiment that the flags were drawn and colored by students in an art class, with all students being assigned the same drawing task each day by the teacher. The context, our main manipulation, was also introduced in the instructions. Similarly to Romoli et al. 2024, the context was manipulated between participants (existential vs. universal) in both experiments: In the existential context, students were instructed to create original flags, using any color except one specific color (e.g., blue), which was the same for all students. In the universal context, students were instructed to create original flags using only one specific color (blue), and no other. The teacher's instructions (i.e., the contextual manipulation) were also presented at the top of every experimental trial as a single sentence to remind participants of the context they were in (e.g., existential: "Teacher's instructions on Tuesday: Use only orange for your flag.", universal: "Teacher's instructions on Tuesday: Use any color but orange for your flag."). In both experiments, each experimental trial consisted of the teacher's instructions at the top of the display, followed by the pictorial situation and the critical statement to be verified against the depicted situation. Before making their judgement, participants were required to complete a secondary task: a yes/no question asking whether the instructions had been followed given the depicted situation. This helped ensure that participants understood the context. See supplementary materials on OSF presents example trial displays for each experiment.

Participants were randomly assigned to one of the two Context conditions, and received

the corresponding instructions. After reading the instructions, participants were presented with two practice items to familiarize themselves with the task. Both experiments had three manipulations: the Polarity of the statement (Positive/Negative), the Condition type (True Control/False Control/Target), and Context (Existential/Universal). Unlike Context, Polarity and Condition type were within-participant factors.

For each experiment, we constructed 24 experimental items: 3 flag shapes (rectangular, square, triangular) × 8 colors (blue, red, green, purple, orange, yellow, black, pink). Experimental items appeared in all 6 conditions (2 Polarity × 3 Condition type) and were rotated through six lists in a Latin square design. Every list had four items per condition. Half of the lists displayed the response button labels in one sequence (see supplementary materials on OSF), whereas the other half presented them in the reverse order. Each experiment had 12 lists in total: 6 lists with existential instructions and 6 lists with universal instructions. In addition, each list had the same 32 filler trials: 8 with color terms and 24 with other predicates such as *triangular, 5-sided, rectangular, square, circular, spotted, striped, chequered.* Both filler types were balanced for polarity (positive/negative) and truth value (true/false). These filler trials were also designed to balance the yes/no responses to the secondary question, resulting in as even a distribution as possible for both context types. Both experiments were programmed in Gorilla and administered online.

In each experiment, we recruited 288 native speakers of English raised with their native language only through Prolifc, and paid £1.50 for their participation. All participants gave their informed consent before participation. We excluded data from color-blind participants based on their response to a relevant question at the beginning of the experiment. We further excluded data from participants who had an accuracy rate of 75% or lower on the principal task in the Control items (with ratings higher than 3 for False Controls and ratings lower than 3 for True Controls being considered errors). This resulted in a final dataset of 276 participants in Experiment 1 and 278 participants in Experiment 2.

4.2 Results

We excluded trials where the secondary question was answered incorrectly from the analysis. Note that following Romoli et al. 2024, we recoded the Context factor into Lax (positive sentences in existential context, negative sentences in universal context) and Strict (positive sentences in universal context, negative sentences in existential context), with the former encouraging non-maximal interpretations and the latter discouraging them. Figure 3 shows mean verification ratings per condition in each experiment. In both experiments, the True Control conditions received high mean ratings overall, while False Control conditions received low ratings. This suggests the availability of homogeneity effects for both positive and negative sentences.

For the data analysis of both experiments, we fitted cumulative link mixed-effects models to analyze verification ratings for Target items. The models included Context (sum-coded), Polarity (treatment-coded, with negative as the reference level), and their interaction as fixed effects. They also included random intercepts for Items and for Participants, and random by-Participant slopes for Polarity. The analyses were conducted using the ordinal package (Christensen 2023) in R (R Core Team 2022).

In Experiment 1, the statistical analysis revealed a strong positive effect of Context ($\beta = 3.21$, $SE = 0.39$, $z = 8.25, p < 0.0001$), with Lax contexts receiving higher ratings than Strict ones overall. This contextual modulation was significantly greater for negative sentences compared to positive ones (significant negative Polarity×Context interaction: $\beta = -0.97$, $SE = 0.35$, $z = -2.79, p < 0.01$). In Experiment 2, the analysis revealed a positive effect of Polarity ($\beta = 0.63$, $SE = 0.30$, $z = 2.12, p < 0.05$) such that positive conditions received generally higher ratings than negative ones, as well as a robust positive effect of Context ($\beta = 3.17$, $SE = 0.38$,

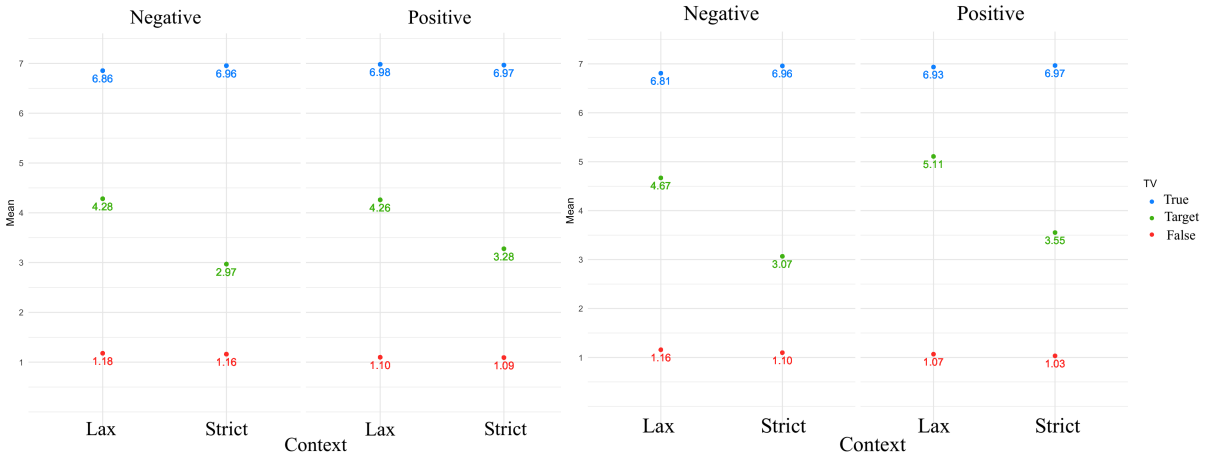


Figure 3: Mean verification ratings per condition in Experiments 1 (left) and 2 (right).

$z = 8.41, p < 0.0001$). Crucially, we found no evidence of greater context-sensitivity for non-maximal interpretations in positive sentences compared to negative ones (Polarity×Context interaction: $p = 0.23$), suggesting that the effect of Context is comparable for positive and negative sentences.

5 General Discussion

Our overall results are symmetric and hence more in line with the predictions of the symmetric view. Note in particular that the interaction effect observed in Experiment 1 was in the opposite direction of the prediction of the asymmetric view, and its effect size is relatively small.

It should also be mentioned that it is unlikely that Experiment 2 was underpowered, given that using a parallel experimental method and design Romoli et al. 2024 observed a robust interaction effect with definite plurals under *every* and *no* such that definite plurals exhibited a larger context effect under *every* (they did not test definite plurals in simple positive and negative sentences because of potential confounds with scope; cf. Križ and Chemla 2015). In order to reinforce this conclusion further, we plan to compare plural definites and summative predicates directly in a single experiment.

Assuming that the difference between summative predicates and definite plurals is real, it can be taken as suggesting that different mechanisms for non-maximality need to be postulated for summative predicates and plural definites. One possible way of theoretically implementing it is, for example, that non-maximality of summative predicates is due to Pexh, thereby predicting the symmetry, while non-maximality of plural definites is due to alternative-pruning, thereby predicting the asymmetry. More generally, our results point towards the conclusion that the class of homogeneity phenomena in natural language is heterogeneous, calling for further investigations that compare different homogeneity phenomena.

Acknowledgements. The research in this paper was supported by an AHRC-DFG grant (AHRC: AH/V003526/1; DFG: EB 523/2-1, STE 958/12-1, STE 2555/3-1).

References

Bar-Lev, Moshe (2018). "Free choice, homogeneity and innocent inclusion". PhD thesis. The Hebrew University of Jerusalem.

Bar-Lev, Moshe E. (2021). "An Implicature Account of Homogeneity and Non-Maximality". In: *Linguistics and Philosophy* 44.5, pp. 1045–1097. DOI: 10.1007/s10988-020-09308-5.

Bassi, Itai, Guillermo Del Pinal, and Uli Sauerland (2021). "Presuppositional exhaustification". In: *Semantics and Pragmatics* 14.11.

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (2012). "Scalar implicature as a grammatical phenomenon". In: *Semantics: An International Handbook of Natural Language Meaning*. Ed. by C. Maienborn, K. von Heusinger, and P. Portner. Vol. 3. Mouton de Gruyter, pp. 2297–2331.

Christensen, Rune H. B. (2023). *ordinal—Regression Models for Ordinal Data*. R package version 2023.12-4.1. URL: https://CRAN.R-project.org/package=ordinal.

Fox, Danny (2007). "Free choice and the theory of scalar implicatures". In: *Presupposition and Implicature in Compositional Semantics*. Ed. by Uli Sauerland and Penka Stateva. London: Palgrave Macmillan UK, pp. 71–120.

Krifka, Manfred (1990). "Boolean and non-boolean 'and'". In: *Papers from the second symposium on logic and language*. Akadémiai Kiadó Budapest, pp. 161–188.

Križ, Manuel (2016). "Homogeneity, maximality, and *all*". In: *Journal of Semantics* 33, pp. 493–539.

Križ, Manuel and Emmanuel Chemla (2015). "Two methods to find truth value gaps and their application to the projection problem of homogeneity". In: *Natural Language Semantics* 23, pp. 205–248. DOI: 10.1007/s11050-015-9114-z.

Križ, Manuel and Benjamin Spector (2021). "Interpreting plural predication: homogeneity and non-maximality". In: *Linguistics & Philosophy* 44.5, pp. 1131–1178. DOI: 10.1007/s11050-015-9114-z.

Löbner, Sebastian (2000). "Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences". In: *Linguistics and Philosophy* 23.3, pp. 213–308.

Malamud, Sophia A. (2012). "The meaning of plural definites: A decision-theoretic approach". In: *Semantics and Pragmatics* 5, pp. 1–58.

Paillé, Mathieu (2022). *Strengthening predicates*. McGill University (Canada).

— (2023a). "Alternatives and jurisdiction in predication". In: *Proceedings of Sinn und Bedeutung*. Vol. 27, pp. 483–499.

— (2023b). "Trivalent exh and summative predicates". In: *Semantics and Linguistic Theory*, pp. 421–438.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Romoli, Jacopo et al. (2024). "Putting plural definites into context". submitted.

Ignorance and attenuation

Anamaria Fălăuș
CNRS & Nantes Université
LLING UMR 6310
Nantes, France
anamaria.falaus@cnrs.fr

Andreea C. Nicolae
Leibniz-Zentrum Allgemeine
Sprachwissenschaft (ZAS)
Berlin, Germany
nicolae@leibniz-zas.de

Abstract

We introduce the existential modifier *cumva* in Romanian, which is ambiguous between a manner adverbial and an attenuating modifier. We describe its distribution and interpretation in declaratives and show that *cumva* has two possible meaning contributions, depending on the element it modifies: (i) ignorance, when it acts as a manner adverb, and (ii) attenuation, when it combines with a gradable predicate.

1 Introduction

In this paper, we introduce the existential element *cumva* in Romanian. Descriptively, there are two distinct uses of *cumva* in declaratives. The first use is that of a manner adverbial with a non-specific interpretation, akin to *somehow*, illustrated in (1):

(1) a. Rezolv asta *cumva*.
solve.1SG this CUMVA
'I'll solve this somehow.'
b. Găsește el *cumva* cazare, nu-i problemă.
finds he CUMVA accommodation not-is problem
'He'll find a place to stay somehow, don't worry.'

This use is expected given its morphological decomposition. Specifically, like other existential quantifiers in the language, *cumva* is built off of the manner *wh*-word *cum* 'how' and the suffix *-va* (for details on *wh*-based paradigms and constructions in Romanian, see e.g., Farkas 2013; Fălăuş and Nicolae 2022; Caponigro and Fălăuş 2023).

On its second use, illustrated in (2), *cumva* acts as an attenuator, akin to *somewhat*. It conveys a moderate degree of the property with which it combines, e.g., tiredness in (2-a).

(2) a. Sunt *cumva* obosită.
am CUMVA tired
'I'm somewhat tired.'
b. Situația e *cumva* mai complicată decât credeam.
situation is CUMVA more complicated than thought.1sg
'The situation is somewhat more complicated than I thought.'

Our main goal in this paper is to provide a detailed description of the two uses in (1) and (2) and understand the meaning contribution of *cumva*. We will show that the presence *cumva* associates with different pragmatic inferences. The nature of the observed inference correlates with its use: ignorance when it is used as a manner adverb (Section 3.1) and attenuation when combining with an adjective (Section 3.2). We also discuss ways in which these interpretations can be derived.

2 The distribution and interpretation of *cumva*

In this section, we describe in more detail the distribution and interpretation of *cumva* in each of the two uses mentioned above.

2.1 Manner adverbial use

On its manner adverbial use, *cumva* typically follows the verb, as in (1) above. In these examples, it can be replaced with other manner adverbs like *repede* 'quickly'. *Cumva* can also occur elsewhere in the sentence, like other adverbials in Romanian, a language with a relatively free word order:

(3) a. *Cumva* a aflat cine l-a dat de gol.
CUMVA has found out who him-has given away
'Somehow, he found out who gave him away.'
b. Mama *cumva* știe mereu ce să zică.
mom CUMVA knows always what SUBJ say
'Somehow Mom always knows what to say.'

Intuitively, the presence of *cumva* signals that the speaker is unwilling or unable to be specific about the manner in which the action under discussion will be carried out. It is quite common to find continuations that make this ignorance component explicit, as illustrated in (4-a). On the other hand, ignorance renders *cumva* incompatible with *namely* continuations, which explicitly identify the manner in question (4-b):

(4) a. *Cumva*, nu știu cum, a aflat cine l-a dat de gol.
CUMVA not know.1SG how has found out who him-has given away
'Somehow, I don't know how, he found out who gave him away.'
b. Rezolv asta *cumva*, #și anume cu bani.
solve I CUMVA and namely with money
'I'll solve this somehow, #namely with money.'

2.2 Degree modifier use

On its degree modifier use, *cumva* typically combines with adjectives like *tired* or *complicated*, as shown in (2) above. It can also combine with other property-denoting expressions, i.e., adverbs, PPs and nouns.[1]

(5) a. Lucrurile arată *cumva* mai bine ca ieri.
things look CUMVA more well than yesterday
'Things are looking somewhat better than yesterday.'
b. Marius a fost mereu *cumva* de partea mea.
Marius has been always CUMVA on side my
'Marius has always been somewhat/sort of on my side.'
c. Mai nou suntem *cumva* prieteni.
more new are.1PL CUMVA friends
'Lately we have been somewhat/sort of friends.'

Intuitively, with gradable adjectives like *tired*, it conveys a moderate degree of the property with which it combines. This renders *cumva*-sentences incompatible with contexts favoring high degree interpretations. For example, the sentence in (6) cannot be felicitously followed by the

[1] In this use, *cumva* can be interchanged with the degree modifier *oarecum* 'somewhat/sort of', to convey the same meaning. Interestingly, *oarecum* cannot be used adverbially, unlike *cumva* (see Section 2.1 above).

continuation in (7-a), which suggests a high degree of difficulty. Note, crucially, that it also appears to convey the same meaning as the bare form without *cumva*, namely that the problem is difficult, as evidenced by the unacceptability of the continuation in (7-b).

(6) Problema asta e *cumva* dificilă.
problem.def this is CUMVA difficult
'This problem is somewhat difficult.'

(7) a. *#it's impossible to solve.*
b. *#I'll solve it in no time.*

In the case of non-gradable expressions, like *friends* in (5-c), *cumva* does not signal moderate degree, but rather a property that is related or similar in some relevant sense to that of being friends. English *sorta* and *kinda* are better approximations of that interpretation. Similar uses are provided by the examples in (8). In the following section we will argue that this approximation use can be viewed on a par with the attenuation use described above.

(8) a. Schimbarea asta e *cumva* ilegală.
change.DEF this is CUMVA illegal
'This change is kind of illegal.'
b. Bateria e *cumva* moartă.
battery.DEF is CUMVA dead
'The battery is kind of dead.'
c. Chiar dacă erai singur, casa era *cumva* vie.
even if were alone, house.DEF was CUMVA alive
'Even though you were alone, the house was kind of alive.'

3 Deriving the different interpretations

In this section we focus on the meaning contributions associated with the uses of *cumva* outlined in the previous section and discuss in more detail their common interpretive components.

3.1 Ignorance use

Possibly most straightforward of its uses is when it acts as a manner adverbial, as in (1) and (3), in which ignorance or indifference about the manner associated with the relevant action is conveyed. In this use, the contribution of *cumva* aligns with what one might expect from a regular existential quantifier. This effect arises via the usual Gricean reasoning over a more informative alternative, that the speaker chose not to employ, either because they don't know or because the alternative is irrelevant for the purposes of the conversation. For instance, in (1-b), the speaker could have specified ways in which the person under discussion can find accommodation: *booking a place to stay, asking a friend, couch surfing* etc. Since none of these alternatives was employed, it is reasonable to conclude they may be false or not relevant.

In the case of an existential quantifier such as *somebody*, this ignorance or indifference effect is optional. This is evidenced by the fact that a sentence like *Somebody called.* allows a specific interpretation of the existential and thus can be felicitously followed-up by *namely John* (see Aloni and Degano 2022 for recent discussion). Adverbials like *cumva* or *somehow* are different in this respect, as they disallow continuations that signal knowledge (as shown in (4-b) above). Where does this difference come from? We believe that the obligatory ignorance effect observed with *cumva* is due to the fact that it is an adjunct, which by definition is optional. Truth-conditionally, there is no difference between (1-a) (*I'll solve this somehow*) and its version

without *cumva* (i.e. *I'll solve this*).[2] General conversational principles prevent such uses, so its insertion must be associated with some meaning contribution. The actual meaning effect depends on the context. In some cases, the manner in which an action will be carried out may be left unspecified because the speaker is ignorant with respect to it. In other cases, it may be because it is irrelevant in the context. The sentences in (1) for example are compatible with both continuations in (9):

(9) a. *I don't know how.*
b. *Let's not get into the details.*

One can also find cases where *cumva* conveys ignorance with respect to the reason why a certain situation holds, as in (10):

(10) *Cumva* mă îndoiesc că e o veste bună.
CUMVA me doubt.1SG that is a news good
'Somehow/For some reason, I doubt this is good news.'

This may be due to the fact that in addition to manner, the *wh*-word *cum* 'how' can also refer to reasons (much like in *how come*-interrogatives in English).

3.2 Attenuator use

3.2.1 Gradable adjective modifier

In most instances where *cumva* acts as an adjectival modifier, as in (2) and (6), it acts specifically as a moderate degree modifier, similarly to English *somewhat*. The interpretation that results is equivalent to that of the bare adjective, as argued by Solt and Wilson 2021 for *somewhat* (see the discussion around examples (6) and (7-b) above). Adopting their semantics of *somewhat*, we take *cumva* to be a degree quantifier that introduces existential quantification over a scalar interval. The resulting interpretation is equivalent to that of the bare adjective, namely that there is a non-zero degree to which the relevant property holds. However, *somewhat* and in turn, Romanian *cumva*, do make a contribution, namely one of attenuation of the degree in question.

Solt and Wilson 2021 propose that the attenuator use of a moderate degree modifier like *somewhat* can be analyzed as arising via competition with an alternative proposition containing a modifier expressing a higher degree, e.g., *foarte* 'very' as in (11). Specifically, the use of (2-a) conveys that the alternative in (11) is not true. Under the assumption that a more complex form, namely one with *somewhat* or *cumva*, should not be used if it does not make any contribution beyond that of the bare counterpart in (12), the competition with the stronger modifier can be taken to be obligatory since it delivers a strengthened interpretation beyond that which is conveyed by the bare form.[3]

(11) Sunt *foarte* obosită.
am very tired
'I'm very tired.'

(12) Sunt obosită.
am tired
'I'm tired.'

Having discussed how the moderate degree contribution of *cumva* is derived when it acts as

[2]Something similar can be said about other existential adverbials, e.g., *I'll come by sometime next week* or *Maria lives somewhere downtown*, whose presence doesn't seem to have any truth-conditional meaning contribution.

[3]There are a number of ways this could be implemented. One possibility is via a form of the Manner Maxim, as proposed by Rett 2020 in the domain of degree modification. This is ultimately no different than the proposals made in the literature for polarity sensitivity items which take their distribution to be guided by a requirement that their use contribute strengthening; see, e.g., Chierchia 2013; Nicolae 2012; Spector 2014; Nicolae 2017 among many others. Note, crucially, that such an account carries the assumption that the bare form does not have as an alternative a structurally more complex alternative involving a high-degree modifier (Katzir 2007).

a gradable adjective modifier, we next turn to its contribution when it modifies non-gradable adjectives.

3.2.2 Non-gradable property modifier

We already introduced some such examples above in (5) where we showed that it can combine with prepositional phrases and nouns which are not inherently gradable. Other uses of *cumva* where it seems to associate with a non-gradable predicate were provided in (8). In this use, the role of *cumva* appears to be similar to that of *sorta/kinda* in English, and that is to indicate that the degree of precision in the context is close to the standard degree, namely the one associated with the bare form (Anderson 2013, building on the work of Morzycki 2011). Consider, for example, how Anderson 2013 implements this for *sorta*: (i) the modifier imposes a coercion of the sister predicate into something gradable, followed by (ii) a comparison between the modified predicate *sorta P* and the bare predicate *P* such that *sorta P* is associated with a degree close to the one associated with *P*.[4] Since the gradability is over degrees of precision, and since the bare form is assumed to be associated with the highest such degree, it follows from the steps outlined above that the use of *sorta* would be associated with a close yet lower degree of precision.

While maintaining the spirit of his proposal, here we would like to argue for an account more closely aligned to the one presented above for the attenuator uses of *cumva*. Specifically, we want to argue that while *cumva*, like *sorta*, can coerce its sister property into something gradable, once it does so it enters into a competition not with the bare form, as proposed by Anderson 2013, but rather with another modified form, namely one involving a high degree modifier. For example, taking (8-a) as an example, we propose that the relevant competitor is (13).

(13) Schimbarea asta e *complet* ilegală.
change.DEF this is completely illegal
'This change is completely illegal.'

(14) Schimbarea asta e ilegală.
change.DEF this is illegal
'This change is illegal.'

Since on a non-coerced interpretation (14) amounts to (13), the same interpretation is achieved as in Anderson's (2013) account, albeit via competition with a different alternative, as per Solt and Wilson's (2021) account.

4 Conclusion and open issues

In this contribution we presented the distribution of the Romanian existential modifier *cumva* in declarative constructions. We showed that it can act as a modifier of nouns, adjectives, verbs and prepositional phrases. In discussing its distribution we also aimed to characterize the possible interpretations associated with it. We identified two main possible contributions, one of ignorance and another of attenuation, with the attenuation interpretation being characteristic of those contexts where a gradable predicate is available, directly or via coercion. The ignorance interpretation was shown to arise mostly in cases where *cumva* took on a manner-like interpretation. This characterization leaves open a number of issues, one major one being a unified theory of *cumva* that can predict its use depending on the nature of its argument. Specifically, we would like to better understand what it is about being a manner adverbial that is more likely to be associated with an ignorance interpretation than with an attenuation one. Note that we see a similar divide even in the use of the simple existential quantifier *some*, with it oscillating between conveying ignorance, (15-a), or a scalar inference, (15-b) (see for example Mihoc 2019 on what governs this divide).

[4] Morzycki 2011, for example, argues for a covert typeshifter PREC which introduces gradability by creating different possible levels of (im)precision with respect to the relevant standard of comparison.

(15) Ari read some books yesterday.
a. ⇝ *I don't know what books.*
b. ⇝ *She didn't read them all.*

There are also instances where *cumva* seems to act as a modifier at a higher level than what we discussed so far. Specifically, in examples like (16), *cumva* appears to do something other than convey ignorance with respect to the manner in which or reasons for why something takes place, unlike the uses of *cumva* in (1), (3) and (10).

(16) a. *Cumva* lumea regretă ce s-a întâmplat.
CUMVA people.DEF regrets what REFL-has happened
'Somehow people regret what happened.'
b. *Cumva* banii nu aduc fericirea.
CUMVA money.DEF not bring happiness.DEF
'Somehow money doesn't bring happiness.'

Note crucially that ignorance-type continuations like those in (9) would not be felicitous in these examples, showing that the presence of *cumva* does not make the same contribution as in the other adverbial uses we discussed previously. Instead, *cumva* acts at the propositional level and its use signals that the proposition with which it combines is true in one, possibly non-standard, way of looking at things. By adding *cumva* in sentences like (16), p becomes true only by expanding the set of situations in which p is taken to be true. A similar effect would obtain in English with something like *in some sense*. Observe that this description of the interpretation strongly suggests a parallel account with the attenuation use of non-gradable predicates discussed previously: a deviation from the standard of precision in the examples in (8), and a deviation from the standard of interpretation in the examples in (16). Since this meaning is mostly evidenced by sentence-initial instances of *cumva* like in (16), this can be taken to further suggest that what is actually modified sits higher in the left periphery.

In fact, *cumva* is also found in polar questions, yet another instance where its contribution appears to be at the speech act level rather than lower in the clause. While we cannot pursue this use of *cumva* here, we nonetheless want to point out that its contribution, at least in these examples, can be seen as one involving attenuation as well, with the only difference being that the attenuation is at the level of the speaker's bias for the proposition highlighted by the question, namely that the addressee got a haircut in (17) and that it seems to be raining outside in (18).

(17) Te-ai tuns *cumva*?
self-have hair-cut CUMVA
'Did you get a haircut?'

(18) Plouă *cumva*?
raining CUMVA
'Is it raining?'

We are not aware of any work describing the use of *cumva* in questions. We leave open the discussion of what *cumva* contributes outside declaratives, but are hopeful that a possible integration of its uses across speech acts can ultimately be achieved.

Acknowledgments. We are deeply grateful to the audience at ZAS who listened to initial versions of this work and offered invaluable comments. We also thank the reviewers for their helpful suggestions. A.N. received support from the ERC Synergy Grant 856421. A.F. received support from the CNRS grant LogicalFun. The names of the authors are listed alphabetically.

References

Aloni, Maria and Marco Degano (2022). "(Non-)specificity across languages: constancy, variation, v-variation". In: *Semantics and Linguistic Theory (SALT) 22*. Ed. by Juhyae Kim John R. Starr and Burak Öney, pp. 185–205. DOI: https://doi.org/10.3765/salt.v1i0.5337.

Anderson, Curt (2013). "Inherent and coerced gradability across categories: manipulating pragmatic halos with *sorta*". In: *Semantics and Linguistic Theory (SALT) 23*. Ed. by Todd Snider, pp. 81–96.

Caponigro, Ivano and Anamaria Fălăuş (2023). "Investigating clausal wh-constructions in Romanian". In: *Isogloss. Open Journal of Romance Linguistics* 9.4, pp. 1–32. DOI: 10.5565/rev/isogloss.325.

Chierchia, Gennaro (2013). *Logic in Grammar*. Oxford University Press. DOI: 10.1093/acprof:oso/9780199697977.001.0001.

Fălăuş, Anamaria and Andreea C. Nicolae (2022). "Additive free choice items". In: *Natural Language Semantics* 30.2, pp. 185–214. DOI: 10.1007/s11050-022-09192-8.

Farkas, Donka (2013). "The semantics of determiners". In: *A Reference Grammar of Romanian*. Ed. by Carmen Dobrovie-Sorin and Ion Giurgea. Amsterdam: John Benjamins, pp. 175–230.

Katzir, Roni (2007). "Structurally defined alternatives". In: *Linguistics and Philosophy* 30.6, pp. 669–690.

Mihoc, Teodora (2019). "Decomposing logic: Modified numerals, polarity, and exhaustification". PhD thesis. Harvard University.

Morzycki, Marcin (2011). "Metalinguistic comparison in an alternative semantics for imprecision". In: *Natural Language Semantics* 19.1, pp. 39–86.

Nicolae, Andreea C. (2012). "Negation-resistant polarity items". In: *Empirical Issues in Syntax and Semantics 9*. Ed. by Christopher Piñón, pp. 225–242.

— (2017). "Deriving the positive polarity behavior of plain disjunction". In: *Semantics and Pragmatics* 10.5, pp. 1–21. DOI: http://dx.doi.org/10.3765/sp.10.5.

Rett, Jessica (2020). "Manner implicatures and how to spot them". In: *International Review of Pragmatics* 12.1, pp. 44–79.

Solt, Stephanie and E. Cameron Wilson (2021). "M-modifiers, attenuation and polarity sensitivity". In: *Sinn und Bedeutung (SuB) 25*. Ed. by Patrick Grosz et al., pp. 787–804.

Spector, Benjamin (2014). "Global positive polarity items and obligatory exhaustivity". In: *Semantics and Pragmatics* 7.11, pp. 1–61. DOI: 10.3765/sp.7.11.

Crossing different ontological domains: the case of *appena*

Valeria Gradimondo
Université Paris Cité, LLF & CNRS
Paris, France
valeria.gradimondo@etu.u-paris.fr

Lucia M. Tovena
Université Paris Cité, UMR7110 CNRS
Paris, France
tovena@linguist.univ-paris-diderot.fr

Timothée Bernard
Université Paris Cité, LLF & CNRS
Paris, France
timothee.bernard@u-paris.fr

Abstract

This paper investigates the semantics of the Italian multi-faceted particle *appena*, which can be interpreted as an adverb (namely *barely, only, just*) and as a temporal connective (*as soon as*). We argue for the presence of the same underlying pattern for all its interpretations, and develop its formalisation in a (neo) Davidsonian event-based semantics.

1 Introduction

Appena admits various interpretations with different syntactic configurations, from an adverb to a temporal connective. Its first meaning coincides with *barely, hardly* (1) (Magni 2017). However, it also translates as *just*, both with quantitative (2) and temporal (3) meanings. It can also be used as a temporal complementiser, namely, *as soon as* (4).

(1) Alice cammina appena.
Alice walk.PRES.3SG appena.
'Alice *barely* walks.'

(2) Alice mangia appena una fetta di torta.
Alice eat.PRES.3SG appena a slice of cake.
'Alice eats *only/just* a slice of cake.'

(3) Alice è appena arrivata.
Alice aux.PRES.3SG appena arrive.PTCP.PST.F.SG.
'Alice has *just* arrived.'

(4) Alice cucina appena Luisa esce di casa.
Alice cook.PRES.3SG appena Luisa leave.PRES.3SG of house.
'Alice starts cooking *as soon as* Luisa leaves the house.'

In (1)–(3), *appena* is an adverb, but it displays both semantic and syntactic differences: e.g. in (1) it appears after an intransitive verb, and it intuitively affects the quality of the event predicate; in (2), similarly, it appears post-verbally but before the complement, and it specifies the quantity of the theme of the predicated event; in (3), it is a temporal adverb and it is placed between the auxiliary and the participle, a typical environment for these adverbs (Cinque 1999), and it specifies a proximity between the event time and the topic/utterance time. In (4), *appena* is a complementizer and expresses the fact that the matrix event is in the immediate posteriority of the embedded one.

Our goal is to capture the pattern that underlines all of *appena*'s interpretations, and to provide a coherent formalisation. This will give us a better understanding of its semantic evolution, from an adverb of quality to a temporal connective.

2 Background: other works on *appena*

Tovena and Schaden (2009) discussed a rich range of readings and properties of *appena*, noting first of all that it expresses a temporal relation of immediate anteriority, with a verb in the perfect form, like the German *gerade*.[1] In such interpretations, as in (3), they argue that the contribution of the particle is to narrow down the possible temporal location of an eventuality relatively to a time preceding and close to the Reference Time (RT), introduced by the perfect form. *Appena* can also be a degree modifier, like in (1). In general, they argue that *appena* works like an approximator by excess, the opposite of *almost*, an approximator by defect, as for Jayez and Tovena (2008). This means that, given a reference value that can be a minimal degree of a gradable property or a point in time, the element under discussion is positioned higher but close to this reference, within a range of indiscernible values. In fact, they argue that the reference and the element under discussion are so close that they are indistinguishable in the context. We share most of their intuitions but propose a new formalisation intended to cover all of *appena*'s interpretations in a systematic way.

2.1 Similar particles

The case of *appena* is not isolated among Romance languages. There are studies on particles with similar etymology and evolution, from manner adverbs to negative evaluative adverbs to particles of temporal proximity, c.f. (Bat-Zeev Shyldkrot 2010, i.a.) on French *à peine*, (Maldonado and Guzmán 2014) on Mexican Spanish *apenas*, and (Amaral 2006) on the counterpart *mal* in European Portuguese. Generalizing, it can be observed that the semantic change undergone by the particles consistently involves a change in their syntactic status, from adverbs to adverbial subordinators. The positions that correspond to the more recent meanings have scope over the position corresponding to the oldest and original ones, with the temporal interpretations as the latest ones. Interestingly, none of the meanings was lost through this process, resulting in particles that are multifaceted and versatile. In our work on *appena*, we attempt to account for this complexity by showing what is the at-issue component that unifies all its senses and that has hypothetically survived through its evolution.

3 Our proposal: an underlying pattern

As the examples (1)–(4) show, *appena* is a particle that can convey information about elements of different ontological domains, namely events and time intervals. We think that all interpretations point to the following pattern. *Appena* works as a particle that connects an event or an interval under discussion b with a reference a_1. a_1 is a set of events or an interval positioned at the beginning of a bigger and abstract ordered structure A, which is a superset of a_1. *Appena* states that b occurs after the elements of a_1 in the ordered structure A. The set A is ordered along a specific dimension, with respect to which it can be subdivided into ordered units. Assuming that the set a_1 identifies the first unit, b is then placed after it, in the second unit, to ensure its proximity to a_1.

To associate intervals/events to a position on an abstract structure, we assume the existence of the contextually determined degree function *deg*, in (5).

(5) *deg* is a function of signature $D \times Dim \rightarrow \mathbb{R}^+$, where D is either the domain of events D_v or intervals D_i^2 $(D_i \times D_i)$, and Dim is a set of dimensions along which A can be ordered.

[1] Note that *appena* does not exhibit the emphatic reading associated to *gerade*, and also the English *just*, and which would be an instance of approximating a high value. For instance, 'This dress is *just* beautiful'.

We assume that, for each unit a_n in A, the elements $e \in a_n$ are such that $deg(e, dim) \in [n\text{-}1, n)$. We argue that this function is at the core of the at-issue component of *appena*, and is present in all its interpretations. We argue, however, that the not-at-issue component is different. We will not deal with this aspect in detail in this paper, but we will briefly present two conversational implicatures (CI) observed with the non-temporal readings. In the next sections, our primary goal is to show how the described pattern works for all the provided interpretations.

3.1 A function targeting *events*

3.1.1 Appena as *barely*

In example (1), A is the set of possible walking events by Alice. They are ordered on a degree-like dimension (e.g., efficiency). Which dimension is relevant in a given sentence depends on the context. Assuming that the speaker cares about efficiency (an efficient walk being one that successful leads the walker to their intended location), a_1 is the subset of A that contains events that have the lowest value of efficiency in A (i.e., the first unit), so low that they can be regarded as not efficient at all (e.g., all the walking events where the agent is not able to reach the target of their walking). b is the event under discussion (the event described by the prejacent *Alice walks*). *Appena* then states that this event is positioned right after the elements in a_1, i.e., in the second unit. In our example, this could mean that Alice's walking is such that she is *barely* able to reach a target, because she takes a lot of time and has to stop multiple times. A further specification is needed as the sentence in (1) can have two readings. The first, *episodic*, reading is about a specific walking event performed by Alice; the second, *habitual*, reading is about Alice's general ability to walk. In this paper, we will consider just the first reading, leaving the analysis of the habitual one for future works. As for the episodic one, we argue for the existence of a conversational implicature (CI), which states that there is an expectation towards the performance of the subject along all the dimensions in which we can evaluate the event under discussion. In fact, for every dimension, the predicated event should display a low degree, in the proximity of a_1. In (6), we display the cancellation of such CI, in a scenario where Alice is an old woman that walks very slowly (i.e., the relevant dimension is speed) but elegantly.

(6) Guarda, Alice cammina appena, *anche se lo fa con molta eleganza.*
'Look, Alice barely walks, *even if she does it with very much elegance.*'

Formalisation In our work, we adopt a minimalist approach to grammar (Chomsky 1995), and we assume that the semantic composition is fed in by the Logical Form (LF) of a sentence. For the purposes of this work, we will not go into details about the transformational rules needed to derive the actual Phonetic Form (PF). To derive (1), we assume that *appena* merges with a full vP-shell, which contains also the subject of the predication. The semantic contribution of *appena* is to state that the event described by the clause is positioned in the second unit of A with respect to the salient dimension. The function *deg* in (7), of signature $D_v \times Dim \to \mathbb{R}^+$, is applied to the event described by the vP-shell, and the salient dimension, *dim*. Adopting a standard (neo) Davidsonian semantics, we propose (7) as the semantic contribution the particle makes in (1). Its CI is provided in (8). In these formulas, P varies over a predicate of events and *dim* is a placeholder for the salient dimension ordering A, e.g., efficiency. The semantic representation of (1) is in (9).

(7) $[\![\text{ Appena}_1]\!] = \lambda P \lambda e[P(e) \wedge deg(e, dim) \in [1, 2)]$

(8) CI: $\underline{\forall dim[dim \in Dim \to \neg[deg(e, dim) \geq 2]]}$

(9) $[\![\text{ Alice cammina appena }]\!] = \exists e[walk(e) \wedge ag(e) = a \wedge deg(e, dim) \in [1, 2) \wedge$
$\underline{\forall dim[dim \in Dim \to \neg[deg(e, dim) \geq 2]]}$

3.1.2 Appena as *only/just*

In examples like (2), similarly, *deg* applies to events. A is the set of possible events of eating by Alice where the theme are slices of cake. This set is ordered according to the size/quantity of the themes, which sets the granularity of the units, from eating events with zero slices to eating events with an increasing number of slices. The event under discussion, b, is described by the prejacent (*Alice mangia una fetta di torta*). *Appena* states that b is right after the first unit a_1 (corresponding to eating less than one slice); it is thus an event of eating one slice of cake by Alice.
Uttering *appena* has a double role: to state (i) that the subject eats a precise quantity and (ii) that such a quantity is not much with respect to what it could have been. A speaker can also felicitously utter *Alice ha mangiato appena quattro pasticcini* (Alice ate just four pastries), if they knows that this is really a low number for Alice, who has a remarkable sweet tooth and was never seen eating less than ten pastries at the time.[2] With (2), the speaker is also implying that Alice didn't perform any other eating event with whatsoever theme. This implicature is very strong, and it is not trivial to cancel. In (10), we propose a scenario in which a cancellation is still acceptable, leading us to analyze it as a conversational implicature.

(10) Alla festa, Alice ha mangiato appena una fetta di torta, *anche se per fortuna ha mangiato anche un panino.*
'At the party, Alice ate just a slice of cake, *although luckily she also had a sandwich.*'

Formalisation We assume that the events in each unit are characterized by the same amount of theme. Function *deg* applied to elements in the first unit will return the same value, zero, that coincides with eating no slices; *deg* applied to the elements in the second one will return the same value, one, that coincides with *eating a slice*. This is due to A's structure, but the function introduced by *appena* remains the same, just reducing to $deg : D_v \times Dim \to \mathbb{N}$; in (2), the salient dimension along which we order events in A is the the quantity of the theme.
As before, *appena* merges with a full vP-shell. The semantic contribution of the particle in (2) is shown in (11), with the CI in (12). The formula in (11) is the same as that in (7), with the only change being the salient dimension taken into account. In these formulas, P varies over a predicate of events (which is of type $\langle v, t\rangle$), *dim* is a placeholder for the salient dimension ordering A and P' is identical to P but with an unspecified theme (i.e., the set of eating events by Alice). The meaning of (2) is formalized in (13).

(11) $[\![\text{ Appena}_2]\!] = \lambda P \lambda e[P(e) \wedge deg(e, dim) \in [1, 2)]$

(12) CI: $\forall x[x \in D_e \setminus \{th(e)\} \to \neg\exists e'[P'(e') \wedge th(e') = x]$

(13) $[\![$ Alice mangia appena una fetta di torta $]\!] = \exists e[eat(e) \wedge ag(e) = a \wedge \exists y[slice(y) \wedge th(e) = y] \wedge deg(e, dim) \in [1, 2) \wedge \forall x[x \in D_e \setminus \{th(e)\} \to \neg\exists e'[eat(e') \wedge ag(e') = a \wedge th(e') = x]]$

3.2 A function targeting *intervals*

3.2.1 Appena as *just*

In (3), A is an ordered set of instants oriented towards the past dimension; its right boundary is given by the end of the topic time (TT) (Klein 1994) and extends toward $-\infty$. The first unit of A, a_1, is the TT itself (which we assumed to be an interval). The interval under discussion, b, is

[2] For this work, we leave aside ironic uses of *appena*. These uses differ from the standard ones just for the presence of a intonational focus on the particle. For instance, uttering *Alice ha mangiato* appena *otto pasticcini* (Alice ate *just* eight pastries) can be used to mark the fact that, from the speaker's perspective, she ate a very high amount of pastries.

the event time (ET) of the arrival event. *Appena* states that b is placed right after the interval a_1, providing information about the relationship between the ET and the TT, in a *retrospective* way. We assume this information is computed at the AspP projection, as for (Cinque 1999). We argue that no CI is involved with these uses of *appena*.
At the beginning of section 3, we mentioned that A is divided into ordered *units*. In temporal interpretations, units are intervals with a specific length of time. We argue that this length depends on the potential duration of the resultant state of the event one consider. For instance, in (3), the resultant state (i.e., Alice's being here) should not be very long, typically no more than a few hours. The sentence is then felicitous if it is uttered at most, say, 20 minutes after the arrival event happened; units are then of 10 minutes each. If one utters *ho appena preso la patente* (I've just got my driving license), the resultant state (i.e., the state of having a license) will potentially last for many years. As a result, a speaker could felicitously utter this sentence even months after the acquisition of the license (units of, say, at least one month each).

Formalisation In this case, *appena* merges with a full AspP-shell, resulting in another AspP projection, where the relationship between the ET and the TT is further specified.The function *deg* is of signature $D_i^2 \times Dim \rightarrow \mathbb{R}^+$; the salient dimension along which we order A in (3) is temporal precedence. We use the function τ to associate each event with its runtime: a pair $\langle t_1, t_2 \rangle \in {D_i}^2$ such that $t_1 \leq t_2$. For convenience, $\langle t_1, t_2 \rangle$ is also noted $t_{1,2}$. The contribution that *appena* makes in (3) is in (14). P is a variable of type $\langle vi^2, t \rangle$, which is the semantic type of complete AspP-shells; *dim* is the dimension along which A is ordered, and $t_{t'}$ is the right boundary of the TT. The semantic formalisation of (3) is in (15). *now* is a special indexical constant that stands for the TU, an interval, while t_t is a variable that stands for the TT, an interval as well. As said above, no CI is involved with these temporal uses.

(14) $[\![\text{ Appena}_3]\!] = \lambda P \lambda e \lambda t_{1,2}[P(e)(t_{1,2}) \wedge deg(t_{1,2}, dim) \in [1, 2)]$

(15) $[\![\text{ Alice è appena arrivata }]\!] = \exists e \exists t_t \exists t_{1,2}[t_t = now \wedge t_{1,2} < t_t \wedge arrive(e) \wedge th(e) = a \wedge \tau(e) = t_{1,2} \wedge deg(t_{1,2}, dim) \in [1, 2)]$

3.2.2 Appena as *as soon as*

In (4), similarly to (3), *deg* applies to an interval. A is an ordered set of instant oriented towards the future. Its left boundary is given by the onset (Landman 2008) of the embedded event — the leaving by Luisa — and it extends toward $+\infty$. The first unit a_1 is identified with the TT, which is the ET of the leaving event. Interval b is the ET of the event predicated in the matrix clause. *Appena* states that b is right after a_1, in the second unit, hence providing information about the temporal relationship between two different events. Its semantic role is also mirrored in its syntactic status as a temporal complementizer, from which it follows that *appena* merges at the CP-level. All the considerations about the (absence of) CI and the concept of proximity in the previous subsection apply to this reading as well.

Formalisation As a complementizer, *appena* is saturated by two full TP-shells. It specifies a relationship of proximity between the TT (the embedded ET) and the ET of the matrix clause. The function *deg* maintains the signature of the previous reading; the salient dimension along which we order A in (4) is temporal succession. The semantic contribution that the particle makes in (4) is in (16). P and Q are variables of type $\langle \langle vi^2 \rangle, t \rangle$ (for the interpretation of TP-shells, respectively of the matrix and the embedded clause); *dim* is the dimension along which A is ordered. The semantic formalisation of (4) is in (17), where *now* is as specified in the previous reading. Note that this is a simplified derivation, where we specified the temporal information (both the event times are after the TU), but we left aside the aspectual information.

(16) $[\![\text{Appena}_4]\!] = \lambda Q \lambda e' \lambda t_{3,4} \lambda P \lambda e \lambda t_{1,2} [P(e)(t_{1,2}) \wedge Q(e')(t_{3,4}) \wedge deg(t_{1,2}, dim) \in [1, 2)]$

(17) $[\![$ Alice cucina appena Luisa esce di casa $]\!] = \exists e \exists t_{3,4} \exists e' \exists t_{1,2} [now < t_{3,4} \wedge [exit(e') \wedge ag(e') = l \wedge \tau(e') = t_{3,4}] \wedge [cook(e) \wedge ag(e) = a \wedge \tau(e) = t_{1,2}] \wedge deg(t_{1,2}, dim) \in [1, 2)]$

4 Conclusions

The behaviour of *appena* has been represented as a pattern where the element under discussion b is positioned in the proximity of a set of elements a_1 of a bigger structure A. This pattern is present in all of its semantic contributions and has been formalised with a degree function *deg* that can apply to elements of different ontological domains, namely events and intervals. In our work, the variability of meaning depends on such ontological difference and, consequently, on the dimension we are identifying a_1 against and placing b on. Eventually, we have proposed a formalisation for all readings in a (neo) Davidsonian event-based semantics. Further work has to be done, especially regarding the not-at-issue component of the particle, the cross-linguistic comparison with other languages, and other readings of *appena*, such as spatial ones like "*L'ufficio è appena dietro l'angolo*" (The office is just around the corner).

References

Amaral, Patrícia (2006). "The polysemy of mal in European Portuguese: A diachronic analysis". In: *Journal of historical pragmatics* 7.1, pp. 1–37.

Bat-Zeev Shyldkrot, Hava (2010). "Polysémie de à peine". In: *L'information grammaticale* 125.1, pp. 12–16.

Chomsky, Noam (1995). *The Minimalist Program*. MIT Press.

Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press.

Jayez, Jacques and Lucia M. Tovena (2008). "*Presque* and *almost*: how argumentation derives from comparative meaning". In: *Empirical issues in Syntax and Semantics 7*, pp. 217–239.

Klein, W. (1994). *Time in Language*. Germanic linguistics. Routledge. ISBN: 9780415104128. URL: https://books.google.fr/books?id=wnBAs-2KzvYC.

Landman, Fred (2008). "On the Differences Between the Tense-Perspective-Aspect Systems of English and Dutch". In: *Theoretical and Crosslinguistic Approaches to the Semantics of Aspect*. Ed. by Susan Deborah Rothstein. John Benjamins, pp. 107–167.

Magni, Elisabetta (2017). "Une grammaticalisation avec peine. Le cas de l'italien *appena*". In: *Mots de liaison et d'intégration: Prépositions, conjonctions et connecteurs*. Ed. by Thierry Ponchon, Hava Bat-Zeev Shyldkrot, and Annie Bertin. John Benjamins, pp. 103–118.

Maldonado, Ricardo and Rocío Guzmán (2014). "*Apenas*. Aspectual reduction and subjective evaluation in Spanish". In: *Review of Cognitive Linguistics* 12.2, pp. 443–469.

Tovena, Lucia M. and Gerhard Schaden (2009). *Immediate anteriority with and without scales*. Handout of the talk delivered at Chronos 9, Paris International Conference on Tense, Aspect, and Modality.

What is special about *special*?

Yael Greenberg
Bar-Ilan University
Ramat Gan, Israel
yaelgree@gmail.com

Hedde Zeijlstra
Goettingen University
Goettingen, Germany
hzeijls@uni-goettingen.de

Abstract

This paper concerns the semantics of the adjective special as in *This is a special book about Stalingrad.* To the best of our knowledge no discussion of the semantics of special exists in the current literature. There is, however, some research on *(a)typical* which takes it to be a subsective, multidimensional and intensional adjective, involving comparison to stereotypes. We first hypothesize that special could be modeled as an antonym of typical, i.e. as atypical. We observe, however, that there are distributional and interpretational differences between *atypical* and *special*, and propose to capture them by taking special to involve non-stereotypical properties which hold of the individuals in question in the speaker's / protagonist's teleologically / bouletically accessible worlds.

1 Introduction

This paper concerns the semantics of the adjective *special* as in (1):

(1) a. This is a special book about Stalingrad
b. This is a special stone
c. John is a special lawyer

To the best of our knowledge no discussion of the semantics of *special* exists in the current literature. There is, however, some research on *typical* which takes it to be a subsective, multidimensional and intensional adjective, involving comparison to stereotypes. We first hypothesize that *special* could be modeled as an antonym of *typical*, i.e. as *atypical*. We observe, however, that there are distributional and interpretational differences between *atypical* and *special*, and propose to capture them by taking *special* to involve non-stereotypical properties which hold of the individuals in question in the worlds speaker's / protagonist's wishes (in a wide sense of the term) are fulfilled.

2 Background: The semantics of *(a)typical*

Consider the adjectives *typical* and *atypical,* as in (1a):

(2) a. John is a typical lawyer b. John is an atypical lawyer

We follow ideas in e.g. Sassoon (2012), McReady & Ogata (2007), del Pinal (2015, 2018) and Guerrini (2024), and take *(a)typical* to be an adjective which is non-

intersective/subsective, multidimensional and intensional. Let us look at each of these properties in turn.

First, unlike an adjective like *American*, the adjective *(a)typical* is not intersective:

(3) a. Fido is a typical dog – NOT: Fido *is typical and Fido is a dog*
b. Fido is an atypical dog – NOT: Fido *is an atypical and Fido is a dog*

Instead, both *typical* and *atypical* appear subsective. For example, in (4a) and (4b) the sets of typical / atypical labradors seem to be subsets of the set of dogs. This is illustrated by the following example (inspired by Guerrini (2024)):

(4) A: That is a typical labrador. B: #No, it isn't a labrador!

In addition, unlike one-dimensional adjectives like *tall,* both *typical* and *atypical* are multidimensional (see e.g. Sassoon 2012, Bylinina 2016, Solt 2018, D'Ambrosio & Hedden 2024). For example, John can be considered a typical lawyer (as well as an atypical lawyer) in various respects, e.g. the way he dresses to work, the number of hours he works, the amount of money he makes etc. Indeed, both *typical* and *atypical* can be modified by phrases like *in many / all respects.*

More specifically, Sassoon (2012) points out that in their positive form (as in (2a)) and (2b)), *typical* and *atypical* behave as conjunctive and disjunctive, respectively[1]. That is, by default for (1a) to be true, John needs to be typical with respect to all relevant dimensions of being a lawyer, whereas for (1b) it is enough that John is atypical with respect to one such dimension. Sassoon supports this claim by observing that continuing such sentences with exceptive phrases, which are independently known to be compatible with universal quantifiers and not with existential ones, is perfectly felicitous with sentences like (2a), but is less common (and significantly less attested) with sentences like (2b):

(4) a. John is a typical lawyer, except for the way he dresses
b. John is an atypical lawyer, #except for the way he dresses

Finally, both *typical* and *atypical* are intensional. For example, for (3a,b) to be true it is not enough to consider just the properties of actual dogs. To illustrate that, consider the following scenario (inspired by an example from Cohen 1999). Suppose that it just so happens that all dogs existing right now have an identifying tag ending with an odd number. Then we will not consider (3a) to be true if Fido also happens to have such an identifying tag, and we will not consider (3b) false if it doesn't. Instead, following ideas in McReady and Ogata (2007), del Pinal (2015, 2018) and Guerrini (2024), we assume that to be considered a typical or an atypical dog we need to look into more inherent properties of dogs, and in particular, into their stereotypical properties.

To capture these properties, we follow Guerrini's (2024) proposal regarding *fake* (pace del Pinal's 2015, 2018 analysis in terms of Dual Content Semantics), and capture the intensionality of *(a)typical* by integrating into its semantics a generic

[1] Similarly to *healthy* and *ill*, respectively.

quantifier, which, we take to be a universal quantifier over individuals and accessible worlds. In the case of the case of *(a)typical* we assume that these worlds are those which are stereotypically accessible from the world of evaluation.

Integrating this with Sassoon's (2012) claims regarding the multidimensionality of *typical* and *atypical* we take the sentence in the positive form *John is a typical lawyer* (2a) to be true in w iff John is a lawyer in w and for all properties Q in all contextually relevant dimensions **D** such that in all stereotypically accessible worlds w', lawyers have Q in w', John has Q in w. Given this approach we can now take (2b) to indicate that there is at least one **D** where John does not share every property Q in **D** with stereotypical lawyers.[2]

Finally, we note that *typical* and *atypical* require the property they modify to be 'natural' (in Chierchia's 2013 terminology, to show a 'sufficiently regular behavior').[3] We take this to be a presupposition, given the infelicity of *(a)typical* with 'unnatural properties' (5a), and the projection of this inference under e.g. *possible* (5b):

(5) a. John is a(n) (a)typical {American / rich / ??162m tall} lawyer
b. It is possible that John is a(n) (a)typical {American / rich / ??162m tall} lawyer

We propose to capture this observation by taking *(a)typical* to (minimally) presuppose that there is at least one dimension where there is at least one property Q that all P members have in all accessible worlds. We thus take the sentences in the positive forms *John is a(n) (a)typical lawyer* to be interpreted as in (6):

(6) a. *John is a typical lawyer* is defined iff:
$\exists$**D.C(D).**$\exists$Q$\in$**D.**$\forall$x,w'[w'R_{sr}w $\wedge$ x is a lawyer in w'] $\rightarrow$Q(x)(w').
When defined, it's true iff Lawyer(j)(w) $\wedge$ $\forall$**D.C(D).**$\forall$Q$\in$**D.**$\forall$x,w'[[w'R_{sr}w $\wedge$ x is a lawyer in w'] $\rightarrow$ Q(x)(w')] $\rightarrow$ Q(j)(w)].
b. *John is an atypical lawyer* is defined iff:
$\exists$**D.C(D).**$\exists$Q$\in$**D.**$\forall$x,w'[w'R_{sr}w $\wedge$ x is a lawyer in w'] $\rightarrow$Q(x)(w').
When defined, it's true iff Lawyer(j)(w) $\wedge$ $\neg\forall$**D.C(D).**$\forall$Q$\in$**D.**$\forall$x,w'[[w'R_{sr}w $\wedge$ x is a lawyer in w'] $\rightarrow$ Q(x)(w')] $\rightarrow$ Q(j)(w)].

3 Novel observations: *(A)typical* vs *special*

A reasonable hypothesis to entertain now is that *special* is the antonym of *typical,* and that it has the semantics of *atypical*, e.g. that (7b) is the same as (7a):

---

[2] These proposals may be oversimplistic compared to more sophisticated views which integrate quantification over dimensions, and / or weights of dimensions in different contexts (e.g. McReady & Ogata 2007 analysis of *typical*-like adjectives in Japanese, see also Sassoon 2012, Bylinina 2016, Solt 2018, D'Ambrosio & Hedden 2024). But at this stage the distinctions between these views is orthogonal to our main claims regarding the semantics of *special*.

[3] Cf. McReady & Ogata (2007) for a similar observation for Japanese *typical*-like adjectives.

(7) a. John is an atypical lawyer b. John is a special lawyer

Indeed, similarly to *(a)typical, special* is also subsective, multidimensional and intensional. In addition, it also appears to presuppose that its modified property is 'natural'. It pattern in the same way as *(a)typical* in (5a,b).

However, we can also observe at least four interpretational and distributional differences between *special* and *(a)typical,* which indicate that understanding the contribution of the former requires more thought.

Observation A: Unlike *atypical, special* feels 'subjective', and even experiential (cf. Bylinina 2016, Solt 2018) as indicated by its felicity as a comparative as a complement to *find* (8a), and with the 'judge PPs' *to x* phrases (8b). This is unexpected if *special* had the semantics of *atypical*.

(8) a. I find it more {special /??atypical}
b. This watch is {special / ??atypical} to me

Observation B: Unlike *atypical, special* appears to have a 'positive' flavor (9):

(9) Context: *The regular amount of cans made by a typical worker in this factory is around 100 per day. But Mary is very quick – she makes 250 per day. And John is very slow – he only makes 25 per day.*
a. {Mary/John} is an atypical worker b. {Mary/??John} is a special worker.

Again, this would be unexpected if *special* had the semantics of *atypical*.

Notice that 'positive' is used here in a very wide sense. In (10), for example, it reflects feelings of excitement or awe. This seems to distinguish *special* from more traditional adjectives which were characterized as 'evaluative' like *good*):

(10) a. This is a special book! Newton himself held it and read it.
b. What makes this fossil special is that it is the first one which was discovered in America

Observation C: Related to the previous observations, we note that *typical* and *special* are completely independent, i.e. that there can be attested cases for any logically possible combination of values of *typical* and *special*. Suppose, for example, that stereotypical lawyers have the following properties, they: (i) dress formally to work, (ii) work many hours a day, (iii) charge a lot per hour, (iv) make lots of money and (v) are very sharp. Now consider the following combinations:

(11) **+Typical +Special** – a lawyer which has all stereotypical properties but also has a 'positive' property not typical of lawyers e.g. she is the only one who won a prestigious prize for a contribution to the field;
+Typical -Special – A lawyer with all stereotypical properties, and who does not have any 'positive' property that makes her different from other stereotypical properties;
-Typical +Special – A lawyer who lacks stereotypical properties (e.g. does not work many hours a day), but who does not charge a lot (which is considered 'positive');

-Typical -Special – A lawyer who lacks stereotypical properties (e.g. does not make lots of money money), and does not have any 'positive' non-stereotypical property either.

Taking *atypical* to be the antonym of *typical,* would not predict this attested independence of *atypical* and *special.*

Observation D: With proper names *special* can easily appear in predicative position in 'out of the blue' contexts (12a), whereas *(a)typical* needs more contextual support (12b):

(12) a. John is {special / ??(a)typical}!
b. Context: *Lawyers are known to charge too much per hour).*
But John is special / atypical!

4 Proposal

We propose that, like *John is a(n) (a)typical lawyer, John is a special lawyer* is defined iff its modified property is 'natural', i.e. if it meets the presuppositions in (6a) and (6b) above, and that part of its truth conditions is that John is a lawyer in the world of evaluation, *lawyer (j)(w).* In addition, it requires that there is a property Q' that John has in w, which is a member of a contextually relevant dimension **D'.C(D'),** and which meets the two conditions in (13):

(13) a. $\forall x,w' [w' R_{ster} w \wedge \text{lawyer}(x)(w')] \rightarrow \neg Q'(x)$
b. $\forall x,w'' [w'' R_{\textbf{bouletic*}} w \wedge \text{lawyer}(x)(w'')] \rightarrow Q'(x)(w'')$

In prose, the property Q' which makes John 'a special lawyer' must meet two conditions. The first requires Q' to be a non-stereotypical property of lawyers, i.e. a property that no lawyer in the stereotypically accessible worlds has. This condition can be met in two scenarios. In the first Q' is non-stereotypical of lawyers because the dimension **D'** of which it is a member is itself not stereotypically associates with lawyers. This happens, for example, in the **+Typical +Special** condition above (11a), assuming that that winning a prize is not a dimension which holds of lawyers in stereotypically accessible worlds. A second scenario is illustrated in the **-Typical +Special** where the dimension **D'** is stereotypically associated with lawyers (e.g. charging a certain amount of money), but **Q'** is not (e.g. charging very little).

The second condition on Q' that all lawyers have this property Q' in all worlds which are accessible from w given the judge's perspective and an accessibility relation ***bouletic**. We take this accessibility relation to yield a set of worlds which are bouletically accessible from w in a wide sense, i.e. not only where concrete wishes of the judge are fulfilled, but also those where she experiences positive feelings.

5 Accounting for the observations

Observation A follows directly through component (10b) of our proposal given that the property which makes John special should hold in the worlds which are accessible from the point of view of the judge.

Component (10b) also can explain observation B regarding the 'positiveness' of *special,* given the characterization of the accessible worlds as bouletically accessible in a wide sense. Given this accessibility relation the property Q' that one should have in order to be considered a special P is a property that all P members have in the worlds where some concrete wishes of the judge are fulfilled (e.g. the wish for a high productivity of workers (as in (7)) but also in worlds where the judge feels excitement or awe (as in (8)).

Observation C is explained by the fact that given our proposal, the meanings of *special* and *(a)typical* do not stand in an entailment relation. More specifically, unlike *atypical* which conveys that x lacks stereotypical properties, to be *special* we need to add a non-stereotypical property which characterizes x, as seen in (10a). While in many cases the two operations lead to identical truth conditions (making *atypical* and *special* look synonymous), they are distinct due to (e.g. due to condition (10b) on *special*).

Finally, observation D can be explained in terms of informativity: With no contextual support we tend to take *John is (a)typical to* involve the very general modified property *PERSON,* i.e. to be interpreted as *John is a(n) (a)typical* person. However, the contribution of *(a)typical* in such a sentence seems to be trivial, since it is met with *John is a person* as well: Virtually every person has enough stereotypical properties of a human being (a person), hence being a *typical* person, and virtually every human individual is not 100% a stereotypical person, as every person must have certain properties that makes it unique, hence being an *atypical* person. This also explains why *John is (a)typical* is improved in contexts which support a more specific modified property (e.g. *lawyer*), where such inference is not trivially met. In contrast, given our proposal *John is (a) special (person)* is informative because it conveys that the speaker / protagonist takes John to have a particular non-stereotypical property, which is, furthermore considered 'positive'.

6 Conclusion

In this paper we provided a semantics for sentences with the adjective *special* in the positive form, and compared it to the adjective *atypical*, which, unlike *special*, has been already discussed in the formal semantics literature.

Our analysis supports analyzing the semantics of at least some multidimensional adjectives as making reference to stereotypical properties, and as involving modal quantification. At the same time, it shows that such multidimensional adjectives can differ from each other along at least two parameters: whether or not they are subjective, and in terms of the accessibility relations that restricts the modal quantification, e.g. just a stereotypical accessibility relation (as with *(a)typical*), or also in terms of a bouletic-like one (as with *typical*).

A more specific implication of our proposal concerns the semantics of sentences like (14), discussed in Greenberg (2023):

(14) This is (not) just another book about Stalingrad!

Greenberg argues that *just* in such cases associates with the focused *another*, and negates alternatives as in (15):

(15) This is a special book about Stalingrad.

Our analysis of *special* correctly explains Greenberg's observation that similarly to the sentence with *special*, a construction like (16) is degraded with 'unnatural properties':

(16) ??This is (not) just another book with 236 pages

As said above we focused in this paper on sentences with *special* in the positive form. In future research we would like to give a compositional analysis to such sentences, which will involve a lexical entry for *special,* and one for a covert *POS* degree modifiers for multidimensional adjectives (following ideas in Sassoon (2012), Greenberg ( )). This will allow us to analyze also sentences with other degree modifiers, e.g. *John is a more special than Bill / John is a bit special*, etc.

7 References

Bylinina, Lisa 2016. *The grammar of standards*. LOT Dissertation Series.

Cohen, Ariel 1999. Generics, Frequency Adverbs and Probability. *Linguistics and Philosophy 22* : 221–253.

D'Ambrosio, Justin & Brian Hedden 2024. Multidimensional Adjectives. *Australian Journal of Philosophy 2* : 253–277.

Del Pinal, Guillermo 2015. Dual content semantics, privative adjectives, and dynamic compositionality. *Semantics and Pragmatics 8*: 7–1–53.

Del Pinal, Guillermo 2018. Meaning, modulation, and context: a multidimensional semantics for truth-conditional pragmatics. *Linguistics and Philosophy 41*: 165–207

Greenberg Yael 2020. An Overt *Even* Operator over Covert-Based Focus Alternatives: The Case of Hebrew *BIXLAL*. *Journal of Semantics 37*: 1–42.

Greenberg, Yael 2023. *Overt and Covert Exclusive Operators over Non-overlap Relations*. Poster presented at SuB28.

Guerrini, Janek 2024. Keeping fake simple. *Journal of Semantics 42*: 175–210.

McReady, Eric & Nori Ogata 2007. Adjectives, stereotypicality, and comparison. *Natural Language Semantics 15*: 35–63

Sassoon, Galit 2013. A typology of multidimensional adjectives. *Journal of Semantics 30*: 335–380.

Scalarity, Information Structure and Relevance in varieties of Hurford Conditionals

Adèle Hénot-Mortier
Department of Linguistics and Philosophy, Massachusetts Institute of Technology
Cambridge, United States
adele.mortier@gmail.com

Abstract

Hurford Conditionals (HCs) involving scalemates appear felicitous, despite the fact that *exh* is not predicted to rescue such structures from Super Redundancy, a principle introduced in Kalomoiros (2024) to capture HCs without scalemates. We thus propose an alternative to Super Redundancy based on two main ideas: (1) expressions evoke QuDs locally, and when they combine via logical operators so do their respective QuDs; (2) a local QuD Q that gets incrementally combined with an existing QuD Q' should have its maximal true answers "fit" the information structure already introduced by Q'. This predicts that the consequent of a conditional has to evoke a question that properly refines some question evoked by the antecedent. The pattern is then captured with the additional assumption that scalar items can evoke fine-grained enough questions (generated by their scalemates) out-of-the-blue, while non-scalar items with different granularities cannot.

1 Introduction

Hurford Disjunctions (Hurford 1974) are disjunctions involving entailing disjuncts, and generally appear infelicitous (1), regardless of the linear order of the disjuncts.

(1) a. # Jo studied in France or Paris. $\mathbf{p} \vee \mathbf{p}^+$
b. # Jo studied in Paris or France. $\mathbf{p}^+ \vee \mathbf{p}$

Gazdar (1979) observed that infelicity disappears when (i) the Hurford disjuncts are scalemates, and (ii) the **weak** disjunct precedes the **stronger** one, as in (2-a). However, when the order of the two disjuncts is reversed, as in (2-b), infelicity remains (Singh 2008).

(2) a. Jo read some or all of the books. $\mathbf{s} \vee \mathbf{s}^+$
b. ??Jo read all or some of the books. $\mathbf{s}^+ \vee \mathbf{s}$

This linear asymmetry has received several accounts (Singh 2008; Fox and Spector 2018; Tomioka 2021; Hénot-Mortier 2022 i.a.), all of which capitalize on the idea that (2-a) can be rescued *via* a local scalar implicature within the first disjunct (allowed by the covert operator *exh*, Fox 2007; Spector, Fox, and Chierchia 2008); while (2-b) cannot, due to an interaction between the first disjunct, and the licensing/timing of *exh* in the second disjunct. In particular, Fox and Spector (2018) suggest *exh* should not be applied to an expression E if it turns out to be Incrementally Weakening (IW), i.e., if it leads to an equivalent/weaker meaning no matter the continuation Γ. The contrast in (2) then boils down to the fact *exh* is not IW in the first disjunct of (2-a) (cf. (3-a)), while it is in the second disjunct of (2-b) (cf. (3-b)).[1]

(3) a. $\exists\Gamma.\ \text{exh}(\mathbf{s})\ \Gamma \equiv (\mathbf{s} \wedge \neg \mathbf{s}^+)\ \Gamma \not\equiv \mathbf{s}\ \Gamma$ (e.g., take Γ to be the empty continuation)
b. $\forall\Gamma.\ (\mathbf{s}^+ \vee \text{exh}(\mathbf{s}))\ \Gamma \equiv (\mathbf{s}^+ \vee (\mathbf{s} \wedge \neg \mathbf{s}^+))\ \Gamma \equiv (\mathbf{s}^+ \vee \mathbf{s})\ \Gamma$

[1] (1-a) cannot be rescued like (2-a), either because *Paris* is not a natural alternative to *France* out-of-the blue, or because exhaustifying *France* by-city would lead to a symmetry problem (Kroch 1972; Fox 2007).

Does this pattern extend to structures isomorphic to HDs assuming material implication? Mandelkern and Romoli (2018) observed that an asymmetry arises in so-called Hurford Conditionals (henceforth HC), when the antecedent and consequent are *not* natural scalemates, as in (4). Interestingly, we observe that the asymmetry *disappears* in HCs involving scalemates (5).

(4) a. If Jo studied in France she did not study in Paris. $\mathbf{p} \rightarrow \neg \mathbf{p}^{+}$
b. #If Jo did not study in Paris she studied in France. $\neg \mathbf{p}^{+} \rightarrow \mathbf{p}$

(5) a. If Jo has read some of the books she hasn't read all. $\mathbf{s} \rightarrow \neg \mathbf{s}^{+}$
b. If Jo hasn't read all of the books she has read some. $\neg \mathbf{s}^{+} \rightarrow \mathbf{s}$

HDs and HCs therefore pattern differently, in both the scalar and the non-scalar case. Kalomoiros (2024) proposed a constraint called Super Redundancy accounting for (4), and spelled out in (6). (4-b) is Super Redundant (henceforth SR), because any local strengthening of its antecedent yields an expression equivalent to its consequent. (4-a) on the other hand, is not SR. SR can also cover (1), and, together with IW, (2).

(6) Super Redundancy. A sentence S is infelicitous if it contains a subconstituent C s.t. $(S)^{-}_{C}$ is defined and for all D, $(S)^{-}_{C} \equiv S_{Str(C,D)}$. In this definition, $(S)^{-}_{C}$ designates S where C got deleted. $Str(C, D)$ refers to a strengthening of C with D, which commutes with negation ($Str(\neg\alpha, D) = \neg(Str(\alpha, D))$) and with binary operators ($Str(O(\alpha, \beta), D) = O(Str(\alpha, D), Str(\beta, D))$). $S_{Str(C,D)}$ designates S where C is replaced by $Str(C, D)$.

What about (5-a) vs. (5-b)? (7) shows that (5-a) is isomorphic to (4-a), given that *exh* is IW in its antecedent and consequent.[2] So (5-a) is correctly predicted to be not SR.[3]

(7) a. *exh* is IW in the antecedent of (5-a).
$\forall\Gamma.\ \text{exh}(\mathbf{s}) \rightarrow \Gamma \equiv \neg(\mathbf{s} \wedge \neg \mathbf{s}^{+}) \vee \Gamma \equiv \neg \mathbf{s} \vee \mathbf{s}^{+} \vee \Gamma \dashv \neg \mathbf{s} \vee \Gamma \equiv \mathbf{s} \rightarrow \Gamma$
b. *exh* is IW in the consequent of (5-a).
$\forall\Gamma.\ (\mathbf{s} \rightarrow \text{exh}(\neg \mathbf{s}^{+}))\ \Gamma \equiv (\neg \mathbf{s} \vee (\neg \mathbf{s}^{+} \wedge \mathbf{s}))\ \Gamma \equiv (\neg \mathbf{s} \vee \neg \mathbf{s}^{+})\ \Gamma \equiv (\mathbf{s} \rightarrow \neg \mathbf{s}^{+})\ \Gamma$

(8) shows that this reasoning extends to (5-b): *exh* is IW in both the antecedent and the consequent of (5-b), so SR incorrectly predicts (5-b) to pattern like (4-b), i.e. to be infelicitous.

(8) a. *exh* is IW in the consequent of (5-b).
$\forall\Gamma.\ (\neg \mathbf{s}^{+} \rightarrow \text{exh}(\mathbf{s}))\ \Gamma \equiv (\mathbf{s}^{+} \vee (\mathbf{s} \wedge \neg \mathbf{s}^{+}))\ \Gamma \equiv (\mathbf{s}^{+} \vee \mathbf{s})\ \Gamma \equiv (\neg \mathbf{s}^{+} \rightarrow \mathbf{s})\ \Gamma$
b. *exh* is IW in the antecedent of (5-b).
$\forall\Gamma.\ (\text{exh}(\neg \mathbf{s}^{+}) \rightarrow \mathbf{s})\ \Gamma \equiv (\neg(\neg \mathbf{s}^{+} \wedge \mathbf{s}) \vee \mathbf{s})\ \Gamma \equiv (\mathbf{s}^{+} \vee \neg \mathbf{s} \vee \mathbf{s})\ \Gamma \dashv (\neg \mathbf{s}^{+} \rightarrow \mathbf{s})\ \Gamma$

At this point, one might want to revise IW, or SR. If SR is maintained and IW is assumed to be inactive in conditionals, then both HCs in (5) would be predicted to be felicitous, due to *exh* being licensed in the consequent of (5-b). An argument against this view comes from Long-Distance HCs (henceforth LDHC). Such structures, inspired from Long-Distance HDs (Marty and Romoli 2022) and exemplified in (9), are derived from (5) by further disjoining $\mathbf{p}^{+}$ with a proposition $\mathbf{r}$ taken to be incompatible with $\mathbf{p}$. Felicity-wise, LDHCs seem quite degraded: (9-a) seems to convey the same kind of information twice (because having studied in France contextually entails not having studied in Brussels), while (9-b)'s consequent seems independent from the subject matter raised by the antecedent. Yet neither (9-a) nor (9-b) are predicted to

[2] We show it for material conditionals here, but this extends to strict conditionals.

[3] Some speakers I consulted reported that (5-a) was hard to make sense of in English (it is fine in my French). One reason for this, in my framework, might be that negated expressions (e.g. *not all*) more saliently evoke "polar" QuDs (e.g. $\forall/\neg\forall$) as opposed to other QuDs (e.g. $\forall/\exists \wedge \neg\forall/\neg\exists$). And, when combined with an antecedent QuD for *some*, a polar QuD for *not all* happens to violate Relevance.

be Super Redundant, due to the presence of **r**. Crucially, this incorrect prediction persists if we maintain SR and assume IW is not active in the consequent of conditionals.

(9) a. ? If Jo studied in France she didn't study in Paris or Brussels. $\mathbf{p} \rightarrow \neg(\mathbf{p}^{+} \vee \mathbf{r})$
b. # If Jo didn't study in Paris or Brussels she studied in France. $\neg(\mathbf{p}^{+} \vee \mathbf{r}) \rightarrow \mathbf{p}$

To capture (5) and (9) while retaining the right predictions for (1), (2) and (4), we thus suggest to maintain IW, and propose an alternative to SR based on three ideas: Questions under Discussion (**QuD**, Roberts 1996) are compositionally accommodated when processing out-of-the-blue declaratives; scalemates may answer same-granularity QuDs; and QuD computation is constrained by Q-Relevance (Hénot-Mortier 2024a). The scalar HCs in (5) can then escape a violation of Q-Relevance, because their consequent can be understood to evoke a question that is fine-grained enough to "fit" the question introduced by their antecedent. In the non-scalar case, (4-a) can do the same, but crucially not (4-b).

2 A compositional theory of accommodated QuDs

We argue that out-of-the-blue declaratives evoke the potential QuDs they may answer (Katzir and Singh 2015; Zhang 2024), and that the derivation of such implicit QuDs is compositional. A sentence compatible with no reasonable implicit QuD is deemed odd. We start by showing that the felicity of disjunctions and conditionals is sensitive to *overt* QuDs – but in different ways. We take this as evidence that out-of-the-blue disjunctions and conditionals accommodate different kinds of implicit QuDs.

If a context contrasting *Paris* and *France but not Paris* is set as in (10), (4-b) and (9-b) improve (cf. Haslinger 2023 for similar effects on disjunctions and conjunctions). This is strange: even if the context and question made *Paris* (but no other French city) a relevant alternative to *France*, *exh* would remain IW in the consequent of (4-b): *if Jo did not grow up in Paris, she grew up in France but not Paris*, is equivalent to *if Jo did not grow up in Paris, she grew up in France*. In other words, *exh* (as constrained by IW) cannot leverage the contextually provided alternatives to make (4-b) escape SR in (10). The same applies to (9-b).

(10) *Context: French accents vary across countries and between Paris the rest of France.*
Al: I'm wondering where Jo learned French.
Lu: I'm not completely sure but... (4-b) ✓(9-b) ✓

This suggests that a purely LF-based view of redundancy such as SR, may be insufficient to capture the interaction between HCs and how their context of utterance packages information. Rather, it seems that the context of (10) makes a specific partition of the CS salient, and that this partition can be used to make otherwise infelicitous assertions accommodate a different question than the one they would evoke out-of-the-blue.

Additionally, conditionals and disjunctions seem to accommodate distinct QuDs. To show this, we use the construction *depending on Q, p* (Karttunen 1977; Kaufmann 2016), where Q is a question and p a proposition. This construction has been argued to force the partition conveyed by Q to match specific live issues raised by p. We understand such "live issues" as the maximal true answers of the QuD evoked by p. The contrast between (11-a) and (11-b) then suggests that the *France* and *Belgium* answers can be matched against Q in the disjunctive, but not in the conditional case. This in turn means that a disjunction introduces a QuD making both disjuncts maximal true answers, while a conditional does not do the same with its consequent and the negation of its antecedent.

(11) Depending on [how her accent sounds like]$_Q$...

a. Jo grew up in France **or** in Belgium. $\mathbf{p}\vee\mathbf{q}$

b. ??**if** Jo didn't grow up in France she grew up in Belgium. $\neg\mathbf{p}\rightarrow\mathbf{q}$

c. ?**if** Jo didn't grow up in France, she grew up in Belgium **or** in Québec. $\neg\mathbf{p}\rightarrow(\mathbf{q}\vee\mathbf{r})$

d. ??**if** Jo didn't grow up in France **or** Belgium, she grew up in Québec. $\neg(\mathbf{p}\vee\mathbf{q})\rightarrow\mathbf{r}$

The existence of an improvement between (11-b) and (11-c), and the absence of a similar improvement in between (11-b) and (11-d), also implies that the answers targeted by *depending on Q*, when p is conditional, are the ones made available by the consequent of p (which is appropriately disjunctive in (11-c), but not (11-d)). Building on these observations and on the formalism introduced in Hénot-Mortier (2024a) and Hénot-Mortier (2024b) to account for non-scalar HCs and HDs, we now focus on explaining the novel datapoints (5).

The formalism we present here summarizes the more complete model set out in Hénot-Mortier (2024a). Building on Büring 2003; Riester 2019; Onea 2016; Zhang 2024, we take QuDs to be trees (**Qtrees**), that have the Context Set (**CS**, Stalnaker 1974) as their root, and are s.t. each intermediate node is a subset of the CS, partitioned by its children nodes. Thus, the leaves of a Qtree partition the CS, and correspond to the standard denotation of questions (Hamblin 1958; Groenendijk 1999). Any subtree rooted in N can be seen as a conditional question, granted N. A proposition answers a Qtree if it can be identified with the union of a strict subset of the Qtree's nodes. Building on Katzir and Singh (2015), Hénot-Mortier (2024a), and Hénot-Mortier (2024b), we assume that any out-of-the-blue declarative sentence denoting a proposition p gets paired with the set of salient Qtrees p may answer. An LF that cannot be paired with any well-formed Qtree is deemed odd. Such Qtrees additionally carry information about how p answers them, in the form of specific nodes entailing p (**verifying nodes**). We will refer to the structure formed by Qtrees, along with their verifying nodes, as "flagged Qtrees". The pairing between LF and flagged Qtrees is compositional, meaning, the flagged Qtrees evoked by a complex LF, are derived from the flagged Qtrees derived from its parts, and given how these parts combine.

We start with a simplex LF X denoting p. We assume here that a Qtree for X may be a depth-1 Qtree whose leaves denote p and $\neg p$; or a depth-1 Qtree whose leaves correspond to the Hamblin partition of the CS generated by p and same-granularity alternatives to p.[4] We take granularity as a primitive, but assume that scalemates such as *some* and *all may* be seen as same granularity alternatives to each other, while non-scalemates, like *Paris* and *France*, *cannot* be considered being so, at least out-of-the blue.[5] In any case, Qtrees derived from simplex LFs get "flagged" by defining their verifying nodes as the set of nodes entailing p.

According to this definition, *Jo read all of the books* gets paired with a "polar" Qtree corresponding to whether or not she read all the books (cf. Fig. 1); and a "wh" Qtree corresponding to whether she read none, only some, or all of the books (generated by $\text{Alt}(\exists) = \{\exists, \forall\}$, cf. Fig. 2). Same can be done for *Jo read some of the books*, except the "polar" Qtree is different (cf. Fig. 3). For *Jo lives in Paris* (resp. *France*), *wh*-Qtrees are generated by city (resp. country) alternatives, cf. Fig. 4 and 5. Verifying nodes are boxed.

[4] Hénot-Mortier (2024a) proposes that simplex LFs may also evoke "tiered" Qtrees with more than one layer. We do not reject this assumption here, but omit it for brevity. This does not affect our main prediction about scalar HCs, which is based on the granularity distinction between *all*/*some* vs. *Paris*/*France*.

[5] This might relate to the symmetry problem with such alternatives: if *Paris* and *France* are considered same-granularity, then *France* and any city in France should be considered to be same-granularity. This seems counter-intuitive, given that at a certain level of abstraction, all French cities *together* cover France.



Figure 1: "Polar" Qtree for *Jo read **all** of the books*

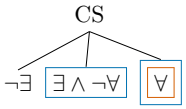


Figure 2: "Wh" Qtree for *Jo read {**some**/**all**} of the books.*



Figure 3: "Polar" Qtree for *Jo read **some** of the books*



Figure 4: Qtrees for *Jo studied in **Paris**.*



Figure 5: Qtrees for *Jo studied in **France**.*

Just like the meanings of simple sentences are incrementally composed, their sets of candidate Qtrees get incrementally combined. The Qtrees compatible with a negated LF $\neg X$, are Qtrees for X in which the set of compatible nodes is "flipped" on a layer-by-layer basis. *Jo did not read all of the books* is thus linked to the Qtrees in Fig. 6 and *Jo didn't study in Paris* to those in Fig. 7. The Qtrees compatible a disjunctive LF $X \vee Y$, are all the Qtrees that result from the union of a tree for X, and a tree for Y. The union operation – understood as union over sets of nodes, sets of edges, and sets of verifying nodes – ensures that the Qtree of a disjunction addresses the QuDs evoked by *both* disjuncts in parallel (Simons 2001; Zhang 2024). *Jo read some or all of the books* is therefore only compatible with Tree 2 because all the other unions obtained from of Trees 1, 3 and 2 fail to generate proper Qtrees. The HDs (1-a)-(1-b) are not compatible with any Qtree, because the Qtrees for *Paris* and those for *France* always subdivide the CS differently.[6] The Qtrees compatible with a conditional LF $X \rightarrow Y$ are Qtrees for X, where each verifying node is replaced by its intersection with a Qtree for Y. Verifying nodes are inherited from the consequent Qtree (in line with the observations in (11)). (5-a) is then compatible with the Tree in Fig. 8; (5-b), with Fig. 9, (4-a) with Fig. 10 and (4-b) with Fig. 11. We proceed to show that both trees associated with (4-b) violate some notion of relevance; while no trees associated with (5-a), (5-b), and (4-a) do. Roughly, the issue is that none of the trees evoked by (4-b) preserve the answer conveyed by its consequent (the *France*-node); while those evoked by (5-a), (5-b) and (4-a) do.

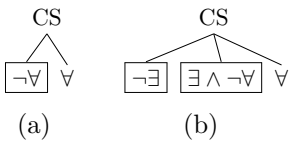


Figure 6: Qtrees for *Jo did**n't** read **all** of the books*, derived from Fig. 1&2

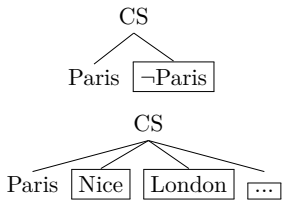


Figure 7: Qtrees for *Jo did**n't** study in **Paris***, derived from Fig. 4.

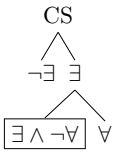


Figure 8: A Qtree compatible with (5-a) derived from Fig. 3&6a/6b

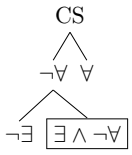


Figure 9: A Qtree compatible with (5-b) derived from Fig. 6a&2

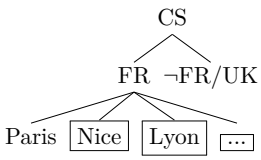


Figure 10: Qtree for (4-a), derived from Fig. 5&7.

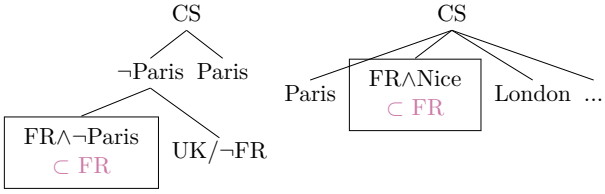


Figure 11: Qtree for (4-b), derived from Fig. 7&5.

[6] Hénot-Mortier (2024a) and Hénot-Mortier (2024b) predict that (1-a)-(1-b) do create proper Qtrees, but that such Qtrees are redundant.

3 Relevance as a constraint on QuD computation

Hénot-Mortier (2024a) defined Q-Relevance as a constraint on Qtree computation: when combining Qtrees incrementally, none of the verifying nodes of the input Qtree should be cut across (i.e. be strictly entailed by some node) in the output Qtree. This allowed to account for the contrast in (4). Let us briefly summarize the argument. (4-a) corresponds to the Qtree in Fig. 10, which is obtained from a country-level antecedent Qtree and a city-level consequent Qtree; therefore, all verifying leaves of the consequent (city nodes different from Paris) are contained in some leaf of the antecedent Qtree, and can thus "fit" into the output Qtree without being cut across. Q-Relevance is thus satisfied. (4-b) corresponds to the Qtrees in Fig. 11, which are obtained from a city-level antecedent Qtree and a country-level consequent Qtree; in such trees, the *France* verifying leaves are always cut across, either by *not Paris*, or by individual city-nodes different from *Paris*. Q-Relevance is thus violated.

The same kind of reasoning shows that the Qtrees corresponding to (5-a) and (5-b), in resp. Fig. 8 and 9, verify Q-Relevance. Starting with Qtree 8: it can be built by incrementally combining Qtree 3 (antecedent Qtree), with Qtree 6b (consequent Qtree). Qtree 6b has $\neg\exists$ and $\exists \wedge \neg\forall$ as verifying nodes; in the output Qtree 8, both nodes are fully preserved. So (5-a) is compatible with a Qtree and is thus felicitous. As for (5-b), its Qtree 9 can be built by incrementally combining Qtree 6b (antecedent Qtree), with Qtree 2 (consequent Qtree). Qtree 2 has $\forall$ and $\exists \wedge \neg\forall$ as verifying nodes; in the output Qtree 8, both nodes are fully preserved. So (5-b) is compatible with a Qtree and is thus felicitous. In brief, (4-b) and (4-a) are both rescued by the fact their consequent can evoke a Qtree whose verifying nodes are fine-grained enough to properly "fit" the structure already introduced by the antecedent Qtree. This extends to the LDHCs (9-a) and (9-b), whose trees are given in Fig. 12 and 13. For (9-a), Q-Relevance is satisfied because none of the verifying leaves from the consequent Qtree (cities different from Paris and Brussels) are cut across in the output Qtree. The feeling of redundancy in (9-a) may come from the fact removing *Brussels* from the sentence leads to the same overall meaning and Qtree.[7] For (9-b), Q-Relevance is violated because the compatible *France* leaf from the consequent Qtree cannot "fit" into the city-level nodes introduced by the antecedent.

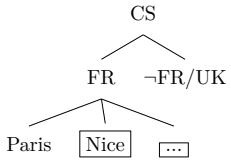


Figure 12: Qtree for (9-a).

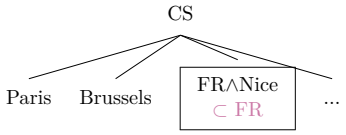


Figure 13: Qtree for (9-b).

4 Conclusion and outlook

We proposed an account of (scalar) HCs exploiting the intuitive idea that conditionals evoke "restricted" questions whose composition is constrained by a new notion of relevance, Q-Relevance (Hénot-Mortier 2024a). The contrast between scalar and non-scalar HCs was thus captured, not *via exh per se*, but instead by appealing to how scalar vs. non-scalar pairs of items differ information-structurally. Specifically, it was assumed scalar items could evoke fine-grained enough questions (generated by their scalemates) out-of-the-blue, while non-scalar items with different granularities could not.

Acknowledgements. Many thanks to Amir Anvari, Athulya Aravind, Omri Doron, Danny Fox, Martin Hackl, Nina Haslinger, Viola Schmitt, and Jad Wehbe for their feedback and advice.

[7] Following Hénot-Mortier (2024a) and Hénot-Mortier (2024b), this implies that (9-a) is Q-Redundant.

References

Büring, Daniel (2003). "On D-Trees, Beans, and B-Accents". In: *Linguistics & Philosophy*.

Fox, Danny (2007). "Free choice and the theory of scalar implicatures". In: *Presupposition and implicature in compositional semantics*. Springer.

Fox, Danny and Benjamin Spector (2018). "Economy and embedded exhaustification". In: *Natural Language Semantics*.

Gazdar, Gerald (1979). *Implicature, Presupposition and Logical Form*.

Groenendijk, Jeroen (1999). "The Logic of Interrogation: Classical Version". In: *SALT*.

Hamblin, Charles L. (1958). "Questions". In: *Australasian Journal of Philosophy*.

Haslinger, Nina (2023). "Pragmatic constraints on imprecision and homogeneity". PhD thesis. Georg-August-Universität Göttingen.

Hénot-Mortier, Adèle (2022). "Alternatives are blind to some but not all kinds of context: the view from Hurford Disjunctions". In: *Proceedings of the 27th Sinn und Bedeutung*.

— (2024a). ""One tool to rule them all"? An integrated model of the QuD for Hurford sentences". In: *Proceedings of the 29th Sinn und Bedeutung*. Draft (`https://adelemortier.github.io/files/SuB_2024_TMT_paper.pdf`), to appear.

— (2024b). "Redundancy under Discussion". In: *Proceedings of the 29th Sinn un Bedeutung*. To appear.

Hurford, James R. (1974). "Exclusive or Inclusive Disjunction". In: *Foundations of Language*.

Kalomoiros, Alexandros (2024). "An approach to Hurford Conditionals". In: *SALT*.

Karttunen, Lauri (1977). "Syntax and Semantics of Questions". In: *L&P*.

Katzir, Roni and Raj Singh (2015). "Economy of structure and information: Oddness, questions, and answers". In: *SuB 19*.

Kaufmann, Magdalena (2016). "Free Choice is a Form of Dependence". In: *Natural Language Semantics*.

Kroch, A. (1972). "Lexical and inferred meanings for some time adverbs". In: *Quarterly Progress Report of the Research Laboratory of Electronics 104*.

Mandelkern, Matthew and Jacopo Romoli (2018). "Hurford Conditionals". In: *Journal of Semantics*.

Marty, Paul and Jacopo Romoli (2022). "Varieties of Hurford disjunctions". In: *Semantics and Pragmatics*.

Onea, E. (2016). *Potential questions at the semantics-pragmatics interface*. Brill.

Riester, Arndt (2019). "Constructing QUD Trees". In: *Questions in Discourse*.

Roberts, Craige (1996). "Information structure in discourse: Towards an integrated formal theory of pragmatics". In: *Semantics and Pragmatics*.

Simons, Mandy (2001). "Disjunction and Alternativeness". In: *L&P*.

Singh, Raj (2008). "On the Interpretation of Disjunction: Asymmetric, Incremental, and Eager for Inconsistency". In: *L& P*.

Spector, Benjamin, Danny Fox, and Gennaro Chierchia (2008). "Hurford's Constraint and the Theory of Scalar Implicatures". In: *Ms.*

Stalnaker, Robert (1974). "Pragmatic Presuppositions". In: *Context and Content*.

Tomioka, Satoshi (2021). "Scalar Implicature, Hurford's Constraint, Contrastiveness and How They All Come Together". In: *Frontiers in Communication*.

Zhang, Yichi (2024). "A Questions-Under-Discussion-Based Account of Redundancy". In: *Unpublished manuscript*.

Additive Presuppositions and Logical Strength

Elena Herburger
Georgetown University
Washington, DC USA
herburge@georgetown.edu

Paul Portner
Georgetown University
Washington, DC USA
portnerp@georgetown.edu

Abstract

This paper deals with the presuppositions of additive *either*. Our aim is to link the negative character of *either*'s presupposition to its being an NPI. We adopt the view that presuppositions translate as silent conjuncts whose content is taken by the interlocutors to follow from the local context and to strengthen the at-issue meaning. To explain why the presupposition of additive *either* features a negation we propose that the presuppositions of additive *either* actually originate as silent disjuncts. On the resulting analysis the negation that licenses additive *either* as an NPI is also responsible for the negation we see in its presupposition. Independent support for a disjunctive analysis of additive *either* is argued to come from the overtly disjunctive uses of *either*.

1 Introduction: Presupposition as conjunction

Presuppositions allow language to convey extra information, beyond what is spelled out explicitly. This sense of "extra information" means, in essence, that presuppositions always strengthen the at-issue meaning. One way to express the strengthening power of presuppositions is to view them as special conjuncts in logical form (e.g. Beaver and Krahmer 2001, Schlenker 2009). (Later we will say more about what "special" means.) We can informally represent this way of looking at presupposition by translating (1-a) as (1-b), where ~~strikeout~~ represents phonological silence and semantic presuppositionality.

(1) a. Anita danced again.
b. $\exists e(\sout{Dance(e,a) \wedge Past(e)}) \wedge \exists e'(Dance(e',a) \wedge Past(e') \wedge e < e')$

Apparently, we never have logical forms like (2), with a backgrounded **disjunct**:

(2) $\exists e(\sout{Dance(e,a) \wedge Past(e)}) \vee \exists e'(Dance(e',a) \wedge Past(e') \wedge e < e')$

We argue that the essential function of presuppositions to strengthen meaning is integrated into the compositional system in a way that rules out (2), but which actually — and correctly — predicts the existence of disjunctive presuppositions in certain instances.

2 The issue: Polarity in additive presuppositions

Additive expressions like *too* and the NPI *either* are known to carry focus-sensitive presuppositions, (cf. e.g. Rullmann 2003, Ahn 2015, Szabolcsi 2017). Whereas (3-b) presupposes that there is another, contextually relevant place besides Paris that Anita visited (here New York), (4-b) carries the negative presupposition that there is another contextually relevant place besides Paris to which Anita did **not** go:

(3) a. Anita went to New York.
b. She went to $[\text{Paris}]_F$ **too**. [additive *too*]
c. $\exists e(\sout{Go(e,a) \wedge To(e,NY) \wedge Past(e)}) \wedge \exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e'))$

(4) a. Anita didn't go to New York.
b. She didn*(n't) go to [Paris]$_F$ **either**. [NPI *either*]
c. ~~$\neg\exists e(Go(e,a) \wedge To(e,NY) \wedge Past(e))$~~ $\wedge$ $\neg\exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e'))$

The question arises as to how the focus alternative is identified and how it is used to compute the observed presuppositions. In the case of (3-b), *too* has scope over *went to [Paris]*$_F$, and so it seeks an alternative of the form *went to* LOCATION$\neq NY$. Identifying *went to New York* in the context, we produce the logical form (3-c).

Turning to (4-b), it is not immediately clear whether or not *either* has scope over the negation. If it has scope over negation, the derivation works in a way completely parallel to (3-b), identifying *not go to New York* in the context and conjoining this with the overt material to produce (4-c). However, there is a problem with assuming this derivation. NPIs always have scope under their licensors, which means that the alternatives should be non-negative (see Rullmann 2003 for arguments). In the case at hand, we should assume that *either* in (4-b) seeks an alternative of the form of the same form that *too* does, i.e. *went to* LOCATION$\neq NY$. It then produces the presupposition that this alternative is false, yielding (4-c).

From this perspective, it seems that *too* and *either* generate opposite presuppositions (cf. Rullmann 2003). But this seems that we may be missing a generalization. The derivation with the NPI *either* involves negation in two ways; it requires a licensing negation, and it negates its focus alternative. Could it be this is the same negation? We propose that it is, but it is not immediately clear how to make this proposal compatible with the idea that the additive presupposition is a backgrounded conjunct. It we introduce the conjunct at the same point in the derivation as *either*, we get the logical form (5).

(5) $\neg(\exists e(Go(e,a) \wedge To(e,NY) \wedge Past(e)) \wedge \exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e')))$

This is clearly incorrect. If we assume that the presupposition projects, it becomes equivalent to (6), which is also incorrect.

(6) ~~$\exists e(Go(e,a) \wedge To(e,NY) \wedge Past(e))$~~ $\wedge$ $\neg\exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e'))$

What would work better would be to assume that *either* involves a silent disjunction, as proposed by Ahn (2015). We add to Ahn's proposal the idea that the the disjunct is a unpronounced copy of the focus alternative and not part of the at-issue meaning, leading to (7):

(7) $\neg(\exists e(Go(e,a) \wedge To(e,NY) \wedge Past(e)) \vee \exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e')))$

If de Morgan's Law applies to (7), (7) can be seen as equivalent to (8), which is the correct logical form within the notation we have been employing.

(8) ~~$\neg\exists e(Go(e,a) \wedge To(e,NY) \wedge Past(e))$~~ $\wedge$ $\neg\exists e'(Go(e',a) \wedge To(e',Paris) \wedge Past(e'))$

In sum, if additive *either* introduces its not-at-issue component as a silent disjunct provided by a distinct focus alternative we can connect the negative character of its presupposition to its status as an NPI.

3 Additional arguments for a disjunctive analysis

We can also observe additional, supporting arguments for a disjunctive analysis of additive *either*. One argument is that a disjunctive analysis is consistent with the fact that disjunction has a low scalar semantics, the kind of semantics that is typical of NPIs (cf. Ahn 2015). This argument only goes so far, however. Not all NPIs are low-scalar expressions (cf. the high-scalar NPIs *at all*, adverbial *much*), nor are all low-scalar expressions are NPIs, e.g. *a, or* are not. NPI-hood on

the analysis adopted here is an arbitrary lexical feature that semantically predisposed expression can come to bear (cf. Chierchia 2013, Herburger 2024). In other words, a low-scalar, disjunctive analysis of additive *either* fits the fact that it has the distribution of an NPI, but does not fully explain it.

A disjunctive analysis of *either* is also appealing because there is an *either* that can appear with overt *or* (cf. also Thomas 2021). When *either* appears to the left of *or* it marks the scope of the disjunction as illustrated by the different interpretations that we observe for (e.g. Larson 1985). Whereas (9-a) is a disjunction over which place they were told to go, (9-b) is a disjunction over what they were told.

(9) a. They were told to **either** go to New York **or** Paris.
b. They **either** were told to go to New York **or** Paris.

The second disjunct in instances where *either* pairs up with *or* can involve syntactic ellipsis, in particular Gapping of the non-focused material in the second disjunct, if that material is the same as in the first disjunct (Schwarz 1999). The difference in meaning we see between (9-a) and (9-b) then stems from the size of the gapped material:

(10) a. They were told to **either** go to New York **or** ~~go-to~~ Paris.
b. They were **either** told to go to New York **or** ~~told-to-go-to~~ Paris.

On the proposed analysis, the logical form of sentences with additive *either* ($either_{addNPI}$) and the overtly disjunctive *either* is rather close:

(11) a. *Anita did*(***n't***) go to New York* **either**$_{addNPI}$ iff
$\neg(\exists e Went(e,a) \wedge Past(e) \wedge To(e, NewYork_j)$
~~$\vee \exists e Went(e,a) \wedge Past(e) \wedge To(e,x_i)$~~)
b. *Anita* **either** *went to New York* **or** ~~*went to Paris*~~ iff
$\exists e Went(e,a) \wedge Past(e) \wedge To(e, NewYork_j)$
$\vee \exists e$ ~~$Went(e,a) \wedge Past(e)$~~ $\wedge To(e, Paris_i)$

They differ in that additive *either* is lexically marked as an NPI, while *either...or* is not. Another difference concerns the size and origin of the silent, presupposed material. In the case of additive *either*, the entire second disjunct is not made explicit; it is a distinct focal alternative that must be specified by the context. In the case of *either...or*, by contrast, gapping optionally extends over the non-focused part of the second disjunct when it is identical to the non-focused part of the first disjunct; it does not include the focus or the connective *or*.

Another set of structures where we see a connection between an instance of *either* and disjunction are those where *either* functions as a determiner for a dual DP. In (13), *either* functions as an NPI, (12-a). Limiting the girls to Frances and Beatrice we find that (13-a) is equivalent to the overt disjunctions (13-b).

(12) a. I didn't see either girl.
b. I didn't see either Frances or Beatrice.

In addition it can also appear as the determiner of a free choice quantifier, where it can also be paraphrased with a disjunction, (13-b).

(13) a. Either girl could have won the dance competition.
b. Either Frances or Beatrice could have won the dance competition.

In a few fixed expressions, dual *either* is better paraphrased with *and* than *or*, (14). These cases are remnants from Old English, where the determiner *either* was universal, e.g. Gast (2013).

(14) There are trees on either side of the house.

In sum, there is evidence from the broader distribution of *either* that it is closely associated with disjunction. These arguments of course do not prove that additive *either* introduces a disjunction, but in our view, they greatly add to the plausibility.

4 A context-sensitive presupposition operator

The basic idea of our analysis is that NPI *either* introduces what will become its presupposition (once it is negated) as a disjunction, and that it is turned into a negated conjunct by the licensing negation. However, there is a puzzle. In order for de Morgan's law to apply, the silent disjunct introduced by *either* must be treated as part of the propositional content that is operated on by negation. That is, at the point that negation applies, the disjunct counts as at-issue meaning. By contrast, the negation of this same material is ultimately projective; in other words, outside the syntactic scope of its licensor, it is treated as a presupposition. Thus we seem to have a contradiction: the silent disjunct is both at-issue and (once negated) presuppositional.

We propose that this equivocal behavior is a consequence of three facts:

1. The basic meaning of *either* is disjunction.
2. As *either* is an additive word, one of the disjuncts is anaphoric, unpronounced, and presuppositional.
3. Presuppositions strengthen the meaning they are a part of.

The silent disjunct of *either* must strengthen the overall meaning, but given that it is introduced as a disjunct, this is only possible once the NPI has been licensed. Thus, we need a compositional process that defers treating the silent disjunct as a presupposition until the point where it strengthens the meaning.

We can imagine several ways of formalizing the process of composition that would model this way of looking at things. One approach is to assume an operator π that applies to the disjunct, storing its argument p and treating $\neg p$ as a presupposition as soon as $\neg p$ is an entailment.

In (15), we specify how *either* introduces the silent disjunct and stores it:

(15) $\psi, (either\ \phi) =$
a. $(\pi(p) \vee q) =$
b. $\langle (p \vee q), p \rangle$

(The standard meanings of ϕ and ψ are p and q, respectively.)

The stored material p is essentially a presupposition waiting to happen; it "waits" until either it or its negation strengthens the overall meaning, and then it triggers a presupposition. This "waiting" and "triggering" process can be incorporated into the process of composition. Let us assume that function application is the only relevant rule.

(16) Function Application (with pending presupposition)
For any node $[_\alpha\ \beta, \gamma]$ and local context c, where $[[\gamma]]^c = \langle d, p \rangle$ and d is in the domain of $[[\beta]]^c$:
a. $[[\ \alpha\ ]]^c = \langle [[\beta]]^c(d), p \rangle$, if $[[\beta]]^c(d)$ entails neither p nor $\neg p$.
b. If $[[\beta]]^c(d)$ entails p:
(i) $[[\alpha]]^c$ is only felicitous if p is entailed by c; and
(ii) if felicitous, $[[\alpha]]^c = [[\beta]]^c(d)$.
c. If $[[\beta]]^c(d)$ entails $\neg p$:

(i) $[[\alpha]]^c$ is only felicitous if $\neg p$ is entailed by c; and
(ii) if felicitous, $[[\alpha]]^c = [[\beta]]^c(d)$.

In the case of *either*, it is (16)(c) that introduces the presupposition that p is false when negation applies and licenses the NPI. Note that, in (16), we treat presuppositions as pragmatic felicity conditions, though for Schlenker (2009), they are definedness conditions.

Let us apply this reasoning to (4-b).

(17) She didn't go to Paris either.
a. $[[$ either ϕ_{NY} $]]^c = \lambda q\langle(NY \vee q), NY\rangle$
b. $[[$ she went to Paris either ϕ_{NY} $]]^c = \langle(NY \vee Paris), NY\rangle$
c. $[[$ not (she went to Paris either ϕ_{NY}) $]]^c$: Apply (16-c)
$\neg(NY \vee Paris)$ entails $\neg NY$, so
i. $[[$ not (she went to Paris either ϕ_{NY}) $]]^c$ is only felicitous if c entails $\neg NY$; and
ii. if felicitous, $[[$ not (she went to Paris either $\phi_{NY})]]^c = \neg(NY \vee Paris)$

The same general form of explanation applies to additive *too* as well. In this case, the additive introduces a conjunction, and the presupposition is triggered as soon as the overt conjunct is combined with the covert one:

(18) She went to Paris too.
a. $[[$ too ϕ_{NY} $]]^c = \lambda q\langle(NY \wedge q), NY\rangle$
b. $[[$ she went to Paris too ϕ_{NY} $]]^c$: Apply (16-b)
$(NY \wedge Paris)$ entails NY, so
i. $[[$ she went to Paris too ϕ_{NY} $]]^c$ is only felicitous if c entails NY; and
ii. if felicitous, $[[$ she went to Paris too ϕ_{NY} $]]^c = (NY \wedge Paris)$

Another way to understand our basic idea has a more syntactic, "constructional" character. Suppose we think of the an NPI and its licensor as two parts of one meaningful object, so that (in the case at hand), we interpret the *not... either* combination in stages. At the lower level, *either* introduces p into the derivation, but it cannot fully incorporate it into the meaning yet, because the the construction is not yet complete. When negation is introduced, it negates p along with the overt material; then, since the construction is complete, $\neg p$ can be fully integrated into the overall meaning. Since it is presuppositional, it must strengthen, and so it is incorporated as a conjunct. We do not attempt to formalize this way of looking at the relation between *either* and its licensor, for reasons of space.

5 Conclusion and open issues

We have argued that presuppositions are backgrounded propositions that have the core function of strengthening the at-issue meaning. In the case of additives, lexical information determines whether the presupposition will be present in a conjunction or a disjunction. When it is a disjunction, as a practical matter the item must occur in a negative context, so that it can strengthen the overall meaning. In other words, we explain the connection between additivity, disjunctive lexical meaning, and NPI-hood. These findings therefore contribute to the larger goal of explaining how presuppositions arise and how they are related to both lexical and grammatical meaning.

We have also sketched a compositional process that instantiates the analysis. It is, of course, an open question how to extend this approach to other types of presuppositions. One obvious

place to start is with the related additive element *neither*. This element has a very specific syntax in mainstream (prescriptive) English, requiring negative inversion, as in (19). In some colloquial varieties, it is a negative concord term.

(19) (Anita didn't go to New York.) Neither did she go to [Paris]F.

In addition to being an additive, *neither* also occurs as the scope marker in *neither...nor*. Thus (20-a) and (20-b) are similar to the corresponding cases with *either*, (10-a) and (10-b).

(20) a. They were told to **neither** go to New York **nor** ~~go-to~~ Paris.
b. They were **neither** told to go to New York **nor** ~~told-to-go-to~~ Paris.

In terms of our analysis, additive *neither* could be seen as introducing either a conjunction of two negative propositions or a disjunction of non-negative ones that needs a higher negation. Both approaches face challenges in assigning the correct polarity to each piece. If we choose the conjunctive analysis, we must explain why the two conjuncts are negated. If we assume the disjunctive analysis, we must explain why the negative sentence $\neg p$ in the context counts as an antecedent for the non-negative disjunct p under the scope of negation $\neg(p \vee q)$. The puzzle similar to the one encountered with *either*. We would like to be able to formalize the intuition that the semantic negation contributed by *neither* takes scope over the disjunction.

Other languages divide the space of additives differently. We find for instance that Spanish *tampoco* can translate as *neither*, which it does in preverbal position, or as additive *either*, as we can observe in postverbal position.

(21) a. Tampoco fue a Paris.
TAMPOCO went to Paris
'Neither did s/he go to Paris.'
b. *(No) fue a Paris tampoco.
not went to Paris TAMPOCO
'S/he didn't go to Paris either.'

The apparent ambiguity of *tampoco* makes it part of a more general phenomenon, negative concord. We believe that the ambiguity we observe with *tampoco* may offer a clue as to how to treat the relation between *neither* and the NPI *either*, and more generally about how polarity and presupposition interact in additive expressions.

References

Ahn, Dorothy (2015). "The semantics of additive either". In: Proceedings of Sinn und Bedeutung 19, pp. 20–35.
Beaver, David and Emiel Krahmer (2001). "A partial account of presupposition projection". In: Journal of Logic, Language and Information 10, pp. 147–182.
Chierchia, Gennaro (2013). Logic in Grammar. Oxford University Press.
Gast, Volker (2013). "From æghwæðer to either: The distribution of a negative polarity item in historical perspective". In: Beyond Any and Ever. Ed. by Eva Csipak, Regine Eckardt, and Manfred Liu Maia an Sailer. : de Gruyter Mouton, pp. 19–101.
Herburger, Elena (2024). "On the history of NPIs and Negative Concord". In: Canadian Journal of Linguistics 68.4. Special issue on historical semantics 2, pp. 555–589.
Larson, Richard (1985). "On the syntax of disjunction scope". In: Natural Language and Linguistic Theory 3, pp. 217–264.

Rullmann, Hotze (2003). "Additive particles and polarity". In: Journal of Semantics 20.4, pp. 329–401.
Schlenker, Philippe (July 2009). "Local Contexts". In: Semantics and Pragmatics 2.3, pp. 1–78.
Schwarz, Bernhard (1999). "On the syntax of either...or". In: Natural Language and Linguistic Theory 17, pp. 339–370.
Szabolcsi, Anna (2017). "Additive presuppositions are derived through activating focus alternatives". In: Proceedings of the 21st Amsterdam Colloquium.
Thomas, William C (2021). "Toward a unified semantics for English either". In: Proceedings of Semantics and Linguistic Theory 32, pp. 446–465.

The Case for an Anchored CAUSE

Angelica Hill
University of Massachusetts Amherst
Amherst, U.S.A.
amhill@umass.edu

Abstract

In this paper I propose the meta-predicate CAUSE can be reduced to a circumstantial modal assuming an anchor semantics. To show this I focus on the semantics of strong overt causative expressions, in English, e.g., *made, forced, got*, and propose they are circumstantial modals with a syntactically explicit modal anchor. This proposal endorses a semantics in which modal components are independently provided by syntax, straightforwardly predicts the distinction between strong and weak overt causatives, eliminates vagueness from our semantic terminology, and accounts for the close relation between linguistic expressions that have received independent treatments in the literature.

1 Introduction

This paper introduces a project that explores the linking of syntax and semantics as a way to construct building blocks of modal meaning. Here I focus on strong overt causatives as they relate to circumstantial modal auxiliaries shown in (1).

(1) a. Jane made George go to the store. b. George had to go to the store.

The goal is to have a way of building causation in the syntax, which is not by any means a novel goal, but done in a way that relates modals and causatives explicitly in a framework that appreciates their syntactic and semantic similarity. From this proposal we predict the close relationship shared between the two expressions with regards to meaning, morphology, and processing, which is not straightforwardly achieved by assuming a meta-predicate CAUSE. Such a decomposed analysis of causation allows us to express causal relations uniformly as circumstantial modality and links differences in meaning to the difference between smaller and larger modals each with varying modal components. In Section 2 I present semantic properties of overt causatives and emphasize how their external argument contributes to their meaning. In Section 3 I highlight various similarities between overt causatives and circumstantial modals and present the basics of anchor semantics. It is also in this section that I extend anchor semantics to overt causatives and describe its advantages and predictions. Lastly, in Section 4 I discuss open questions which are the focus of ongoing and future work.

2 Overt Causatives

In English, overt causatives are verbs such as, *make, have, cause, force, get.*[1]

(2) Jane *made/had/caused/forced/got* George (to) go to the store.

In the classic analysis, overt causatives are analyzed as relating two events[2] in which the causer event is somehow responsible for bringing about the caused event. This causal relation

[1] Overt causatives have also been called 'periphrastic causatives' in the literature.

[2] The kind of relata that stand in a causal relation varies in the literature (e.g., events, situations, individuals, etc.). For an overview of proposals that differ in terms of causal relata see Bar-Asher Siegal and Boneh (2020).

is often evoked in the literature by the meta-predicate CAUSE, where 'e_1CAUSE e_2' refers to the relation whereby the event e_1 somehow brought about event e_2 (Dowty 1979; Levin and Rappaport Hovav 1995; Kratzer 2005; Pylkkänen 2000; Pylkkänen 2008, a.o.). CAUSE is generally understood to describe a counterfactual relation between the two events, following Lewis (1973), however, there have been arguments against CAUSE denoting counterfactual causation (Copley and Harley 2015; Copley, Wolff, and Shepard 2015; Talmy 1988; Wolff and Song 2003; Wolff 2007). A uniform analysis of overt causatives is challenging since there seems to be considerable variation with respect to ways in which the caused event is brought about by the causer event.

2.1 Flavor Variation

In the literature it's been reported that the ways in which a caused event is understood to come about varies (Lauer 2010; Luo 2024; Nadathur and Lauer 2020). I propose to classify these observations along three dimensions for organizational purposes. These dimensions can be considered kinds of causal "flavors," which are not meant to be exclusive. In fact, depending on the overt causative's subject, multiple flavors may be available.

A caused event can come about as the result of persuasion/ordering. For example, sentence (1-a) is true in a context in which Jane convinced, ordered, or pushed George to go to the store. However, a caused event can also come about by means of natural laws, such as in the *made*-sentence below.

(3) The flowers made Jane sneeze.

Sentence (3) is true in a context in which someone put roses on the table in front of extremely allergic Jane, the particles of which entered through her nose sparking a reaction. Additionally, overt causatives can also denote an instrumental use of its subject. Sentence (3) is true in a context in which miserable George, seeking any excuse to leave the party, tried to trigger an allergic reaction in Jane by putting the roses directly in front of her. In this context, the flowers are used as an instrument to achieve some goal. This instrumental interpretation holds of other overt causative forms, such as *got*.

(4) The drink got Jane to dance at the party.

This sentence is true in a context in which a drink has intentions and pulls Jane onto the dance floor (think: an adult cartoon), or one in which a drink was used as an instrument to achieve some goal, maybe by a friend to get miserable Jane dancing at the party. Again, the natural law interpretation is also available as was the case with the previous example.

Discussions concerning these different causal flavors (persuasion/ordering, natural laws, instrument/goal), have remained rather isolated in the literature. I would like to emphasize that the availability of a given flavor seems to be determined by the kind of external argument a overt causative takes. For example, an overt causative sentence with an animate external argument can be true in a context that involves persuasion or ordering (sentence (1-a)), whereas overt causative sentences with an inanimate external argument can be true in contexts in which natural laws are highlighted (sentence (3)) or the external argument is understood to be an instrument for an implicit goal (sentences (3) and (4)). It would be advantageous to have a semantics of overt causatives in which the subject and its properties derive the available causal flavors, thereby avoiding polysemous analyses of flavor variation.

2.2 Form Variation

It's known that overt causative forms have different lexical semantic restrictions (Baglini and Bar-Asher Siegal 2019; Copley, Wolff, and Shepard 2015; Doron 1999; Eckardt 2000; Maienborn

and Herdtfelder 2017; Nadathur and Lauer 2020; Neeleman and Van de Koot 2012; Reinhart 2002). For example, *got* requires animate external arguments with goals; if it lacks one, an instrumental interpretation is conveyed,

(5) George/The flowers got Jane to sneeze.

Had is also restricted to an animate subject with a goal. However, unlike *got*, an instrumental interpretation is not available.

(6) George/The drink had Jane entertain the guests.

The sentence with *the drink* as subject is only true in a context in which the drink is personified (again, think: adult cartoon). The reading where the drink is used as an instrument for a goal is not available as it was with *got* in sentence (4). Compare these to *forced*, which merely requires some resistance on the part of the argument in the caused event. Sentence (7-b) is true only if Jane is at least somewhat reluctant to cycle to Zandvoort.

(7) a. George/The wind forced the windows open.
b. George forced Jane to cycle to Zandvoort.

In this section I've shown that overt causatives come in a variety of flavors, and that they vary in terms of their lexical semantic restrictions. However, what I hope to have clearly demonstrated is that the subject puts restrictions on available flavors, and that lexicalization distinctions may reflect this subject/flavor landscape. A bi-eventive analysis, such as the one that makes use of CAUSE, does not straightforwardly offer insight into why the properties of the subject affect the available "flavors" of the causative. If we assume overt causatives are modals, an anchor-semantics analysis would provide a way to derive modal properties directly from the modal's arguments. I motivate this decision by highlighting several similarities between overt causatives and circumstantial modals. The main idea is that if we assume that overt causatives are circumstantial modals, then we can use the notion of modal anchors to build causation into the syntax just as we have used them to construct modal meanings in the syntax.

3 Overt Causatives as Circumstantial Modals

The link between modality and causatives is well-established in the literature (Ilič 2014; Koenig and Davis 2001; Martin and Schäfer 2012; Martin and Schäfer 2017; Palucci 2023; Privoznov 2023, a.o.). In a classic proposal for modality, circumstantial modals quantify over accessible possible situations projected from their modal base and embed a proposition that is true in the actual world (Kratzer 1977; Kratzer 1981; Kratzer 1991). More recently, circumstantial modals have been understood as the expressions we use when we engage in actuality-directed modal thought, i.e., the various ways we figure out what the actual world is like by considering and evaluating possible extensions of an accessible part of it (Phillips and Kratzer 2024).

We find that there are many similarities between circumstantial modal expressions and overt causatives. This isn't surprising given that both expressions refer to some set of facts in the actual world that necessitates the truth of a proposition. Similar to (circumstantial) modals, overt causatives vary in terms of force. There are strong overt causatives such as, *made, cause, forced,* and weak overt causatives such as *enable* and *let* (Wolff, Song, and Driscoll 2002; Wolff and Song 2003; Wolff 2007), and as is the case with modals this difference in force leads to entailment relations whereby strong overt causative sentences asymmetrically entail their weaker counterparts. There is the observation that some overt causative sentences asymmetrically entail circumstantial modal sentences. For example, sentence (1-b) is true in all the situations in which sentence (1-a) is true, though further work in required in order to determine if this observation

generalizes to other overt causative sentences with respect to their circumstantial modal auxiliary counterparts. Moreover, cross-linguistically we find linguistic expressions that are used to denote both circumstantial and overt causative meaning. For example, Privoznov (2023) argues that in Balkar the suffix *-al* is used to mark deontic, circumstantial, and causal necessity, and we find Mandarin's *yao* can be used for both circumstantial and causal necessity.

(8) a. George yao Jane mai qiaokeli
George **yao** Jane buy chocolate
'George made Jane buy chocolate.'

b. Jane yao mai qiaokeli
Jane **yao** buy chocolate
'Jane needed to buy chocolate.'

There is also psycholinguistic evidence that overt causative *made* primes the production of circumstantial modal *had to*, suggesting that the two are processed similarly and may share abstract representations which are accessed during priming (Hill and Momma 2024). Taken together, it is clear that overt causatives and circumstantial modals have some shared core. In the next section I review anchor semantics as it was proposed in Kratzer (2013) and Phillips and Kratzer (2024), and then explore extensions to the analysis of overt causative and reconstruct their interpretation as circumstantial modals with a syntactically explicit modal anchor.

3.1 Anchor Semantics for Modals

Modal predicates, such as *be fragile*, have been used to motivate the shift from a modal semantics in which a modal's domain of quantification is projected from its modal base, i.e., a set of contextually provided propositions, to an anchor semantics in which a modal's domain of quantification is projected from the modal's argument, e.g., an individual or an individual and its circumstances. To appreciate this shift, consider sentences below, originally presented in Lewis (1997), and discussed in Kratzer (2013) and Phillips and Kratzer (2024).

(9) a. The glass is fragile.
b. The glass can break easily.

On the face of it, the two sentences seem to mean more or less the same thing and are more or less interchangeable. However, we can find a context in which one is true and the other isn't (necessarily) true: the context of the diligent Sorcerer, introduced by Lewis.

> "A sorcerer takes a liking to a fragile glass...He does nothing at all to change the dispositional character of his glass. He only watches and waits, resolved that if ever his glass is struck, then, quick as a flash, he will cast a spell that changes the glass, renders it no longer fragile, and thereby aborts the process of breaking."
>
> Lewis (1997, p.147)

The general intuition[3] is that sentence (9-a) is true in the Sorcerer context, but sentence (9-b) doesn't have to be because there is an interpretation of the sentence in which although being-fragile is an intrinsic property of being a glass, the fragile glass will never break so long as the Sorcerer is present. That is, depending on how much of the situation is kept stable across the domain of quantification (i.e., just the glass, or the glass together with the fact that there is the diligent Sorcerer present), one could judge sentence (9-b) as either true or false.

Appealing to an anchor semantics allows us to link the syntactic differences between modals to explain semantic differences between modals in a way that a classic analysis which relies on

[3]I say "general intuition" because judgments about these sentences are not unanimous. For example, some speakers prefer different tense on the modal auxiliary (e.g., *The glass could break easily*). Though these observations deserve further investigation, they don't directly affect the argument of this paper, which only relies on there being a truth conditional distinction between the modal predicate and the modal auxiliary.

modal bases for domain projection, doesn't. Kratzer (2013) and Phillips and Kratzer (2024) propose an anchor semantics for modal predicates and modal auxiliaries, though the idea of a modal anchor is traced back to Arregui (2005), Arregui (2007), Arregui (2009), and Hacquard (2006). The framework requires richer formal models than possible worlds, and so makes use of possible situations, which are linked by mereological part-of relations where worlds are assumed to be maximal situations. (see Kratzer (2021) for an overview of situation semantics for natural language). In the framework, a modal's domain of quantification is anchored to and therefore projected from its argument (the Modal Anchor Hypothesis (Kratzer 2013; Phillips and Kratzer 2024)). The denotation for (9-b) is the following,

(10) $[\![\textbf{The glass can break easily}]\!]^{f,s_0} = \exists w \in f_{fact}(s_0) \wedge N(w) \wedge \exists s'[s' \leq w \wedge \textsc{match}(s_0, s') \wedge$ the glass breaks easily in $s']$

f_{fact} is the function responsible for factual domain projection – it creates a domain for which every modal alternative has a counterpart to the anchor given by the topic situation; $N(w_i)$ refers to normality restrictions which govern world w_i (e.g., natural laws, stereotypical behavior); $\leq$ denotes the part-of relation between situations where worlds are maximal situations; and $\textsc{match}(s_i, s_n)$ maps similarity relations between situations via counterparts (Lewis 1986). Sentence (9-b) is true if and only if there is a normal world in the factual domain projection of the topic situation s_0, such that there exists an accessible possible situation in which there is a counterpart of the topic situation and the glass breaks in that accessible possible situation. The truth conditions for sentence (9-a) will be identical to those in (10), except instead of the topic situation as the modal anchor, a time-slice of the subject, i.e., the glass at the topic situation will be the anchor. Using modal anchors allows us to account for the truth conditional difference between sentences (9-a) and (9-b) since there is flexibility with respect to the anchor situation. For (9-b), if we take the topic situation as containing the Sorcerer, the sentence will come out false, if we restrict the topic situation to just the glass, the sentence is true. With sentence (9-a), the anchor situation is restricted to just the glass, which means so long as the accessible counterpart situation matches the time-slice of the glass in the topic situation the sentence is true, even if there is also a diligent Sorcerer, it's raining outside, the glass is on the table, etc.

3.2 Anchor Semantics for Causatives

I propose that overt causatives are the phonological spell-out of a circumstantial modal with a syntactically explicit modal anchor. Consider the LF and truth conditions for sentence (1-b). The embedded clause introduces a situation argument, λs_1, which eventually will co-refer with the topic situation, s_0.

(11) a. [**had to** λs_1 [**George** [**go-store**]]]
b. $[\![(1\text{-b})]\!]^{f,s_0} = \forall w \in f_{fact}(s_0) \wedge N(w) \wedge \exists s'[s' \leq w \wedge \textsc{match}(s_0, s') \wedge$ George goes to the store in $s']$

The truth conditions for the overt causative sentence (1-a) will be identical to those for (1-b) except the anchor situation will be a past time-slice of Jane picked out by the topic situation, which can include her desires, her commands, any actions she's doing, etc.

(12) a. [**Jane**$_{s_1}$ **made** λs_1 [**George** [**go-store**]]]
b. $[\![(1\text{-a})]\!]^{f,s_0} = \forall w \in f_{fact}(\text{Jane}_{s_0}) \wedge N(w) \wedge \exists s'[s' \leq w \wedge \textsc{match}(\text{Jane}_{s_0}, s') \wedge$ George goes to the store in $s']$

Sentence (1-a) is true if and only if in all possible normal worlds in the factual domain projection of Jane as picked out by the topic situation (which can include laws such as George's tendency to

obey Jane's wishes), there exists a possible situation in which there is an accessible counterpart of that time-slice of Jane and in that possible situation George goes to the store. Intuitively, the meaning of a strong overt causative sentence is that given whatever facts pertain to the causative's subject in the topic situation, they are responsible for the argument of the resultant situation having to do something and actually doing it. This follows straightforwardly from the semantics since *made* universally quantifies over worlds in which there is a situation that matches the anchor situation in the actual world, meaning that the proposition, *George goes to the store,* will necessarily be true in the actual world.

There are a number of consequences and predictions that follow from this proposal. One result, which isn't unique to this proposal but any quantificational modal account of overt causatives, is that we can attribute the force difference between strong and weak overt causatives to the difference in modal necessity and possibility. An advantage of anchors is that the kind of external argument the causative takes will affect the kind of facts that hold in the topic situation and thus the kind of relation that holds between the anchor and resultant situation. Furthermore, since overt causatives are analyzed as essentially circumstantial modal auxiliaries with something extra, run-of-the-mill circumstantial modals such as *had to*, which always have an implicit modal anchor, will also always have an implicit cause. This seems to be true. Out of the blue, *George had to go to the store*, because of something, be it his hunger, or Jane's desire, etc. Moreover, given this consequence we predict that circumstantial modals such as *had to* should also be sensitive to the same kind of causal flavors available to overt causatives. Again, this seems to be the case. *The vase had to fall,* can be true according to natural physical laws. For example, given the angle and force of Jane's elbow with respect to the weight and stability of the vase, there was no way the vase couldn't have fallen, but a persuasion/ordering relation is only available if we consider a contrived, personified, context. Lastly, given the shared structure that results from overt causatives containing a circumstantial modal core, we predict that cross-linguistically there will be overlap between morphology that is used to denote circumstantial, and therefore causal, necessity, which seems to be the case in Balkar. Moreover, we predict that it's possible for a language to use the same lexical entry, such as Mandarin's *yao*, to convey circumstantial modality either with an explicit modal anchor or not.

4 Towards a Complete Framework

I've sketched out a way to extend a proposal that ties modal interpretations to the syntax-semantics interface in the case of overt causatives. This proposal makes claims about the "building blocks" of causative meanings in line with the building blocks of modal interpretations. The extension accounts for the cross-linguistic, semantic, and processing similarities between causatives and circumstantial modals. A main selling-point of anchor semantics is the role the syntax-semantics interface plays in providing modal components, yet as it currently stands it's unclear how this interface unfolds. For example, how does the denotation for modal predicates compose with structure such as Voice and Aspect? Or how exactly are anchors composed, i.e., how do we construct a time-slice of the external argument? How can we integrate features such as animacy into our derivations? How do the arguments of embedded clauses also affect available flavors as is the case with, *The vase had to fall*? Accounting for the interface is crucial for then extending the framework to overt causatives. Yet there are still loose ends even with the proposal as I've sketched it here. I point out that the fact that overt causative sentences entail actuality follows directly from the fact we are universally quantifying of counterparts of the actual world, meaning the actual world will necessarily be included in that set. Of course, if overt causatives and circumstantial modals share similar truth conditions, a question remains as to why circumstantial necessity modal auxiliaries such as *had to* do not entail actuality. These open questions and (not so small) details are the focus of ongoing and future work.

Acknowledgments. I am indebted to my supervisors Ana Arregui and Shota Momma for all their guidance. The reviewers for the Amsterdam Colloquium provided invaluable feedback, for which I am grateful. A special thanks to my Mandarin consultants, Jia Ren and Suet-Ying Lam. I'd also like the thank the people with whom I've discussed this project at its various stages, specifically, Faruk Akkuş, Valentine Hacquard, Vincent Homer, Kyle Johnson, Eva Neu, Guillermo del Pinal, Giorgio Sbardolini, and the attendees of UMass' Semantics Workshop.

References

Arregui, Ana (2005). *On the Accessiblity of Possible Worlds: The Role of Tense and Aspect*. University of Massachusetts Amherst.

— (2007). "When aspect matters: The case of would-conditionals". In: *Natural Language Semantics* 15, pp. 221–264.

— (2009). "On similarity in counterfactuals". In: *Linguistics and Philosophy* 32, pp. 245–278.

Baglini, Rebekah and Elitzur Bar-Asher Siegal (2019). "Direct causation: A new approach to an old question". In: *U. Penn Working Papers in Linguistics 26*. University of Pennsylvania.

Bar-Asher Siegal, Elitzur and Nora Boneh (2020). "Causation: From metaphysics to semantics and back". In: *Perspectives on causation: Selected papers from the jerusalem 2017 workshop*. Springer, pp. 3–51.

Copley, Bridget and Heidi Harley (2015). "A force-theoretic framework for event structure". In: *Linguistics and Philosophy* 38.2, pp. 103–158.

Copley, Bridget, Phillip Wolff, and Jason Shepard (2015). "Force interaction in the expression of causation". In: *Semantics and Linguistic Theory*, pp. 433–451.

Doron, Edit (1999). "The semantics of transitivity alternations". In: *Proceedings of the twelfth Amsterdam colloquium*, pp. 103–108.

Dowty, David R (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Eckardt, Regine (2000). "Causation, contexts, and event individuation". In: *Speaking of Events*. NewYork/Oxford: Oxford University Press, pp. 105–121.

Hacquard, Valentine (2006). "Aspects of modality". Dissertation. PhD thesis. M.I.T.

Hill, Angelica and Shota Momma (2024). "Priming abstract modal mechanisms in modals with causatives". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 46.

Ilič, Tatjana (2014). "Modality and causation: Two sides of the same coin". In: *Causation in Grammatical Structures*. Ed. by Bridget Copley and Fabienne Martin. Oxford Studies in Theoretical Linguistics. Oxford University Press, pp. 72–86.

Koenig, Jean-Pierre and Anthony R Davis (2001). "Sublexical modality and the structure of lexical semantic representations". In: *Linguistics and Philosophy* 24, pp. 71–124.

Kratzer, Angelika (1977). "What 'must'and 'can'must and can mean". In: *Linguistics and philosophy* 1, pp. 337–355.

— (1981). "The notional category of modality". In: *Words, worlds, and contexts* 38, p. 74.

— (1991). "Modality". In: *Semantik*, pp. 639–650.

— (2005). "Building resultatives". In: *Event arguments: Foundations and applications* 177, p. 212.

— (2013). "Modality for the 21st century". In: *19th International Congress of Linguists*, pp. 181–201.

— (2021). "Situations in Natural Language Semantics". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2023. Metaphysics Research Lab, Stanford University.

Lauer, Sven (2010). "Periphrastic causative verbs in English: What do they mean". In: *Manuscript, Stanford University*.

Levin, Beth and Malka Rappaport Hovav (1995). *Unaccusativity*. Cambridge MA: MIT Press.

Lewis, David (1973). "Causation". In: *The journal of philosophy* 70.17, pp. 556–567.
— (1986). *On the plurality of worlds*. Oxford, Blackwell Publishing.
— (1997). "Finkish dispositions". In: *The Philosophical Quarterly* 47.187, pp. 143–158.
Luo, Zhuosi (2024). "Causality and modality: A case study on Teochew periphrastic causatives". In: *Proceedings of the Linguistic Society of America* 9.1, pp. 5650–5650.
Maienborn, Claudia and Johanna Herdtfelder (2017). "Eventive versus stative causation: the case of German causal von-modifiers". In: *Linguistics and Philosophy* 40, pp. 279–320.
Martin, Fabienne and Florian Schäfer (2012). "The modality of offer and other defeasible causative verbs". In: *West Coast Conference on Formal Linguistics (WCCFL)*. Vol. 30. Citeseer, pp. 248–258.
— (2017). "Sublexical modality in defeasible causative verbs". In: *Modality across syntactic categories*. Ed. by Ana Arregui, María Luisa Rivero, and Andrés Salanova. Vol. 63. Oxford University Press, p. 87.
Nadathur, Prerna and Sven Lauer (2020). "Causal necessity, causal sufficiency, and the implications of causative verbs". In: *Glossa: A journal of general linguistics* 5.1.
Neeleman, Ad and Hans Van de Koot (2012). "The linguistic expression of causation". In: *The theta system: Argument structure at the interface* 20, pp. 20–51.
Palucci, Jonathan (2023). "Modal ingredients of causative *have*". In: *Proceedings of the 53rd North East Linguistics Society*. Vol. 53.
Phillips, Jonathan and Angelika Kratzer (2024). "Decomposing modal thought." In: *Psychological Review* 131.4, p. 966.
Privoznov, Dmitry (2023). "On Actuality Entailments, Causation, and Telicity in Balkar". In: *Languages* 8.3, p. 178.
Pylkkänen, Liina (2000). "Representing causatives". In: *Semantics and Linguistic Theory*. Vol. 10, pp. 132–148.
— (2008). *Introducing arguments*. Vol. 49. MIT press.
Reinhart, Tanya (2002). "The theta system–an overview". In: *Theoretical linguistics* 28.3, pp. 229–290.
Talmy, Leonard (1988). "Force dynamics in language and cognition". In: *Cognitive science* 12.1, pp. 49–100.
Wolff, Phillip (2007). "Representing causation." In: *Journal of experimental psychology: General* 136.1, p. 82.
Wolff, Phillip and Grace Song (2003). "Models of causation and the semantics of causal verbs". In: *Cognitive psychology* 47.3, pp. 276–332.
Wolff, Phillip, Grace Song, and David Driscoll (2002). "Models of causation and causal verbs". In: *Papers from the 37th meeting of the Chicago Linguistics Society, Main session*. Vol. 1. Chicago Linguistics Society Chicago, pp. 607–622.

A Pragma-Semantic Account for Negative Island Obviation by *Wa* in Japanese

Hitomi Hirayama
Keio University
Kanagawa, Japan
hhirayam@keio.jp

Abstract

This paper provides a pragma-semantic account for negative island obviation by *wa* in Japanese. I propose that negative island obviation by *wa* is achieved by special discourse effects carried by this particle. In particular, the particle *wa* loosens the requirement of questions that there should be a maximally informative answer. The proposal does not rely on dense scales and suggests that discourse effects of a particle could play a significant role in the interface between semantics and pragmatics.

1 Introduction

Negative islands and obviation effects have revealed how informativeness is encoded in languages and how we calculate the measurement (Abrusán and Spector 2011; Abrusán 2014; Fox and Hackl 2006). In Japanese, a particle *wa* makes obviation possible, as reported by Schwarz and Shimoyama (2010) and shown in (1).

(1) Taro-wa doredake nagaku-*(wa) doitu-ni taizaisi nak-atta no?
Taro-TOP how long-*wa* Germany-in stay NEG-PAST Q
'How long-*(wa) did Taro not stay in Germany?'

This paper gives a pragma-semantic analysis of negative island obviation by *wa* in Japanese, combining the interval semantics by Abrusán and Spector (2011) and special discourse effects of the particle *wa*. It is revealed that special discourse effects associated with certain marked questions play important roles at the interface of semantics and pragmatics. Furthermore, unlike Schwarz and Shimoyama's (2010) analysis, it is not necessary to assume dense scales.

The rest of paper is organized as follows. Section 2 provides background on negative island obviation focusing on how the interval semantics can derive obviation in an intricate context. Section 3 illustrates the problems on negative degree questions in Japanese. The pragma-semantic analysis is provided in Section 4. In Section 5, I demonstrate how the proposed analysis works and compare it with the analysis in the literature. Section 6 concludes the paper and consider theoretical implications of the analysis.

2 Background

Degree questions with negation are infelicitous with the amount reading (Heycock 1995). For example, (2) is interpretable as a question that asks the cardinality of the (contextually available) set of books John did not buy. This cannot be interpreted as a question that asks for n such that John did not buy n books.

(2) How many books did John not buy?

Syntactic (Rizzi 1990), pragmatic (Kroch 1989), and semantic analyses have been made to explain why this is the case. In semantic analyses, there also have been discussions whether this phenomenon tells anything about the scale we use in computing the meaning (Fox and Hackl 2006).

Among different semantic approaches, I adopt the analysis by Abrusán and Spector (2011) as the baseline of my analysis. In this analysis, degrees are treated as intervals. Combining the interval semantics and maximal informativity principle (MIP), shown in (4), their analysis correctly predicts that questions such as (3) are infelicitous in an out-of-the-blue context but can be rescued under a certain context.

(3) How many rabbits does none of these women have?
   # Context 1: There are nine women. Five women have five rabbits.
   Four women have nine.
   ▷ Possible intervals that can be true answers: $[1,4]$, $[6,9[$, $[7,7]$, $[10,+\infty[$
   ✓ Context 2: There are nine women.
   For all $0 < n < 10$, one woman has exactly n rabbits.
   ▷ Possible intervals that can be true answers: $[10,16]$, $[12,12]$, $[10,+\infty[$

(4) Maximal Informativity Principle (MIP): A question is supposed to have a maximally informative answer, which is a true answer that entails all the other possible answers.

When there is no special context, (3) is judged infelicitous because there cannot be an interval that contains all the other intervals that would give rise to a true answer. As an example of a context that is not special and close to an out-of-the-blue context, let us imagine Context 1. In this context, there cannot be an interval that includes all true answers. For instance, $[10,+\infty[$ does not include $[7,7]$, which can also be a possible true answer.

By contrast, when an intricate context such as Context 2 is available, (3) is felicitous. What is special about Context 2 is that for every number n such that $0 < n < 10$, there is one woman such that she has exactly n rabbit. Such a situation is not readily available in a normal context. Nevertheless, as long as such a context is available, (3) can get the answer 10. This is possible because we can get an interval $[10,+\infty[$, which contains all the other intervals that would give rise to a true answer.

Note that the process of deriving the (in)felicity of the questions does not rely on dense scales. The scale is assumed to be discrete in the example above because the number of entities is under discussion. Even with discrete scales, the interval semantics provides us with the same results and predictions.

3 Problem

In principle, the semantic account outlined above would predict that similar effects would be observed in other languages. However, Japanese shows a different picture; even under Context 2, which can ameliorate the English question, the Japanese counterparts (5a-b) are infelicitous.[1] The questions sound equally odd across contexts. One way to make obviation possible is by adding *wa* to the *wh*-phrase, as observed in Schwarz and Shimoyama (2010). A notable feature of these questions is that they do not require complicated context. The possible answer refers to a number that exceeds the maximum number of rabbits owned by at least one woman. In other words, in both Context 1 and 2, the answer would be 10. Furthermore, the answer to this question also has to be marked with *wa*, as shown in (6).

[1] I appreciate an anonymous reviewer suggesting more appropriate Japanese counterparts of English sentences. The original version used *dono hito-mo*, which could be parsed differently from what is intended here.

(5) "How many rabbits does nobody have?"[2]

a. Dare-mo usagi-o nan-biki-#(wa) kattei-nai no?
every person rabbit-ACC what-CL-WA have-NEG Q

b. Dare-mo nan-biki-#(wa) usagi-o kattei-nai no?
every person what-CL-WA rabbit-ACC have-NEG Q

(6) Zip-piki-#(wa) kattei-nai.
ten-CL-WA have-NEG
'(Every person) does not have 10-*wa* (rabbits).'

In Schwarz and Shimoyama (2010), negative island obviation by *wa* is explained by assigning *at least* semantics to *wa* while giving *exactly* semantics to numerals. Without *wa*, it is not possible to get the maximally informative answer because the possible answers are neither downward nor upward scalar due to the *exactly* semantics. By contrast, adding *wa* makes the possible answers aligned by semantic strength, and consequently, there is a maximally informative answer available. However, if changing the *exactly* semantics is the key, not only adding *wa* but also adding a lexical item meaning *at least* (*sukunakutomo* or *izyoo* in Japanese) to the numeral should be able to rescue sentences with negative islands. However, this is not borne out, as shown by (7).

(7) Adding *at least* cannot rescue negative island questions

a. # Dare-mo usagi-o sukunakutomo nan-biki kattei-nai no?
every person rabbit-ACC at least what-CL have-NEG Q

b. # Dare-mo usagi-o nan-biki-izyoo kattei-nai no?
every person rabbit-ACC what-CL-more than have-NEG Q

The above examples show that aligning alternatives cannot make obviation possible. In other words, using this particular particle, *wa* is necessary for obviation. As a result, it is necessary to say that something special about *wa* is the key.

4 Proposal

I argue that the presupposition of *wa* plays an important role in the obviation effect. In particular, I analyze *wa*-questions as a marked question with special discourse effects. The interaction between the presupposition of questions — namely, the Maximal Informativity Principle (MIP)

[2]The two Japanese examples both use floating quantifiers, and there is a reason behind this. There are at least four ways to express numeral+noun such as *how many rabbits* in Japanese (Watanabe 2006). Out of the four, two are ambiguous with respect to specificity. The other two are unambiguously non-specific, and the examples used in (5) are categorized into this type. It needs further investigation whether this specificity correlates with referentiality, but the amount reading under discussion here should go along with non-specific noun phrases. The point is the intricate context cannot rescue the negative island obviation regardless.

Furthermore, as an anonymous reviewer pointed out, it is possible to place numerals prenominally, as shown in (i). This noun phrase (or DP) is ambiguous between specific and non-specific, so in principle, it is predicted to behave in the same way as (5). With an accusative marker *-o*, the question is infelicitous even under Context 2. However, to my ears, a version with *wa* does not sound as good as (5a-b) do.

(i) Dare-mo nan-biki-no usagi-#o/??wa kattei-nai no?
every person what-CL-GEN rabbit-ACC/WA have-NEG Q

At this point, I do not know why this is the case. *Wa* does not have to be associated with an adjacent lexical item, so if *nanbiki* 'how many' bears focus, there should be no problem with it getting an association with *wa*. For completeness, *usagi nan-biki-o* can express *how many rabbits* in the object position. This is infelicitous with a negative island as the other three patterns. However, *wa* cannot rescue this pattern; With *wa*, the sentence is ungrammatical regardless of whether there is a negative island.

(Dayal 1996) — and discourse effects of contrastive *wa* enables the question to yield interpretable answers.

4.1 Special Discourse Effects of *Wa*

First, *wa* in (5) is treated as contrastive *wa* (Kuno 1973) because it is attached to a *wh*-phrase and hence cannot be a topic *wa*. The intuition shared by the analyses of contrastive *wa* (Hara 2006; Tomioka 2009; Yabushita 2017) is that this *wa* utilizes an alternative set derived by focus (Rooth 1985) in the discourse. For example, an assertion with contrastive *wa* (8) can bring about an ignorance inference. This inference can be obtained from the fact that the speaker did not use stronger alternatives such as *Taro and Jiro* or *Taro, Jiro, and Saburo*.

When *wa* is used in a polar question (9), evoking a set of alternatives brings about an inference that other questions can be asked. In other words, there is an overarching question that resolves the question under discussion in the immediate context (i.e., *Who came?*). As a result, it is possible to infer that the speaker asks (9) because they expect that the addressee can at least answer the question even though it just partly resolves the issue.

(8) [$_F$Taroo-wa] kita.
Taro-WA came.
'Taro came.'
⇝ The speaker is not sure about the others.

(9) [$_F$Taroo-wa] kita?
Taro-WA came.
'Did Taro come?'
⇝ The speaker assumes the addressee at least knows whether Taro came.

How does *wa* work with constituent questions such as those in (5)? It is true that *wh*-questions with *wa* are marked and rare. However, as Hirayama (2019, p. 120) notes, there does not have to be a contrasted question available in the context as claimed by Miyagawa (1987). When *wa* is used with a *wh*-phrase, in the same way as (9), it projects other questions that can be asked in the immediate discourse. For example, when the speaker utters (10), they signal that there could be other questions about the same domain, which is the domain of individuals in this case. The possible questions are given after ⇝ in (10).

(10) [$_F$Dare-wa] kita?
Taro-WA came.
'Who come?'
⇝ The speaker might want to know who did not come
⇝ The speaker might want to know who the addressee is not sure but might have come...

The second possible question illustrated in (10) shows that the speaker is loosening a normal assumption of questioning acts, which is the speaker expects the addressee to be fully knowledgeable and give a maximally informative answer. In other words, the expected answer can be approximate and does not have to be exact. When *wa* is used in this way, I argue that *wa* is playing a role as a discourse effect modifier, and the question itself has special discourse effects as marked questions (Farkas and Roelofsen 2017). The special discourse effect provided by *wa* is loosening the MIP given in (4).

4.2 The Interaction of *Wa* and the MIP

Given the discourse effects of *wa*, let us combine them with degree questions. When *wa* is used in the degree *wh*-question, as it does with *who*-questions, it signals that the speaker does not expect the maximally informative answer. The overall effects of degree questions with *wa* are illustrated in (11).

(11) Discourse effects of degree questions with *wa*
Degree questions with *wa* (Q_{DegWa}) presuppose that for every world w compatible with

common knowledge, there is an interval I_{Ans} that corresponds to the answer to Q_{DegWa} such that

(i) I_{Ans} gives rise to a possible true answer to the question in w

(ii) I_{Ans} specifies the interval $[m, n]$ such that either of its complement intervals gives rise to true answers.

(i) is the presupposition that degree questions have. The effect of *wa* is illustrated in (ii). Why do we have a complication about the complements? As stated earlier, when *wa* is used, the speaker signals that they can ask questions on the same domain, which is intervals with a degree question. What is different between the *who*-questions we saw earlier and degree questions is that the alternatives have a semantic relationship with one another. It is this semantic relation that limits possible questions that the speaker can ask in the subsequent discourse.

Let us look at a simpler example without negation. (12) is a felicitous question regardless of whether there is *wa*.

(12) Taro-wa usagi-o nan-biki-(wa) kattei-ru no?
Taro-TOP rabbit-ACC what-CL-WA have-PRES Q
'How many rabbits-(wa) does Taro have?'

Without *wa*, there is no special discourse effect, and the addressee is expected to give a maximally informative answer to this question. If the addressee answers by saying "five," that is interpreted as the maximally informative answer. By contrast, when the speaker uses *wa*, they signal that they do not expect the maximally informative answer and it is possible for them to ask follow-up questions if necessary. That means the answer does not have to be exact; it can be approximate. For example, if the addressee answers by saying *go-hiki-wa kattei-ru* 'He has five-wa,' the speaker can follow up by referring to alternatives in the same domain, intervals. In doing so, not all intervals are targeted because of informativeness. For instance, the interval $[4, 4]$ would not be targeted because the answer by the addressee entails its truth. Only non-trivial intervals are targeted. From the speaker's perspective, the degree question with *wa* is used to understand how they should partition the intervals to ask further questions only if necessary. To put it differently, by uttering Q_{DegWa}, the speaker signals that there is at least a threshold that the hearer can use to give an answer to the question.

5 Discussion

Let us see how the proposal accounts for the negative island obviation. Remember that an intricate context is needed to rescue negative island questions in English because otherwise, the maximally informative answer does not exist. That is, with a negation, there is normally no interval that includes all the intervals that can be a true answer. In Context 2, however, the complicated context can guarantee that the addressee can find a maximally informative answer, and the question is felicitous in the end.

In Japanese, as proposed earlier, Q_{DegWa} ask for the interval that is the upper or lower bound of the intervals that give rise to true answers. In Context 1, for instance, one possible answer is $[10, 10]$. In this context, the interval $[6, 6]$, for example, cannot be I_{Ans} since neither of its complements can be a set of possible answers because of $[5, 5]$ and $[9, 9]$. As a result, this question asks the number that exceeds the largest number of rabbits somebody has, which accords with the interpretation of Q_{DegWa} reported by Schwarz and Shimoyama (2010). The discourse effects give the same answer under Context 2. Under this context, too, $[10, 10]$ is the threshold that partitions true answers and false answers. In sum, regardless of whether there is intricate context, by using *wa* the speaker can draw some contextually sufficient answer from

the addressee. This explains why the addition of *wa* can obviate a negative island. Note that from the addressee's side, it is necessary to acknowledge that they are responding to non-default questions. This is why the answer to *wa* questions must also have *wa*. By using *wa* in the answer, the addressee indicates an answer to a non-canonical question and the existence of the alternative set evoked in the discourse.

The proposed mechanism of obviation relies on presuppositions or discourse effects of the particular particle, *wa*. This explains why adding *izyoo* and *sukunakutomo*, which roughly mean *at least* in English, to the *wh*-degree phrase cannot do the same trick in Japanese. These lexical items do not carry a special presupposition or discourse effects that weaken the default assumptions of questioning acts: There is a maximally informative answer. In other words, the question with *at least* still requires the maximally informative answer, which does not exist in Context 1.

6 Conclusion and Implication

The proposed analysis does not explain why questions without *wa* are infelicitous even under Context 2, which can ameliorate the English counterpart. I argue that this is because of pragmatic competition between unmarked questions and marked ones. Because there is a tool that can be used to weaken the MIP in Japanese, not using the tool, namely the absence of *wa*, would suggest that the hearer should be able to figure out an answer without special assumptions about the context.

Unlike Schwarz and Shimoyama's (2010) analysis, the proposed analysis on negative island obviation avoids reliance on dense scales. Further research is necessary to determine whether Japanese differs from English with respect to the scales used in computations, but obviation by *wa* is also observed in other negative island environments without degrees, as shown in (13).

(13) Taro-wa donna-huu-ni-*(wa) eego-o hanasemasen ka?
Taro-TOP which-way-in-*wa* English-ACC cannot speak Q
'How can Taro not speak English?' [Hirayama (2019, p. 143)]

There is no evident interval or scale in (13), but a possible answer to (13) is '*He cannot speak like a native speaker*,' which is based on a scale of proficiency. Considering such a scale is likely to be discrete, at least in Japanese, it is not necessary to assume scales are dense to account for (in)felicity of questions with negative islands.

In this paper, I analyzed *wa*-questions as marked questions equipped with special discourse effects. Japanese is known to have a variety of biased questions (Sudo 2013; Ito and Oshima 2014). By nature, biased questions are polar questions; It is not possible to have some bias with *wh*-questions. The negative island obviation by *wa* suggests that Japanese has special *wh*-questions, in addition to biased questions. Special *wh*-questions and their pragmatic effects have been investigated in Romanian (Farkas 2022) and German (Eckardt 2020). Japanese *wa* questions are an interesting additional data point to cross-linguistic research of special questions.

Acknowledgements. I am grateful Yu Tomita and three anonymous AC 2024 reviewers for their helpful comments. This study is supported by JSPS KAKENHI Grant Number 21K12985 (PI: Hitomi Hirayama).

References

Abrusán, Márta (2014). *Weak Island Semantics*. Oxford University Press.

Abrusán, Márta and Benjamin Spector (2011). "A Semantics for Degree Questions Based on Intervals: Negative Islands and Their Obviation". In: *Journal of Semantics* 28, pp. 107–147.

Dayal, Veneeta (1996). *Locality in wh quantification*. Dordrecht: Kluwer.

Eckardt, Regine (2020). "Conjectual questions: The case of German verb-final *wohl*-quesitons". In: *Semantics & Pragmatics* 13.9, pp. 1–54.

Farkas, Donka F. (Feb. 2022). "Non-Intrusive Questions as a Special Type of Non-Canonical Questions". In: *Journal of Semantics* 39.2, pp. 295–337. ISSN: 0167-5133. DOI: 10.1093/jos/ffac001. eprint: https://academic.oup.com/jos/article-pdf/39/2/295/43661506/ffac001.pdf. URL: https://doi.org/10.1093/jos/ffac001.

Farkas, Donka F. and Floris Roelofsen (2017). "Division of labor in the interpretation of declaratives and interrogatives". In: *Journal of Semantics* 34.2, pp. 237–289.

Fox, Danny and Martin Hackl (2006). "The universal density of measurement". In: *Linguistics and Philosophy* 29, pp. 537–586.

Hara, Yurie (2006). "Japanese Discourse Items at Interfaces". PhD thesis. University of Delaware.

Heycock, Caroline (1995). "Asymmetries in Reconstruction". In: *Linguistic Inquiry* 26.4, pp. 547–570.

Hirayama, Hitomi (2019). "Asking and Answering Questions: Discourse Strategies in Japanese". PhD thesis. University of California, Santa Cruz.

Ito, Satoshi and David Y. Oshima (2014). "On Two Varieties of Negative Polar Interrogatives in Japanese". In: *Japanese/Korean Linguistics 23*. Ed. by Michael Kenstowicz, Ted Levin, and Ryo Masuda.

Kroch, Anthony (1989). "Amount Quantification, Referentiality, and Long Wh-Movement". In: *University of Pennsylvania Working Papers in Linguistics* 5.2.

Kuno, Susumu (1973). *The Structure of the Japanese Language*. Cambridge: MIT Press.

Miyagawa, Shigeru (1987). "*Wa* and the WH phrase". In: *Perspectives on Topicalization: The case of Japanese wa*. Ed. by John Hinds, Shoichi Iwasaki, and Senko K. Maynard. John Benjamins, pp. 185–217.

Rizzi, Luigi (1990). *Relativized minimality*. Cambridge: MIT Press.

Rooth, Mats (1985). "Association with Focus". PhD thesis. University of Massachusetts, Amherst.

Schwarz, Bernhard and Junko Shimoyama (2010). "Negative Islands and Obviation by *Wa* in Japanese Degree Questions". In: *Proceedings of SALT* 20, pp. 1–18.

Sudo, Yasutada (2013). "Biased Polar questions in English and Japanese". In: *Beyond Expressives: Explorations in Use-Conditional Meaning*. Ed. by Daniel Gutzmann and Hans-Martin Gaertner. Vol. Current Research in the Semantics/Pragmatics Interface (CRiSPI) 28. Brill.

Tomioka, Satoshi (2009). "Contrastive topics operate on speech acts". In: *Information Structure: Theoretical, Typological, and Experimental Perspectives*. Ed. by Malte Zimmermann and Caroline Fèry. Oxford University Press, pp. 115–138.

Watanabe, Akira (2006). "Functional Projections of Nominals in Japanese: Syntax of Classifiers". In: *Natural Language and Linguistic Theory* 24, pp. 241–306.

Yabushita, Katsuhiko (2017). "Partition Semantics and Pragmatics of Contrastive Topic". In: *Contrastiveness in Information Structure, Alternatives and Scalar Implicatures*. Ed. by Chungmin Lee, Ferenc Kiefer, and Manfred Krifka. Vol. 91. Springer, pp. 23–45.

Same Type of Comparatives, But with Different Syntax and Semantics

Yaqing Hu
University of Southern California
yaqinghu@usc.edu

Abstract

Two aspects of Chinese adjectival *bi*-comparatives (*bi*-comparatives) remain underexplored. First, a critical morphosyntactic difference between bare (without differential phrases) and differential *bi*-comparatives (with differential phrases) is that the morpheme *chu*, literally meaning 'beyond/exceed', is exclusively licensed in the latter, but not in the former. Second, unlike their English counterparts, *bi*-comparatives do not freely allow measure phrases (MPs) as the standard of comparison. These observations suggest that degrees may be accessible in some constructions, but not in others. I propose that (i) bare *bi*-comparatives do not characterize an ordering of degrees, but a directed scale segment (Schwarzschild 2020); and (ii) differential *bi*-comparatives also characterize a directed segment, but it allows the mapping of the segment onto a degree specified by the differential phrase, a role fulfilled by *chu* (à la Wellwood 2015). Taken together, Chinese adjectival *bi*-comparatives may constitute a case where degrees are not encoded in the lexical semantics of gradable adjectives (GAs), but introduced via a functional morpheme (Wellwood 2015; Bochnak et al. 2020).

1 Introduction

One prominent way of constructing comparatives in Chinese involves using the morpheme *bi*, which typically signifies the comparative nature of such constructions, as schematized in (1).

(1) X *bi* Y Dimension (Differential Phrase)

Descriptively, X represents items being compared and Y represents the standard of comparison, with the dimension of comparison indicated by either gradable adjectives (GAs) or verbs. On the surface, differential phrases are generally optional. This paper will specifically focus on *bi*-comparatives that utilize GAs as the main predicate, such as (2).

(2) a. ZS bi LS **gao**.
ZS BI LS tall
'ZS is taller than LS.'

b. ZS bi LS **gao** 5-gongfen.
ZS BI LS tall 5-cm
'ZS is 5 cm taller than LS.'

I designate (2a) as '**bare *bi*-comparatives**', referring to comparative constructions without differential phrases, and (2b) as '**differential *bi*-comparatives**', which are comparative constructions that include a differential phrase indicating the extent of difference between the two compared items along a specific dimension.

Two aspects of Chinese adjectival *bi*-comparatives remain underexplored in the current literature. First, canonically, both bare and differential *bi*-comparatives do not freely allow an MP to be the standard of comparison, as in (3b) and (4b). Cross-linguistically, Chinese, like its counterpart English, is categorized as having a positive setting for the 'Degree Semantics Parameter' (DSP), which posits that 'a language {does/does not} have gradable predicate (type $\langle d, \langle e, t\rangle\rangle$ and related), i.e. lexical items that introduce degree arguments' (Beck et al. 2009:19). This parameter setting predicts that both languages exhibit expressions that refer to degrees and combine with degree operators. Specifically, (i) both English and Chinese are predicted to allow a comparison with a degree denoted by an MP, as in 'Sheldon is taller than **180 cm**'; and (ii) both are predicted to freely allow comparatives with a differential phrases, as in 'Sheldon is **5 cm** taller than Leonard'. However, the [+DSP] setting for Chinese incorrectly predicts the grammaticality of (3b) and (4b), while it correctly predicts (4a).

(3) a. ZS bi **LS** gao.
ZS BI LS tall
'ZS is taller than LS.'

b. ??/*ZS bi **180-gongfen** gao.
ZS BI 180-cm tall
Intended: 'ZS is taller than 180 cm.'

(4) a. ZS bi **LS** gao 5-gongfen.
ZS BI LS tall 5-cm
'ZS is 5 cm taller than LS.'

b. ??/*ZS bi **180-gongfen** gao 5-gongfen.
ZS BI 180-cm tall 5-cm
Intended: 'ZS is 5 cm taller than 180 cm.'

In (3a) and (4a), if we replace LS with an MP, the result is ungrammatical or seriously degraded, as in (3b) and (4b) [1]. This raises a question: **How to explain the canonical unavailability of MPs as the standard of comparison in Chinese adjectival *bi*-comparatives?** Second, despite the general surface structure sketched in (1), a critical morphosyntactic difference exists between bare and differential *bi*-comparatives: the morpheme *chu*, literally meaning 'beyond/exceed', is exclusively licensed in the latter, but not in the former, as in (5).

(5) a. *ZS bi LS gao-**chu**.
ZS BI LS tall-CHU
Intended: 'ZS is taller than LS.'

b. ZS bi LS gao-**chu** *(5-gongfen).
ZS BI LS tall-CHU 5-cm
'ZS is 5 cm taller than LS.'

The ungrammaticality of (5a) shows that *chu* is disallowed when there is no MP, whereas it can combine with a GA only when an MP is present, as in (5b), whose meaning is the same as (4a). This raises another question: **Why is the morpheme *chu* permitted in differential *bi*-comparatives but not in bare *bi*-comparatives?**

The observations above seem to suggest that degrees are accessible in some constructions, but not in others. Specifically, I propose that (i) bare *bi*-comparatives do not characterize an ordering of degrees, but a directed scale segment (Schwarzschild 2020); and (ii) differential *bi*-comparatives also characterize a directed segment, but it allows the mapping of the segment onto a degree specified by the differential phrase, a role fulfilled by the morpheme *chu* (à la Wellwood 2015). Taken together, Chinese adjectival *bi*-comparatives may constitute a case where degrees are not encoded in the lexical semantics of GAs, but introduced via a functional morpheme (Wellwood 2015; Bochnak et al. 2020).

2 Ingredients

This paper makes use of three established approaches to account for the semantics of Chinese adjectival *bi*-comparatives. First, I follow Wellwood (2015) and Cariani, Santorio, and Wellwood (2023), which interpret bare GAs as invoking state structures and expressing contextually-determined threshold properties, as in (6a). When a GA appears in comparatives, a silent morpheme DEG invokes a **bg**(**background**) function and discards the threshold properties, as in (6b). A canonical comparative construction in English is illustrated in (7).

(6) a. $[\![\text{tall}]\!]^c = \lambda s_v.s \in Dom(\langle D_{\textbf{height}}, \succcurlyeq_{\textbf{height}}\rangle).\textbf{tall}_C(s)$ b. $[\![\text{DEG}]\!]^c = \lambda P_{vt}\lambda v_v.\textbf{bg}(P,v)$

(7) a. $\text{Joe}_{[+\text{ho}]}$ is taller than $\text{Katie}_{[+\text{ho}]}$.
b. $[\![(7a)]\!]^c$ = T iff $(\exists s)(\textbf{holder}(s, j)$ & $\textbf{bg}(\textbf{tall}_C)$ & $A(\mu)(s) > \textbf{max}(\lambda d'.(\exists s')(\textbf{holder}(s', k)$ & $\textbf{bg}(\textbf{tall}_C)$ & $A(\mu)(s') > d')))$

θ-marked syntactic arguments are interpreted as properties of eventualities, which flows from the neodavidsonian framework (Champollion 2015). (7b) says that Joe is in a state in the domain of the background structure associated with the contextually-determined property of being tall, the measurement of which exceeds any such state of Katie.

[1] Concerning the availability of MPs as the standard of comparison in Chinese adjectival *bi*-comparatives, complications typically arise depending on the specific GA involved. First, for GAs associated with traditional measurement units, such as *gao* 'tall', *chang* 'long', and *kuan* 'wide', MPs may be licensed as the standard of comparison if preceding discourse provides the same MPs. Second, for GAs like *re* 'hot' and *leng* 'cold,' MPs can serve as the standard of comparison only if they represent a temperature value that meets or exceeds the contextually determined threshold. Finally, for GAs such as *kuai* 'fast' and *man* 'slow,' MPs are categorically disallowed as the standard of comparison, irrespective of the discourse context. The analysis of this restriction will be reserved for a separate research.

Second, I adopt the segmental semantics by Schwarzschild (2020), which argues that a comparative characterizes a directed segment σ, i.e., a segment of a line that has directionality, defined by a scale and two points on that scale – **START** and **END**, as illustrated by (8).

(8) a. Raj is taller than Leonard.

b. $[\![(8a)]\!]^c = (\exists\sigma)(\mathbf{END}(\sigma) = \mu_\sigma(r)\ \&\ \mu_\sigma = \text{HEIGHT}\ \&\ \nearrow(\sigma)\ \&\ \mathbf{START}(\sigma) = \mu_\sigma(l))$

(8b) says there is a rising segment starting with Leonard's height and ending with Raj's height.

Third, I take the equivalency relation by Bale (2006) as the definition for the equivalency relation for states, as in (9).

(9) $(\forall s \in D_A)((s_x \succcurlyeq_A s \Leftrightarrow s_y \succcurlyeq_A s)\ \&\ (s \succcurlyeq_A s_x \Leftrightarrow s \succcurlyeq_A s_y))$ (A for adjectives)

According to this definition, two states s_x and s_y are equivalent to each other if and only if every state to which s_x is related, s_y is also related and vice versa. Next, after forming equivalency classes by grouping all the states that are equivalent to each other into the same set, an ordering can be achieved by ranking equivalency classes according to how their members are related to one another. Namely, if the members of the equivalency class X are greater than the members of the equivalency class Y, then X is ordered above Y. Such an ordering is connected, transitive, reflexive, and anti-symmetric (cf. Bale 2006).

Overall, I take it that a segment is conceptualized as a connected region that orders equivalency classes of states. When a comparative characterizes a directed scale segment, the start and end points of this segment correspond, respectively, to two equivalency classes—each consisting of a set of states. The comparative 'a is taller than b' is to characterize a rising segment whose start point is the equivalency class containing b's height state and end point is the equivalency class containing a's height state.

3 Two Fundamental Issues

First, GAs in Chinese, despite the same surface forms, may receive different interpretations.

(10) a. ZS juede LS **gao**.
ZS believe LS tall
'ZS believes that LS is tall.'
→ LS is tall.

b. ZS bi LS **gao**.
ZS BI LS tall
'ZS is taller than LS.'
↛ ZS or LS is tall.

GA *gao* 'tall' in (10a) receives a positive reading, implying that Lisi is tall. In contrast, despite the same surface GA, (10b) does not entail that ZS or LS is tall. To capture such a property of GAs, based on Cariani, Santorio, and Wellwood (2023), I assume that there is a silent morpheme BG (background) immediately preceding the GA in *bi*-comparatives that invokes a **bg** function and helps the GA discard the threshold property, as shown in (11).

(11) a. $[\![\text{BG}]\!]^c = \lambda P_{vt}\lambda v_v.\mathbf{bg}(P,v)$ b. $[\![\text{BG } gao]\!]^c = [\![\text{BG}]\!]^c([\![gao]\!]^c) = \lambda s.\mathbf{bg}(\mathbf{tall}_C)(s)$

When a GA appears in (10b), BG inputs a property and outputs its background structure (i.e., an overall ordering of states of height, and states of being tall are part of this broader ordering).

Second, building on Liu (2018), I argue that Chinese adjectival *bi*-comparatives have an overt comparative morpheme *jiao*, as exemplified in (11).

(12) a. ZS **jiao** gao.
ZS JIAO tall
'ZS is taller than somebody.'

b. ...jiage **jiao** bi yiqian gui...
price JIAO BI previous expensive...
'...the price is more expensive than before...'

The morpheme *jiao* contributes a *greater-than* meaning to (12a), and it may occur overtly in (12b), an example extracted from the CCL corpus at Peking University. I further assume that it may be covert in examples like (2a) or (2b), represented as JIAO. It is formally defined in (13).

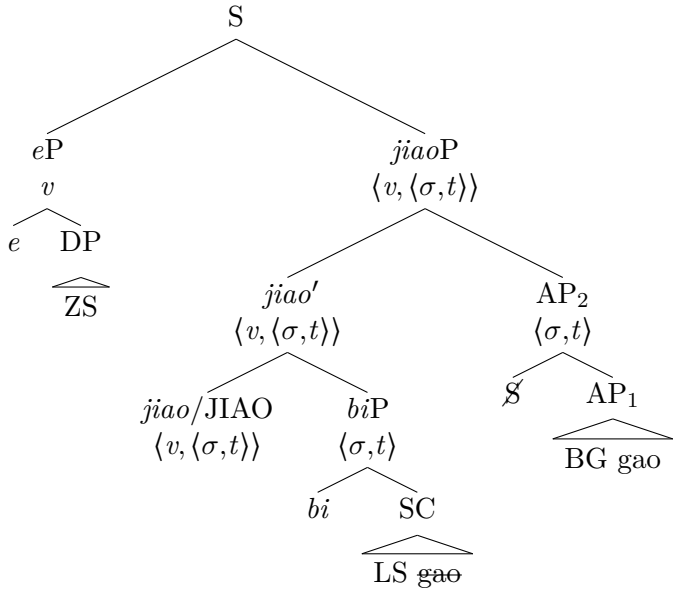
(13) $[\![jiao/\text{JIAO}]\!]^c = \lambda s\lambda\sigma.\mathbf{END}(\sigma) = s^{-\zeta}$ $\langle v, \langle\sigma,t\rangle\rangle$

In essence, *jiao* and its covert allomorph JIAO are semantically equivalent to *more*/*-er* in English comparatives in the sense of Schwarzschild (2020), which introduces the end point of a directed segment. According to (13), the end point of the segment corresponds to an equivalency class containing a height state held by an individual, represented as $s^{-\zeta}$ following Bale (2006).

4 Chinese Bare *bi*-Comparatives

Based on the two fundamental issues outlined above, I assign the structure in (14) to (2a).

(14)



Structurally, two points require clarification here. First, I take it that the comparative morpheme *jiao*/JIAO takes both *bi*P and AP as its arguments (cf. Bresnan 1973; Heim 2000; Bhatt and Pancheva 2004). Second, with the respect to the internal structure of *bi*P, following Liu (1996, 2011) and Erlewine (2018), I posit that there is an instance of AP within it, which forms a small clause with DP but is subject to obligatory deletion (Hsieh 2017). Before we proceed to interpret the structure in (14), two points are also in order. First, for *bi*, I synthesize the semantics of *than* in English comparatives, as proposed in degree-based approaches (Stechow 1984; Rullmann 1995; Kennedy 1997), with the semantics assigned to *than* in segmental semantics by Schwarzschild (2020) – the standard marker introducing the start point of a segment. Specifically, I propose that the standard marker *bi* performs two functions: (a) introducing the start point of a segment characterized by an adjectival *bi*-comparative; and (b) taking the characteristic function of a set of states and outputting the maximal state that the set maps to an equivalency class corresponding to the start point of the segment. The semantics of *bi* is formalized in (15).

(15) $[\![bi]\!]^c = \lambda g_{vt}\lambda\sigma_\sigma.\mathbf{START}(\sigma) = (\mathbf{max}(g))^{\text{-}\zeta}$

The internal structure of the *bi*P is given schematically in (16a), which is interpreted in (16b-c).

(16) a. $[_{biP}\ bi\ [_{SC}\ \text{LS}_{[+ho]}\ \sout{\text{BG } gao}]]$

b. $[\![\text{SC}]\!]^c = [\![[\text{LS}_{[+ho]}]]\!]^c\ \&\ [\![[\text{BG } gao]]\!]^c = \lambda s.\mathbf{holder}(s,\ l)\ \&\ \mathbf{bg}(\mathbf{tall}_C)(s)$

c. $[\![biP]\!]^c = [\![bi]\!]^c([\![\text{SC}]\!]^c) = \lambda\sigma.\mathbf{START}(\sigma) = (\mathbf{max}(\lambda s.\mathbf{holder}(s,\ l)\ \&\ \mathbf{bg}(\mathbf{tall}_C)(s)))^{\text{-}H}$

Second, following Schwarzschild (2020), I posit the existence of an operator Ś (composed of an 'S' for scale and a line through it representing a segment) that combines with a GA, resulting in a predicate of segments. The operator is defined in (17).

(17) $[\![Ś]\!]^c = \lambda f_{vt}\lambda\sigma.\sigma \sqsubseteq \{S \subseteq f : \exists s \in f.(S = s^{\text{-}\zeta})\}\ \&\ \sigma_C$

Given a set of states, a segment is constructed from a set of equivalency classes formed out of this set of states. This can be represented as $\sigma \sqsubseteq \{S \subseteq f : \exists s \in f.(S = s^{\text{-}\zeta})\}$, which is essentially equivalent to $\sigma \sqsubseteq Dom(D_{\text{-}\zeta,\succeq_{\text{-}\zeta}})$, namely the domain of a segment consists of a set of equivalency classes of states and each equivalency class is linearly ordered. The subset relation $\sqsubseteq$ is adopted from Schwarzschild and Wilkinson (2002). σ_C is used to represent the directionality of the segment determined by contextually supplied GAs: when the operator combines with a positive GA, it introduces a predicate of a rising segment; conversely, when it combines with a negative GA, it introduces a predicate of a falling segment. Thus, AP_2 in (14) is interpreted in (18).

(18) $[\![\text{AP}_2]\!]^c = [\![\not{S}]\!]^c([\![[\![_{\text{AP1}}\text{BG gao}]]\!]^c) = \lambda\sigma.\sigma \sqsubseteq Dom(D_{\text{-}H,\succeq_{\text{-}H}}) \;\&\; \nearrow(\sigma)$ $\langle\sigma, t\rangle$

With all the requisite assumptions in position, the interpretation of (14) is derived in (19).

(19) a. $[\![jiao\text{P}]\!]^c = ([\![\text{JIAO}]\!]^c \;\&\; [\![bi\text{P}]\!]^c) \;\&\; [\![\text{AP}_2]\!]^c = \lambda s_1\lambda\sigma.\mathbf{END} = s_1{}^{\text{-}H} \;\&\; \mathbf{START}(\sigma) = (\mathbf{max}(\lambda s.\mathbf{holder}(s,\, l) \;\&\; \mathbf{bg}(\mathbf{tall}_C)(s)))^{\text{-}H} \;\&\; \sigma \sqsubseteq Dom(D_{\text{-}H,\succeq_{\text{-}H}}) \;\&\; \nearrow(\sigma)$

b. $[\![e\text{P}]\!]^c = [\![e]\!]^c([\![\text{ZS}_{[+\text{ho}]}]\!]^c) = [\lambda P_{vt}.\epsilon s_1(P(s_1)](\lambda s.\mathbf{holder}(s,\, z)) = (\epsilon s_1)(\mathbf{holder}(s_1,\, z))$

c. $[\![\text{S}]\!]^c = [\![jiao\text{P}]\!]^c([\![e\text{P}]\!]^c) = \lambda\sigma.\mathbf{END} = ((\epsilon s_1)(\mathbf{holder}(s_1,\, z)))^{\text{-}H} \;\&\; \mathbf{START}(\sigma) = (\mathbf{max}(\lambda s.\mathbf{holder}(s,\, l) \;\&\; \mathbf{bg}(\mathbf{tall}_C)(s)))^{\text{-}H} \;\&\; \sigma \sqsubseteq Dom(D_{\text{-}H,\succeq_{\text{-}H}}) \;\&\; \nearrow(\sigma)$

Some clarification is required at this point. First, there exists a type mismatch between *jiao* and *bi*P, as well as between *jiao′* and AP_2. This mismatch can be resolved by applying the 'Segment Identification' rule (Schwarzschild 2020:237), which is reformulated in (20).

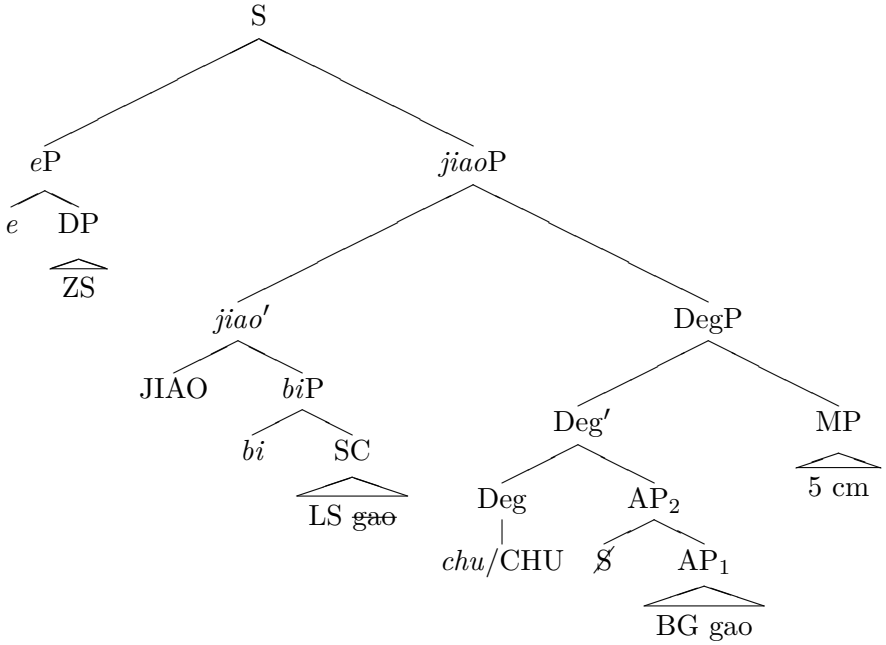
(20) **Segment Identification** (states-based verison)
'Let α be a node with two daughters, β and γ. Let β be of type $\langle v,\langle\sigma,t\rangle\rangle$ and γ be of type $\langle\sigma,t\rangle$, then $[\![\alpha]\!]^{g,c} = \lambda v\lambda\sigma.[\![\beta]\!]^{g,c}(v)(\sigma) \;\&\; [\![\gamma]\!]^{g,c}(\sigma)$.'

Following this rule, *jiao* and *bi*P, as well as *jiao′* and AP_2, combine intersectively, as illustrated in (19a) and (19b), respectively. Second, I assume that DP *ZS* combines with a silent indefinite determiner *e* (interpreted using ϵ operator), which is the indefinite counterpart of ι and does not presuppose uniqueness (see Heusinger 1997). This is explicated in (19b). It is crucial to note that the state held by ZS must be commensurate with the state held by LS in (19c), namely the start and end points of the segment characterized by the *bi*-comparative must both correspond to equivalency classes consisting solely of height states. Putting everything together, (14) has the logical form in (19c). This interpretation is true just in case there is a rising segment that begins with an equivalency class containing the height state held by LS and ends with an equivalency class containing the height state held by ZS. Thus, bare adjectival *bi*-comparatives characterize a directed segment by accessing equivalency classes of sets of states held by individuals, which directly blocks degree-denoting MPs as the standard of comparison.

5 Differential *bi*-Comparatives

One crucial morphosyntactic distinction between bare and differential *bi*-comparatives lies in the possible occurrence of the morpheme *chu* 'beyond/exceed' in the latter. To characterize the role of *chu* in differential *bi*-comparatives, I follow Gu and Guo (2015) and argue that *chu* (or its covert allomorph CHU) is a functional morpheme acting as Deg head that projects a Degree Phrase in the structure. I assign the structure in (21) to (2b).

(21)



To interpret the structure in (21), I follow Grano and Kennedy (2012) and argue that the morpheme *chu*/CHU is a functional morpheme that introduces a degree argument, namely mapping a segment onto a degree, based on the idea that degrees are intervals (Schwarzschild and Wilkinson 2002). *chu*/CHU is formally defined in (22) below in which Σ is a variable ranging over a set of segments.

(22) $[\![chu/\text{CHU}]\!]^c = \lambda\Sigma_{\sigma t}\lambda \boldsymbol{d}\lambda\sigma.\Sigma(\sigma)\ \&\ \mu(\sigma) \geq \boldsymbol{d}$

Based on the definition in (22), (21) is interpreted in the following.

(23) a. $[\![\text{Deg}']\!]^c = [\![\text{Deg}]\!]^c([\![\text{AP}_2]\!]^c) = \lambda d\lambda\sigma.\sigma \sqsubseteq Dom(D_{\text{-}H,\succeq\text{-}H})\ \&\ \nearrow(\sigma)\ \&\ \mu(\sigma) \geq d$

b. $[\![\text{DegP}]\!]^c = [\![\text{Deg}']\!]^c([\![\text{MP}]\!]^c) = \lambda\sigma.\sigma \sqsubseteq Dom(D_{\text{-}H,\succeq\text{-}H})\ \&\ \nearrow(\sigma)\ \&\ \mu(\sigma) \geq \textbf{5 cm}$

c. $[\![\text{S}]\!]^c = [\![jiao\text{P}]\!]^c([\![e\text{P}]\!]^c)$
$= \lambda\sigma.\textbf{END} = ((\epsilon s_1)(\textbf{holder}(s_1,\ z)))^{\text{-}H}\ \&\ \textbf{START}(\sigma) = (\textbf{max}(\lambda s.\textbf{holder}(s,\ l)\ \&\ \textbf{bg}(\textbf{tall}_C)(s)))^{\text{-}H}\ \&\ \sigma \sqsubseteq Dom(D_{\text{-}H,\succeq\text{-}H})\ \&\ \nearrow(\sigma)\ \&\ \mu(\sigma) \geq \textbf{5 cm}$

Note that there is a type mismatch between *jiao*′ ($\langle v, \langle\sigma,\ t\ \rangle\rangle$) and DegP ($\langle\sigma,\ t\rangle$), which would be resolved by the 'Segment Identification' rule described in (20). Given the projection of DegP headed by the morpheme *chu*/CHU in the syntax of differential *bi*-comparatives, as in (21), the interpretation of differential *bi*-comparatives, to some extent, diverges from that of bare *bi*-comparatives. This distinction is elucidated in (23a-b), in which it is demonstrated that the Deg head introduces a degree argument. The syntactic structure above DegP is interpreted in a manner consistent with the analysis provided in (19). In a word, the differential *bi*-comparative in (21) has the logical form in (23c), which says that it is true just in case there is a rising segment that starts with an equivalency class containing LS's height state, ends with an equivalency class containing ZS's height state, and is measured 5 cm. The core idea is that a pivotal difference between bare and differential *bi*-comparatives lies in the accessing of degrees facilitated by a functional morpheme in the latter. That is why *chu*/CHU can only occur in differential *bi*-comparatives. Overall, differential *bi*-comparatives build a directed segment based on equivalency classes of sets of states and access degrees indirectly via the functional morpheme which maps the segment onto a degree. This rules out the permissibility of having a direct comparison with a degree, given that MPs directly predicate of degrees (Schwarzschild 2005).

6 Conclusion

From a descriptive point of view, two defining features of Chinese adjectival *bi*-comparatives have been underexplored: (a) the canonical prohibition against MPs from serving as the standard of comparison both in bare and differential *bi*-comparatives; and (b) the exclusive licensing of the morpheme *chu* in differential *bi*-comparatives. Building on these observations, I propose that the *bi*-phrase, which functions as the standard phrase, characterizes an ordering of equivalency classes whose ordering base is states (Wellwood 2015). This semantic constraint inherently precludes degree-denoting MPs from serving as the standard of comparison in both bare and differential *bi*-comparatives. I further argue that degrees in Chinese adjectival *bi*-comparatives are accessed indirectly, i.e., via a functional morpheme. This may potentially offer support for some cross-linguistic patterns observed by Bochnak et al. (2020) which argues that a range of functional morphemes, such as comparative morphemes and degree modifiers, introduce degrees. However, as briefly noted in footnote 1, there are cases where MPs can serve as the standard of comparison when certain GAs function as the main predicate of comparison, provided appropriate discourse contexts are established. The semantics proposed here aims to provide a foundation for future research to explore the sources of variation in the restrictions on MPs as standards of comparison.

References

Bale, Alan (2006). "The universal scale and the semantics of comparison". PhD thesis. McGill University.

Beck, Sigrid et al. (2009). "Crosslinguistic Variation in Comparison constructions". In: Linguistic Variation Yearbook 9, pp. 1–66.

Bhatt, Rajesh and Roumyana Pancheva (2004). "Late merger of degree clauses". In: Linguistic Inquiry 35.1, pp. 1–46.

Bochnak, M.Ryan et al. (2020). "Degreefulness is the result of functional inventory, not a parameter". Handout from *Sinn und Bedeutung* 25, QMUL.

Bresnan, Joan (1973). "Syntax of the comparative clause construction in English". In: Linguistic Inquiry 4.3, pp. 275–343.

Cariani, Fabrizio, Paolo Santorio, and Alexis Wellwood (2023). "Positive gradable adjective ascriptions without positive morphemes". In: Proceedings of Sinn Und Bedeutung. Vol. 27, pp. 96–113.

Champollion, Lucas (2015). "The interaction of compositional semantics and event semantics". In: Linguistics and Philosophy 38, pp. 31–66.

Erlewine, Michael Yoshitaka (2018). "Clausal comparison without degree abstraction in Mandarin Chinese". In: Natural Language and Linguistic Theory 36.2, pp. 445–482.

Grano, Thomas and Chris Kennedy (2012). "Mandarin transitive comparatives and the grammar of measurement". In: Journal of East Asian Linguistics 21.3, pp. 219–266.

Gu, Yang and Jie Guo (2015). "On the Internal Structure of Comparative Constructions: From Chinese to Japanese to English". In: Chinese Syntax in a Crosslinguistic Perspective. Ed. by Audrey Li, Andrew Simpson, and Wei-Tien Tsai. Oxford: Oxford University Press, pp. 334–374.

Heim, Irene (2000). "Degree Operator and Scope". In: Proceedings of SALT X. Ed. by B. Jackson and T. Matthews. Cornell University. CLC Publications, pp. 40–64.

Heusinger, Klaus von (1997). "Definite descriptions and choice functions". In: Logic, Language and Computation 5, pp. 61–91.

Hsieh, I-Ta Chris (2017). "On Arguments against Comparative Deletion in Mandarin". In: Tsing Hua Journal of Chinese Studies 47.2, pp. 255–287.

Kennedy, Chris (1997). "Projecting the adjective: The syntax and semantics of gradability and comparison". PhD thesis. University of California, Santa Cruz.

Liu, Cheng-Sheng (1996). "A note on Chinese comparatives". In: Studies in The Linguistic Science 26, pp. 217–235.

— (2011). "The Chinese bi Comparatives". In: Lingua 121.12, pp. 1767–1795.

— (2018). "Projecting adjectives in Chinese". In: Journal of East Asian Linguistics 27.1, pp. 67–109.

Rullmann, Hotze (1995). "Maximality in the Semantics of Wh-constructions". PhD thesis. University of Massachusetts, Amherst.

Schwarzschild, Roger (2005). "Measure phrases as modifiers of adjectives". In: Recherches linguistiques de Vincennes 34, pp. 207–228.

— (2020). "From Possible Individuals to Scalar Segments". In: Interactions of degree and quantification. Ed. by Peter Hallman. Brill, pp. 231–270.

Schwarzschild, Roger and Karina Wilkinson (2002). "Quantifiers in Comparatives: a Semantics of Degree Based on Intervals". In: Natural Language Semantics 10, pp. 1–41.

Stechow, Arnim von (1984). "Comparing Semantic Theories of Comparison". In: Journal of Semantics 3.1-2, pp. 1–77.

Wellwood, Alexis (2015). "On the semantics of comparison across categories". In: Linguistics and Philosophy 38, pp. 67–101.

DUAL as a core concept and the pronounceability of alternatives

Paloma Jeretič
University of Pennsylvania
Philadelphia, U.S.A
pjeretic@upenn.edu

Aurore Gonzalez
University of Milano-Bicocca
Milan, Italy
aurore.gonzalez@unimib.it

Itai Bassi
ZAS & Ben Gurion University
Berlin, Germany & Israel
bassi@leibniz-zas.de

Kazuko Yatsushiro
ZAS
Berlin, Germany
Yatsushiro@leibniz-zas.de

Uli Sauerland
ZAS
Berlin, Germany
sauerland@leibniz-zas.de

Abstract

Chemla (2007) observed that the French universal quantifier *tous* ('all') is anti-dual, even though French has no word for 'both' to feed a *Maximize Presupposition* competition.[1] This has become one of the better-known examples suggesting the need for 'conceptual alternatives' (Buccola, Križ, and Chemla 2018), but no detailed account of it has been put forth. Furthermore, we show that a naive implementation of the idea overgenerates anti-duality inferences in other quantifiers, such as *each*, *which* and *one* in English and French. This paper proposes an account of Chemla's (2007) puzzle where French *tous* has a dual universal alternative built from a dual core concept. That alternative is blocked from being realized due to a principle that we call *Avoid Ambiguity*. In addition to accounting for *tous*'s anti-duality, this proposal accounts for the lack of anti-duality in other quantifiers.

1 The puzzle

The English universal quantifiers *all* and *every* are 'anti-dual', i.e., cannot be used if their domain is known to contain only two individuals, as shown in (1-a). Instead, the dual universal quantifier *both* can be used in those contexts (1-b).

(1) a. #Lea broke all her arms. / #Lea broke every arm of hers.
b. Lea broke both her arms.

Percus (2006) and Sauerland (2008) argue that the anti-duality of *all* and *every* arises as an implicated presupposition due to competition with the dual universal lexical item *both*, which, via *Maximize Presupposition* (MP), makes a universal quantifier with no size restriction anti-dual.

Chemla (2007) raises a puzzle for this analysis. In French, the universal quantifier *tous* is also anti-dual (2-a). But French does not have a lexical item for 'both'. In fact, the most direct translation of 'both' is the complex definite numeral expression *les deux* ('the two') in (2-b).

(2) a. #Léa s'est cassé tous les bras.
Léa REFL.is broke all the arms
#'Léa broke all her arms.'

b. Léa s'est cassé les deux bras.
Léa REFL.is broke the two arms
'Léa broke both her arms.'

The French data in (2) constitute a problem for the competition-based account. First, *les deux* cannot be obtained as an alternative to *tous* using a standard Katzirian algorithm that assumes replacements of nodes of the same syntactic category (Katzir 2007). Second, as noted

[1] This paper addresses the same puzzles as Jeretič et al. (2024). We propose a reformulation of the theoretical proposal that, while minimally different, suggests a different answer to the question raised by Buccola, Križ, and Chemla (2018), namely whether there must exist conceptual alternatives.

by Chemla (2007), if we extend it to allow for *tous* to compete with *les deux*, it is a puzzle why it does not compete with another identically structured numeral expression like *les trois* ('the three'), which would produce an unattested anti-triality inference of *tous*.

Chemla suggests that the explanation of *tous*'s anti-duality lies in the existence of a dual 'core concept' that can participate in competition with *tous*. This observation has become one of the better-known examples suggesting the need for 'conceptual alternatives', that is, non-utterable meaningful objects that can compete with pronounceable linguistic material and be included in operations over alternatives (Buccola, Križ, and Chemla 2021). Nevertheless, no full-fledged account of the anti-duality of *tous* has yet been proposed. This paper takes on this challenge.

Chemla's (2007) puzzle suggests that there is a parallel between the anti-duality of *all* in languages which have a lexicalized dual counterpart *both*, like English, and the anti-duality of *all* in languages that don't, like French. However, we find a number of other items that are anti-dual in some languages, apparently due to the presence of a lexicalized dual counterpart, but are not anti-dual in other languages. A minimally different example to universal quantifiers involves negative quantifiers, which also have universal force. English negative quantifier *no* (also *none*) is anti-dual, as shown in (3-a), which can easily be explained by MP competition with its dual counterpart *neither*. French does not have a counterpart to *neither*, just like it does not have a counterpart to *both*; but French *aucun* ('no'/'none'), unlike *tous* ('all'), does not give rise to an anti-duality inference, as shown in (3-b). So whatever makes *tous* anti-dual in the absence of an overt dual counterpart does not apply to *aucun*.

(3) {#None, Neither} of the sides of this sheet of paper has been used.

(4) Aucun des côtés de cette feuille n'a été utilisé. (French)
no of.the sides of this paper neg.has been used
'Neither of the sides of this sheet of paper has been used.'

This type of observation extends to other quantificational items, such as interrogative *which*, which has dual counterparts in Icelandic and Japanese, and distributive *each* and existential *one*, which have dual counterparts in Japanese. Their plural equivalents exhibit anti-duality in those languages. For reasons of space we show the relevant data below only for *which* in Icelandic ((5); Jordan Chark, p.c.) and in Japanese ((6); native speaker intuition of one of the authors).

(5) a. Á hvor-um handlegg-num brotna-ði hún? (Icelandic)
On which.DUAL-DAT arm-DAT.DEF break.INT-PST she
'Which arm did she break?'
b. ?Á hvaða handlegg brotna-ði hún?
On which arm.DAT broke.INT-PST she

(6) a. Taroo-wa dotti-no ude-o o-tta-no? (Japanese)
Taro-TOP IND.DUAL-GEN arm-ACC break-past-Q
'Which arm did Taro break?'
b. #Taroo-wa dono ude-o o-tta-no?
Taroo-TOP IND arm-ACC break-past-Q

But in both English (7) and French (8), no anti-duality arises in any of these cases. Once again, this is illustrated below only for *which*, but the observations extend to *each* and *one*.

(7) Which arm hurts you?

(8) Quel bras te fait mal? (French)
which arm you cause pain
'Which arm hurts you?'

On a Chemla-inspired account in which dual is a core concept, we would expect that the

dual quantifiers in Japanese and Icelandic are lexicalizations of that core concept together with corresponding quantifiers. But then we would expect the possibility of that combination to occur in all languages, even those that do not lexicalize it, and generate anti-duality in the corresponding quantifiers, just like it does with French *tous*. Therefore, a proposal for the anti-duality of French *tous* should also be equipped to explain the absence of anti-duality with French and English *quel/which*, *chaque/each*, *un/one*.

As a final data point to consider, one may wonder specifically about the differences between universal quantifiers, where *all*, *every*, and *tous* are anti-dual, but not *each*, mentioned above, or definite plurals, the latter shown in (9) and (10) for English and French respectively.

(9) My arms hurt.

(10) Mes bras me font mal. (French)
my arm me cause pain
'My arms hurt.'

An account that simply has the dual as a core concept would struggle to explain the lack of anti-duality in all these cases. In this work, we provide an account that captures (i) the anti-duality of French *tous*, and its lack of anti-n-ality for $n > 2$, and (ii) the absence of anti-duality in other quantifiers, which might be expected to arise due to its presence in other languages.

2 The proposal

Our account will maintain Chemla's (2007) original idea that the anti-duality of *tous* is an implicated presupposition.[2] However, since French does not have a word for 'both', we need to find a competitor to *tous* which carries a dual presupposition. Once this competitor is found, the use of *tous* is predicted, via Heim's MP in (19), to implicate anti-duality, as observed.

A dual core concept. We follow Chemla's intuition that the number concept DUAL is special and provides the basis for an alternative to French *tous* to derive anti-duality.[3] We assume that dual number features are grammatically present in every language, but that they are sometimes homophonous with each other. In a language like Slovenian, it arises as morphologically distinct number features (11), whereas in others like French, it is not morphologically differentiated from plural (12).

(11) računalnik-a (Slovenian)
computer-DUAL
'the computers'

(12) les ordinateurs (French)
the computer.DUAL/PL
'the computers'

Furthermore, DUAL can combine with quantifiers to modify their domain. Sometimes, the result is lexicalized: English *both* is the lexicalization of ∀+DUAL. English *neither* is the lexicalization of ¬+∃+DUAL. Japanese *dotti* is an existential with DUAL, available in combination with other morphemes to create dual *one*, *which* and *each*.

[2]Evidence for anti-duality of *tous* as an implicated presupposition (as opposed to lexical specification), is that it needs not project universally in the scope of a universal quantifier (Sauerland 2008), in contrast to a typical presupposition (Chemla 2009; Heim 1983; Schlenker 2008, a.o.).

(i) *Context: Some candidates to a job have written only one paper, others two, the rest more than two.*
J'ai vérifié, chaque candidat a envoyé tous ses articles.
I.AUX checked each candidate AUX sent all their articles
'I checked, every candidate sent every paper of theirs.'

[3]See Aravind 2018 (section 4.5.3, pp. 135-139) for a similar solution. As far as we can see, however, Aravind's analysis would not predict the absence of anti-duality with other quantifiers as we saw in (7)-(8).

We assume that alternatives can be obtained by replacements of lexical items (Katzir 2007).[4] As a result, a structure with DUAL features is generated as an alternative to a structure with other number features, in any language. For French, the dual structures in (13-b) and (14-b) are alternatives to their plural counterparts in (13-a) and (14-a).

(13) a. les NP.PL
b. les NP.DUAL

(14) a. tous les NP.PL
b. tous les NP.DUAL

In the case of the definites in (13), MP applies to (13-a), resulting in an anti-dual expression. (13-b) is still available in contexts in which the referent of the definite consists of two individuals. The same process applies to the universally quantified expressions in (14). But, recalling our central puzzle, this means we need to block the parse in (14-b) (when the domain consists of just 2 individuals) from being realized as *tous les NP* (cf. (2-a)). We turn to this next.

Blocking the universal dual expression. We propose a principle in (15) that blocks the dual expression in (14-b).

(15) **Avoid Ambiguity.** Let S be a string which is ambiguous between two parses P1 and P2, and let S' be a string such that:
a. S' has a parse P1' equivalent to P1
b. S' has no parse P2' equivalent to P2
c. S' is syntactically at most as complex as S (complexity := number of nodes)
Then P1 is blocked from being pronounced as S (i.e. string S cannot realize P1).

Avoid Ambiguity blocks the parse in (14-b) due to the presence of the unambiguous string *les deux NP.DUAL/PL*, which has a meaning equivalent to (14-b) (see below), no meaning equivalent to (14-a), and is at most as complex as it. First, *les deux NP* and *tous les NP* have identical complexity, with (at least) three overt terminal nodes each.

(16) a. $[_{QP}\ [_{Q}$ tous $]\ [_{DP}\ \ [_{D}$ les $]\ \ [_{NP}$ NP $]]]$
b. $[_{DP}\ [_{D}$ les $]\ \ [_{NumP}\ [_{Num}$ deux $]\ [_{NP}$ NP $]]]$

As anticipated, our proposal crucially relies on there being semantic equivalence between sentences where *tous les NP.DUAL* (14-b) and *les deux NP* are replaced with each other. To illustrate how this semantic equivalence comes about, we consider in turn the strings *les deux verres sont pleins* ('the two cups are full') in (17) and *tous les verres.DUAL sont pleins* ('all cups.DUAL are full') in (18). To start with, we adopt the familiar view of the definite article to denote the maximal sum σ of individuals present in the extension of its argument (Sharvy 1980). We further take the extension of a predicate to contain both atoms and pluralities. Even though numerals are typically assumed to indicate a lower bound (e.g., 'at least 2'), and achieve an exact number reading (e.g., 'exactly 2') through exhaustification of higher-number alternatives, for the sake of simplicity we here shortcut to *deux* ('two') having the meaning 'exactly 2'. We

[4] See below for a complete definition (this merges definitions (18), (19), (20) and (41) of Katzir's (2007) paper).

(i) A **structural alternative** of a parse tree ϕ is a parse tree ψ obtained from ϕ by a finite series of deletions (removing edges and nodes), contractions (removing an edge and identifying its end nodes), and replacements of constituents in ϕ with lexical items of the same syntactic category or subtrees of ϕ.

The notion 'of the same syntactic category' is crucial for our purposes: this restriction entails that *les n NP*, for any numeral n, is not generated as an alternative to *tous les NP* because *tous*, a quantifier, is not replaceable by *les*, a determiner, and *les* is not replaceable by a numeral. Thus, *les deux NP* is not directly generated as an alternative by the Katzirian algorithm.

obtain (17-b) as the denotation of *les deux verres*. In (17-b) it composes with the VP *sont pleins* ('are full') via a distributivity operator D.

(17) a. $[\![$ les deux verres $]\!] = \sigma x.[cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2]$
b. $[\![$ les deux verres D sont pleins $]\!] =$
$\forall z.atom(z) \wedge z \sqsubset \sigma x.[cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2] \rightarrow full(z)$

Moving to the composition of *tous les NP.*DUAL, we take French *tous* to be a universal quantifier over contextual covers of a definite plurality (here we look at the special case of quantification over atoms). This semantics is based on the selectional properties of *tous*, which can only compose with an NP if it is plural-marked, like in English, and via a definite article, unlike in English. We further assume that DUAL means 'exactly 2' (to align with Harbour's (2014) proposal). We obtain (18-a) as the denotation of *tous les verres.*DUAL, and (18-b) for the entire string, which is equivalent in meaning to (17-b).

(18) a. $[\![$ tous les DUAL verres $]\!] =$
$\lambda Q.\forall z.atom(z) \wedge z \sqsubset [\sigma x.cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2] \rightarrow Q(z)$
b. $[\![$ tous les DUAL verres sont pleins $]\!] =$
$\forall z.z \sqsubset [\sigma x.cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2] \rightarrow [\forall y.atom(y) \wedge y \sqsubset z \rightarrow full(y)]$
$\equiv \forall z.atom(z) \wedge z \sqsubset [\sigma x.cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2] \rightarrow full(z)$

Deriving anti-duality. We have shown that a dual universal expression is generated as an alternative to the plural (number-neutral) structure, but that it is blocked from being realized due to Avoid Ambiguity. Now we make a crucial stipulation, namely that expressions generated through the Katzirian algorithm count as alternatives, even when they are blocked from being realized. This allows *tous les NP.*DUAL *VP* to be an alternative to *tous les NP.*PL *VP*, which then feeds MP competition between the two alternatives to derive anti-duality for the latter.

(19) ***Maximize Presupposition***
Do not use ϕ in context c if ψ is an alternative to ϕ such that:
a. ψ has a stronger presupposition than ϕ.
b. ϕ and ψ are contextually equivalent in c.

We show that a sentence containing *tous les NP.*PL and one containing *tous les NP.*DUAL in the same position are contextually equivalent in a context where their presuppositions are satisfied, and only differ in their presupposed content, thus allowing for MP to apply.

The meaning of *tous les NP VP* is in (20), and that of *les deux NP VP* repeated in (21).

(20) $[\![$ tous les verres.PL sont pleins $]\!] = \forall z.atom(z) \wedge z \sqsubset [\sigma x.cup(x)] \rightarrow full(z)$

(21) $[\![$ tous les verres.DUAL sont pleins $]\!] =$
$\forall z.atom(z) \wedge z \sqsubset \sigma x.[cup(x) \wedge |\{y|atom(y) \wedge y \sqsubset x\}| = 2] \rightarrow full(z)$

It is easy to see that the at-issue meanings between these two expressions are the same, i.e. that all cup atoms are full. The only difference between them is found in the restrictor of the σ operator, and thus the presuppositional component: (20) presupposes there is a maximal sum of cups, and (21), that there is a maximal sum of cups which contains exactly two atoms. MP applies: (20) and (21) are contextually equivalent when presuppositions are satisfied, and (21) has a stronger presupposition than (20). Therefore, if the number of atoms is known to be exactly two, the presupposition of the definite numeral expression is met, and the presuppositionally weaker *tous* expression is blocked.

3 Anti-Duality and its Absence with Other Quantifiers

We now show why anti-duality does not arise with other quantifiers. The dual is always generated as an alternative to the non-dual version, but in most cases, it is not blocked by *Avoid Ambiguity* from being realized. The reason is that in order for the dual expression to be blocked, there needs to be an expression equivalent in meaning and at most as complex as it. However, very often there is none.

Let us start with plural definites. *The NP.*PL generates *the NP.*DUAL as an alternative. However, they are not anti-dual in French nor English. This is because there is no expression equivalent to *the NP.*DUAL that is at most as complex as it. The closest equivalence would be the definite numeral expression *the two NP*, but that is already too complex.

Turning to negative quantifiers, English has a dual negative existential quantifier *neither*, which induces an anti-duality inference in the non-dual counterparts *no(ne)* (3). In contrast, negative quantifiers in French do not seem to carry anti-duality (4). This follows from our proposal. Indeed, while there is a dual alternative generated by applying DUAL to the domain of the quantifier, there is no unambiguous expression equivalent in meaning to it that is at most as complex. The smallest expressions with this meaning are *aucun des deux* ('none of the two'), or *ni l'un ni l'autre* ('neither one nor the other'), which are more complex than *aucun*.

As far as Icelandic and Japanese are concerned, recall that these languages express duality with quantifiers other than *all*, namely, *some*, *which*, and *each*, shown for *which* in (5) and (6). In addition, we observed that the non-dual counterparts to those quantifiers are anti-dual. But this anti-duality is not observed in English and French *some*, *which*, and *each*, which do not have dual counterparts. The answer is simple: even though the dual counterparts to *some*, *which* and *each* are generated, they are not blocked from pronunciation due to the absence of simple enough expressions meaning *some of the two*, *which of the two* and *each of the two*.

4 Conclusion

We propose a solution for French *tous*' anti-duality by proposing that the dual is a core concept that is grammatically present in all languages and serves as the basis for the dual competitor to *tous*. No anti-triality is generated, because there is no trial core concept. A dual expression is generated as an alternative to a non-dual one, but sometimes blocked by the Avoid Ambiguity principle, which posits that an ambiguous string can be rendered unambiguous by blocking one of its parses if there is another unambiguous string in the language equivalent in meaning and at most as complex. This is what happens to *tous les NP.*DUAL: there is the equivalent and simple enough expression *les deux NP* that blocks it from being pronounced. However, we also assume that MP competition with *tous les NP.*DUAL is still licensed, despite it being blocked from being realized, which allows *tous les NP.*PL to be anti-dual. Other expressions have their dual counterparts generated as alternatives, but they are not blocked by Avoid Ambiguity when the language lacks a simple enough equivalent expression. This is the case, for example, for *the, each, which* in English and French, and French (but not English) *aucun* ('no').

Zooming out, this appears to corroborate the claim that Conceptual Alternatives must exist as defined by Buccola, Križ, and Chemla 2018. However we would like to remain cautious about this conclusion: note that the 'conceptual' (i.e. unpronounceable due to Avoid Ambiguity) alternative in our analysis is in some sense replaced by an equivalent pronounceable expression in the language. In other words, the meaning denoted by the conceptual alternative is still expressible. Could it be that competition is only licensed when there is such an expression, and thus maintaining the Gricean intuition that alternatives are pronounceable? This is a question that can only be answered by looking at further cases of conceptual alternatives.

References

Aravind, Athulya (2018). "Presuppositions in context". PhD thesis. Massachusetts Institute of Technology.

Buccola, Brian, Manuel Križ, and Emmanuel Chemla (2018). "Conceptual alternatives: Competition in language and beyond". In: *Paris: ENS*.

— (2021). "Conceptual Alternatives: Competition in Language and Beyond". In: *Linguistics and Philosophy* 45.2, pp. 265–291. DOI: 10.1007/s10988-021-09327-w.

Chemla, Emmanuel (2007). "French *both*: A gap in the theory of antipresupposition". In: *Snippets* 15, pp. 4–5.

— (2009). "Presuppositions of quantified sentences: Experimental data". In: *Natural language semantics* 17, pp. 299–340.

Harbour, Daniel (2014). "Paucity, abundance, and the theory of number". In: *Language*, pp. 185–229.

Heim, Irene (1983). "On the projection problem for presuppositions". In: *Formal semantics–the essential readings*, pp. 249–260.

Jeretič, Paloma et al. (2024). *Core Concepts and Indirect Alternatives: On the Anti-Duality of Quantifiers*. URL: https://ling.auf.net/lingbuzz/007694.

Katzir, Roni (2007). "Structurally-defined alternatives". In: *Linguistics and Philosophy* 30.6, pp. 669–690. DOI: 10.1007/s10988-008-9029-y. URL: https://doi.org/10.1007/s10988-008-9029-y.

Percus, Orin (2006). "Antipresuppositions". In: *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science* 52, p. 73.

Sauerland, Uli (2008). "Implicated Presuppositions". In: *Sentence and Context*. Ed. by Anita Steube. Berlin, Germany: Mouton de Gruyter.

Schlenker, Philippe (2008). "Be Articulate: A pragmatic theory of presupposition projection". In: *Theoretical Linguistics* 34.3, pp. 157–212. DOI: doi:10.1515/THLI.2008.013. URL: https://doi.org/10.1515/THLI.2008.013.

Sharvy, Richard (1980). "A more general theory of definite descriptions". In: *The philosophical review* 89.4, pp. 607–624.

A Semantics for Weak, Question-Sensitive Belief

Andrej Jovićević
Institute of Philosophy, KU Leuven
Leuven, Belgium
andrej.jovicevic@student.kuleuven.be

Abstract

Recent work in epistemology defends the unorthodox theses that belief is (1) an evidentially weak, and (2) question-sensitive attitude, and (3) that forming beliefs is sometimes a matter of guessing. What motivates these theses are examples of rationally permissible belief-ascriptions that exhibit these traits. The main aim of this paper is to outline a semantic account of categorical and conditional belief-ascriptions that captures the motivating data. We then survey some consequences of the proposed semantics, particularly with respect to the question of whether closure under rules of inference is rationally required for weak, question-sensitive belief.

1 Introduction

Recent work in epistemology defends the unorthodox theses that belief is (1) an evidentially weak, (2) question-sensitive attitude, and that (3) rationally permitted belief is sometimes a matter of guessing. (1) Belief is weak: the evidential standards for believing a proposition are undemanding; specifically, one can believe p without knowing p and without being highly confident in p [13, 26, 16]. (2) Belief is question-sensitive: agents believe propositions against the backdrop of salient partitions of the logical space provided by a question under discussion (QUD, [25]) [34, 7, 14, 15]. (3) Beliefs are best guesses: in situations of uncertainty, one's belief just is one's best guess, where one's guess concerns a QUD, and a particular guess can dominate other guesses despite one's evidence assigning it an arbitrarily low probability [16, 5].[1]

First, we outline a semantics for belief-ascriptions with categorical and conditional contents (Section 3). On our approach, belief is treated as a modal sensitive to information states with probabilistic and inquisitive structure. A categorical φ is believed relative an information state just in case φ is the most probable complete answer to the underlying question. A categorical φ is believed *conditionally* on ψ w.r.t. an information state just in case φ is believed w.r.t. an information state appropriately updated with ψ. Finally, a conditional $\varphi \to \psi$ is believed w.r.t. an information state just in case it is more probable than any alternative $\varphi \to \chi$, where χ is any other complete answer to the underlying question.

Second, we outline some consequences of the semantics pertaining to the question of whether rational weak, question-sensitive belief is closed under rules of inference (Section 4). Our semantics allows us to reinforce a recent finding [24] that the target notion of belief is not closed under rules of inference like Modus Tollens, Disjunctive Syllogism, and others. We argue that this prediction is not unwarranted due to the semantic equivalence of believing $\varphi \to \chi$ and believing χ conditional on φ, and conjecture that any relevant failure of closure under rule X matches the failure of X to be truth-preserving in the presence of modalised sentences.

2 Data and Desiderata

We first introduce the theses that belief is weak, question-sensitive, and a matter of guessing via examples, and outline some desiderata for a semantic account of belief-ascriptions.

[1] Theses (1)-(3) fit naturally, but not necessarily. See [22] for a partition-sensitive account of *strong* belief.

Consider an urn containing 100 marbles, with 40 blue, 30 green, 15 red, and 15 yellow marbles inside. A marble is randomly chosen from the urn. Supposing one is in a context where questions demand only complete answers, (1) appears fine, while (2) appears odd w.r.t. Q:

Q: What is the colour of the chosen marble?

(1) I {believe}{think} that the chosen marble is blue.
(2) I {believe}{think} that the chosen marble is non-blue.

Q asks for a complete guess concerning the colour of the chosen marble, and only (1) delivers on this while respecting the chances of picking each colour. Relative to Q', however, the situation is reversed:

Q': Is the chosen marble blue or not?

(1) I {believe}{think} that the chosen marble is blue.
(2) I {believe}{think} that the chosen marble is non-blue.

Q' asks whether the chosen marble is blue, and in the reversed situation only (2) delivers on this in accordance with the chances. As one's credence that the chosen marble is blue should be 0.4 in light of the evidence, the proposition believed w.r.t. Q is not the most probable on one's evidence *overall*. In both scenarios, moreover, one permissibly believes a proposition despite not knowing it. Finally, one's answer to both questions seems to be an (informed) guess.

Agents hold beliefs not only toward propositions with categorical content, but also propositions with conditional content. Consider the case of rolling a fair, seven-sided die. Relative to Q", both (3) and (4) appear fine, while (5) appears odd:

Q": Did the die land on even or odd?

(3) I {believe}{think} that the die landed on odd.
(4) I {believe}{think} that if the die landed on a composite, it landed on even.
(5) I {believe}{think} that if the die landed on a composite, it landed on odd.

As the chance of the die landing on an even [odd] given that it lands on a composite are $\frac{2}{3}$ [$\frac{1}{3}$], (4) [(5)] appears permissible [impermissible]. Moreover, note that (4) and (5) appear to communicate the same as (4') and (5'):[2]

(4') If the die landed on a composite, I {believe}{think} that it landed on even.

(5') If the die landed on a composite, I {believe}{think} that it landed on odd.

Call ascriptions in the form of (4) beliefs in conditionals, and those in the form of (4') conditional beliefs.

We seek to develop a semantics for belief-ascriptions of the kind captured by these examples. We thus intend to capture the predictions of the theses surveyed in Section 1. Some desiderata for our semantics, judging by the examples above:

- The semantics should capture belief-ascriptions with both categorical and conditional contents.
- Belief-ascriptions should be sensitive to QUDs and contextually salient 'chances'.
- Beliefs in conditionals and conditional beliefs should be semantically equivalent.

[2]Variants of sentences above with 'maybe' or 'might' also sound pairwise indistinguishable [9]. The same goes for 'probably', and plausibly other epistemic modals and similar expressions [4, pp. 603-5]. Conditionals are also said to 'commute' with negation [30, Ch. 3]. Nevertheless, note that embedding non-doxastic attitudes like 'hope', 'hate', and 'regret' is not pairwise indistinguishable from holding those attitudes to the corresponding conditionals [6, 2].

3 The Semantics

Let $\mathscr{L}$ be a propositional language with negation, conjunction, and a belief operator (B). Let P be the set of atomic sentences of $\mathscr{L}$, W the set of possible worlds (LOGICAL SPACE), and $[\![\cdot]\!]$ be the function $P \longrightarrow \mathscr{P}(W)$. (In what follows, we use $\phi, \psi, \ldots$ as placeholders for atomic sentences.) Some definitions needed for the semantics:

Definition 3.1. An INFORMATION STATE s is a set of possible worlds, i.e., $s \in \mathscr{P}(W)$. Let $S = \mathscr{P}(W)$ be the set of all information states on W.

Definition 3.2. A QUESTION is a partition of s, i.e., a division of $s \in S$ into subsets $\{X_i\}_{i\in I}$ s.t. (a) $\bigcup_{i\in I} X_i = s$, (b) $X_i \cap X_j = \emptyset$ for all $i \neq j \in I$, and (c) $X_i \neq \emptyset$ for all $i \in I$. [12, 19, 11]

Definition 3.3. A QUESTION SPACE on s, $\mathscr{Q}(s)$ is the set of all admissible questions on s; we reserve the formal clarification of $\mathscr{Q}(s)$ for a footnote.[3]

Definition 3.4. A PROBABILITY DISTRIBUTION is a function Pr on the elements of a Boolean algebra $\mathscr{A}$ over W s.t. (a) $Pr(W) = 1$ and (b) for any $s, s' \in S$ s.t. $s \neq s'$, $Pr(s \cup s') = Pr(s) + Pr(s')$.

Definition 3.5. A PROBABILISTIC-INQUISITIVE (PI) INFORMATION STATE i is a triple $\langle s, Pr(\cdot \mid s), Q\rangle$, where $Q \in \mathscr{Q}(s)$. (Cf. sharp information states in [32].)

Sentences of $\mathscr{L}$ are evaluated w.r.t. worlds and PI information states. The truth-conditions for the Boolean fragment of $\mathscr{L}$ are standard and only sensitive to worlds. The truth-conditions for categorical belief ascriptions, or sentences of the form $B\phi$ are as follows:[4]

$$[\![B\phi]\!]^{w,i} \text{ is true iff } s \cap [\![\phi]\!] \in Q \text{ and } Pr(s \cap [\![\phi]\!]) > \max_{s\cap[\![\phi]\!] \neq X \in Q} Pr(X) \tag{6}$$

Intuitively, our account of categorical belief-ascriptions checks that the believed proposition is a cell in a partition of the logical space (a complete answer to some Q) and that Pr assigns a higher probability to the believed proposition than to other cells in the partition (with Q and Pr specified in the relevant i). Onto conditional contents.

Let $\mathscr{L}^+$ be an extension of $\mathscr{L}$ that adds an indicative conditional operator $\rightarrow$. We adopt a path semantic account of the indicative conditional, as given in [27]. A path semantic account provides a combination of an informational [21, 31, 8] and a selectional [28] account of conditionals. Intuitively, each path comes pre-equipped with an ordering on worlds; as such, paths break up the logical space into finer units than worlds. Some background on path semantics:

Definition 3.6 (Paths)**.**

1. A PATH is a sequence of worlds without repetitions.
2. A path is of LENGTH n iff the path is a sequence of n members; let p_n be the n^{th} member of a path.
3. The SET OF PATHS on a state $s = \{w_1, \ldots, w_n\}$ is $\text{PATH}(s) = \{\langle w'_1, \ldots, w'_n\rangle \mid \{w'_1, \ldots, w'_n\} = s, w'_1 \neq w'_n \text{ for all } i \neq j\}$.
4. p' is a PERMUTATION of p ($p' * p$) iff $p' \in \text{PATH}(s)$.
5. A PATH p'S STATE $\mathbf{S}(p)$ is $\{p'_1 \mid p' * p, p'_1 \neq p_1\}$
6. p' is a SUBSEQUENCE of p ($p' \leq p$) iff $\forall w \in \mathbf{S}(p)$, if $w \in \mathbf{S}(p')$, then $n(p', w) \leq n(p, w)$ (where $n(p, w) = i$ iff $p_i = w$).

[3] There exists a bijection between the set of partitions of S and the set of equivalence classes on S. For any equivalence relation $\approx_i$, the set of all equivalence classes in S generated by $\approx_i$ is $\{x \in S | [x]_{\approx_i}\}$, or the quotient set of S for $\approx_i$, $S/\approx_i$. Where $\mathscr{E}(S)$ is the set of all equivalence relations on S, $\mathscr{Q}(S) = \{S/\approx_i \mid \approx_i \in \mathscr{E}(S)\}$ is the set of all quotient sets on S. $\mathscr{Q}(S)$ just is the set of possible partitions on S. See [17] for a similar account.

[4] In case $\{X \in Q \mid X \neq s \cap [\![P]\!]\} = \emptyset$, we set $Pr(X) = 0$.

We also need definitions of update to capture the semantics of indicative conditionals:

Definition 3.7 (Update)**.**

1. A PATH UPDATE of p with q, $p+q$, is the largest member of the set $\{p' \leq p \mid \forall p''$ if $p'' * p'$, then $p''_1 \in [\![q]\!]\}$.
2. An INFORMATION STATE UPDATE of s with q, s_q, is $s \cap [\![q]\!]$.
3. A QUESTION UPDATE of Q with q, Q_q, is $\{X_i \cap [\![q]\!] \mid X_i \in Q$ and $X_i \cap [\![q]\!] \neq \emptyset\}$.
4. A PROBABILITY DISTRIBUTION UPDATE of Pr with q, Pr_q, is $Pr(\cdot \mid q)$.

Updating a PI information state $i = \langle s, Pr, Q\rangle$ with q results in $i_q = \{s_q, Pr_q, Q_q\}$.

In path semantics, antecedents of indicative conditionals update the path and PI information state at which the consequent is evaluated. If the consequent is Boolean, the truth-value of the conditional depends on the truth-value of the consequent at the first world in the updated path; if the consequent is modal, the truth-value of the conditional depends on the truth-value of the consequent at the updated PI information state. Let P and $[\![\cdot]\!]$ be as with $\mathscr{L}$.

Definition 3.8. A PATH interpretation function $\|\cdot\|$ maps sentences and paths to truth-values so that, for any $q \in P$, $\|q\|^p$ is true iff $[\![q]\!]^{p_1}$ is true.

We extend $\|\cdot\|$ so that it assigns truth-values to indicative conditionals as follows:

$$\|\phi \rightarrow \psi\|^{p_1,p} \text{ is true iff } \|\psi\|^{p_1,p+q} \text{ is true iff } [\![\psi]\!]^{(p+q)_1} \text{ is true.} \tag{7}$$

Let $I(p)$ map p to some $i = \langle s, Pr(\cdot \mid s), Q\rangle$ s.t. $s = \mathbf{S}(p)$. We account for categorical belief-ascriptions on analogy with (6):

$$\|B\phi\|^{p_1,p,I(p)} \text{ is true iff } s \cap [\![\phi]\!] \in Q \text{ and } Pr(s \cap [\![\phi]\!]) > \max_{s \cap [\![\phi]\!] \neq X \in Q} Pr(X) \tag{8}$$

Conditional beliefs are captured by updating the parameters at which the consequent is evaluated:

$$\|\phi \rightarrow B\psi\|^{p_1,p,I(p)} \text{ is true iff } \|B\psi\|^{p_1,p+q,I(p)_q} \text{ is true} \tag{9}$$

Accounting for beliefs in conditionals is more challenging. Conditionals, as per (7), are not true at possible worlds but at paths, and as such express sets of paths and not sets of possible worlds. This implies that the semantic content of conditional statements cannot serve the purpose of possible-world propositions in the foregoing account of belief. One way of sidestepping the issue is incorporating more fine-grained information states, and along with them more fine-grained partitions and probability distributions (as in [10]). We sidestep the issue in another way: when $A \rightarrow C$ is believed w.r.t. an i, we check that the probability assigned to $A \rightarrow C$ is greater than that assigned to $A \rightarrow X$, where X is any alternative to C in Q. To capture this, we define alternatives and probabilities of conditionals.

Definition 3.9. The set of ALTERNATIVES to ϕ in $I(p)$ is $\mathscr{A}_\phi^{I(p)} = \{X \in Q \mid X \neq [\![\phi]\!] \cap s\}$ (provided $[\![\phi]\!] \cap s \in Q$).

Definition 3.10. The probability of a conditional $A \rightarrow C$ w.r.t. an information state s, $C_s(A \rightarrow C)$, is $Pr(a \mid c \cap s)$, provided A and C are true at sets of possible worlds (a and c, respectively).[5]

[5]That the probability of a conditional is the probability of the consequent conditional on the antecedent is notoriously contested [23]. Selectional renditions of this thesis [29, 3, 1, 20] nevertheless fare better than other accounts. As the proofs in [10] show, the equation of Definition 3.10 provably holds on a path semantic account. We assume the equality in what follows, without spelling out the details on the lifting procedure that independently specifies the value of $C(A \rightarrow C)$ w.r.t. some s.

This allows us to account for beliefs in conditionals:

$$\|B(\phi \to \psi)\|^{p_1,p,I(p)} \text{ is true iff } s \cap [\![\psi]\!] \in Q \text{ and } C(\phi \to \psi) > \max_{X_i \in \mathscr{A}_r^{I(p)}} C(\phi \to \chi_i) \tag{10}$$

The core of our semantic proposal is given in (8)-(10). The semantics should predict the following:

Fact 3.11. (1) is true and (2) is false w.r.t. Q, and vice versa w.r.t. Q'.

Fact 3.12. (4) and (4') are true, and (5) and (5') are false w.r.t. Q".

Fact 3.13. $\|B(A \to C)\|^{p_1,p,p(I)}$ is true iff $\|A \to B(C)\|^{p_1,p,p(I)}$ is true.

We reserve the proofs for Appendix A.

4 Closure and Rules of Inference

One's beliefs can be used productively as premises in further reasoning. For instance, if I believe that it's raining if the ground is wet, and I believe that the ground is wet, I should plausibly also believe it's raining. Say that an attitude is *closed* under a rule of inference 'From Δ, infer γ' iff one cannot rationally hold an attitude to all sentences in Δ without holding the attitude to γ. If one's beliefs are to be used as premises in further reasoning, it is to be expected that belief obeys some such properties of closure. This section highlights some closure-related predictions of the foregoing semantics; a fuller treatment is left for future work.

Take as paradigmatic the rule of inference Modus Tollens (MT): 'From $\phi \to \psi$ and $\neg\psi$, infer $\neg\phi$.' For belief to be closed under MT, then, means that whenever $B(\phi \to \psi)$ and $B(\neg\psi)$ hold w.r.t to some $p_1, p, I(p)$, it also holds that $B(\neg\phi)$. For concreteness, consider a scenario (from [24]) in which Alice and Bob roll a fair, six-sided die, and Alice wins if the die lands 1–4, and Bob wins otherwise. A natural model of this scenario predicts that (11) and (12) hold w.r.t. Q, 'Who will win?'. (We omit the formal model for brevity.)

(11) I {believe}{think} that Alice will win.

(12) I {believe}{think} that, if the die lands high [> 3], Bob will win.

Let the sentences one believes w.r.t. Q be A and $H \to \neg A$ (i.e., $H \to B$). These sentences licence an inference to $\neg H$ via MT, i.e., the sentence 'The die will not land high'. Provided belief is closed under MT, the fact that $B(H \to \neg A)$ and $B(A)$ hold should imply that $B(\neg H)$ also holds. Nevertheless, the following is not permitted on our semantics:

(13) I {believe}{think} that the die will not land high.

As the chances of a fair die landing high are the same as those of it landing low, (13) seems to be a paradigm of *im*permissibility. As our theory predicts the permissibility of (11) and (12) and the impermissibility of (13), it implies that weak, question-sensitive belief is not closed under MT.[6] Is this a desirable prediction?

Pending further engagement, we offer the following perspective. Closure (or a weaker principle in the vicinity) is usually taken as a desideratum for attitudes like knowledge and belief as failures to follow up on the consequences of what one already knows or believes appears irrational. Nevertheless, once we admit a weaker sense of belief – one on which beliefs are tied to available information, partitions, and assignments of probabilities – it might result that failing to follow up on the consequences of what

[6] As the discussion in [24] suggests, our theory is expected to predict further failures of closure, for instance under Disjunctive Syllogism ('From $P \vee \neg P$, $P \to Q$, and $\neg P \to Q$, infer Q'), under Modus Ponens when the consequent is itself a conditional ('From $P \to (Q \to R)$, and P, infer $Q \to R$'), etc.

one so believes is not always irrational. Concretely, and without immediate reference to the foregoing account, beliefs (11) and (12) are pre-theoretically permissible in the given scenario, while (13) is not. If rationality is what counts, then these weak beliefs should not be jettisoned for not obeying closure. Rather, it is closure that should be jettisoned, at least in the context of weak belief, provided it makes unwarranted predictions.

Parenthetically, note that reasoning via a schema *resembling* MT *is* warranted in this scenario. In the initial setup, no outcome of the die-roll can be excluded from one's information state, so that one's information state includes worlds in which the die lands on any number from 1 to 6. Suppose, now, that a trusted informant asserts that Bob will not win (i.e., $\neg B$ / A). After updating, one's information state includes worlds in which the die lands $1-4$. But now, relative to the question 'Will the die land high or low?', our semantics predicts (13) to be permissible. In brief: after asserting (12) and being informed that Bob did not win, one *can* permissibly come to believe the MT-licensed conclusion that the die will not land high. As such, while the probabilistic guesses (11) and (12) by themselves do not permit one to guess whether the die lands high or low, relevant non-probabilistic input can make the guess in (13) permissible.

Even if these justifications are found wanting, it is useful to draw a parallel between failures of closure for weak belief and failures of truth-preservation for the corresponding rules of inference in the presence of modal vocabulary. We conjecture that each failure of closure under MT for weak, question-sensitive belief maps onto a failure of MT. The case above can serve as an example. Note that $B(A \rightarrow C)$ and $A \rightarrow B(C)$ are equivalent on our account (see Appendix A). We have that $H \rightarrow B(\neg A)$ and $\neg B(\neg A)$ are true at some p and $I(p)$. Let p be $\langle w_1, w_2, w_3, w_4, w_5, w_6 \rangle$, and let $b(w_1) = p$ without loss of generality. As $w_1 \notin [\![H]\!]$, H is false at w_1 while both $H \rightarrow B(\neg A)$ and $\neg B(\neg A)$ are true at $p = b(w_1)$. Intuitively speaking, this is a failure of MT.[7]

The parallel suggests the following. (We focus on MT for concreteness.) Failures of closure under MT for weak, question-sensitive belief share a common core with failures of truth-preservation for MT. In both cases considered above, what the conditionals do is restrict the consequent's domain of evaluation to antecedent-verifying worlds, so that when a belief is embedded in the consequent or when a conditional itself is believed, one's beliefs only hold *given some restriction*. However, one's beliefs with respect to restricted and unrestricted portions of the logical space do not entail the truth or falsity of the various available restrictions, as the restrictions are entertained suppositionally in the first place.

5 Conclusion

One aim of this paper was to outline a semantic account of belief that makes good on the claims that belief is weak and question-sensitive. On our approach, belief is a modal sensitive to information states with probabilistic and inquisitive structure. For ϕ to be believed w.r.t. such a state is, intuitively, for ϕ to be a probabilistically dominant complete answer to some question (where both the question and the probability distribution are specified in the information state).

Another aim of the paper was to outline some verdicts of the semantics with respect to the question of whether closure under valid rules of inference is rationally required for the attitude of weak, question-sensitive belief. First, we indicated that our semantics features failures of belief to be closed under valid rules of inference (e.g. Modus Tollens, Disjunctive Syllogism, etc.). Second, we argued that this prediction is warranted on account of the close confluence of each failure of closure under X for our target notion of belief and failures of X to be truth-preserving in the presence of modalised sentences.

Acknowledgements. Thanks to Ben Holguín for helpful comments on a previous draft.

[7]The failure is akin to those observed with probability modals [33], desire-related modals [18], and other information-sensitive vocabulary embedded in consequents of conditionals. Other superficial failures of otherwise valid inference rules are expected, moreover, when dealing with attitudes embedded in consequents of conditionals [2].

References

[1] Andrew Bacon. Stalnakers thesis in context. *The Review of Symbolic Logic*, 8(1):131–163, 2015.

[2] Kyle Blumberg and Ben Holguín. Embedded attitudes. *Journal of Semantics*, 36(3):377–406, 2019.

[3] Richard Bradley. Multidimensional possible-world semantics for conditionals. *Philosophical Review*, 121(4):539–571, 2012.

[4] Ivano Ciardelli. Restriction without quantification: Embedding and probability for indicative conditionals. *Ergo*, 8, 2021.

[5] Kevin Dorst and Matthew Mandelkern. Good Guesses. *Philosophy and Phenomenological Research*, 105(3):581–618, 2022.

[6] Daniel Drucker. Policy externalism. *Philosophy and Phenomenological Research*, 98(2):261–285, 2017.

[7] Daniel Drucker. The attitudes we can have. *Philosophical Review*, 129(4):591–642, 2020.

[8] Anthony S. Gillies. Iffiness. *Semantics and Pragmatics*, 3:4:1–42, 2010.

[9] Anthony S Gillies. Updating Data Semantics. *Mind*, 129(513):1–41, 2020.

[10] Simon Goldstein and Paolo Santorio. Probability for Epistemic Modalities. *Philosophers' Imprint*, 21(33), 2021.

[11] Jeroen Groenendijk and Martin Stokhof. *Studies on the semantics of questions and the pragmatics of answers*. PhD thesis, Univ. Amsterdam, 1984.

[12] C. L. Hamblin. Questions in Montague English. *Foundations of Language*, 10(1):41–53, 1973.

[13] John Hawthorne, Daniel Rothschild, and Levi Spectre. Belief is Weak. *Philosophical Studies*, 173(5):1393–1404, 2016.

[14] Daniel Hoek. Questions in Action. *Journal of Philosophy*, 119(3):113–143, 2022.

[15] Daniel Hoek. Minimal Rationality and the Web of Questions. In Peter van Elswyk, Dirk Kindermann, Cameron Domenico Kirk-Giannini, and Andy Egan, editors, *Unstructured Content*. Oxford University Press, 2024.

[16] Ben Holguin. Thinking, Guessing, and Believing. *Philosophers' Imprint*, 22(1):1–34, 2022.

[17] Joris Hulstijn. Structured information states: Raising and resolving issues. In *Proceedings of MunDial*, volume 97, pages 99–117. University of Munich Munich, Germany, 1997.

[18] Ethan Jerzak. Two ways to want? *Journal of Philosophy*, 116(2):65–98, 2019.

[19] Lauri Karttunen. Syntax and semantics of questions. *Linguistics and Philosophy*, 1(1):3–44, 1977.

[20] Justin Khoo. *The meaning of if*. Oxford University Press, Oxford, 2022.

[21] Angelika Kratzer. Conditionals. In *Modals and Conditionals*, pages 86–108. Oxford University Press, Oxford, 2012.

[22] Hannes Leitgeb. *The stability of belief: How rational belief coheres with probability*. Oxford University Press, Oxford, 2017.

[23] David Lewis. Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, 85(3):297–315, 1976.

[24] Joshua Edward Pearson. A puzzle about weak belief. *Analysis*, 2024.

[25] Craige Roberts. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and pragmatics*, 5:6–1, 2012.

[26] Daniel Rothschild. What It Takes to Believe. *Philosophical Studies*, 177(5):1345–1362, 2020.

[27] Paolo Santorio. Path Semantics for Indicative Conditionals. *Mind*, 131(521):59–98, 2022.

[28] Robert Stalnaker. A Theory of Conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell, 1968.

[29] Bas C van Fraassen. Probabilities of conditionals. In William L. Harper and Clifford Alan Hooker, editors, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, pages 261–308. 1976.

[30] Timothy Williamson. *Suppose and tell: the semantics and heuristics of conditionals*. Oxford University Press, Oxford, 2020.

[31] Seth Yalcin. Epistemic Modals. *Mind*, 116(464):983–1026, 2007.

[32] Seth Yalcin. Context probabilism. In *Logic, Language and Meaning: 18th Amsterdam Colloquium, Amsterdam, The Netherlands, December 19-21, 2011, Revised Selected Papers*, pages 12–21. Springer, 2012.

[33] Seth Yalcin. A Counterexample to Modus Tollens. *Journal of Philosophical Logic*, 41(6):1001–1024, 2012.

[34] Seth Yalcin. Belief as Question-Sensitive. *Philosophy and Phenomenological Research*, 97(1):23–47, 2018.

A Semantic Predictions

Model A.1. Let $W = s = \{b,g,r,y\}$, and the sentence B [G/R/Y/N] – 'The chosen marble is blue [green/red/yellow/non-blue]' – express the proposition b $[g/r/y/n]$ (with $n = g \cup r \cup y$). Let Q $= \{\{b\},\{g\},\{r\},\{y\}\}$ and Q' $= \{\{b\},\{n\}\}$. Let $Pr(b) = 0.4$, $Pr(g) = 0.3$, $Pr(r) = Pr(y) = 0.15$. Without loss of generality, let $p = \langle b,g,r,y\rangle$. Let $I(p) = \langle s, Pr, Q\rangle$ and $I'(p) = \langle s, Pr, Q'\rangle$.

Fact 3.11. In Model A.1, (1) is true [false] and (2) is false [true] relative to $p_1, p, I(p)[I'(p)]$.

$\|$I believe that the chosen marble is blue$\|^{p_1,p,I(p)}$is true
iff $\|$I believe that B$\|^{p_1,p,I(p)}$ is true
iff $\|Bb\|^{p_1,p,I(p)}$ is true
iff $s \cap b \in Q$ and $Pr(s \cap b) > \max_{s\cap b \neq X \in Q} Pr(X)$
iff $b \in Q$ and $0.4 > 0.3$ iff $\top$

$\|$I believe that the chosen marble is non-blue$\|^{p_1,p,I(p)}$is true
iff $\|$I believe that N$\|^{p_1,p,I(p)}$ is true
iff $\|Bn\|^{p_1,p,I(p)}$ is true
iff $s \cap n \in Q$ and $Pr(s \cap n) > \max_{s\cap n \neq X \in Q} Pr(X)$
iff $\bot$

$\|$I believe that the chosen marble is blue$\|^{p_1,p,I'(p)}$is true
iff $\|$I believe that B$\|^{p_1,p,I'(p)}$ is true
iff $\|Bb\|^{p_1,p,I'(p)}$ is true
iff $s \cap b \in Q$ and $Pr(s \cap b) > \max_{s\cap b \neq X \in Q} Pr(X)$
iff $b \in Q'$ and $0.4 > 0.6$ iff $\bot$

$\|$I believe that the chosen marble is non-blue$\|^{p_1,p,I'(p)}$is true
iff $\|$I believe that N$\|^{p_1,p,I'(p)}$ is true
iff $s \cap n \in Q'$ and $Pr(s \cap n) > \max_{s\cap n \neq X \in Q'} Pr(X)$
iff $n \in Q'$ and $0.6 > 0.4$
iff $\top$

Model A.2. Let $s' = \{w_1, \ldots w_7\}$, and the sentence O [E/C/P] – 'The die landed on an odd [even/composite/prime]' – express the proposition $o[e/c/p] = \{w_n \in W \mid n$ is odd [even/composite/prime]$\}$. Let Q" $= \{\{e\},\{o\}\}$, $Pr(w_n) = \frac{1}{7}$ for $1 \leq n \leq 7$, $Pr(o) = Pr(p) = \frac{4}{7}$, and $Pr(e) = Pr(c) = \frac{3}{7}$. Without loss of generality, let $p = \langle w_1, \ldots, w_7\rangle$. Let $I(p) = \langle s, Pr, Q''\rangle$.

Fact 3.12. In Model A.2, (4) and (4') are true, and (5) and (5') are false relative to i_3.

$\|$I believe that C $\rightarrow$ E$\|^{p_1,p,I(p)}$ is true
iff $\|B(c \rightarrow e)\|^{p_1,p,I(p)}$ is true
iff $s \cap e \in Q''$ and $C(c \rightarrow e) > C(c \rightarrow o)$
iff $e \in Q''$ and $Pr(e \mid c) > Pr(o \mid c)$
iff $e \in Q''$ and
iff $\top$

$\|$If C, I believe that E$\|^{p_1,p,I(p)}$ is true
iff $\|c \rightarrow Be\|^{p_1,p,I(p)}$ is true
iff $\|Be\|^{p_1,p+c,I(p)_c}$ is true
iff $s_c \cap e \in Q_c''$ and $Pr_c(e) > \max_{s_c\cap e \neq X \in Q_c''} Pr_c(X)$
iff $\{2,4\} \in Q_c''$ and $\frac{2}{3} > \frac{1}{3}$
iff $\top$

$\|$I believe that C $\rightarrow$ O$\|^{p_1,p,I(p)}$ is true
iff $\|B(c \rightarrow o)\|^{p_1,p,I(p)}$ is true
iff $e \in Q''$ and $\frac{1}{3} > \frac{2}{3}$
iff $\bot$

$\|$If C, I believe that O$\|^{p_1,p,I(p)}$ is true
iff $s_c \cap o \in Q_c''$ and $Pr(o \mid c) > \max_{s_c\cap o \neq X \in Q_c''} Pr(X)$
iff $\{1\} \in \{\{2,4\},\{1\}\}$ and $\frac{1}{3} > \frac{2}{3}$
iff $\bot$

Fact 3.13. $\|B(A \rightarrow C)\|^{p_1,p,p(I)}$ is true iff $\|A \rightarrow B(C)\|^{p_1,p,p(I)}$ is true.

Proof. $\Rightarrow$ Suppose $\|B(A \rightarrow C)\|$ is true at some p_1, p, $I(p) = i = \langle s, Pr, Q\rangle$. By (10), $s \cap c \in Q$ and $C(a \rightarrow c) > \max_{X_i \in \mathscr{A}c^i} C(a \rightarrow Xi)$. By Definition 3.10, $Pr(c \mid a) > \max_{X_i \in \mathscr{A}c^i} Pr(Xi \mid a)$. By Definition 3.9, for all $X \in \mathscr{A}c^i$, it holds that $X \neq s \cap c$ and $X \in Q$. As such, $\max X_i \in \mathscr{A}c^i Pr(Xi \mid a) = \max_{s\cap c \neq X \in Q} Pr(X \mid a) = \max_{s_a \cap c \neq X \in Q_a} Pr(X)$. By Definition 3.7.4, $Pr(c \mid a) = Pr_a(c)$. By Definition 3.7.3, if $s \cap c \in Q$, then $s \cap c \cap a \in Q_a$. From this it follows that $\|Bc\|$ is true at $p_1, p+a, i_a = \langle s \cap a, Pr(\cdot \mid a), Q_a\rangle$. But then it holds that $\|a \rightarrow Bc\|$ is true at w, p, i. □

Proof. $\Leftarrow$ Suppose $\|a \rightarrow Bc\|$ is true at some p_1, p, $I(p) = i = \langle s, Pr, Q\rangle$, i.e., $\|Bc\|$ is true at some $p_1, p+a, I(p)a = \langle s_a, Pr(\cdot \mid a), Q_a\rangle$. Then, $sa \cap c \in Q_a$ and $Pr_a(c) > \max_{s_a\cap c \neq X \in Q_a} Pr_a(X)$. As above, $Pr_a(c) = Pr(c \mid a)$ and $\max_{s_a\cap c \neq X \in Q_a} Pr_a(X) = \max_{s\cap c \neq X \in Q} Pr(X \mid a) = \max_{X_i \in \mathscr{A}c^i} Pr(Xi \mid a)$. Then, $Cr(a \rightarrow c) > \max_{X_i \in \mathscr{A}c^i} Cr(a \rightarrow xi)$. By Definition 3.7.3, if $s_a \cap c \in Q_a$, then $s \cap c \in Q$. But then it holds that $\|B(a \rightarrow c)\|$ is true at w, p, i. □

Presuppositions project asymmetrically, *unless* they don't

Alexandros Kalomoiros
Institute for Linguistics, HHU, Düsseldorf
Düsseldorf, Germany
akalomoiros@gmail.com

Matthew Mandelkern
Department of Philosophy, NYU
New York, USA
mandelkern@nyu.com

Jacopo Romoli
Institute for Linguistics, HHU, Düsseldorf
Düsseldorf, Germany
jacopo.romoli@gmail.com

Florian Schwarz
Department of Linguistics, University of Pennsylvania
Philadelphia, USA
florians@sas.upenn.edu

Abstract

We present two experiments that shed new light on the old (but ongoing) debate on filtering (a-)symmetries with presuppositions (Karttunen 1974; Heim 1983; Soames 1989; Schlenker 2008 a.o.). The experiments establish that *unless* and *if not* exhibit symmetric filtering. In this respect, they pattern together with *disjunction*, and contrast with *conjunction*. We end by examining the ramifications of our results for possible theories of presupposition.

1 Introduction

A core choice point for suggested solutions to the projection problem is whether linear order should play a role in filtering calculations. The main relevant puzzle can be illustrated by comparing conjunction and disjunction:

(1) Conjunction
   a. #John ate carbs again today and he ate carbs yesterday.
   b. ✓John ate carbs yesterday and he ate carbs again today.

(2) Disjunction
   a. (#) Either John ate carbs again today or he didn't eat carbs yesterday.
   b. ✓Either John didn't eat carbs yesterday or he ate carbs again today.

In conjunction we observe an asymmetry: while it is always possible to filter a presupposition in the right conjunct if the left conjunct entails it ((1b); *Left-to-Right filtering*), the reverse ((1a); *Right-to-Left filtering*) is not possible (Karttunen 1973; Stalnaker 1974).

Whether disjunction shows the same asymmetry has been much more controversial; while there has been an intuition that disjunctions show symmetric behavior in allowing both L-R and R-L filtering (e.g. Evans 1977's 'bathroom' sentences; Karttunen 1973, (ft. 11), Karttunen 1974, (ft. 3), Hausser 1976; Schlenker 2008 a.o.), it has also been argued that this intuition is linked to processes other than filtering (e.g., local accommodation/cancellation; Gazdar 1979 a.o.).[1]

In view of this landscape, two types of theoretical approaches have been developed. The first, dating in conception to the dynamic systems of Karttunen 1974 and Heim 1983, essentially lexicalized filtering: one could write down rules that allow only L-R filtering, or only R-L filtering, or both. Such systems can easily predict differences in (a-)symmetry between operators by individually specifying the availability of R-L and L-R filtering; but this comes at the cost of

[1]Similarly, there have been attempts to blame the asymmetry of conjunction on processes beyond filtering, e.g. redundancy considerations (cf. Rothschild 2011 a.o.). However, the results of Mandelkern, Zehr, et al. 2020 speak against this, as redundancy was explicitly controlled for in their design.

stipulation and corresponding loss of explanatoriness (Rooth in a letter to Heim, Soames 1989; Heim 1990; Schlenker 2008).

The desire for more explanatory theories led to the second kind of approach, which starts from the assumption that all operators fundamentally allow both L-R and R-L filtering (i.e. they are fully symmetric in this respect). However, a general preference for L-R filtering is also posited, assumed to be rooted in general cognitive mechanisms behind incremental processing (following an intuition going back to Stalnaker 1974). Essentially, because humans have to process language incrementally, their filtering mechanism reflects this incrementality in the form of an L-R preference. A final assumption in this kind of theory (which we can call the *processing approach*) is that the default processing preference for L-R filtering can be overridden, but at a processing cost (Schlenker 2008; Schlenker 2009; Rothschild 2011).

While more explanatory, this processing approach faces at least two potential challenges. ***Challenge 1*** concerns the prediction that R-L filtering is uniformly costly across all operators, relative to L-R filtering. Finding genuine instances of R-L filtering that do not exhibit such relative costs would then present a challenge. And indeed, recent experimental evidence suggests that disjunctions like (2a) constitute such cases (Kalomoiros and Schwarz 2024).

Challenge 2 revolves around the issue whether it is indeed **linear** order that is relevant for the filtering system. The challenge comes from data involving conditionals:

(3) a. If John is happy that he is in France, then he isn't in Paris.
b. John isn't in Paris, if he's happy that he's in France.

The intuition reported in the literature (Mandelkern and Romoli 2017; Mandelkern and Romoli 2018) is that both of these conditionals presuppose *John is in France*. But processing theories predict that in (3b) we should observe L-R filtering, since the antecedent is ***linearly preceded*** by material that should filter its presupposition that *John is in France* (i.e., the negation of the consequent *John is not in Paris*). So, to the extent that the whole conditional presupposes that *John is in France*, we might conclude that the notion of *linear precedence* doesn't make the right cut.[2] Under a more abstract structure-based notion of precedence, antecedents can be seen as always preceding consequents (regardless of linear order), and hence their presupposition is predicted to project in both (3a) and (3b) (Mandelkern and Romoli 2018).

Against this backdrop, we present two experiments that establish the following empirical points: **1)** *unless* and *if not* allow equally for both L-R and R-L filtering. This strengthens *Challenge 1* above for processing approaches. **2)** In our data, we have genuine instances of a presupposition in an antecedent being filtered by material in the linearly preceding consequent. This adds new nuance to *Challenge 2* by establishing that it is not a general feature of conditionals that presuppositions from their antecedents always project. We end by discussing the consequences of these findings for theories of presupposition.

2 The Experiments

We conducted two experiments, testing the filtering behavior of *unless* (Exp 1) and *if not* (Exp 2). We consider these near equivalents for present purposes. They were chosen due to a long-standing intuition that they are like disjunction in terms of filtering (Soames 1989; Schlenker 2009).

Our experiments follow closely a design established in two previous studies: Mandelkern, Zehr, et al. 2020, which tested the filtering behavior of conjunction, arguing that it is asymmetric; and Kalomoiros and Schwarz 2024, which applied the logic of Mandelkern, Zehr, et al. 2020 to disjunction, arguing for its symmetry.

[2] Another route is to try to derive this pattern by considering more complex semantics for conditionals (e.g. strict and variably strict semantics). See Mandelkern and Romoli 2017 for this approach and challenges it faces.

2.1 Design

The stimuli were built around 3 triggers (*too, again, re-*). First, we had presuppositional sentences (Ps) that implemented an Order manipulation across the two connectives. These were presented in E(xplicit) I(gnorance) contexts that denied knowledge of the relevant presupposition (cf. Simons 2001; Abusch 2010):[3,4]

(4) a. EI Context: John and Mary are siblings and want to study abroad. Options include Tokyo and Kyoto in Japan, or Beijing and Shanghai in China. Mary is interested in studying in Japan: she would go to Kyoto on her own, but she doesn't want to go to an enormous city like Tokyo, unless John also comes with her to Japan (if not to Tokyo, then at least to Kyoto). I don't know what they ended up deciding so ***I have no idea whether Mary is currently studying in Tokyo or whether she even decided to go to Japan.*** However, given the above, I know that:
b. **Unless** John is studying in Japan too, Mary is not studying in Tokyo. Un-PsFirst
c. Mary is not studying in Tokyo, **unless** John is studying in Japan too. Un-PsSecond
d. **If** John is**n't** studying in Japan too, Mary is not studying in Tokyo. If-PsFirst
e. Mary is not studying in Tokyo, **if** John is**n't** studying in Japan too. If-PsSecond

Non-presuppositional (NoPs) variants that simply removed the relevant trigger were included to control for any order effects that are not due to the presence of a presupposition. For consistency, these were also presented in EI contexts (omitted here for space).

(5) a. Unless/If John is(n't) studying in Japan, Mary is not studying in Tokyo. NoPsFirst
b. Mary is not studying in Tokyo, unless/if John is(n't) studying in Japan. NoPsSecond

Finally, we included controls measuring the cost of local accommodation in (6), by contrasting, across EI vs. S(upport) contexts, presuppositional sentences whose consequent is unrelated to the presupposition and thus don't allow for filtering. EI contexts deny knowledge of the presupposition at the global level; thus, the only way to avoid a clash with the context is through local accommodation. In contrast, S contexts do not give rise to a contextual clash to begin with. Given that local accommodation is a costly operation (cf. Chemla and Bott 2013), the decrease in acceptability of such NoFiltering sentences in EI contexts, compared to S contexts, provides a baseline measure for the cost of local accommodation across our triggers.

(6) NoFilt
a. EI Context: ...***I have no idea whether Mary is studying in Japan***...
b. S Context: ...***I know that in the end Mary went to Japan, but I have no idea what John did***...
c. Unless/If John is(n't) studying in Japan too, Mary is unhappy. NoFiltFirst
d. Mary is unhappy, unless/if John is(n't) studying in Japan too. NoFiltSecond

2.2 Predictions

Processing accounts predict that PsSecond cases should be more acceptable than PsFirst. The reason is that in the latter case, using the material in the consequent to filter the presupposition of the antecedent requires recourse to R-L filtering. No such recourse is required in the PsSecond

[3]For a full list of stimuli, see the OSF page for this paper. Click *here* for a link.

[4]Due to space limitations, we cannot explicitly contrast our studies to Chemla and Schlenker 2012, the only previous study on filtering with *unless*, but we briefly note two important differences: 1) Our order manipulation keeps the presupposition trigger in the antecedent. 2) The negation of our consequent always asymmetrically entails the presupposition in our antecedents, as a further safe-guard against potential explanations that invoke local accommodation to explain symmetric effects (see Hirsch and Hackl 2014).

case, since the relevant material precedes the antecedent. Since this preference for PsFirst over PsSecond is driven solely by the presupposition, the cost it imposes should be ***in addition to*** any potential order-based preferences measured in the NoPs controls. As such, any effects of the Order manipulation should be more pronounced in the Ps cases compared to the NoPs ones. In statistical terms, this means that processing accounts predict an interaction between presuppositionality (Ps: Ps vs. NoPs) and order (Order: First vs. Second).

At the same time, processing accounts predict that the cost imposed by R-L filtering should be less than that of local accommodation.[5] This means that any decrease in acceptability between PsFirst vs. PsSecond (which measures the cost of R-L filtering) should be less than the difference between EI-NoFiltFirst/Second vs. S-NoFiltFirst/Second (which measures the cost of local accommodation). To think about this statistically, we need to set up the following factors: Filt categorizes conditions on the basis of whether they involve filtering or local accommodation. Thus PsFirst and PsSecond are categorized as Filt, and EI-NoFiltFirst/Second and S-NoFiltFirst/Second as NoFilt. Another factor, PriorS categorizes conditions on the basis of whether they exhibit linearly preceding support: PsSecond and S-NoFiltFirst/Second are coded as PriorS(upport) since they exhibit such prior support (because they are in an S context, and because of the preceding consequent respectively); PsFirst and EI-NoFiltFirst/Second are categorized as NoPriorS. Given this, the prediction that R-L filtering is less costly than local accommodation means that the effect of PriorS should be stronger in the NoFilt cases compared to the Filt cases: in other words, a Filt × PriorS interaction is predicted.

On a lexical view, the predictions depend on how *unless/if* is stipulated to project the presuppositions of its arguments.[6] The standard view in prior literature (e.g. Heim 1983 a.o.) is that presuppositions always project from the antecedent, with no option of filtering based on material in the consequent. Alternatively, we could assume that the consequent can always be used to filter the presuppositions of the antecedent, to capture apparent symmetric effects. Either way, we wouldn't expect any difference between PsFirst vs. PsSecond (going beyond any order-based differences already present in the NoPs controls); hence there should be no Ps × Order interaction. A Filt × PriorS interaction is still expected. The reason is that going from PriorS to NoPriorS should have a larger effect in the NoFilt conditions, and no effect in the Filt conditions (which should be equal).

If we stipulate that for both *unless* and *if* only L-R filtering is available, and assume that the R-L option is never available, then we predict a Ps × Order interaction (for the same reason this prediction is made on processing accounts). We also predict no Filt × PriorS interaction: since no R-L filtering is available, the only way to save PsFirst from projection is to apply local accommodation. Hence the difference between PsFirst vs. PsSecond should parallel fully the difference between EI-NoFiltFirst/Second and S-NoFiltFirst/Second.

2.3 Procedure

24 critical items were created in total (8 per trigger). There were also 24 fillers. For each connective the stimuli were divided into two lists: the Ps list contained the PsFirst, PsSecond and EI-NoFiltFirst/Second conditions; the NoPs list contained the NoPsFirst, NoPsSecond and S-NoFiltFirst/Second conditions. Each list then formed its own sub-experiment, and was counterbalanced in a Latin Square design, with 6 items per trigger appearing in each condition. Once assigned to one of the Latin square groups in a given list, each participant saw 48 items

[5] It's technically possible to state a processing account where R-L filtering is as costly as local accommodation. But such an account isn't effectively distinguishable from an account where there is no R-L filtering, but only local accommodation, at least on the acceptability measure we are using here.

[6] On all accounts negation projects its presuppositions. So projection for the antecedent of *if not* depends solely on the projection rule for *if*.

(24 critical and 24 filler) in random order. They were presented with a context and a sentence and were asked to judge how natural the sentence sounds in the given context on a 1-7 scale. For *unless* we collected responses from 106 people in the Ps list, and from 95 people in the NoPs list (participants were recruited from the UPenn subject pool). For *if not* we collected responses from 100 people in the Ps list, and from 99 people in the NoPs list (participants were recruited on Prolific).

2.4 Results

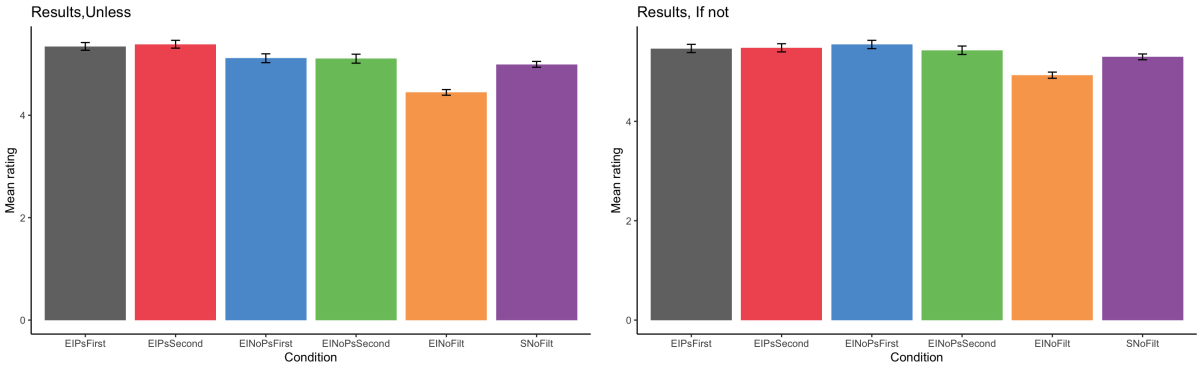


Figure 1: Results. Mean rating by condition, for each connective. Error bars represent standard error. EI-NoFilt collapses the results for EI-NoFiltFirst and EI-NoFiltSecond. Similarly for S-NoFilt.

As is clear from Fig 1, *unless* and *if not* showed parallel behavior. Firstly, for both connectives Order has no effect on acceptability either in the Ps or the NoPs conditions, suggesting that there is no Order × Ps interaction. Indeed fitting a cumulative mixed-effects model predicting Rating from Order, Ps and their interaction (including random intercepts for participant and item), is consistent with this (Unless: $\beta = -0.02,\ SE = 0.15,\ z = -0.19,\ p = 0.8$, If not: $\beta = -0.05,\ SE = 0.16,\ z = -0.3,\ p = 0.7$).[7]

Given that this is a null result, we used a Bayesian approach to quantify our credence in the model without the Order × Ps interaction. We conducted Bayesian analyses of the data for both connectives, using the data from Mandelkern, Zehr, et al. 2020 in order to extract priors.[8] We then proceeded to conduct a Bayes factor analysis. For both connectives, we found extremely strong evidence for the model without the interaction ($BF_{10} < 0.01$, Jeffreys 1939).

To evaluate the Filt × PriorS interaction, we fit (for both connectives) a cumulative mixed-effects model predicting Rating from Filt, PriorS and their interaction (including random intercepts for participant and item, as well as a by-participant random slope for Filt, and by-item random slopes for Filt and PriorS). The interaction comes out significant for both connectives (Unless: $\beta = -0.53,\ SE = 0.17,\ z = -3.05,\ p = 0.002$, If not: $\beta = -0.45,\ SE = 0.21,\ z = -2.07,\ p = 0.03$). This is driven by the fact that there is a difference between EI-NoFilt vs. S-NoFilt, but no difference between PsFirst vs. PsSecond.

3 General Discussion

The summary in Table 1 points to R-L filtering for *unless/If not* without any additional cost (relative to L-R filtering), parallel to disjunction (Kalomoiros and Schwarz 2024). Thus, this strengthens the evidence that a processing approach that postulates a uniform L-R bias across connectives is ultimately untenable.

[7] In all the mixed-effects models reported in this paper, all factors were sum-coded.
[8] This was possible because our design is parallel to Mandelkern, Zehr, et al. 2020 design.

Prediction	Processing	Lexical (asym)	Lexical (sym)	Results
Order × Ps	✓	✓	✘	✘
Filt × PriorS	✓	✘	✓	✓

Table 1: Predictions + Results.

Another important question was whether filtering presuppositions was allowed even in sentence-final antecedents in the L-R direction. Our results show that at least for negated antecedents, filtering based on preceding consequents is indeed possible.[9]

From the types of theories we have reviewed in this paper, only a lexical view where antecedents are assumed to always allow for both L-R and R-L filtering is consistent with our results. Of course, even if such an approach turns out to be descriptively adequate (more on this below), it still doesn't solve the explanatoriness problem that lexical approaches face. On the contrary, the fact that two essentially equivalent connectives (*unless* and *if not*) show the same filtering behavior raises the suspicion that filtering properties derive from truth conditions, and aren't just a matter of stipulation.

At this point, two other, more recent, approaches should be mentioned which are compatible with our data: the *'disappointment'* approach in George 2008, and the *Limited Symmetry* approach in Kalomoiros 2023. Both are designed to predictively derive asymmetric filtering for conjunction, and symmetric filtering for disjunction and (at least some) conditionals. George 2008 is a trivalent theory, while Kalomoiros 2023 takes the processing approach of Schlenker 2008 as its starting point, and modifies it to predict the relevant (a-)symmetries.

Can we further differentiate between these three approaches? Yes. Their predictions come apart when looking at **conditionals with positive antecedents**. The symmetric lexical approach and the trivalent *'disappointment'* approach predict that presuppositions in antecedents can always be filtered by the negation of the consequent, regardless of the polarity of the antecedent. This prediction is a general feature of the compositional way in which these approaches are stated: the presuppositions of an operator derive solely from the presuppositions of its arguments. Since presuppositions project from negation, a positive antecedent A and its negation $\neg A$ are always presuppositionally equivalent; hence, $A \rightarrow B$ and $\neg A \rightarrow B$ are predicted to have the same presuppositions.

However, the *Limited Symmetry* approach predicts that conditionals with negated antecedents (and hence also *unless*) should be symmetric, but once the negation is removed, asymmetric L-R filtering becomes the default (see Kalomoiros 2023).

Looking forward, our plan is to to test this fascinating split in predictions, hopefully shedding further light into the thorny issues of (a-)symmetries in conditional constructions.[10]

References

Abusch, Dorit (2010). "Presupposition triggering from alternatives". In: *Journal of Semantics* 27.1, pp. 37–80.

Chemla, Emmanuel and Lewis Bott (2013). "Processing presuppositions: Dynamic semantics vs pragmatic enrichment". In: *Language and Cognitive Processes* 38.3, pp. 241–260. ISSN: 0169-0965.

[9]The triggers we tested did not include factives or change-of-state verbs, as in the examples discussed in Mandelkern and Romoli 2018, hence we cannot rule out at this stage that those type of triggers might behave differently in these environments.

[10]As a preview, initial results from deploying our experimental design on conditionals with positive antecedents point to an asymmetry, with R-L filtering being more costly than L-R.

Chemla, Emmanuel and Philippe Schlenker (2012). "Incremental vs. symmetric accounts of presupposition projection: An experimental approach". In: *Natural Language Semantics* 20.2, pp. 177–226.

Evans, Gareth (1977). "Pronouns, Quantifiers and Relative Clauses". In: *Canadian Journal of Philosophy* 7, pp. 467–536.

Gazdar, Gerald (1979). *Pragmatics, Implicature, Presuposition and Lógical Form*. New York: Academic Press.

George, Benjamin R (2008). "A new predictive theory of presupposition projection". In: *Semantics and Linguistic Theory*. Vol. 18, pp. 358–375.

Hausser, Roland (1976). "Presuppositions in Montague Grammar". In: *Theoretical Linguistics* 3, pp. 245–280.

Heim, Irene (1983). "On the projection problem for presuppositions". In: *Proceedings of WCCFL 2*. Ed. by M. Barlow, D. Flickinger, and N. Wiegand, pp. 114–125.

— (1990). "Presupposition projection". In: *Reader for the Nijmegen workshop on presupposition, lexical meaning, and discourse processes*. Ed. by R. van der Sandt. University of Nijmegen.

Hirsch, Aaron and Martin Hackl (2014). "Incremental presupposition evaluation in disjunction". In: *Proceedings of the 44th annual meeting of the North East Linguistic Society*. Vol. 1, pp. 177–190.

Jeffreys, Harrold (1939). *Theory of Probability*. Oxford: Clarendon Press.

Kalomoiros, Alexandros (2023). "Presupposition and its (A-)symmetries". PhD thesis. University of Pennsylvania.

Kalomoiros, Alexandros and Florian Schwarz (2024). "Presupposition projection from 'and' vs 'or': Experimental data and theoretical implications". In: *Journal of Semantics*.

Karttunen, Lauri (1973). "Presuppositions of compound sentences". In: *Linguistic Inquiry* 4.2, pp. 169–193.

— (1974). "Presupposition and linguistic context". In: *Theoretical Linguistics* 1.1-3, pp. 181–194.

Mandelkern, Matthew and Jacopo Romoli (2017). "Parsing and presuppositions in the calculation of local contexts". In: *Semantics and Pragmatics* 10.

— (2018). "Hurford Conditionals". In: *Journal of Semantics* 35.2. _eprint: https://academic.oup.com/jos/article-pdf/35/2/357/24786773/ffx022.pdf, pp. 357–367. ISSN: 0167-5133. DOI: 10.1093/jos/ffx022.

Mandelkern, Matthew, Jérémy Zehr, et al. (2020). "We've discovered that projection across conjunction is asymmetric (and it is!)" In: *Linguistics and Philosophy* 43.5, pp. 473–514.

Rothschild, Daniel (2011). "Explaining presupposition projection with dynamic semantics". In: *Semantics and Pragmatics* 4, pp. 1–43.

Schlenker, Philippe (2008). "Be articulate: A pragmatic theory of presupposition projection". In: *Theoretical Linguistics*, pp. 157–212.

— (2009). "Local contexts". In: *Semantics and Pragmatics* 2, pp. 1–78.

Simons, Mandy (2001). "On the Conversational Basis of Some Presuppositions". In: *Semantics and Linguistic Theory* 11.

Soames, Scott (1989). "Presupposition". In: *Handbook of philosophical logic*. Springer, pp. 553–616.

Stalnaker, Robert (1974). "Pragmatic Presuppositions". In: *Semantics and Philosophy*. Ed. by Milton K. Munitz and Peter K. Unger. New York University Press, pp. 197–213.

Relative quantification and equative scope-taking

Andrew Kato
Dept. of Linguistics, UC Santa Cruz
Santa Cruz, CA, USA
anmkato@ucsc.edu

Abstract

Relative measure (RM) expressions, e.g., *60% of the olives*, are known to be amenable to a non-conservative interpretation, e.g., *60% olives*$_F$, since Ahn & Sauerland (2015a). Herein, I propose an analysis for modified RM phrases with equatives (*as much as 60%*) and with non-increasing restriction (*exactly 60%*) that combines a quantificational semantics for proportions (Pasternak & Sauerland 2022) with entity negation for pluralities (Bledin 2024; Elliott 2024). I apply this theory to the novel observation that non-conservative RM phrases can yield a cumulative reading (*exactly 2 suppliers sold exactly 60% olives*$_F$), as in van Benthem's puzzle (Benthem 1986; Brasoveanu 2013; Charlow 2021; a.o.).

1 Interpreting proportions

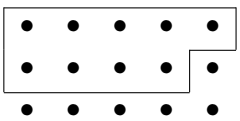
Similar in character to the proportional (1-a) and relative proportional (1-b) readings of vague *many* and *few* (Westerståhl 1985, p. 403; Herburger 1997; Partee 1989), relative measure (RM) expressions (2) can admit non-conservative readings (following Ahn & Sauerland 2015a,b, 2017).

(1) a. Many [$_S$ Scandinavians] [$_N$ have won the Nobel Prize in literature].
$\rightsquigarrow |S \cap N| \geqslant n$ (cardinal)
$\rightsquigarrow {}^{|S \cap N|}/_{|S|} \geqslant n$ (proportional)
b. Many [$_S$ Scandinavians$_F$] [$_N$ have won the Nobel Prize in literature].
$\rightsquigarrow {}^{|S \cap N|}/_{|N|} \geqslant n$ (relative proportional)

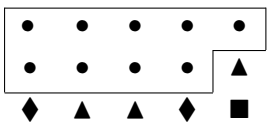
(2) a. The fruit supplier sold 60% of the olives.
b. The fruit supplier sold 60% olives$_F$.

With the removal of overt partitive material and the addition of focus-marking on the substance noun (*olives*), (2-b) yields reversed quantification. It asserts that, out of all things a particular fruit supplier sold (restrictor), 60% are olives (scope). We might visualize the state of affairs between the two readings as follows:

(3) a. 60% of the olives



b. 60% olives$_F$



Descriptively, what is partitioned in the former is the set denoted by the genitive substance noun, while it is the set denoting the sold items (of which, olives are included) for the latter.

In this paper, I address RM constructions with modified numerals in contexts that have been argued to potentially implicate other degree-related, scope-bearing elements — namely, as phrasal standards in quantity equatives and as components of non-increasing cardinality in cumulative readings (van Benthem's puzzle, Benthem 1986). The solution undertaken for the latter, which will inform the former, comes by way of extending the nascent proposal of negative individuals (Bledin 2024; Elliott 2024) to precise proportions.

2 A quantificational semantics

2.1 Pluralities with negation

Bledin (2024) observes that coordination with non-upward entailing quantifiers (*Mary and nobody else*) can follow from a compositional semantics where entity domains are 'polarized' to differentiate between positive and negative elements (Def. 2.1). Elliott (2024) picks up this proposal to develop a candidate theory of modified numerals, in comparison to post-suppositional (Brasoveanu 2010, 2013) and dynamic update (Charlow 2021) accounts.

Meta-language expressions and types are recorded in sans serif, with Function Application ($\ulcorner \alpha\ y\ x \urcorner$) left-associative and type complexity ($\ulcorner \alpha :: \phi \to \psi \urcorner$) right-associative by default.

Definition 2.1. Following Bledin (2024), we say that the *orthogonal counterpart* to an entity $x :: \mathsf{e}$ is $\neg x :: \mathsf{e}$ ('not-x'), to which *entity negation* has applied. Entity negation is assumed to apply on the atoms of the domain of entities D_{e}, but sum-closure of D_{e} excludes those with both an atom and its orthogonal counterpart as subparts, e.g., $x \oplus \neg x$ ('incoherence', Elliott 2024).

Given entity negation, it follows that the conventional denotation of a plural noun, e.g., $[\![\text{olives}]\!] = \lambda x.{}^*\mathsf{olive}\ x$, includes negative, non-negative, and mixed-polarity sums (4). What is the effect of composing a mixed-polarity predicate with a numeral n? We'll obtain the set of relevant entities with at least n-many non-negative atoms (5). If parthood and the sum relation are themselves set-based, $|x| = 3$ for $x = a \oplus b \oplus c$. Similarly, $|\cdot|^+$ counts only non-negative atoms.

(4) If $[\![\mathsf{olive}]\!] = \{a, b, c, \neg a, \neg b, \neg c\}$, $[\![{}^*\mathsf{olive}]\!] = \{a, b, c, \neg a, \neg b, \neg c, a \oplus \neg b, \ldots, a \oplus b \oplus \neg c, \ldots\}$

(5) a. $[\![\text{one olive}]\!] = \lambda x.{}^*\mathsf{olive}\ x \wedge |x|^+ \geqslant 1 \rightsquigarrow \{a, b, c, a \oplus b, a \oplus \neg b, \ldots\}$ $\mathsf{e} \to \mathsf{t}$

b. $[\![\text{two olives}]\!] = \lambda x.{}^*\mathsf{olive}\ x \wedge |x|^+ \geqslant 2 \rightsquigarrow \{a \oplus b, \ldots, a \oplus b \oplus \neg c, \ldots\}$ $\mathsf{e} \to \mathsf{t}$

2.2 *Percent* semantics

With this basic sketch, we can return to a semantics for proportions. The RM noun *percent* is typed as a degree quantifier given the denotation in (6) (adapted from Pasternak & Sauerland 2022, p. 251; Spathas 2022, p. 279), where $\mathsf{D} ::= (\mathsf{d} \to \mathsf{t}) \to \mathsf{t}$.[1] The **max** operator (7-a) returns the highest degree from a given degree predicate (Heim 2000; Buccola & Spector 2016; a.o.).

(6) $[\![\text{percent}]\!] := \lambda d \lambda D. \dfrac{\mathbf{max}\ D}{\mathbf{max}(\mathbf{dom}\ D)} \geqslant \dfrac{d}{100}$ $\mathsf{d} \to \mathsf{D}$

(7) a. $[\![\mathbf{max}]\!] := \lambda D \iota d.D\ d \wedge \forall d'[D\ d' \to d' \leqslant d]$ $(\mathsf{d} \to \mathsf{t}) \to \mathsf{d}$

b. $[\![\mathbf{size}]\!] := \lambda d \lambda x.|x|^+ \geqslant d$ $\mathsf{d} \to \mathsf{e} \to \mathsf{t}$

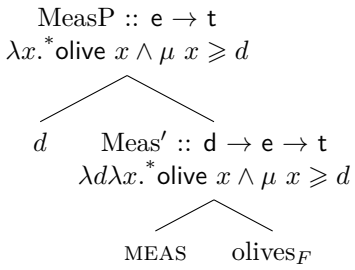
To mesh an atomic-degree denotation for numerals (as needed for (6)) with their intersective meaning from (5), I'll assume a type-shifter from the former to the latter (7-b). I notate it as **size**, though it goes under different guises across the literature (see Bylinina & Nouwen 2020 for an overview). Arriving at the Predicate Modification in (5) is thus preceded by applying (7-b).

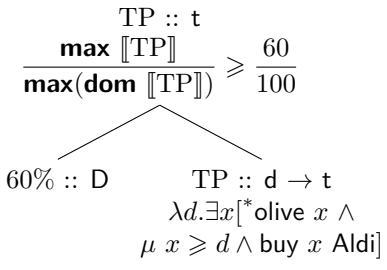
(8) 60% λd the fruit supplier sold d-MEAS olives$_F$ (cf. (2-b))

With *percent* and its numeral forming a separate constituent, attempting to compose with a gradable predicate ($\mathsf{d} \to \mathsf{e} \to \mathsf{t}$) would fail. The natural solution is for the RM noun to occupy a higher scope (QR) and saturate the degree argument with its trace, i.e, (8). The substance noun is shifted to a gradable denotation with an off-the-shelf measure operator (Solt's (2015)

[1] For space, I restrict attention to *percent* (which I use interchangeably with %) and set aside other RM nouns, e.g., *half*, *third*, *quarter*, which may warrant separate semantic (cf. Benbaji-Elhadad & Wehbe 2024) or morphosyntactic (cf. Pasternak & Sauerland 2022, pp. 236–237) treatments.

MEAS/Rett's (2014) M-OP). The basic picture of (2-b) is depicted as follows, where (9) shows the launch site of the degree quantifier, and (10), the landing site (intensional descriptions omitted; subject a proper name for simplicity).[2]

(9)


(10)


For count nouns, e.g., units of fruit, the (contextual) measure function μ will amount to the non-negative cardinality from earlier.[3] For the regular, conservative use (e.g., (2-a)), I assume *percent* is shifted to a function that remains in-situ and expresses a parthood relation with the e-type genitive noun (Pasternak & Sauerland 2022).

3 Scopal interactions

We now turn to the issue of cumulativity (11) that §2's machinery has built up to. (11) is true when the maximum number of interviewing recruiters is 2, and the maximum proportion of women interviewed by recruiters (out of all interviewees) is 60%.

(11) Exactly two recruiters interviewed exactly 60% women$_F$ (between them).

This is, when excluding *between them* in (11), in addition to a distributive reading with surface scope — such that two (and only two) recruiters each interviewed 60% women and 40% non-women. The cumulative intuition does not fall out from the assumption that *exactly two recruiters* is simply existential, i.e., that there exists a two-member plurality of recruiters (since (11) should be false if there are, e.g., three interviewing recruiters). The challenge, moreover, is how to prevent (11) from yielding truth when there are multiple different combinations of interviewing recruiters such that each combination has the 60%-40% split.

This challenge is 'van Benthem's problem' (following Benthem 1986), which has elicited a variety of proposals about the scopal properties of non-increasing modified numerals (Krifka 1999; Brasoveanu 2013; Alxatib & Ivlieva 2018; Haslinger & Schmitt 2020; Charlow 2021; a.o.). RM nouns serve as a unique angle into this problem — since they can be employed in modified numeral constructions (unlike, e.g., **{exactly, less than, at most} many recruiters*) while also being degree-quantificational.

Enrichment with entity negation as sketched in §2 allows for a straightforward approach to this phenomenon following Elliott (2024): *exactly two recruiters* will denote a maximal plurality of the set of recruiters P with two non-negative members and all other elements from P encoded as negative members. To get there, Elliott (2024) leans crucially on an assumption that plural nouns must themselves only denote maximal sums, e.g., where *olive in (4) would only contain the three-member elements. Without it, we'd be back at the original challenge, since sums without negative members for all other recruiters would also be admitted (no maximality effect would be predicted). Rather than generalize such an assumption to all plural nouns, I'll encode it into the

[2] Accessing Roothian focus alternatives to the substance noun will allow for the denominator to make reference to a domain larger than the noun itself. See Pasternak & Sauerland (2022, pp. 257–261) for additional discussion.

[3] Another strategy for measuring cardinality would be to gather the non-negative subparts (if any) first, and then feed them as a set or distinct sum to a non-polarity-specific $|\cdot|$.

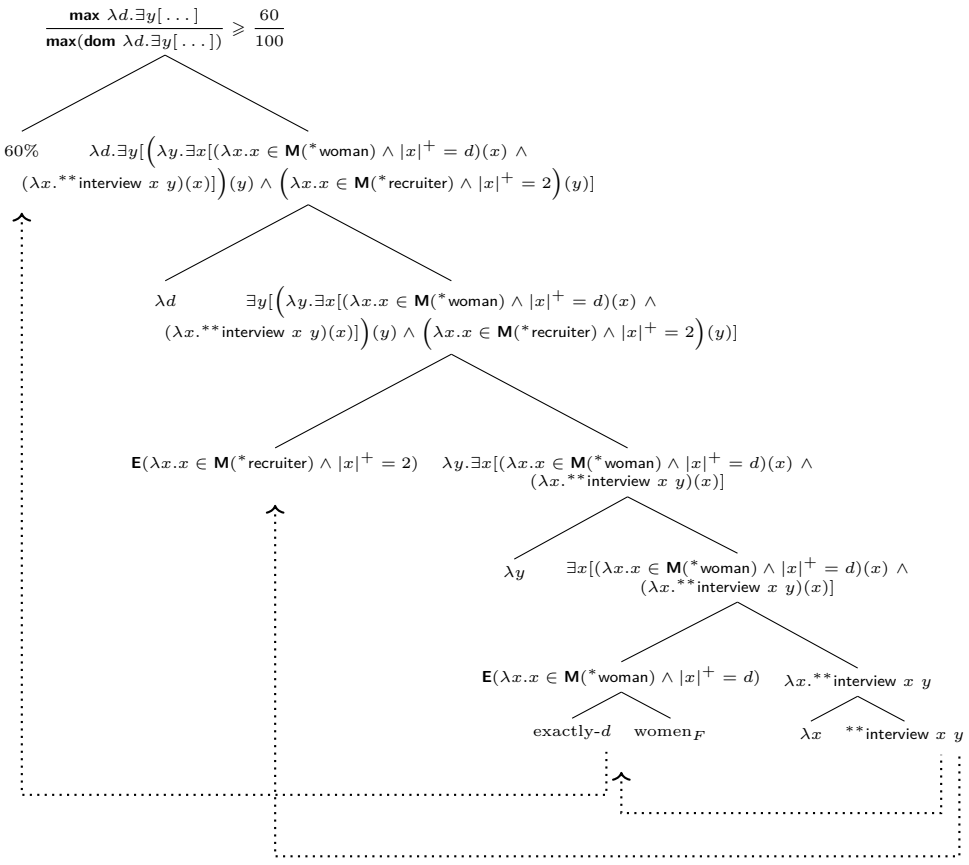


Figure 1: Scoping a modified proportion

denotation of the modifier. Whereas the predicate meaning of *two* is the set of sums with two or more non-negative members (12-a), *exactly two recruiters* denotes the set of maximal sums in *recruiter with exactly two non-negative members (12-b)-(12-c). The operator gathering the maximal sums (notated as **M** to differentiate from **max**) will return a set with multiple elements, since sets with entity negation will not have a unique maximum, e.g., $\mathbf{M}(\{a, \neg a\}) = \{a, \neg a\}$.

(12) a. $[\![\textbf{size}\ \text{two}]\!] = \lambda x.|x|^+ \geqslant 2 \rightsquigarrow \{a \oplus b \oplus c, a \oplus b \oplus \neg c, a \oplus \neg b \oplus c, \neg a \oplus b \oplus c, a \oplus b, a \oplus c, b \oplus c\}$
b. $[\![\text{exactly}]\!]([\![\text{two}]\!]) = \lambda P \lambda x.x \in \mathbf{M}(P) \wedge |x|^+ = 2 \rightsquigarrow \{a \oplus b \oplus \neg c, a \oplus \neg b \oplus c, \neg a \oplus b \oplus c\}$
c. $[\![\text{exactly}]\!]([\![\text{two}]\!])([\![\text{recruiters}]\!]) = \lambda x.x \in \mathbf{M}(*\text{recruiter}) \wedge |x|^+ = 2$

Extending entity negation to modified precise proportions from (11), the degree trace that *60%* abstracts over (due to type-clash) serves as an argument to the modifier (Fig. 1).[4] Aside from this, composition proceeds in the same fashion as (12-c). The modifier undertakes the task that would normally be attributed to a measure operator, so no MEAS applies in the lower e-type quantifier.[5] The object quantifier will need to scope under the RM so the d-trace is bound.[6]

We can also consider cases besides non-monotone quantifiers, as in quantity equatives (13), without a RM modifier that introduces maximality or an 'exact'-reading. As one example,

[4]For a polarity-sensitive variant of the cumulation operator ** for verbal predicates (Sternefeld 1998; Beck & Sauerland 2000), see Elliott (2024). It stipulates a given relation must hold only of non-negative subparts.

[5]See (19) in the Appendix for a denotation that permits MEAS to always apply to the substance noun by accepting a gradable predicate.

[6]**E** is the type-shifter from predicates to quantifiers that introduces existential force (i) (following Partee 1987).

(i) $\mathbf{E}\ P = \lambda Q.\exists x[P\ x \wedge Q\ x]$.

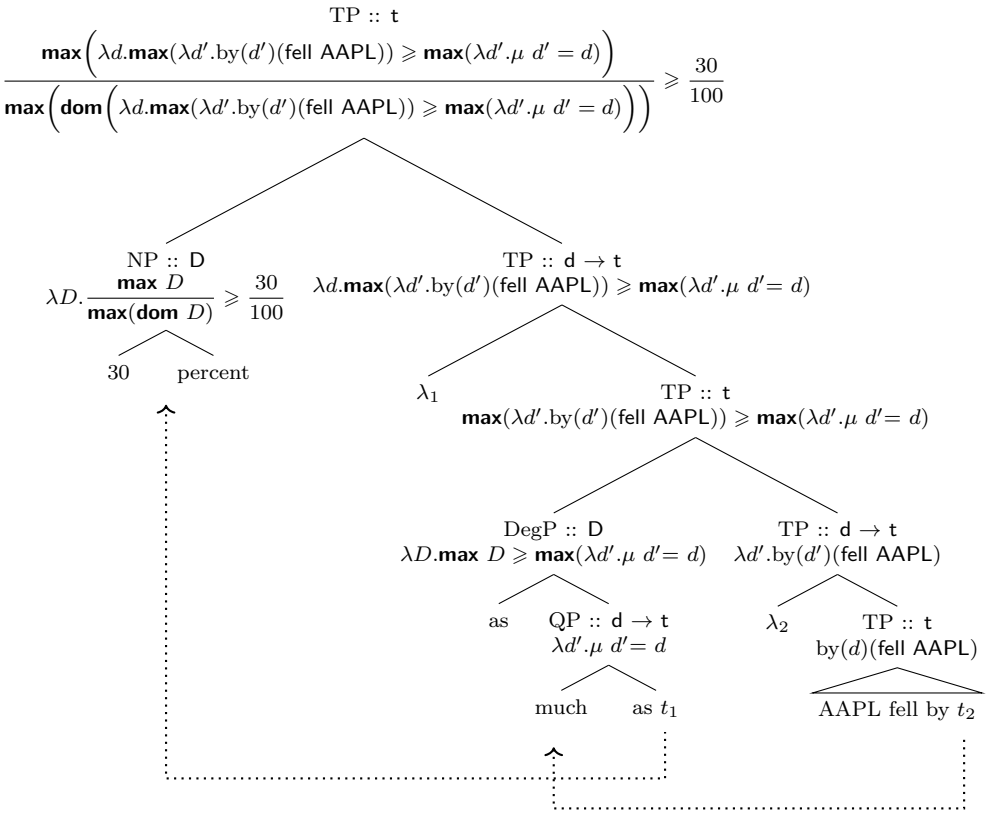


Figure 2: A proportion without a non-monotone quantifier

equatives are themselves typically understood to implicate degree quantification, as with the entry for the equative operator (the first *as* in, e.g., *as much as*) in (14) (e.g., Heim 2000; Rett 2010, 2015, 2020). In this case, the resulting reading for the RM phrase is still an 'at-least' one.

(13) a. Fanta contains as much as 30% sugar$_F$.
b. The price fell by as much as 30%.

(14) $[\![\text{as}]\!] := \lambda D \lambda D'.\textbf{max}\ D' \geqslant \textbf{max}\ D$ $\quad (\mathsf{d} \to \mathsf{t}) \to \mathsf{D}$

Accounting for the LF of (13), then, means considering the scope-taking potential of both the equative operator and the RM noun, as in Fig. 2.[7] While a degree-quantifier approach to percentages is not the only one available to us — e.g., the degree-relational denotation (16-a) from Gobeski & Morzycki (2017), who primarily focus on reduced clausal equatives — we must avoid type-conflicts with the quantity adjective (16-b) (entry adapted from Coppock & Bogal-Allbritten 2018). Evident in (15), however, the modifier is vacuous in its contribution, hence the persistence of the 'at-least' reading.

(15) $[\![\text{Fanta contains as much as 30\% sugar}_F.]\!] \rightsquigarrow$

$$\frac{\textbf{max}\Big(\lambda d.\textbf{max}(\lambda d'.\mathsf{contains}(d'\text{-MEAS sugar})(\mathsf{Fanta})) \geqslant \textbf{max}(\lambda d'.\mu\ d' = d)\Big)}{\textbf{max}\Big(\textbf{dom}\Big(\lambda d.\textbf{max}(\lambda d'.\mathsf{contains}(d'\text{-MEAS sugar})(\mathsf{Fanta})) \geqslant \textbf{max}(\lambda d'.\mu\ d' = d)\Big)\Big)} \geqslant \frac{30}{100}$$

(16) a. $[\![10\%]\!] = \lambda d.10\% \times d$ (Gobeski & Morzycki 2017, p. 727)
b. TP-[by [as $\boxed{\text{[much as 30\%]}}$]] $\rightsquigarrow$ $[\lambda d \lambda d'.\mu\ d' = d]([\![\text{G\&M17 } 30\%]\!])$ :: ✗

[7] For a fine-grained analysis of RM constructions that properly accounts for the degree-of-change meaning that (13-b) reflects, see Spathas (2024).

4 Concluding discussion

This paper provides a first analysis of the novel observation that non-conservative proportions can be modified and receive cumulative and non-cumulative readings. The corresponding analysis combines a quantificational theory of proportions (Pasternak & Sauerland 2022) with a theory of entity negation (Bledin 2024; Elliott 2024) to derive the correct 'exact'-interpretation.

Acknowledgements. For earlier discussion and feedback, many thanks to Dylan Bumford, Jess Law, Haoze Li, Aleksandre Maskharashvili, Maziar Toosarvandani, and the anonymous reviewers for the 2024 Amsterdam Colloquium.

A Bare proportions and degree pluralities

Li (2022) provides the following judgement for Mandarin (17), where the unmodified proportional DP must take narrow scope relative to the modified-numeral subject. This is taken to be an instantiation of what, e.g., Bylinina & Nouwen (2020) and Gajewski (2008) refer to as the Heim-Kennedy generalization (HKG; following Kennedy 1997; Heim 2000), which stipulates that degree quantifiers cannot scope over e-type quantifiers.

(17) Zhenghao you liang-ge xuesheng du-le sanfenzhiyi de xiaoshuo$_F$.
exactly have two-CL students read-PRF one.third DE novels
a. 'There are exactly two students such that they each read one third novels$_F$.'
b. #'One third of the books that exactly two students read were novels$_F$.'
(Li 2022, pp. 10, 26)

The narrow scope in (17) is explained as undefinedness from the failure to gather a unique minimum degree, within the context of a degree-plurality framework (see Li 2022 for analysis; see Dotlačil & Nouwen 2016 and Nouwen & Dotlačil 2017 for degree pluralities). If RM expressions modified by *exactly* indeed take scope, why do we not observe infelicity for the cumulative reading from §3? For space, I simply point out that the modified proportion would *not* violate the variant of the HKG suggested by Crnič (2017) in (18-a), if we 'pluralize' our ontology for degrees and liken *exactly* to SHIFT (both change an 'at-least' reading to 'exactly') in (18-b).

(18) a. Modified HKG: If the scope of an e-type quantifier contains the trace d of a degree quantifier, d must be an argument to SHIFT.
b. $[\![\text{SHIFT}]\!] := \lambda d\lambda A\lambda x.\mathbf{max}(\lambda d'.A\ d'\ x) \sqsubseteq d$ $\quad \mathsf{d} \to (\mathsf{d} \to \mathsf{e} \to \mathsf{t}) \to \mathsf{e} \to \mathsf{t}$
(adapted from Crnič 2017)

Under such assumptions, we predict why the proportion can take scope over the modified-numeral subject specifically when modified by *exactly*. The resulting denotation for a modified RM phrase in (19) also preserves the structural consistency of MEAS applying to the substance noun, though the maximality condition would still need to be enforced.

(19) exactly-d women$_F$ $\rightsquigarrow \lambda x.\mathbf{max}(\lambda d'.(\text{MEAS}\ {}^*\mathsf{woman})\ d'\ x) \sqsubseteq d$ (cf. Fig. 1)

Just as we may pluralize degrees, we can also consider what explanatory possibilities arise from polarizing them (as Bledin 2024 does for entities). I leave this to future work, but note that a potential avenue may be via corrective-*but* applying to degrees (or their $\mathsf{e} \to \mathsf{t}$ equivalent), e.g., (20) — where, under a notion of 'degree negation', *not 20 but 30* might denote the sum $\neg 20 \oplus 30$.

(20) a. The fruit supplier sold not 20 but 30% olives$_F$.
b. The basketball team lost by not 20 but 30 points.

References

Ahn, Dorothy and Uli Sauerland (2015a). "Non-conservative quantification with proportional quantifiers: Crosslinguistic data". In: *Proceedings of NELS 45*. Ed. by Thuy Bui and Deniz Ozyıldıze, pp. 1–11.

— (2015b). "The grammar of relative measurement". In: *Proceedings of SALT 25*. Ed. by Sarah D'Antonio, Mary Moroney, and Carol Rose Little, pp. 125–142.

— (2017). "Measure constructions with relative measurements: Towards a syntax of non-conservative construals". In: *The Linguistic Review* 34, pp. 215–248.

Alxatib, Sam and Natasha Ivlieva (2018). "Van Benthem's problem, exhaustification, and distributivity". In: *Proceedings of Sinn und Bedeutung 21*. Truswell, Robert et al., pp. 1–18.

Beck, Sigrid and Uli Sauerland (2000). "Cumulation is needed: A reply to Winter (2000)". In: *Natural Language Semantics* 8, pp. 349–371.

Benbaji-Elhadad, Ido and Jad Wehbe (2024). "On the semantics of unit fractions". Manuscript (to appear in *Proceedings of Sinn und Bedeutung 28*), Massachusetts Institute of Technology.

Benthem, Johan van (1986). *Essays in logical semantics*. Dordrecht: Springer.

Bledin, Justin (2024). "Composing menus". Manuscript, Johns Hopkins University.

Brasoveanu, Adrian (2010). "Modified numerals as post-suppositions". In: *Logic, language, and meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 16-18, 2009, revised selected papers*. Ed. by Maria Aloni et al. Berlin: Springer, pp. 203–212.

— (2013). "Modified numerals as post-suppositions". In: *Journal of Semantics* 30, pp. 155–209.

Buccola, Brian and Benjamin Spector (2016). "Modified numerals and maximality". In: *Linguistics & Philosophy* 39, pp. 151–199.

Bylinina, Lisa and Rick Nouwen (2020). "Numeral semantics". In: *Language and Linguistics Compass* 14, pp. 1–18.

Charlow, Simon (2021). "Post-suppositions and semantic theory". Manuscript to appear in *Journal of Semantics*.

Coppock, Elizabeth and Elizabeth Bogal-Allbritten (2018). "Forces at the inferface of gradability and quantification". In: *Proceedings of SALT 28*. Ed. by Sireemas Maspong et al., pp. 747–767.

Crnič, Luka (2017). "More on *less*". In: *Proceedings of NELS 47*. Ed. by Andrew Lamont and Katerina Tetzloff. Amherst, MA: GLSA.

Dotlačil, Jakub and Rick Nouwen (2016). "The comparative and degree pluralities". In: *Natural Language Semantics* 24, pp. 45–78.

Elliott, Patrick D. (2024). "Mixed-polarity pluralities: A solution to van Benthem's problem". Paper presented at Sinn und Bedeutung 29 (Noto).

Gajewski, Jon (2008). "More on quantifiers in comparative clauses". In: *Proceedings of SALT XVIII*. Ed. by Tova Friedman and Satoshi Ito, pp. 340–357.

Gobeski, Adam and Marcin Morzycki (2017). "Percentages, relational degrees, and degree constructions". In: *Proceedings of SALT 27*, pp. 721–737.

Haslinger, Nina and Viola Schmitt (2020). "Cumulative readings of modified numerals: A plural projection approach". In: *Proceedings of Sinn und Bedeutung 24*. Ed. by Michael Franke et al., pp. 323–240.

Heim, Irene (2000). "Degree operators and scope". In: *Proceedings of SALT X*. Ed. by Jackson Brendan and Tanya Matthews. Ithaca, NY: Cornell University, pp. 40–64.

Herburger, Elena (1997). "Focus and weak noun phrases". In: *Natural Language Semantics* 5, pp. 53–78.

Kennedy, Christopher (1997). "Projecting the adjective". PhD thesis. University of California, Santa Cruz.

Krifka, Manfred (1999). "At least some determiners aren't determiners". In: *The semantics/pragmatics interface from different points of view*. Ed. by Ken Turner. vol. 1, Amsterdam: Elsevier, pp. 257–291.

Li, Haoze (2022). "Relative measurement and scope in Mandarin". In: *Glossa: A Journal of General Linguistics* 7, pp. 1–40.

Nouwen, Rick and Jakub Dotlačil (2017). "The scope of nominal quantifiers in comparative clauses". In: *Semantics & Pragmatics* 10, pp. 1–22.

Partee, Barbara (1987). "Noun phrase interpretation and type-shifting principles". In: *Studies in discourse representation theories and the theory of generalized quantifiers*. Ed. by Jeroen Groenendijk, Dick de Jongh, and Martin Stokhof. Dordrecht: Foris, pp. 115–144.

— (1989). "Many quantifiers". In: *ESCOL 89: Proceedings of the Eastern States Conference on Linguistics*. Ed. by Joyce Powers and Kenneth de Jong. Ohio State University, pp. 383–402.

Pasternak, Robert and Uli Sauerland (2022). "German measurement structures: Case-marking and non-conservativity". In: *Journal of Germanic Linguistics* 25, pp. 221–272.

Rett, Jessica (2010). "Equatives, measure phrases, and NPIs". In: *Logic, language, and meaning: 17th Amsterdam Colloquium, Amsterdam, The Netherlands, December 16-18, 2009, revised selected papers*. Ed. by Maria Aloni et al. Berlin: Springer, pp. 364–373.

— (2014). "The polysemy of measurement". In: *Lingua* 143, pp. 242–266.

— (2015). "Measure phrase equatives and modified numerals". In: *Journal of Semantics* 32, pp. 425–475.

— (2020). "Separate but equal: A typology of equative constructions". In: *Interactions of degree and quantification*. Ed. by Peter Hallman. Brill, pp. 163–204.

Solt, Stephanie (2015). "Q-adjectives and the semantics of quantity". In: *Journal of Semantics* 32, pp. 221–273.

Spathas, Giorgios (2022). "Domain restricted measure functions and the extent readings of relative measure". In: *Measurements, numerals and scales: Essays in honour of Stephanie Solt*. Ed. by Nicole Gotzner and Uli Sauerland. Palgrave Macmillan, pp. 275–288.

— (2024). "Measuring change with precise proportions". Paper presented at Sinn und Bedeutung 29 (Noto).

Sternefeld, Wolfgang (1998). "Reciprocity and cumulative predication". In: *Natural Language Semantics* 6, pp. 303–337.

Westerståhl, Dag (1985). "Logical constraints in quantifier languages". In: *Linguistics and Philosophy* 8, pp. 387–413.

Measuring event participants

Boram Kim
Department of Linguistics, UCLA
Los Angeles, CA, U.S.A.
bkim21@ucla.edu

Abstract

Measure nouns like *percent* can be interpreted reverse proportionally, as in "The company hired 50% women." The most influential analyses of this phenomenon derive it via association with focus, which is known to invert restrictor-scope relations for adnominal markers like *only* and *few*. I offer three arguments that reverse proportional readings in Korean are not focus-sensitive, and are instead simply adverbial modifiers quantifying over event participants, in the spirit of Nakanishi's (2007) analysis of bare numerals. I suggest that this opens the door to possible re-analyses of the data crosslinguistically.

1 Introduction

In their foundational work on quantification in natural language, Barwise and Cooper (1981) and Keenan and Stavi (1986) famously conjectured that all quantificational determiners are conservative on their second argument.

(1) Conservativity universal: For any determiner D, and sets R and S, $D(R)(S) = D(R)(R \cap S)$.

However, there are apparent counterexamples. Measure nouns like *percent* can be interpreted nonconservatively. If *50%* were conservative in (2), it should be possible to assess the truth by restricting attention to just the set of women. But its most natural reading requires comparing the proportion of women among the hires to that of non-women, ostensibly violating conservativity.

(2) The company hired 50% women
'50% of women were hired by the company.' (✗proportional)
'50% of the hires were women.' (✓reverse proportional)

Ahn and Sauerland (2017) argue that these violations are only apparent. They assume that focus is required on *women* for *50%* to be interpreted reverse proportionally, and propose that the measure phrase undergoes covert focus movement. Therefore, at LF, the restrictor of *50%* is a contextual parameter bound to the focus alternatives of *50%*'s syntactic restrictor, as in (3).

(3) [50% C] [∼C [λx the company hired [x women$_{\mathbf{F}}$]]]

Other proposals on non-conservativity share the key assumption of Ahn and Sauerland's analysis that focus is required on the NP following the measure phrase for reverse proportional readings to arise, and covert focus movement is what derives the reverse proportional interpretation compositionally (e.g., Li 2022; Pasternak and Sauerland 2022).

In support of this nominal displacement derivation, Ahn and Sauerland show that similar readings emerge in Korean, as in (4b) and (5b), where the measure phrases are clearly pronounced outside of the corresponding nominals, overtly reifying the implicit scope-taking in (3).

(4) a. [haksayng-uy **isip** **phulo**-ka] chamsekhayssta
student-GEN twenty percent-NOM attended

'20% of the students attended.' (✓proportional)

b. [haksayng-i] **isip** **phulo** chamsekhayssta
student-NOM twenty percent attended

'20% of the students attended.' (✓proportional)
'20% of the attendees were students.' (✓reverse proportional)

(5) a. hoysa-ka [isiptay ciwenca-uy **isip** **phulo**-lul] chayyonghayssta
company-NOM twenties applicant-GEN twenty percent-ACC hired

'The company hired 20% of the applicants who were in their twenties.' (✓proportional)

b. hoysa-ka [isiptay ciwenca-lul] **isip** **phulo** chayyonghayssta
company-NOM twenties applicant-ACC twenty percent hired

'The company hired 20% of the applicants who were in their twenties.' (✓proportional)
'20% of the hires were in their twenties.' (✓reverse proportional)

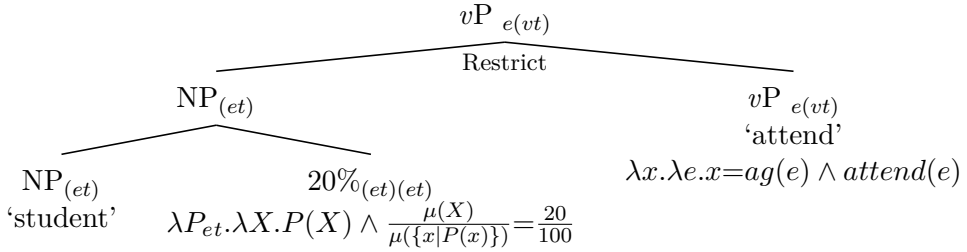
Though the word order is consistent with their analysis, they offer no particular evidence that the reverse proportional measure phrases originate within the nominal. This paper argues that they do not, and furthermore that focus is irrelevant to the reverse proportional interpretation in Korean. Instead, reverse proportional measure phrases are adverbial modifiers that involve quantification over event participants.

2 Analysis

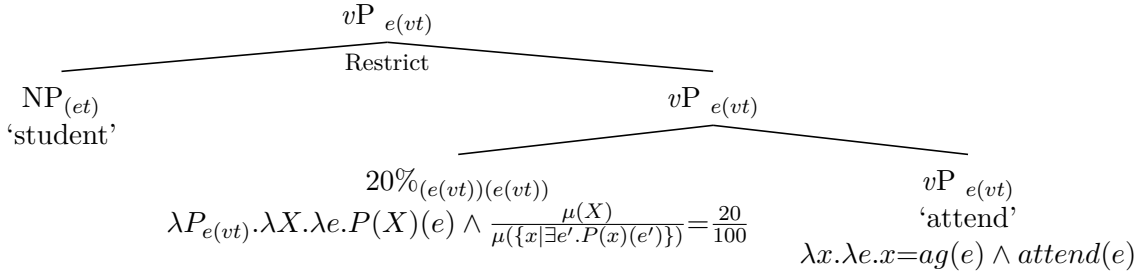
I argue that reverse proportional measure phrases in (4b) and (5b) start out syntactically within the verbal domain rather than the nominal domain, and semantically quantify over event participants rather than focus alternatives. In contrast, measure phrases inside case-marked nominals, (4a) and (5a), are nominal modifiers that quantify over individuals as usual.

(6) and (7) represent the structure and final truth-condition of subjected-oriented measurement sentences, (4a) and (4b), respectively. I assume Chung and Ladusaw's (2004) Restrict for the mode of composition between the properties et and $e(vt)$. The individual and event arguments are later bound by existential closure. x and X are both variables of type e, with the former denoting atomic entities and the latter denoting pluralities.

(6) Proportional measure phrase (4a)
$\exists X.\exists e.student(X) \wedge X{=}ag(e) \wedge attend(e) \wedge \frac{\mu(X)}{\mu(\{x|student(x)\})}{=}\frac{20}{100}$

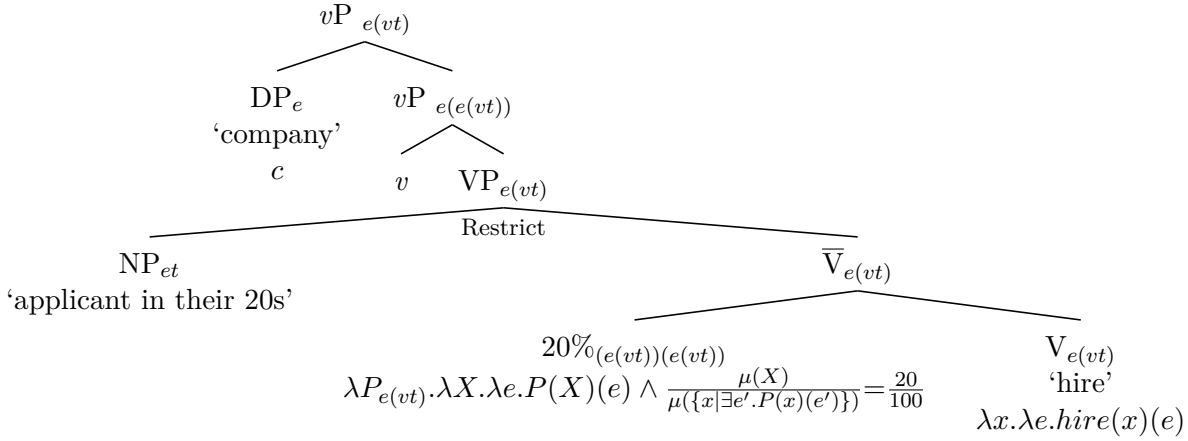


(7) Reverse proportional measure phrase (4b)
$\exists X.\exists e.student(X) \wedge X{=}ag(e) \wedge attend(e) \wedge \frac{\mu(X)}{\mu(\{x|\exists e'.x{=}ag(e')\wedge attend(e')\})}{=}\frac{20}{100}$



For measure phrases in the object position, the relevant event participants are not agents but themes. Therefore, the event participants measured in (5b) are hires.

(8) Reverse proportional measure phrase (5b)
$\exists X.\exists e.20s\text{-}applicant(X) \wedge c{=}ag(e) \wedge hire(X)(e) \wedge \frac{\mu(X)}{\mu(\{x|\exists e'.hire(x)(e')\})}{=}\frac{20}{100}$



20% in (6) and (7) is polysemous, measuring either individuals satisfying the nominal description or individuals participating in an event. The syntactic category of event-participant-measuring *20%* varies depending on the syntactic context: in the subject-oriented sentence (7), it is an adverb adjoined to *v*P, while in the object-oriented sentence (8), it adjoins to a verb head, creating a complex predicate.

The remaining proportional readings in (6b) and (7b) are due to Multiple Nominative Constructions (MNC) and Multiple Accusative Constructions (MAC), productive case configurations in the language. This is evidenced by the fact that sentences with explicit double case-marking allow only proportional readings, as in (9).

(9) haksayng-i **isip** **phulo**-ka chamsekhayssta
student-NOM twenty percent-NOM attended

'20% of the students attended.' (✓proportional)
'20% of the attendees were students.' (✗reverse proportional)

The proposed analysis is inspired by Nakanishi (2007), who argues that bare numerals in Japanese can either be nominal or verbal modifiers. Relevant data for Korean is shown in (10). While numerals do not lead to different interpretations, the idea is that 'two CL' in (10a) is

a nominal modifier, whereas in (10b) it is a verbal modifier—a distinction that can naturally extend to measure phrases, assuming classifiers correspond to measure nouns.

(10) a. [haksayng **twu myeng**-i] chamsekhayssta
student two CL-NOM attended

b. [haksayng-i] **twu myeng** chamsekhayssta
student-NOM two CL attended

'Two students attended'

3 Arguments for the analysis

3.1 Lack of focus intervention effects

Ahn and Sauerland (2017) claim that focus is required on the NP for reverse proportional interpretations to arise. Therefore, in (11), *haksayng* 'student' should be focused.

(11) Focus analysis of reverse proportional measure phrases
[haksayng$_{\mathbf{F}}$ ___$_1$ -i] [**isip** **phulo**]$_1$ chamsekhayssta
student -NOM twenty percent attended

'20% of the attendees were students.' (✓reverse proportional)

The focus analysis predicts that focus intervention effects (Beck and Kim 1997) should arise in wh-questions involving a reverse proportional measure phrase. This prediction is not borne out. With (12a) as a baseline, (12b) can still be interpreted reverse proportionally, despite the putative intervention configuration.

(12) a. *haksayng$_F$-man enu phathi-lul chamsekhayss-ni?
student-only which party-ACC attended-ITR
'Which party did only students attend?'

b. haksayng$_F$-i enu phathi-lul **isip** **phulo** chamsekhayss-ni?
student-NOM which party-ACC twenty percent attended-ITR
'Which party did 20% of the students attend?' (✓proportional)
'Which party had 20% of its attendees as students?' (✓reverse proportional)

3.2 Resistance to narrow scope negation

Second, reverse proportional measure phrases cannot outscope verbal negation.

(13) a. [haksayng-uy **isip** **phulo**-ka] chamsekhaci anhassta
student-GEN twenty percent-NOM attend not.PAST

'20% of students didn't attend.' (✓proportional)

b. [haksayng-i] **isip** **phulo** chamsekhaci anhassta
student-NOM twenty percent attend not.PAST

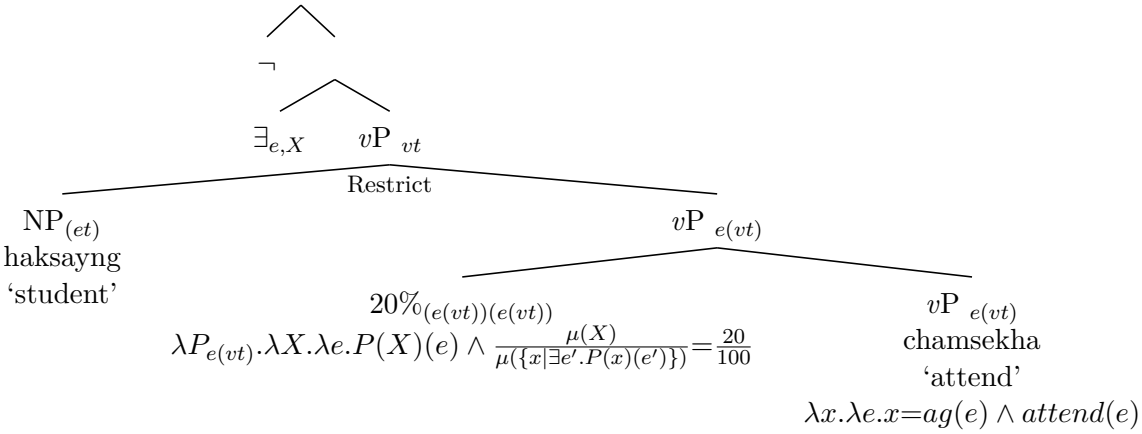
'20% of students didn't attend.' (✓proportional)
'20% of non-attendees were students.' (✗reverse proportional with verbal negation)
'It is not the case that 20% of attendees were students.' (✓reverse proportional with clausal negation)

The resistance to narrow scope negation follows from the proposed analysis, in conjunction with the standard understanding of the placement of negation within the functional hierarchy, as in (14), which is above existential closure of the event argument.

(14) subject T **neg** $\exists e$ [___ v V]

Since negation cannot occur lower than existential closure, there is no way to generate the reading that involves measuring non-attendees. The only possible structure for (13b) is (15).

(15) Reverse proportional measure phrase with negation
$\neg\exists e.\exists X.student(X) \wedge X{=}ag(e) \wedge attend(e) \wedge \frac{\mu(X)}{\mu(\{x|\exists e'.x{=}ag(e')\wedge attend(e')\})}{=}\frac{20}{100}$



In the focus-based nominal displacement analysis, it is not obvious how this restriction with verbal negation could be explained because *20%* moves to the edge of the clause and acts as a propositional operator; it would certainly be well-typed and sensible for it to alternate scopally with clausal negation, yet it does not.

3.3 Similar distributions with adverbs

Adverbs in Korean can occur flexibly either before or after an internal argument, as in (16).

(16) Nari-ka {**kippukey**} phathi-lul {**kippukey**} chamsekhayssta
Nari-NOM happily party-ACC happily attended

'Nari happily attended the party'

If the subject-oriented reverse proportional measure phrase is a verbal modifier attached to vP, as proposed in Section 2, the prediction is that it can also flexibly occur where adverbs can. (17) shows that reverse proportional readings survive regardless of the position of *20%*.

(17) haksayng-i {**isip phulo**} phathi-lul {**isip phulo**} chamsekhayssta
student-NOM twenty percent party-ACC twenty percent attended

'20% of students attended the party' (✓proportional)
'20% of party attendees were students' (✓reverse proportional)

In wh-questions, sentences degrade in acceptability when the adverb *kippukey* 'happily' appears before an internal argument that is a wh-phrase, as shown in (18). The same contrast is observed between (19a) and (19b) under reverse proportional readings, which did not exist in their declarative counterparts.

(18) Nari-ka ?{**kippukey**} enu phathi-lul {**kippukey**} chamsekhayss-ni?
Nari-NOM happily which party-ACC happily attended-ITR

'Which party did Nari happily attend?'

(19) a. haksayng-i enu phathi-lul **isip** **phulo** chamsekhayss-ni?
student-NOM which party-ACC twenty percent attended-ITR

'Which party did 20% of the students attend?' (✓proportional)
'Which party had 20% of its attendees as students?' (✓reverse proportional)

b. haksayng-i **isip** **phulo** enu phathi-lul chamsekhayss-ni?
student-NOM twenty percent which party-ACC attended-ITR

'Which party did 20% of the students attend?' (✓proportional)
'Which party had 20% of its attendees as students?' (?reverse proportional)

4 Conclusion

This paper argued that measure phrase modifiers in Korean are polysemous, involving different types of quantification: one over individuals and the other over event participants. I conclude by suggesting the analytical possibility that reverse proportional measure phrases in German may also be modifiers quantifying over event participants.

(20) **Zwanzig Prozent** Studenten haben für Harris gestimmt
twenty percent students have for Harris voted

'20% of the voters for Harris were students.' (✓reverse proportional)

Pasternak and Sauerland (2022) present convincing arguments that the German reverse proportional measure phrase *zwanzig Prozent* 'twenty percent' in sentences like (20) is not equivalent to *zu zwanzig Prozent* 'at twenty percent,' which also induces a reverse proportional reading. Their arguments are based on case, agreement, and constituency patterns specific to *zu*-phrases. While Pasternak and Sauerland may be correct that preposition-less measure phrases are not covertly *zu*-marked, *zu*-phrases are not the only adverbials that appear next to NPs, seemingly violating the verb-second constraint in German. Following Meyer and Sauerland (2009), they refer to this class as 'DP-adjoinable adverbials', which includes *größtenteils* 'for the most part', *ausschließlich* 'exclusively', and *nur* 'only'.

(21) **Größtenteils** Studenten haben für Harris gestimmt
mostly students have for Harris voted

'Most of the voters for Harris were students.' (✓reverse proportional)

I suggest that this latter class of adverbials is derivationally analogous to the reverse proportional *20%*, quantifying over event participants, given that adverbials such as *größtenteils* can receive an event-based analysis for event predication (Zobel 2021). It is still possible that these adverbials have a distinct derivation from the PP *zu 20%*, even if Pasternak and Sauerland are right that *20%* is not equivalent to *zu 20%*. The only difference from Korean reverse proportional measure phrases is that the German measure phrase in (20) does not syntactically appear in the verbal projection, but rather in the nominal projection, as a DP-adjoined modifier. If this is on the right track, the proposed analysis may provide a semantic model for reverse proportional quantification across a wider range of languages, measuring participants in events.

Acknowledgements. I thank Dylan Bumford, Itamar Francez, Tim Hunter, Chris Kennedy Ethan Poole, Jessica Rett, Malte Willer, Erik Zyman, and the audience at the UCLA Syntax-Semantics Seminar for their helpful discussions and encouraging feedback.

References

Ahn, Dorothy and Uli Sauerland (2017). "Measure constructions with relative measures: Towards a syntax of non-conservative construals". In: *The Linguistic Review* 34.2, pp. 215–248.

Barwise, Jon and Robin Cooper (1981). "Generalized quantifiers and natural language". In: *Philosophy, language, and artificial intelligence: Resources for processing natural language*. Springer, pp. 241–301.

Beck, Sigrid and Shin-Sook Kim (1997). "On wh-and operator scope in Korean". In: *Journal of East Asian Linguistics* 6.4, pp. 339–384.

Chung, Sandra and William A Ladusaw (2004). *Restriction and saturation*. MIT press.

Keenan, Edward L and Jonathan Stavi (1986). "A semantic characterization of natural language determiners". In: *Linguistics and philosophy*, pp. 253–326.

Li, Haoze (2022). "Relative measurement and scope in Mandarin". In: *Glossa: a journal of general linguistics* 7.1.

Meyer, Marie-Christine and Uli Sauerland (2009). "A pragmatic constraint on ambiguity detection: A rejoinder to Büring and Hartmann and to Reis". In: *Natural Language & Linguistic Theory* 27, pp. 139–150.

Nakanishi, Kimiko (2007). "Measurement in the nominal and verbal domains". In: *Linguistics and Philosophy* 30, pp. 235–276.

Pasternak, Robert and Uli Sauerland (2022). "German measurement structures: Case-marking and non-conservativity". In: *The Journal of Comparative Germanic Linguistics* 25.2, pp. 221–272.

Zobel, Sarah (2021). "Quantificational variability effects with German'größtenteils' and implicit agents". In: *Proceedings of Sinn und Bedeutung*. Vol. 25, pp. 939–956.

Counterpossibles, Functional Decision Theory, and Artificial Agents

Alexander W. Kocurek
University of California, San Diego
San Diego, CA, United States
akocurek@ucsd.edu

1 Introduction

Recently, Yudkowsky and Soares (2018) and Levinstein and Soares (2020) have developed a novel decision theory, *Functional Decision Theory* (FDT). They claim FDT outperforms both *Evidential Decision Theory* (EDT) and *Causal Decision Theory* (CDT). Yet FDT faces several challenges. First, it yields some very counterintuitive results (Schwarz 2018; MacAskill 2019). Second, it requires a theory of counterpossibles, for which even Yudkowsky and Soares (2018) and Levinstein and Soares (2020) admit we lack a "full" or "satisfactory" account.

Here, I focus on the latter problem of counterpossibles. My aim is to establish two claims. First, the problem of counterpossibles does not even arise without a fairly strong assumption—one that rarely applies to human agents, but may apply to artificial agents. And second, even given this assumption, the problem is solv*able*, though how best to solve it remains an open question.

2 Background

To begin, recall a familiar problem from decision theory (Nozick 1969):

> *Newcomb.* Before you are two boxes: one transparent and one opaque. You see \$1K in the transparent box, but you can't see inside the opaque box. You're can take either both boxes or just the opaque box. A reliable predictor put \$1M in the opaque box if they predicted you'd take just the opaque box; otherwise, they put \$0 in it.

What should you do in *Newcomb*? There are two standard answers. According to *Evidential Decision Theory* (EDT), you should do whatever provides the best evidence that good outcomes will obtain if you do it (Jeffrey 1965; Price 1991; Ahmed 2014b). For EDTers, this means the expected utility of an act a is calculated using the agent's credence function *conditionalized* on their choosing a. More precisely, let's say a **decision situation** is a tuple $\Delta = \langle \Omega, P, O, V \rangle$, where $\langle \Omega, P \rangle$ is a probability measure representing the agent's credal state, O is a set of options available to the agent, and V is their utility function for various act-state pairs. According to EDT, the expected utility of an option a in Δ is given as follows:

$$\mathrm{EDT}(\Delta, a) = \sum\nolimits_{s \in \Omega} P(s \mid a) V(a, s)$$

EDT recommends one-boxing: if you take one box, it's very likely the predictor put \$1M in it and you're a millionaire, whereas if you take both boxes, it's very likely the predictor put \$0 in it and you can't even pay your rent. So you should take just the opaque box.

According to *Causal Decision Theory* (CDT), you should do whatever would most likely bring about good outcomes were you to do it (Gibbard and Harper 1978; Lewis 1981; Joyce 1999).[1] For CDTers, this means the expected utility of an act a is calculated using the agent's

[1] Some use "counterfactual" in place of "causal"; see Hedden 2023; Gallow 2024 (cf. Collins 1996).

credence function *imaged* on their choosing a. Imaging can be interpreted in a number of ways (Lewis 1981; Gärdenfors 1982; Joyce 1999; Hitchcock 2016; Joyce and Gibbard 2016). The exact interpretation we adopt will not matter below, so for ease of exposition and simplicity, we'll interpret imaging in terms of counterfactuals (Gibbard and Harper 1978):[2]

$$\mathrm{CDT}(\Delta, a) = \sum_{s\in\Omega} P(s \,||\, a)V(a, s), \text{ where (for simplicity) } P(s \,||\, a) \coloneqq P(a \mathrel{\square\!\!\rightarrow} s)$$

CDT recommends two-boxing: nothing you do now can change what's in the opaque box, and so no matter what, if you were to take both boxes, you'd have \$1K more than if you were to take just the opaque box. So you should take both boxes.

Various counterexamples have been proposed to each decision theory. Against EDT, consider the following case (Egan 2007):

> *Smoking Lesion.* Smoking doesn't cause cancer. Instead, there's a brain lesion some people are born with that both causes cancer and causes people to smoke. Smokers are not more likely to develop cancer than non-smokers with the same lesion-status.

What should you do in *Smoking Lesion*? Intuitively, you should smoke: after all, smoking will have no impact on whether you have the lesion, so you might as well get the pleasure of smoking. This is what CDT recommends. By contrast, EDT says you should not smoke: smoking is evidence you have the lesion, and thus evidence that you will develop cancer.

Against CDT, consider the following case (Gibbard and Harper 1978):

> *Death in Damascus.* You are in Damascus when Death knocks on your door. Death says, "I'm coming for you tomorrow". You know Death works from a book, written in advance, that lists the time and place of each person's death. You have two options: stay in Damascus or flee to Aleppo. If you stay in Damascus, you will have a peaceful evening. If you flee to Aleppo, you'll be up all night and exhausted.

What should you do in *Death in Damascus*? Intuitively, you should stay put: no matter where you go, Death will likely follow, so you might as well enjoy your final night in peace. This is what EDT recommends. By contrast, CDT does not give a stable recommendation: the more likely you are to stay in Damascus, the higher the expected utility of fleeing and vice versa.

Whether EDT or CDT can overcome these problems, or whether these predictions are even problematic to begin with, is debated (Eells 1984; Egan 2007; Arntzenius 2008; Joyce 2012; Ahmed 2014a). But Yudkowsky and Soares (2018) and Levinstein and Soares (2020) take cases like the above to show there is something wrong with both EDT and CDT and that we should seek an alternative decision theory that will yield the "right" results. While I myself disagree (as a CDTer), for the sake of discussion, I will grant these intuitive judgments and assume they present at least a prima facie problems for EDT and CDT.

3 Functional Decision Theory

Functional Decision Theory (FDT) starts with the idea that agents employ certain *decision algorithms* when making decisions. Little is said about what this means, but roughly, we can think of these algorithms as decision rules that encode deliberative dispositions: agents are disposed to make decisions in line with what their decision algorithm recommends. Formally, we can represent decision algorithms as mathematical functions taking a decision situation and an

[2] Yudkowsky and Soares (2018) and Levinstein and Soares (2020) both instead use causal models to define imaging. This would amount to adding a causal graph to the definition of a decision situation. As these details won't matter below, I'll stick with the simple formulation in terms of counterfactuals.

option as inputs and outputting an expected value for that option in that situation.[3] Both EDT and CDT can be seen as examples of decision algorithms in this sense.

According to FDT, agents should consider not just what would happen if *they were to choose differently*, but also what would happen if their *decision algorithm were to output different choices*. Informally, you should do whatever would most likely bring about good outcomes *were your decision algorithm to output it*. Formally, FDT is like CDT in that it calculates expected utility using imaging. But FDT images on a different proposition. Whereas CDT images the agent's credence function on the proposition that the agent *chooses* a certain action, FDT images on the proposition that *the agent's decision algorithm δ maximizes its output* on that action.

$$\text{FDT}(\Delta, a) = \sum\nolimits_{s\in\Omega} P(s \,||\, \text{argmax}_x \delta(\Delta, x) = a) V(a, s)$$

In a slogan: *be the kind of agent it would be best to be*. Or, less catchily: employ the decision algorithm that would most likely yield good outcomes.

To see how FDT differs from rivals, consider what it recommends in each of the three decision situations above. In *Newcomb*, FDT agrees with EDT that you should one-box. For if your decision algorithm were to output one-boxing, the predictor would very likely predict this, in which case, there would be \$1M in the opaque box. Yet if your decision algorithm were to output two-boxing, the predictor would very likely predict this, in which case, there would be \$0 in the opaque box. So it would be better if your decision algorithm outputted one-boxing.

Likewise, in *Death in Damascus*, FDT recommends staying in Damascus. For if your decision algorithm were to output staying, Death's book would have you down as staying. And if your decision algorithm were to output fleeing, Death's book would have you down as fleeing. So it would be better if your decision algorithm saved you the trouble and outputted staying.

The reason FDT differs from CDT is that the output of one's decision algorithm is causally (or explanatorily) upstream from the decision itself. In effect, the output of the agent's algorithm is a "common cause" of both the agent's decision and the predictor's prediction, who, we might suppose, has some reliable indicator of how their decision algorithm behaves.

By contrast, in *Smoking Lesion*, FDT agrees with CDT that you should smoke. One's decision algorithm does not cause/explain the presence of the relevant brain lesion from birth. Thus, regardless of what your decision algorithm were to recommend, the chance that you would have the lesion remains the same, and so, you might as well smoke.

Advocates of FDT defend these verdicts: you should one-box in *Newcomb*, smoke in *Smoking Lesion*, and stay in Damascus in *Death in Damascus*. Again, while there is substantial debate over this point, for the sake of discussion, I'm granting advocates these verdicts are indeed the correct ones. Still, FDT faces two main problems.

The first major problem is there are cases where FDT yields highly counterintuitive verdicts. Consider *Transparent Newcomb*, which is exactly like Newcomb except both boxes are transparent: you can see clearly that one box contains \$1M while the other contains \$1K. Intuitively, in this case, you should take both boxes—otherwise, you're leaving money on the table! Indeed, in this case, both EDT and CDT recommend two-boxing, as your action is both evidentially and causally independent of the contents of the box.

By contrast, FDT *still* recommends one-boxing in *Transparent Newcomb*. Even though you can clearly see \$1M inside, FDT says that doesn't matter: what matters is what *would* likely be inside *were* your decision algorithm to output such-and-such. If your decision algorithm were to output two-boxing, the predictor would very likely have predicted this and put \$0 in that box. If your decision algorithm were to output one-box, the predictor would very likely have predicted this and put \$1M in that box. So it would be better if your decision algorithm outputted one-boxing. Thus, according to FDT, you should one-box.

[3] For certain purposes, we may want a more fine-grained notion of algorithm on which multiple algorithms may produce the same input-output pairs. I'll set this complication aside.

Whether examples like this ultimately undermine FDT is a matter of debate. Yudkowsky and Soares (2018) argue that one-boxing in *Transparent Newcomb* is the correct result. By contrast, Schwarz (2018) and MacAskill (2019) both present variants of *Transparent Newcomb* that make the counterintuitiveness more pressing. For lack of space, I set aside concerns about whether FDT makes the intuitively correct verdicts in these cases for another time.

The second major problem for FDT, which is the one I take up here, is its reliance on counterpossibles. FDT's calculation of expected utility requires imaging on a mathematically impossible proposition, viz., that a certain mathematical function δ, given certain inputs, yields different outputs than it actually yields. Yet, by Yudkowsky and Soares's (2018) and Levinstein and Soares's (2020) own admissions, we do not have a "full" or "satisfactory" theory of counterpossibles. Yudkowsky and Soares (2018) describe it as "the main drawback of FDT relative to CDT". Similarly, Levinstein and Soares (2020) admit that "for FDT to be successful, a more worked out theory [of counterpossibles] is necessary". Critics likewise see this as a major issue for the view (Schwarz 2018; MacAskill 2019).

The problem rests on two premises: (1) FDT requires agents to evaluate the probability of counterpossibles; and (2) we have no "full" or "satisfactory" account of counterpossibles. I will challenge each premise and, in so doing, argue the problem is at least solvable, though it remains to be seen which of the solutions on the table are most appealing.

4 Are counterpossibles necessary for FDT?

Why think FDT requires counterpossibles? The idea is this. Let δ be the function representing the agent's decision algorithm. Whether δ maximizes its value on an input a i.e., whether $\text{argmax}_x\delta(\Delta, x) = a$, is a mathematical fact. Mathematical facts are not contingent. So the claim '$\text{argmax}_x\delta(\Delta, x) = a$', if false, is mathematically impossible.

Note, however, when we say "Let δ be the function…", we're treating 'δ' as a rigid designator, denoting a fixed mathematical function that represents the algorithm that the agent *actually* employs. This is how Yudkowsky and Soares (2018) and Levinstein and Soares (2020) formulate FDT. But if we instead treat 'δ' as a non-rigid designator, so that, for each world w where the agent exists, 'δ' denotes the function representing the agent's algorithm at w, '$\text{argmax}_x\delta(x) = a$' becomes a contingent statement since agents may employ different algorithms at different worlds. In other words, if 'the agent's decision algorithm' is interpreted *de dicto* (rather than *de re*) in the relevant counterfactual, it isn't a counterpossible.

Arguably, the de dicto interpretation is enough to make the predictions its advocates want it to make (cf. Schwarz 2018): it still predicts you should one-box in *Newcomb*, smoke in *Smoking Lesion*, and stay home in *Death in Damascus*. Moreover, the de re interpretation does not seem relevant for most agents in most contexts. Most agents do not have introspective access to their own decision-making process (if only!): they are *uncertain* regarding which decision algorithm they do employ or would employ. Reasoning about what would happen if (per impossibile) a certain function f had yielded different outputs is only relevant if the agent is certain that f represents their decision-making process. If an agent doesn't know which function represents their decision algorithm, this uncertainty will need to be reflected in how they calculate expected utility, which is what the de dicto (but not the de re) interpretation does.

More precisely, let $\hat{\delta}$ be a non-rigid term denoting, at each world w, the decision algorithm the agent uses at w. Let δ be a rigid term denoting the decision algorithm the agent actually uses. The de re interpretation of FDT for a decision situation Δ seems to assume the following:[4]

Counterfactual Robustness:

$$P(\hat{\delta} = \delta \,||\, \text{argmax}_x\hat{\delta}(\Delta, x) = a) = P(\hat{\delta} = \delta \,||\, \text{argmax}_x\delta(\Delta, x) = a) = 1$$

[4]This violates the Strangeness of Impossibility Condition from Nolan 1997; see Kocurek 2021a for discussion.

The de dicto and de re interpretations collapse given Counterfactual Robustness together with modest assumptions.[5] But in most contexts, human agents do not obey this principle.

This severely limits the force of the problem of counterpossibles, but it does not eliminate it. For one, there may be limited circumstances where Counterfactual Robustness is correct even for rational human agents. Perhaps a human agent learns what their decision algorithm is after years of therapy or after a series of invasive brain scans. Or perhaps a decision theorist becomes so convinced of the arguments for a certain decision theory that they become certain they would employ that theory regardless of its outputs. Moreover, the initial motivation for developing FDT was to develop a decision theory for *artificial* agents, who may have access to their own source code (Soares and Fallenstein 2015). For such agents, Counterfactual Robustness may be more a reasonable assumption. Indeed, advocates of FDT often talk not about what an agent's decision algorithm recommends, but what *FDT itself* recommends, suggesting they have in mind agents who know their decision algorithm is FDT. So, Counterfactual Robustness may be appropriate in special applications. In that case, advocates of FDT have to tackle counterpossibles head on.

5 Are counterpossibles problematic?

Even if we formulate FDT so as to require counterpossibles, this isn't necessarily problematic. There are many accounts of counterpossibles in the literature (see Kocurek 2021a for overview), including: similarity accounts that extend the classic ordering semantics with impossible worlds (Nolan 1997; Krakauer 2012; Brogaard and Salerno 2013; Kment 2014; Berto et al. 2018); grounding accounts that extend the causal modeling framework with noncontingent nodes (Schaffer 2016; Wilson 2018; Baron, Colyvan, and Ripley 2020; Khoo 2022); and counterconventional accounts that simulate counterpossible reasoning without direct appeal to "worldly" impossibilities (Kocurek and Jerzak 2021; Kocurek 2019, 2021b). Still, there are two related problems that might lead one to think none of these accounts is satisfactory for current purposes.

First, for the expected utility calculation to work in FDT, the probability function imaged on $\operatorname{argmax}_x \delta(\Delta, x) = a$ must still be a *probability* function. But Schwarz (2018) raises the worry that imaging on the impossible could lead to violations of the Kolmogorov axioms. For example, these axioms entail that $P(A \wedge B) \leqslant P(A)$. So if $P(A \wedge \neg A \,||\, A \wedge \neg A) = 1$, then $P(A \,||\, A \wedge \neg A) = P(\neg A \,||\, A \wedge \neg A) = 1$. But this violates another implication of the Kolmogorov axioms: $P(\neg A) = 1 - P(A)$. More generally, the Kolmogorov axioms assume the logical connectives $\neg$, $\wedge$, and so on are interpreted classically: $\neg A$ is true at w iff A is not true at w, $A \wedge B$ is true at w iff A and B are both true at w, and so on. Yet at logically impossible worlds, these assumptions can fail: there can be worlds where both $\neg A$ and A are true (or both false), worlds where $A \wedge B$ is true and yet A and/or B are false, and so on.

Second, as Yudkowsky and Soares (2018) note, while there are methods for discovering causal structure (Pearl 2000), we have no such thing for counterpossible dependencies. How is an agent to assign rational credences to counterpossibles? Arguably, there is no universal similarity ordering, grounding graph, etc. that works in every case and we should instead focus on developing ones tailored to specific applications. We may even have artificial agents *learn*, via standard machine learning techniques, which ordering, graphs, etc. are most effective at maximizing utility in a range of typical scenarios. Still, it would be more satisfactory to have a story regarding how agents should assess the probability of a counterpossible.

[5] Here are the modest assumptions: First, $P(\cdot \,||\, A)$ is a probability function (see section 5). Second, $P(A \,||\, A) = 1$ (success). Third, $P(A \,||\, B) = P(B \,||\, A) = 1$ implies $P(C \,||\, A) = P(C \,||\, B)$ (replacement of counterfactually equivalent antecedents). Finally, perhaps redundantly, $P(a = b \,||\, A) = 1$ implies $P(\phi(a) \,||\, A) = P(\phi(b) \,||\, A)$ (substitution of counterfactual identicals; this may or may not follow from the first assumption, depending on how its formulated). I don't claim these assumptions hold universally (e.g., if the antecedents express the failures of these very rules), but they seem plausible in this context.

To solve the first problem, we need to clarify how we're interpreting logical operators used in probability statements. Suppose we want to characterize the credences of an intuitionistic logician. When we write '$P(\neg A)$', what does '$\neg$' refer to? According to an *object-level* interpretation, '$\neg$' refers to an operation the *agent* uses in their own thought, e.g., intuitionistic negation. According to a *meta-level* interpretation, '$\neg$' refers to an operation the *theorist* uses in the metalanguage, viz., classical negation. The problem arises from conflating these two interpretations. When we, as *theorists*, state the Kolmogorov axioms, we adopt a meta-level interpretation, where the connectives are interpreted as the theorist interprets them in their (meta)language. By contrast, when discussing logically deviant or logically uncertain agents, we often want to represent their credal state from *their* perspective, in *their* language, and thus adopt an object-level interpretation, where the connectives stand for the representations agents themselves use.

One way to avoid this confusion is to introduce a different set of symbols for operators with different interpretations. For example, we may use $-$, $\cap$, $\cup$, and so on, for *classically rigid* connectives, which are always interpreted according to the (classical) metalanguage even under the scope of a counterpossible supposition. We may then reserve $\neg$, $\wedge$, $\vee$, and so on for the operations that the agent themselves use, which they may interpret nonclassically. This allows us to simultaneously state the Kolmogorov axioms in the metalanguage while representing the language the agent uses "from within". For example, for intuitionistic agents, $P(A \cup -A) = 1$ even if $P(A \vee \neg A) \neq 1$. And $P(-A) = 1 - P(A)$ even if $P(\neg A) \neq 1 - P(A)$.[6] Thus, even if an agent images on a classically impossible proposition, the result can still be a probability function. For example, $P^*(\cdot) \coloneqq P(\cdot \,||\, A \wedge \neg A)$ can still satisfy the Kolmogorov axioms stated in the classical metalanguage (e.g., while $P^*(A \wedge \neg A) = 1$, still $P^*(A \cap -A) = 0$).

This resolution to the first problem helps resolve the second. When representing agents entertaining counterpossible suppositions, it seems most appropriate to use the object-level operators that they use to represent those impossibilities. In that case, we can understand their counterpossible suppositional reasoning as a kind of shift in the interpretation of the terms they use to state those suppositions. This is effectively the expressivist approach to counterlogicals adopted by Kocurek and Jerzak (2021) and Kocurek (2021b). Indeed, Kocurek and Jerzak (2021) show that this approach can represent the same hyperintensional phenomena as the standard impossible worlds approach: both approaches generate the same hyperintensional logic over the propositional counterfactual language (though they diverge in more powerful languages; see Kocurek 2021b). By allowing re-interpretation of the object-language terms, it is thus possible to simulate all counterpossible reasoning without appealing to impossible worlds.

At the same time, interpretation-shifting illuminates why certain counterpossibles are true and also how we may be able to know which counterpossible dependencies hold: our ability to know counterpossibles stems from our general ability to consider and reason about alternative interpretations. This approach therefore offers a general method for establishing counterpossible dependencies via the manipulation of interpretations. How is an agent to determine whether a counterpossible holds? By exploring alternative interpretation-world pairings that make the antecedent true and evaluating whether the consequent is true at those pairs.

Of course, this is not a complete theory of counterpossibles. The space of possible interpretations is vast. Yet agents only select a small set of reasonable alternative interpretations when engaged in counterpossible reasoning. What counts as a "reasonable" alternative interpretation pair will be a context-sensitive matter: not anything goes. So, in a sense, the problem of counterpossible lingers. But this approach at least offers a method via which we can start to address some of the thorny questions surrounding context-resolution that would help settle how to uncover counterpossible dependencies in a more objective manner.[7]

[6] It is possible to mix meta-level and object-level connectives in a rigorous and formally consistent manner, as in *hyperlogic* (Kocurek 2019, 2021, 2024a, 2024b).

[7] Thanks to Alan Hájek, Ethan Jerzak, Rachel Rudolph, and the audience members at the Dartmouth Summer

References

Ahmed, Arif (2014a). "Dicing with death". In: *Analysis* 74.4, pp. 587–592.
— (2014b). *Evidence, Decision and Causality*. Cambridge: Cambridge University Press.
Arntzenius, Frank (2008). "No Regrets, or: Edith Piaf Revamps Decision Theory". In: *Erkenntnis* 68, pp. 277–297.
Baron, Sam, Mark Colyvan, and David Ripley (2020). "A Counterfactual Approach to Explanation in Mathematics". In: *Philosophia Mathematica* 28.1, pp. 1–34.
Berto, Francesco et al. (2018). "Williamson on Counterpossibles". In: *Journal of Philosophical Logic* 47, pp. 693–713.
Brogaard, Berit and Joe Salerno (2013). "Remarks on Counterpossibles". In: *Synthese* 190.4, pp. 639–660.
Collins, J. (1996). "Supposition and Choice: Why 'Causal Decision Theory' is a Misnomer". Manuscript.
Eells, Ellery (1984). "Newcomb's Many Solutions". In: *Theory and Decision* 16.1, pp. 59–105.
Egan, Andy (2007). "Some counterexamples to causal decision theory". In: *The Philosophical Review* 116.1, pp. 93–114.
Gallow, J. Dmitri (2024). "Counterfactual Decision Theory is Causal Decision Theory". In: *Pacific Philosophical Quarterly* 105, pp. 115–156.
Gärdenfors, Peter (1982). "Imaging and Conditionalization". In: *Journal of Philosophy* 79.12, pp. 747–760.
Gibbard, Allan and William L Harper (1978). "Counterfactuals and Two Kinds of Expected Utility". In: *Foundations and Applications of Decision Theory*. Ed. by Clifford Alan Hooker, James J Leach, and Edward Francis McClennen. Dordrecht: D. Reidel, pp. 125–162.
Hedden, Brian (2023). "Counterfactual Decision Theory". In: *Mind* 132.527, pp. 730–761.
Hitchcock, Christopher (2016). "Conditioning, Intervening, and Decision". In: *Synthese* 193, pp. 1157–1176.
Jeffrey, Richard C. (1965). *The Logic of Decision*. Chicago: University of Chicago Press.
Joyce, James (1999). *The Foundations of Causal Decision Theory*. Cambridge: Cambridge University Press.
— (2012). "Regret and Instability in Causal Decision Theory". In: *Synthese* 187, pp. 123–145.
Joyce, James and Allan Gibbard (2016). "Causal Decision Theory". In: *Readings in Formal Epistemology*. Ed. by Horacio Arlø-Costa, Vincent F. Hendricks, and Johan van Benthem. Berlin: Springer, pp. 457–491.
Khoo, Justin (2022). *The Meaning of If*. Oxford: Oxford University Press.
Kment, Boris (2014). *Modality and Explanatory Reasoning*. Oxford: Oxford University Press.
Kocurek, Alexander W. (2019). "Hyperlogic: A System for Talking about Logics". In: *Proceedings for the 22nd Amsterdam Colloquium*. Ed. by Julian J. Schlöder, Dean McHugh, and Floris Roelofsen. ILLC. Amsterdam, pp. 238–247.
— (2021a). "Counterpossibles". In: *Philosophy Compass* 16.11, e12787.
— (2021b). "Logic Talk". In: *Synthese* 199.5–6, pp. 13661–13688.
— (2024a). "The Logic of Hyperlogic. Part A: Foundations". In: *Review of Symbolic Logic* 17.1, pp. 244–271.
— (2024b). "The Logic of Hyperlogic. Part B: Extensions". In: *Review of Symbolic Logic* 17.3, pp. 654–681.
Kocurek, Alexander W. and Ethan J. Jerzak (2021). "Counterlogicals as Counterconventionals". In: *Journal of Philosophical Logic* 50.4, pp. 673–704.
Krakauer, Barak (2012). "Counterpossibles". PhD thesis. University of Massachusetts, Amherst.

Workshop on Artificial Intelligence, Language, and Cognition for helpful feedback on this project.

Levinstein, Benjamin A. and Nate Soares (2020). "Cheating Death in Damascus". In: *The Journal of Philosophy* 117.5, pp. 237–266.

Lewis, David K (1981). "Causal Decision Theory". In: *Australasian Journal of Philosophy* 59.1, pp. 5–30.

MacAskill, William D. (2019). "A Critique of Functional Decision Theory". https://www.lesswrong.com/posts/ySLYSsNeFL5CoAQzN/a-critique-of-functional-decision-theory.

Nolan, Daniel (1997). "Impossible Worlds: A Modest Approach". In: *Notre Dame Journal of Formal Logic* 38.4, pp. 535–572.

Nozick, Robert (1969). "Newcomb's Problem and Two Principles of Choice". In: *Essays in Honor of Carl G. Hempel*. Ed. by Nicholas Rescher. Dordrecht: Springer, pp. 114–146.

Pearl, Judea (2000). *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Price, Huw (1991). "Agency and Probabilistic Causality". In: *British Journal for Philosophy of Science* 42, pp. 157–176.

Schaffer, Jonathan (2016). "Grounding in the Image of Causation". In: *Philosophical Studies* 173, pp. 49–100.

Schwarz, Wolfgang (Dec. 2018). "On Functional Decision Theory". https://www.umsu.de/wo/2018/688.

Soares, Nate and Benja Fallenstein (2015). "Toward Idealized Decision Theory". In: *arXiv* 1507.01986[cs.AI], pp. 1–15. URL: https://arxiv.org/abs/1507.01986.

Wilson, Alastair (2018). "Grounding Entails Counterpossible Non-Triviality". In: *Philosophy and Phenomenological Research* 46.3, pp. 716–728.

Yudkowsky, Eliezer and Nate Soares (2018). "Functional Decision Theory: A New Theory of Instrumental Rationality". In: *arXiv* 1710.05060[cs.AI], pp. 1–36. URL: https://arxiv.org/abs/1710.05060.

Hurford's Constraint and Disjunctions over Speech Acts

Manfred Krifka[1]
Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS) and Humboldt-Universität zu Berlin
Berlin, Germany
krifka@leibniz-zas.de

Abstract

The paper focuses on apparent exceptions to Hurford's generalization on disjunctions of the type *John was born in Paris or at least in France* and *John was born in France or even in Paris.* It argues that such cases are disjunctions of speech acts of different strength. For this reason, the assertion with the stronger proposition does not license the other one, nor vice versa, hence Hurford's generalization is not violated. The paper raises the issues of disjunctions of speech acts and of assertive strength, and also reconsiders apparent exceptions of the type *Some or all of the children are sick.*

1 Hurford's Generalization and its Exceptions

Fifty years ago, Hurford (1974) observed that disjunctions as in (1) are infelicitous, and proposed that "the joining of two sentences by *or* is unacceptable if one sentence entails the other" (p. 410).

(1) *#John is an American or a Californian.*

This generalization can be derived from more general pragmatic principles, like the Maxim of Manner (Grice 1975), cf. Meyer 2014 for "Efficiency": The meaning of (1) is already expressed by the shorter *John is an American*, as it is known that everyone who is Californian also is American,.

Gazdar (1979: 81), pointing out exceptions like (2), proposed that Hurford's generalization does not apply when the entailed sentence would potentially implicate the negation of the entailing sentence (as *some were there* would implicate 'not all were there').

(2) *Some or all of them were there.*

Singh (2008) and Chierchia et al. (2009, 2012) argued that cases like (2) do not constitute exceptions to Hurford's generalization because the first disjunct actually is strengthened by a local scalar implicature that generates an exhaustive meaning, 'only some were there'; as a consequence, it is not entailed by the second conjunct, *all were there.* Cases like (2) form the core argument for such implicatures, often associated with a silent operator EXH (cf. Katzir & Singh 2014, Fox & Spector 2018, Hénot-Mortier 2023).

Below I will offer a different proposal. It starts with an observation by Simons (2001) and by Singh (2008), which has generated less attention as a potential exception to Hurford's generalization: While (3)(a) is unacceptable, as predicted by Hurford (1974), (b) is fine.

(3) a. *#John was born in Paris or in France.*
b. *John was born in Paris or at least in France.*

I will proceed as follows. Section 2 discusses existing accounts of the contrast illustrated in (3). Section 3 proposes a new account, according to which the assertion *John was born in Paris* does not conversationally entail *John was born at least in France*, because while the latter is a weaker proposition, it is a stronger assertion. Section 4 returns to cases like (2) and proposes an analysis

[1] The work reported here was funded by ERC 787929, ERC Advanced Grant SPAGAD "Speech Acts in Grammar and Discourse". I thank Anton Benz, Dan Goodhue and Tue Trinh for valuable comments.

in terms of weak and strong assertions as well. Section 5 considers ways to implement disjunctions of speech acts and assertions of different strength, and Section 6 concludes.

2 "At least" Exceptions: Previous Accounts

There are a few existing attempts to explain apparent exceptions to Hurford's generalization of the *at least* type exemplified by (3)(b). We will discuss them in this section.

Singh (2008) suggested, in a footnote, that the second disjunct "retracts" the first one. He states that the sentence comes with a "comma" intonation, i.e. is realized by two prosodic phrases, (*John was born in Paris*) (*or at least in France*). There are problems with this suggestion. First, it is unclear why this intonation is not able to salvage example (3)(a). Second, the comma intonation does not seem to be required for (3)(b) to be acceptable. Third, true retractions as illustrated in (4)(a) cannot be applied to the case at hand, witness the oddity of (4)(b).

(4) a. *John was born in Paris, or rather in Lyon.*
b. *#John was born in Paris, or rather in France.*

Westera (2019) argues that *at least* in (3)(b) indicates that the two disjuncts have different value on some scale and relates the difference to the Question under Discussion (QuD). Building on this work, Zhang (2022) suggests that the QuD in (3)(b) is changed from 'In which city was John born?' to 'In which state was John born?'. As the disjuncts answer different QuDs, they do not stand in an entailment relation to each other. I think that an explanation along this line is problematic, as we can fix the QuD explicitly, and still can recreate the contrast in (3):

(5) A: *In which city was John born?*
B: *He was born in Paris or #in a city in France / at least in a city in France.*

According to Simons (2001), the individual clauses in a regular disjunction should give distinct, non-overlapping answers to the same QuD. For disjunctions with the second disjunct marked by *rather* or *at least*, Simons argues that they have a "metalinguistic" nature and "indicate a particular relation between speech acts" in which "the second 'disjunct' is to fill the same 'discourse slot' as the first, and thus in some sense replaces that act" (p. 613). The second disjunct requires some "qualifying expression" such as "*at least, rather, even, I should say* and *should I say?*" Simons observes that such disjunctions do not allow for *either—or* constructions (6)(a), that they are not symmetric (b), and that *or* can be eliminated (c):

(6) a. *#Either George lives in London, or at least he lives somewhere in England.*
b. *#George lives (at least) somewhere in England, or in London.*
c. *Georges lives in London. At least, he lives somewhere in England.*

The proposal that I will develop is inspired by Simon's impression that the disjunctions combine different speech acts. I will propose that (3)(b) is felicitous because the disjunction can scope over two assertions, where the first assertion does not license the second, as the second is a stronger assertion, as indicated by *at least*. Without *at least* as in (1), the assertions are of the same strength; consequently the first disjunct licenses the second one, and the configuration is ruled out by the pragmatic factors that are responsible for Hurford's generalization.

3 Disjunction of Assertions and Assertive Strength

I propose that disjunctions that appears within assertions can be read as disjunctions of assertions. I assume here the commitment account of assertions (cf. Brandom 1994, Krifka 2015, Geurts 2019, Shapiro 2020) and represent the assertion by speaker S of proposition Φ as "S ⊢ Φ", that S vouches for the truth of Φ. Hence, I propose that the assertion by S in (7)(a) has, in addition to the propositional assertion (7)(b), a reading with a disjunction on the speech act level, (7)(c).

(7) a. S: *John was born in Paris or in Lyon.*
b. S ⊢ ['John was born in Paris' ∨ 'John was born in Lyon']
c. [S ⊢ 'John was born in Paris'] ∨ [S ⊢ 'John was born in Lyon']

Readings involving disjunctions over speech acts are often disfavored because they lead to ambiguous discourse states. After (7)(c), speaker S is either committed to 'John was born in Paris' or to 'John was born in Lyon', a commitment that is difficult to keep track of in discourse. Therefore, disjunctions of speech acts were supposed to be ruled out. For example, Krifka (2001) derived the fact that quantification into constituent questions, as in *Which dish did every guest bring to the party?* is possible only with universal quantifiers like *every guest*, i.e. with generalized conjunctions, but not with other quantifiers like *most guests* that are also based on disjunctions. For (7)(a), this means that the reading with a proposition-scope disjunction (7)(b) is strongly favored, as this results in a non-ambiguous commitment of S to a (disjunctive) proposition.

Hurford's generalization, originally based on entailment between propositions, can be recast for disjunctions of assertions. A speaker that commits to a proposition Φ is also obliged to commit to a proposition Ψ, if Φ entails Ψ in a transparent manner (cf. commitment inheritance or entitlement, Brandom 1994). Hence the commitment to Ψ is redundant, which explains the oddness of (1) and (3)(a). This observation may seem irrelevant, if disjunctions of assertions are disfavored for the reason given above. But now we can develop an argument why (3)(b) is good.

The crucial property of (3)(b) is that the second disjunct contains the modifier *at least.* Such "superlative" modifiers have been argued to be interpreted not on the propositional level, but on the level of epistemic operators (Geurts & Nouwen 2007) or speech acts (Cohen & Krifka 2014). Assuming that *at least* is focus-sensitive and hence relates to alternatives, I propose that it is an assertion modifier that results in a commitment to the asserted proposition that is higher than the commitment to the alternative propositions. In disjoint assertions of the type (3)(b), the proposition of the first disjunct is an explicitly mentioned alternative to the proposition of the second disjunct. 'John was born in France' is logically weaker than 'John was born in Paris', and for this reason can be committed to and hence asserted more strongly. (Note that there are cases that are not based on logical entailment but on argumentative weight, as in *Mary is a full professor or at least an asscociate professor*, cf. Ducrot 1973). Thus, the reading of (3)(b) in which the disjunction scopes over the commitment has the following configuration:

(8) S: [S ⊢ 'John was born in Paris'] ∨ [S $\vdash_{\text{strong}}$ 'John was born in France']

The subscript indicates that the second commitment is stronger than the first. That speech acts may come with various strengths has been proposed before, e.g. Searle & Vanderveken (1986), who include a 5-level strength parameter, and Greenberg & Wolf (2018), who state that assertions come with different degrees. Different strengths of assertions have typically been discussed in terms of mitigation of assertions, cf. McDowell (1987), Sbisà (2001), Incurvati & Schlöder (2019). We make the plausible assumption that S⊢Φ does not oblige S to S$\vdash_{\text{strong}}$Φ; also, S⊢Φ does not oblige S to S$\vdash_{\text{strong}}$Ψ even if Φ transparently entails Ψ. Note that under this analysis, cases like (3)(b) are predicted to be felicitous by Hurford's generalization.

We predict that other modifiers of the strength of assertive commitments have similar effects. *Definitely* and *in any case* occur with the less specific proposition and should be classified as strengtheners, like *at least*, cf. (9)(a). This also holds for *certainly*, which makes the assertion of a proposition easier to defend, hence assertable under more circumstances (cf. Lassiter 2016, Yatsushiro et al. 2021). Assertive commitments with different strength do not require a weakening of the proposition but can also involve hedging by weakening the evidential backup, cf. (9)(b).

(9) a. *John was born in Paris or definitely / in any case / certainly in France.*
b. *John was born in Paris, or at least he told me so.*

Simons (2011), Westera (2019) and Zhang (2022) pointed out that the particle *even* can also lead to apparent exemptions from Hurford's generalization, cf. (10)(a). In contrast to *at least*, *even* occurs with the more specific proposition, which I take to flag a weaker commitment, leading to the analysis (10)(b). Classifying the second assertion as weaker may seem counterintuitive, as *even* seems to mark extreme alternatives. However, this applies to the proposition in comparison with its alternatives, and not to the assertive strength itself (cf. the fact that *perhaps* and *even* happily cooccur).

(10) a. *John was born in France or even in Paris.*
b. S $\vdash$ 'John was born in France' $\vee$ S $\vdash_{\text{weak}}$ 'John was born in Paris'

The order of disjuncts is relevant; assertions like *#John was born at least in France or in Paris* and *#John was born even in Paris or in France* are infelicitous. This order-sensitivity (cf. also Tomioka 2021) holds for focus-sensitive operators in general, as explicitly mentioned alternatives have to be mentioned before the operator in order to be understood as a relevant alternative (e.g. *I met John and even Bill* vs. *#I met even Bill and John*).

Operators like *at least* and *even* also occur in disjunctions of other speech acts than assertions (cf. Ciardelli & Roelofsen 2017 for questions). They can abide by Hurford's generalization due to similar asymmetries as with assertions. With questions like *Was John born in Paris or at least in France?* the second disjunct is a more likely answer. With imperatives like *Go to Paris or at least to France!* the second disjunct is a stronger command. We observe a similar phenomenon in propositional attitude contexts as in (11), where we have, in addition to a commitment-related reading (i), a reading that involves different strengths of belief (ii).

(11) *Mary believes that John was born in Paris or at least in France.*
i. [S $\vdash$ 'M BEL J was born in Paris'] $\vee$ [S $\vdash_{\text{strong}}$ 'M BEL J was born in France']
ii. S $\vdash$ ['M BEL J was born in Paris' $\vee$ 'M $\text{BEL}_{\text{strong}}$ J was born in France']

In interpretation (ii), S vouches for the proposition that Mary has a belief of regular strength that John was born in Paris, and a stronger belief that John was born in France. This description of Mary's belief is not redundant, as we cannot infer from Mary's regular belief in the first proposition to Mary's strong belief in the second.

In concluding this section, I want to point out that in all these cases of apparent violations to Hurford's generalization, there is an additional pragmatic reason for the expression of two disjuncts. With presenting two disjuncts, sentences like like (3)(a) and (9)(a) introduce two concepts (Paris and France) and draw attention to two propositions (that John was born in Paris, that John was born in France). This non-truthconditional aspect in Hurford disjunctions was pointed out by Westera (2019). However, notice that this is not sufficient to explain why apparently redundant disjunctions are pragmatically acceptable, as we would then predict that cases like (3)(a) should be acceptable as well.

4 "Some/All" Exceptions: An Alternative Account

For the apparent exceptions to Hurford's generalization considered so far, it was essential that one disjunct (the second) had an explicit marker, like *at least* or *even*, that arguably indicated a different strength from the other (the first). Disjunctions without this marker, as in (3)(a), do not simply coerce the second disjunct into a stronger or weaker reading.

Turning to apparent exceptions of the *some or all* type like (2) and (12)(a), we see that such explicit markers are not necessary; we do not have to insert *only* before *all* in (12). In fact, there is a slight degradation with overt *only*; on Google Ngrams the string *only some or all* occurs 4 orders of magnitude less frequently than *some or all*, and seems largely confined to texts in logic, computer science, and linguistics.

(12) a. *John read some (most, several) or all novels by Dickens.*
b. *John read three or seven novels by Dickens.*
c. *John read 'Oliver Twist' or 'Bleak House', or both of these novels.*

We observe that exceptions of the *some or all* type can be reproduced with other items, in particular number words as in (12)(b); note that the proposition 'John read 7 novels by Dickens' entails the proposition 'John read 3 novels by Dickens'. (Cases like *John read three or four novels by Dickens*, spoken with one prosodic phrase over the disjunction, have to be treated separately, as approximate number words).

There are several proposals to explain apparent exceptions to Hurford's generalization of the *some or all* type (cf. Marty & Romoli 2022 for a critical overview). I would like to point out that these cases can be subsumed under the exceptions involving different assertive strengths if we assume that the second disjunct makes a weaker assertion, as in (13). This is similar to our analysis of cases with *even* as in (10). Note that we could insert *even* before *some* and *seven* in (12)(a,b); in fact, on Google Ngrams *some or even all* only occurs only 2 orders of magnitude less frequently than *some or all*.

(13) a. S $\vdash$ 'John read some of the novels by D' $\vee$ S $\vdash_{\text{weak}}$ 'John read all of the novels by D'
b. S $\vdash$ 'John read 3 novels by D' $\vee$ S $\vdash_{\text{weak}}$ 'John read 7 novels by D'
c. S $\vdash$ 'John read OT or BH' $\vee$ S $\vdash_{\text{weak}}$ 'John read OT and BT'

The question arises why cases like (12) do not require an explicit marker of assertive strength of the second disjunct, in contrast to (3), (9) and (10). A plausible reason is that these cases involve alternatives that are derived from the structure of the lexicon, or Horn scales. Hence purely linguistic competence will result in the recognition of differences in assertive strength. An additional difference is that quantifiers like *some, all*, numerals like *three, seven* and Boolean connectives like *or, and* [*both*] form total (linear) relations, in contrast to part relations or hyponymy relations like *Paris, France* or *beetle, insect*. The lexical nature of the alternatives, perhaps in addition to their linear semantic order, reduces the need to mark the non-default strength of the assertion explicitly. Merin (2003), without referring to Hurford's generalization explicitly, discusses the use of *if not* in such cases, as in *some if not all* and *three if not four*, and develops an analysis in terms of act-based relevance similar to the one proposed here.

The two competing analyses both make use of the notion of alternatives; one understands *some or all* as 'only some or all', the other as 'some or even all'. There is a structural difference: In the first paraphrase, the alternative-sensitive operator, *only* or EXH, occurs before the disjunct that mentions the alternative; in the second, the operator *even* occurs after the disjunct that mentions the alternative. We have seen that explicitly mentioned alternatives should occur before the alternative-sensitive operator; this is a further argument for the analysis proposed here.

5 Implementing Speech Act Strength and Disjunction

In sections 3 and 4 I sketched a theoretical account from which apparent exceptions to Hurford's generalization lose their exceptional status. It makes two essential assumptions: We can disjoin speech acts, and we have assertions of different strength. In the remainder of this short paper, I will point to research that show how these notions can be formally implemented and tested.

Disjunctions of speech acts can be handled in the framework of Commitment Spaces (CS) introduced in Krifka (2015), where it is central for alternative questions and constituent questions. The CS framework comprises the common ground at a particular state of conversation, the root, together with its possible continuations. The root is ideally a single commitment state, but it can also be ambiguous, with more than one commitment state (cf. Kamali & Krifka 2020). Update of the root {c} with a commitment S⊢Φ to {c + S⊢Φ} triggers further update with the asserted proposition Φ to {c + S⊢Φ + Φ} in case the addressee does not prevent this (cf. Krifka 2022). Update of a CS with root {c} with a disjunction of assertions, like S⊢Φ ∨ S⊢Ψ, results in an ambiguous root {c + S⊢Φ, c + S⊢Ψ}, which should be avoided, as it leads to an ambiguous discourse state at which it is unclear whether Φ or Ψ can be assumed. However, in the case of disjunctions like (8), of the form S⊢Φ ∨ $S\vdash_{strong}\Psi$, where Φ ⇒ Ψ, the resulting discourse state allows for the introduction of the more general proposition Ψ in any case, as it is supported by both $S\vdash_{strong}\Psi$ and, indirectly, by S⊢Φ. Hence the resulting discourse state is considerably less ambiguous now, as it allows for assuming Ψ in any case. In addition, there is partial and weaker support for Φ. Furthermore, the introduction of Φ as a propositional discourse referent makes this proposition salient (in addition to Ψ) so that it can be taken up in subsequent (cf. Westera 2019).

The cases we have considered all involved disjunctions of sub-clausal constituents, such as *in Paris or at least in France* for (3)(b), which appears at odds with the idea that they indicate disjunctions of assertions. However, we can assume that redundant material is not phonologically realized, hence (3)(b) is derived by deletion, *John was born in Paris or ~~John was born~~ at least in France.* Alternatively, we can explore the idea that there is an interpretation to the disjunctive phrase *in Paris or at least in France* that takes a speech act with a trace as an argument, like λxλc[c + S⊢'John was born in x']. Such quantifications into speech acts have been proposed for alternative questions and constituent questions in Krifka (2015).

I want to raise, as a final point, the notion of strength. In the commitment account of assertions, assertive strength can be operationalized by the level of warrant with which the speaker backs up a commitment to the asserted proposition. This idea was introduced by Ch. S. Peirce in terms of the "esteem", or reputation, of the speaker that may be affected if the asserted proposition turns out to be false and the speaker does not have a sufficient excuse (cf. Tuzet 2006). In general, stronger assertions result in an increased threat of one's reputation than weaker ones, where strength can be mitigated or increased by epistemic and evidential operators and by direct modifications of commitments (cf. Krifka & Modarresi 2023). In accounts of assertions that work with the notion of expressed beliefs, one would have to assume different belief strengths.

6 Conclusion

In conclusion, we propose a novel explanation of exceptions to Hurford's generalization like (3)(b) that have not been given a satisfying explanation so far, and tentatively extend this to better-known exceptions (2) that have been explained differently. Our approach also throws light on the possible scope of disjunction over speech acts and the role of speech acts in general. It also provides new evidence for the graded nature of assertive commitments and the mitigation and strengthening of speech acts.

References

Brandom, Robert B. 1994. *Making it explicit. Reasoning, representing, and discourse commitment*. Cambridge, Mass.: Harvard University Press.

Ciardelli, Ivano, & Floris Roelofsen. 2017. Hurford's constraint, the semantics of disjunction, and the nature of alternatives. *Natural Lang Semantics* 25.

Chierchia, Gennaro, Fox, Danny, & Spector, Benjamin. 2009. Hurford's constraint and the theory of scalar implicatures: Evidence for embedded implicatures. Ed. Paul Egré & G. Magri, *Presuppositions and implicatures*. MIT Working Papers in Linguistics 60, 47-62.

Chierchia, Gennaro, Fox, Danny, & Spector, Benjamin. 2012. Scalar implicature as a grammatical phenomenon. Ed. Claudia Maienborn, Klaus von Heusinger, & Paul Portner, *Handbook of Semantics Vol. 3*. Berlin: Walter de Gruyter. 2997-2331.

Cohen, Ariel, & Manfred Krifka. 2014. Superlative quantifiers and meta-speech acts. *Linguistics and Philosophy* 37, 41-90.

Ducrot, Oswald. 1973. *La preuve et le dire*. Paris: Mame.

Fox, Danny, & Benjamin Spector. 2018. Economy and embedded exhaustification. *Natural Language Semantics* 26(1), 1-50.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presuposition and logical form*. New York: Academic Press.

Geurts, Bart, & Rick Nouwen. 2007. 'At least' et al.: The semantics of scalar modifiers. *Language*. 83(3), 533-559.

Geurts, Bart. 2019. Communication as commitment sharing: speech acts, implicatures, common ground. *Theoretical Linguistics* 45(1-2), 1-30.

Greenberg, Yael, & Wolf, Lavi 2018. Gradable assertion speech acts. *NELS* 48.

Grice, H. Paul. 1975. Logic and conversation. Ed. Peter Cole & Jerry L. Morgan, *Syntax and Semantics 3: Speech Acts*. New York: Academic Press. 41-58.

Hénot-Mortier, Adèle 2023. Alternatives are blind to some but all kinds of context: The view from Hurford disjunctions. *Sinn & Bedeutung* 27, 291-308.

Hurford, James R. 1974. Exclusive or Inclusive Disjunction. *Foundations of Language* 11, No. 3, 409-411.

Incurvati, Luca, & Julian J. Schlöder. 2019. Weak assertion. *Philosophical Quarterly* 69, 741-770.

Kamali, Beste, & Manfred Krifka. 2020. Focus and contrastive topic in questions and answers, with particular reference to Turkish. *Theoretical Linguistics* 46, 1-71.

Katzir, Roni, & Singh, Raj 2014. Hurford disjunctions: Embedded exhaustification and structural economy. *Sinn & Bedeutung* 18.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9(1), 1-40.

Krifka, Manfred 2015. Bias in Commitment Space Semantics: Declarative questions, negated questions, and question tags. *Semantics and Linguistic Theory (SALT)* 25, 328-345.

Krifka, Manfred 2022. Adjacency pairs in common ground update: Assertions, questions, greetings, offers, commands. *SemDIAL* 26, 94-105.

Krifka, Manfred. 2023. Layers of assertive clauses: Propositions, judgments, commitments, acts. Ed. Julia Hartmann & Angelika Wöllstein, *Propositionale Argumente im Sprachvergleich: Theorie und Empirie*. Tübingen: Gunter Narr. 115-181.

Krifka, Manfred, & Modarresi, Fereshteh 2023. Modifications of assertive commitments and their effect on trustworthiness. *Construction of Meaning Series, Stanford University*. https://tinyurl.com/TrustCommitment

Lassiter, Daniel. 2016. Must, knowledge, and (in)directness. *Nat Lang Semantics* 24(2), 117-163.

Marty, Paul, & Jacopo Romoli. 2022. Varieties of Hurford disjunctions. *Semantics and Pragmatics* 15(3), 1-25. http://dx.doi.org/10.3765/sp.15.3

McDowell, Joyce P. 1987. *Assertion and modality.* Doctoral dissertation, University of Southern California.

Merin, Arthur. 2003. Replacing 'Horn scales' by act-based relevance orderings to keep negation and numerals meaningful. Ed. *Forschungsberichte der DFG-Forschergruppe 'Logik in der Philosophie'. No.110.* Konstanz. Also in SemanticsArchive.

Meyer, Marie-Christine. 2014. Deriving Hurford's constraint. *SALT* 24, 577-594.

Searle, John R., & Vanderveken, Daniel. 1985. *Foundations of illocutionary logic.* Cambridge: Cambridge University Press.

Sbisà, Marina. 2001. Illocutionary force and degrees of strength in language use. *Journal of Pragmatics* 33, 1791-1814.

Shapiro, Lionel. 2020. Commitment accounts of assertion. Ed. Sanford Goldberg, *Oxford Handbook of Assertion.* Oxford: Oxford University Press. 73-97.

Simons, Mandy. 2001. Disjunction and alternativeness. *Linguistics and Philosophy* 24(5), 597-619.

Singh, Raj. 2008. On the interpretation of disjunction: asymmetric, incremental, and eager for inconsistency. *Linguistics and Philosophy* 31, 245-260.

Tuzet, Giovanni. 2006. Responsible for Truth? Peirce on judgement and assertion. *Cognitio* 7, 317-336.

Tomioka, Satoshi. 2021. Scalar implicature, Hurford's constraint, contrastiveness and how they all come together. *Frontiers in Communication* 5, 1.

Westera, Matthijs. 2019. Hurford disjunctions: an in-depth comparison of the grammar. Semanticsarchive.net.

Yatsushiro, Kazuko et al. 2022. Certainly but not certain: The expression of subjective and objective probability. *Glossa: a journal of general linguistics* 7(1).

Zhang, Yichi. 2022. *New perspectives on inquisitive semantics.* Doctoral dissertation, University of Maryland.

Circumstantial Indirectness: Future Orientation of Modals Revised

Jakob Majdič
Masaryk University Brno
Brno, Czech Republic
jakob.majdic@phil.muni.cz

Abstract

Circumstantially interpreted modals are obligatorily future oriented. Standard accounts of this correlation rely on a condition demanding that the modal base must be *diverse* with respect to the prejacent proposition (e.g. Condoravdi 2002; Rullmann and Matthewson 2018; Williamson 2021). I present evidence against this account by showing that circumstantial *will* is veridical and compatible with full certainty of the prejacent. I develop an account on which future orientation of circumstantial modals results from an 'indirectness presupposition' that has been independently motivated for the epistemic domain (von Fintel and Gillies 2010).

1 Introduction

This paper is concerned with the interaction between the interpretation of a modal auxiliary (its 'flavor') and the temporal interpretation of its embedded proposition (its 'prejacent'). Following Condoravdi (2002) I will call the time of evaluation of a modal its *temporal perspective* and the time of evaluation of the prejacent its *temporal orientation*. The central empirical puzzle of this paper is that circumstantial[1] modality obligatorily gives rise to future temporal orientation relative to the modal perspective. More precisely:

(1) **Puzzle:** For any modal $\mathcal{M}$ with temporal perspective t and temporal orientation t', if $\mathcal{M}$ is interpreted circumstantially, then $t < t'$.

Most existing accounts of this correlation involve DIVERSITY, a condition requiring that the modal base of a modal must be *diverse*, i.e. non-homogeneous, with respect to the prejacent. However, DIVERSITY is at odds with the *veridical* semantic profile of circumstantial (i.e. 'temporal') *will* which is commonly considered a modal auxiliary. A case in point are sentences like (2), which seem to be incompatible with DIVERSITY.

(2) In 1 hour, 60 minutes **will** have passed.

The present paper derives obligatory future orientation of circumstantial modals via an independently motivated mechanism of modal indirectness going back to von Fintel and Gillies (2010) which is compatible with full prejacent certainty and veridicality of the modal.

Section 2, recaps the temporal interpretation of modal auxiliaries and briefly describes the standard account of future orientation of modals. In Section 3, I present three pieces of data involving English circumstantial *will*, which will prove to be incompatible with the standard accounts, which rely on DIVERSITY. Section 4 contains my analysis of future orientation of modals. Section 5 concludes the paper.

[1]Note that this paper is specifically concerned with the *metaphysical* (or 'non-priority'(Thomas 2014)) subkind of circumstantial modality. I.e. circumstantial modality in its most basic form, in which it does not involve any ordering. Circumstantial modality seems to be future oriented quite generally, however (see Skibra 2020; Rullmann and Matthewson 2018).

2 Circumstantial Modality and Temporal Interpretation

Modal flavor is known to correlate with temporal orientation: circumstantial modals are invariably future oriented, while epistemic modals can be past or present oriented.[2] The English auxiliary *will*, which according to the vast majority of accounts is a modal auxiliary (Cariani and Santorio 2018; Condoravdi 2002; Copley 2002; Giannakidou and Mari 2018; Kaufmann 2005; Werner 2006 among others) brings out very clearly the questions regarding the nature of this correlation. A major argument in favor of a modal analysis of *will* is the existence of epistemic readings of *will* as exhibited by the sentence in (3) (see Klecha 2014 for further arguments). The temporal interpretation of *will* exhibited in (4), i.e. the reading usually glossed as a *future tense* in traditional grammars, has been claimed to correspond to the circumstantial reading of *will*.

(3) (*QUD: Where is Meg hiding?*) (epistemic)
She**'ll** be hiding in the attic.

(4) (*QUD: Who will win tomorrow?*) (circumstantial)
The Hens **will** win against the Boars (tomorrow).

Modal analyses of *will* capture the difference between the two readings in (3) and (4) in terms of a shift of the modal base: *will* takes an epistemic modal base in (3) and a circumstantial modal base in (4). This assumption alone leaves unexplained why *will*, and circumstantial modals in general, induce future orientation.[3] Existing accounts resort to a definedness condition demanding DIVERSITY of the modal base with respect to the prejacent proposition, as defined in (5). Together with a branching future conception of circumstantial modality, DIVERSITY assures that the evaluation time of the prejacent follows the evaluation time of the modal (Werner 2006; Rullmann and Matthewson 2018).[4]

(5) **Diversity condition:** (adapted from Rullmann and Matthewson 2018, p.300)
For any modal $\mathcal{M}$, prejacent p_j of type $\langle s,t\rangle$ with temporal orientation j, world w, time t and assignment g:
$[\![\mathcal{M}(MB)(p_j)]\!]^{w,t,g}$ is defined only if: $\exists w', w''[w', w'' \in MB(w,t)\ \&\ p_j(w')=1\ \&\ p_j(w'')=0]$.

A future-branching modal space is modeled as the set of worlds that are identical up to a certain time t and may diverge from each other after t. The relation of historic equivalence relative to a time is defined in possible world semantics as in (6). A circumstantial modal base then corresponds to the set specified in (7) (see Condoravdi 2002; Kaufmann 2005).

(6) **Historic Equivalence**
For all $t \in T$, $\simeq_t$ is an equivalence relation on $W \times W$ such that:
for any w,w' $\in$ W and t,t' $\in$ T, if $w \simeq_t w'$, and $t' < t$, then $w \simeq_{t'} w'$

(7) **Circumstantial Modal Bases**
$MB_{circ}(w)(t) = \{w' \mid w \simeq_t w'\}$.

Future orientation then results from DIVERSITY whenever the modal base is circumstantial: since all the worlds in a circumstantial modal base $MB_{circ}(w)(t)$ are identical to w up to t,

[2]The literature does not provide a definite answer to the question whether epistemic modals can be future oriented too. See Williamson (2021) for the view that they can and Klecha (2016) for the view that they can not.

[3]Positing an additional future-orienting temporal component either in the lexical entries of the modal or in a separate head in its scope (e.g. Cariani and Santorio 2018; Condoravdi 2002; Cariani 2021; Copley 2002; Giannakidou and Mari 2018; Rullmann and Matthewson 2018; Williamson 2021) simply shifts the question to why this temporal component does not induce future orientation with non-circumstantial flavors.

[4]The accounts by Klecha (2016) and Skibra (2020) are notable exceptions to DIVERSITY-approaches. The former account hard wires temporal orientation into the denotation of the respective modal base. The latter pursues an analysis in terms of causation. A comparison with my account is beyond the scope of this paper.

circumstantial modal bases can satisfy DIVERSITY only relative to a time following t.

3 Against Diversity: novel data from temporal *will*

If circumstantial *will* induces future orientation due to DIVERSITY, *will* should be incompatible with full certainty of the prejacent proposition, i.e. it should be infelicitous to assert circumstantial *will p* if all circumstantially accessible worlds support p. This prediction is disproved by examples in which the prejacent of *will* expresses a logical relation between two standardized measures of time, as in (8).

(8) In one hour, 60 minutes **will** have passed.

According to DIVERSITY, (8) should be felicitous only if the circumstantial modal space can be divided into one set of worlds w such that 60 minutes will have passed after one hour in w and one set of worlds w' such that 60 minutes *won't* have passed after one hour in w'. This goes against our basic intuitions regarding (8) and even seems impossible to model using a world-time frame. (8) is therefore a clear counterexample to the claim that *will* depends on DIVERSITY.

Second, *will* is compatible with a prejacent that is logically inferred from a known fact.

(9) **Context:** Al won an Olympic medal last year.
   a. [Al:] I have already won one Olympic medal. And I hope to win one more this year.
   b. [Bo:] In any case, at the end of your life you **will** have won at least one Olympic medal. So that's already very impressive.

In the context of (9), the proposition that *Al has won one gold medal* is true at the time of the utterance. It follows that for *any* world circumstantially accessible at the time of utterance the proposition *Al has won one Olympic medal* is true at any future time. The circumstantial modal base in (9) is hence non-diverse with respect to the prejacent of (9) as it homogenously supports the truth of the prejacent. Yet the *will*-sentence used by Bo in (9-b) is felicitous.

The third argument comes from the behavior of *will* regarding speaker commitment to the prejacent proposition. The scenario in (10) demonstrates that a speaker asserting *will p* commits herself to the truth of p.[5]

(10) a. [A:] The ship **will** arrive tomorrow. (based on von Fintel and Gillies 2010)
   b. [two days later, B:] See the ship did not arrive. You were wrong.
   c. [A:] #I was not wrong. Look, I only said 'the ship *will* arrive tomorrow'.

A's pointing to her use of *will* does not help her attempt to escape B's accusation of having asserted something false. This shows that through the assertion of *will p*, the prejacent p necessarily ends up being asserted as well. Since a modal operator can only assert its prejacent if it is veridical, I conclude that *will* is veridical, as defined in (11).

(11) **Veridicality**
A modal operator $\mathcal{M}$ is veridical iff $[\![\mathcal{M}(p)]\!]^w \rightarrow [\![p]\!]^w$. Else it is non-veridical.

Veridicality and DIVERSITY are incompatible. If the modal base of a universally quantifying modal satisfies DIVERSITY, then the evaluation world of the modal cannot be guaranteed to be

[5] Interestingly, even though *will p* seems to make a weaker statement than *must p* (Cariani and Santorio 2018; Mihoc et al. 2019), both variants come with obligatory speaker commitment to the prejacent, even in the case of epistemic *will* (Majdič 2022).

in the domain of quantification anymore, resulting in non-veridicality.[6] A 'centered' selection function analysis of *will* (Cariani and Santorio 2018; Cariani 2021), on which unembedded *will* essentially passes down its evaluation world to the prejacent, is incompatible with DIVERSITY too.[7] DIVERSITY as a general principle predicts that circumstantial *will p* must *always* signal that *p* is circumstantially uncertain. But due to its lexical semantics, *will p* can only ever be truthfully uttered, if *p* is true in the world of evaluation and hence is predicted to signal the exclusion of $\neg p$ worlds. Therefore, without any additional assumptions, veridicality of a modal is incompatible with a general requirement of DIVERSITY.

To conclude, temporal *will* does not behave as expected from a modal conforming to DIVERSITY. I have shown that temporal *will* is compatible with full circumstantial certainty of the prejacent and that it behaves like a veridical modal with regards to speaker commitment. I conclude that *will* is not subject to DIVERSITY. It follows that DIVERSITY cannot be the source of future orientation of circumstantial *will*. This leads me to reject DIVERSITY as the source of future orientation for *any* circumstantial modal.

4 Proposal: Indirectness as a source of future orientation

I propose an account that derives future orientation of circumstantial modals without resorting to DIVERSITY. Following von Fintel and Gillies (2010), I assume that a modal base is derived from a *kernel*. Intuitively, a circumstantial kernel is the set of propositions for which the past and present is wholly relevant. My account derives future orientation via an indirectness condition demanding that the prejacent must not be an issue in the partition induced by the kernel.

4.1 Setting the stage

I assume that the maximal projection of the prejacent branches into a node denoting a predicate of times (type $\langle i,t\rangle$) and a node denoting a semantically vacuous temporal pronoun Vac_i (type i). The modal combines with the prejacent via intensional function application (von Fintel and Heim 2021).

(12) a. $[_{\text{ModP}}$ *will/might* $[_{\text{PrejacentP}}$ [$\boldsymbol{Vac}_i$] [$\lambda t.p(t)$]]]
b. $[\![Vac_i]\!]^g$ = g(i) if g(i) is a time and undefined otherwise

I follow von Fintel and Gillies in assuming that the modal base of a modal is derived from a set of propositions K (for 'kernel) by intersecting all the propositions in K, i.e. $MB = \bigcap K$.

Which set of propositions corresponds to the kernel K_{circ} of a circumstantial modal base? Intuitively speaking, the set of historically equivalent worlds share a set of propositions p such that *what has already happened* in w at t is wholly relevant for p. To model this set, I first define the set of atomic sentences that are true in w at some time before or at t:

(13) **Atomic sentences about the past**
For any sentence ϕ, $\phi \in [\![\boldsymbol{HIST}^{\boldsymbol{atom}}_{\boldsymbol{t_0}}]\!]^{w,g}$ iff:
a. ϕ is of the form $\boldsymbol{p(k_1,\ldots k_{n-1},t_i)}$ where $\boldsymbol{p}$ is a predicate of n-arity, $\boldsymbol{k_1 \ldots k_{n-1}}$ are terms, $\boldsymbol{t_i}$ is a time constant with index i st. $g(i) \leq g(0)$
b. and $[\![\phi]\!]^{w,g}$=1.

[6] It is quite clear that *will* is not a necessity modal (see Cariani and Santorio 2018; Cariani 2021). For further data that can be interpreted as arguments against a necessity analysis see Mihoc et al. 2019.

[7] Cariani and Santorio (2018) assume that *will* denotes a selection function *s*: $\mathcal{W}\times(\ \mathcal{W} \to \mathcal{P}(\mathcal{W})\ ) \to \mathcal{W}$ such that (a) if $MB(w)$ is non-empty, then $s(w,MB(w)) \in$ MB(w) ('success') and (b) if $w \in MB(w)$, then $s(w,MB(w)) = w$ ('centering').

The set of all atomic sentences true up to t indeed affords a complete description of the history up to t. Still $HIST^{atom}$ is not suitable as a circumstantial kernel. $HIST^{atom}$ only contains *atomic* sentences. But, given non-atomic prejacents, we will need the set of *all* sentences that are entirely about the past – atomic and non-atomic. I propose that a circumstantial kernel $K^{circ}_{t_0}$ corresponds to the set of sentences ϕ of the form $p(t_i)$ such that t_i is not preceded by t_0 and such that ϕ is entailed by an element in $HIST^{atom}_{t_0}$. (See example (18) for an illustration).

(14) **Circumstantial Kernels**
For any sentence ϕ, $\phi \in [\![K^{circ}_{t_0}]\!]^{w,g}$ iff:
a. $\exists\ \psi\ [\psi \in [\![HIST^{atom}_{t_0}]\!]^{w,g}\ \ \&\ \ \psi \rightarrow \phi]$
b. ϕ is of the form $p(t_i)$ where p is any formula and t_i is a time such that $t_i \leq t_0$

From a circumstantial kernel K^{circ} the set of historically equivalent worlds is then derived by intersecting all the propositions in K^{circ} (i.e. $MB^{circ} = \bigcap K^{circ}$).[8]

4.2 Future orientation through indirectness

To account for the epistemic indirectness signaled by epistemic *must*, von Fintel and Gillies (2010) propose, in one of two suggested implementations, that epistemic modals have the presupposition in (15).[9] This presupposition demands that the prejacent proposition must not be an *issue* in the *subject matter* induced by the propositions in the kernel. I assume that a modal takes as its arguments a proposition p, a temporal pronoun t_i and a modal base $\bigcap K_{ti}$, whose time argument is co-indexed with the matrix time t_i.

(15) **Indirectness**
$[\![Mod(p)(\bigcap K_{t_i})(t_i)]\!]^{w,g}$ is defined only if p is not an issue in $S_{K_{t_i}}$
(where $S_{K_{t_i}}$ is the subject matter induced by K_{t_i})

The subject matter, determined by a proposition p can be captured as an equivalence relation between two worlds w, v such that w and v are in the same equivalence class if and only if p has the same truth value in w and v. This partitions logical space into two cells: one cell of worlds in which p is true and one in which p is false, see (16).

(16) A subject matter S is an equivalence relation on W such that
a. $S[P] = \{\langle w, v\rangle \in S : w \in P \text{ iff } v \in P\}$
b. P is an issue in S iff $S[P] = S$ (von Fintel and Gillies 2010, p.377)

The subject matter induced by a set of propositions K is then the set of partitions along the boundaries of the equivalence classes of each of the propositions in K, as defined in (17).

(17) Let $K = \{P_1 \ldots P_n\}$, then: $S_K = S[P_1] \ldots [P_n]$ (adapted from ibid.)

Adopting the presupposition in (15) for circumstantial modals allows us to derive future orientation without the problematic assumption of DIVERSITY: K^{circ}_t corresponds to the set of all propositions that are entirely about the past (of w relative to t). The presupposition in (17) demands that the prejacent must not be an issue in the partition induced by K^{circ}_t. Since *any* proposition entirely about the past is settled by this partition, it follows that the prejacent of a circumstantial modal must be future oriented.

To illustrate, let's recall the *Olympic medal* example in (9). In such a scenario, the circumstantial kernel would contain (minimally) the sentences in (18-b).

[8] Note that I will switch freely between the term 'kernel' (and the symbol K) as standing for a set of sentences and standing for the corresponding set of propositions, which contains for each sentence ϕ in K the intension of ϕ.

[9] My proposal represents additional evidence in favor of von Fintel and Gillies' (2010) *Implementation 2*.

(18) Let $\boldsymbol{t_1}, \boldsymbol{t_2}, \boldsymbol{t_3}, \boldsymbol{t_4}$ denote time constants such that $g(1) < g < (2) < g(3) < g(4) < g(5)$ are times ordered by temporal precedence ($<$) as indicated. Then:

a. $HIST^{atom}_{\boldsymbol{g(3)}} = \{\boldsymbol{Al\text{-}wins\text{-}medal(t_1)}, \ldots\}$

b. $K^{circ}_{\boldsymbol{g(3)}} = \{\boldsymbol{Al\text{-}wins\text{-}medal(t_1)}, \boldsymbol{PERF_i(Al\text{-}wins\text{-}medal(t_i))(t_2)}, \boldsymbol{PERF_i(Al\text{-}wins\text{-}medal(t_i))(t_3)}, \ldots\}$[10]

The circumstantial kernel $K^{circ}_{\boldsymbol{g(3)}}$ contains the sentences listed in (18-b). In accordance with (14), each of the sentences in $K^{circ}_{\boldsymbol{g(3)}}$ is entailed by an element in $HIST^{atom}_{\boldsymbol{g(3)}}$ and has as its temporal argument a time that is not later than $g(4)$. But crucially now, the kernel $K^{circ}_{\boldsymbol{g(3)}}$ does not contain the sentence $\boldsymbol{PERF_i(Al\text{-}wins\text{-}medal(t_i))(t_4)}$ because its time argument follows the time $g(3)$ relative to which the circumstantial kernel is evaluated. Consider now the sentence in (19) with the LF in (19-a), whose prejacent corresponds to the sentence in (19-b).

(19) Al will$_{\text{circ}}$ have won one medal.

a. LF: [... [$_{\text{Mod}}$ *will*] [$_{\text{PrejacentP}}$ [Vac_{i}] [$\lambda t.\ PERF_k(Al\text{-}wins\text{-}medal(t_k))$(t)]]]

b. PrejacentP $\rightsquigarrow$ $\boldsymbol{PERF_m(Al\text{-}wins\text{-}medal(t_m))(Vac_i)}$

Since $\boldsymbol{Vac_i}$ is defined as long as $g(i)$ is any time, the prejacent $\boldsymbol{PERF_k(Al\text{-}wins\text{-}medal(t_k))(Vac_i)}$ does not contain any information with regards to its temporal orientation. It is the modal which by demanding that its prejacent must not be an issue in its kernel only allows for a future oriented interpretation of the prejacent. For example, if $\boldsymbol{Vac_i}$ would be mapped to the time denoted in (18) by $\boldsymbol{t_2}$, then the modal expression would be undefined since the sentence $\boldsymbol{[PERF_m(Al\text{-}wins\text{-}medal\ (t_m))(t_2)}$ is an element in $K^{circ}_{t_n}$. Since $K^{circ}_{t_n}$ contains all the sentences that are entirely about the past, the prejacent is an issue in $S_{K^{circ}_{t_n}}$ on *any* assignment g that maps i to a non-future time. The indirectness presupposition hence can never be satisfied whenever the temporal pronoun $\boldsymbol{Vac_i}$ is mapped to a non-future time.

(20) a. $[\![\boldsymbol{MODAL}(\bigcap K^{circ}_{t_n})(\boldsymbol{PERF_m(Al\text{-}wins\text{-}medal\ (t_m))(Vac_i)})]\!]^{w,g}$ is defined only if $[\![\boldsymbol{PERF_m(Al\text{-}wins\text{-}medal\ (t_m))(Vac_i)}]\!]^g_c$ is not an issue in $S_{K^{circ}_{t_n}}$ (by (15))

b. $[\![\boldsymbol{MODAL}(\bigcap K^{circ}_{t_n})(\boldsymbol{PERF_k(Al\text{-}wins\text{-}medal\ (t_k))(Vac_i)})]\!]^{w,g}$ is defined only on assignments g such that $g(i) > g(n)$ (by (12-b) and (14))

This predicts that the temporal orientation of the prejacent lies in the future whenever the modal flavor of the modal is circumstantial, as desired.

5 Conclusion

I have suggested a novel account of the cross-linguistically observed correlation between future orientation and circumstantial modal flavor. I have shown that DIVERSITY-accounts of future orientation are incompatible with three pieces of data on English circumstantial *will* that have been hitherto ignored in the literature. I have proposed a novel definition of a circumstantial kernel on which future orientation of circumstantial modals can be modeled as a resulting from an independently motivated indirectness presupposition (von Fintel and Gillies 2010). The account predicts that any modal induces future orientation whenever its modal base is circumstantial.

Acknowledgements. I would like to thank three anonymous reviewers. I am indebted to Kathi Görzer, Nina Haslinger, Magda Kaufmann, Viola Schmitt and Marcin Wągiel for their invaluable comments and feedback. All errors are of course my own.

[10] For the meaning of $\boldsymbol{PERF}$ we could simply assume that $[\![\boldsymbol{PERF_i(\phi)(t_k)}]\!]^g = 1$ iff $\boldsymbol{\phi}$ is a sentence of the form $p(t_i)$ and $g(i)$ is any time such that $g(i) < g(k)$. But any other commonly assumed denotation would do.

References

Cariani, Fabrizio (2021). *The modal future: A theory of future-directed thought and talk*. Cambridge University Press.

Cariani, Fabrizio and Paolo Santorio (2018). "Will done Better: Selection Semantics, Future Credence, and Indeterminacy". In: *Mind* 127.505, pp. 129–165. DOI: 10.1093/mind/fzw004. URL: https://academic.oup.com/mind/article/127/505/129/3053474 (visited on 02/21/2021).

Condoravdi, Cleo (2002). "Temporal interpretation of modals: Modals for the present and for the past". In: *The construction of meaning*. Ed. by David Beaver, Stefan Kaufmann, Brady Clark, and Luis Casiillas. CSLI Publications Stanford, CA, pp. 59–88.

Copley, Bridget (2002). "The Semantics of the Future". PhD thesis. Cambridge, MA: Massachusetts Institute of Technology.

Giannakidou, Anastasia and Alda Mari (2018). "A unified analysis of the future as epistemic modality: The view from Greek and Italian". In: *Natural Language & Linguistic Theory* 36.1, pp. 85–129. DOI: 10.1007/s11049-017-9366-z. URL: http://link.springer.com/10.1007/s11049-017-9366-z (visited on 06/16/2020).

Kaufmann, Stefan (2005). "Conditional Truth and Future Reference". In: *Journal of Semantics* 22.3, pp. 231–280. DOI: 10.1093/jos/ffh025. URL: http://academic.oup.com/jos/article/22/3/231/1672017/Conditional-Truth-and-Future-Reference (visited on 06/18/2020).

Klecha, Peter (2014). "Diagnosing Modality in Predictive Expressions". In: *Journal of Semantics* 31.3, pp. 443–455. DOI: 10.1093/jos/fft011. URL: https://academic.oup.com/jos/article-lookup/doi/10.1093/jos/fft011 (visited on 06/19/2020).

— (2016). "Modality and embedded temporal operators". In: *Semantics and Pragmatics* 9, pp. 9–1.

Majdič, Jakob (2022). "Temporal and epistemic readings of English "will" and German "werden"". MA thesis. University of Vienna. URL: https://phaidra.univie.ac.at/detail/o:1544849 (visited on 11/07/2024).

Mihoc, Teodora, Diti Bhadra, and Anamaria Falaus (2019). "Epistemic modals, deduction, and factivity: new insights from the epistemic future". In: *Semantics and Linguistic Theory*. Vol. 29, pp. 351–370.

Rullmann, Hotze and Lisa Matthewson (2018). "Towards a theory of modal-temporal interaction". In: *Language* 94.2, pp. 281–331. DOI: 10.1353/lan.2018.0018. URL: https://muse.jhu.edu/article/697182 (visited on 03/14/2024).

Skibra, Daniel (2020). "Future Orientation on an Event-Relative Semantics for Modals". In: *NELS 49: Proceedings of the Forty-Ninth Annual Meeting of the North East Linguistic Society*. Ed. by Maggie Baird. Vol. 3. Amherst, MA: GLSA, Dept. of Linguistics, pp. 149–162.

Thomas, Guillaume (2014). "Circumstantial modality and the diversity condition". In: *Proceedings of Sinn und Bedeutung*. Vol. 18, pp. 433–450.

von Fintel, Kai and Anthony S. Gillies (2010). "Must . . . stay . . . strong!" In: *Natural Language Semantics* 18.4, pp. 351–383. DOI: 10.1007/s11050-010-9058-2. URL: http://link.springer.com/10.1007/s11050-010-9058-2 (visited on 06/16/2020).

von Fintel, Kai and Irene Heim (2021). *Intensional semantics. Updated Version 2021*. Unpublished manuscript, MIT.

Werner, Tom (2006). "Future and non-future modal sentences". In: *Natural Language Semantics* 14.3, pp. 235–255. DOI: 10.1007/s11050-006-9001-8. URL: http://link.springer.com/10.1007/s11050-006-9001-8 (visited on 06/16/2020).

Williamson, Gregor (2021). "Worlds of the Future: Modality & Future Licensing at the Syntax-Semantics Interface". PhD thesis. UCL (University College London).

Northern Khanty demonstratives: a markedness-based approach

Lera Marinina
Individual researcher
vlrmarinina@gmail.com

Abstract

This paper presents a semantic analysis of the Northern Khanty demonstratives. Northern Khanty has a tripartite demonstrative system: proximal term *tăm*, distal term *tum* and unmarked one *śi*. All terms may be used both exophorically and anaphorically. Based mainly on Ahn's (2022) and Reisinger and Huijsmans' (2021) approaches to the demonstratives, I account for the data postulating allosemy and presuppositions of different strengths for Northern Khanty demonstratives.

1 Introduction

1.1 Language and speakers

The current paper provides descriptive data and analysis of demonstratives in Kazym dialect of Northern Khanty language (Khanty < Uralic; Russia). According to Kaksin (2010) Kazym dialect is spoken in the upper reaches of the Ob river in Russia by about 1700 people. The data were collected during fieldwork in the village of Kazym (Khanty-Mansi Autonomous Region - Yugra, Russia) in 2022-2024.

1.2 Methodology

All the consultants I worked with are bilingual, i.e., they speak Northern Khanty and Russian, and their first acquired language is Northern Khanty. Russian was used as lingua franca during fieldwork. I use several methods in my research:

- elicitation
- for exophoric uses: elicitation from Russian-language stimuli based on Wilkins' (1999b) experimental questionnaire (distinctions in anchoring, distance, and visibility)
- for contrastive exophoric uses: elicitation from Russian-language stimuli based on experimental questionnaires of Rostovtsev-Popel (2009) and Wilkins (1999a)
- analysis of corpus examples
- quasi-experimental task (referential communication: instructing another person to arrange the objects in the proper order)

2 Northern Khanty demonstrative system

2.1 An overview

Northern Khanty is an articleless language. It has 3 demonstrative bases. The discussion in this paper is mainly concerned with the adnominal forms *tăm* (proximal), *tum* (distal) and *śi* (unmarked). There exist a special pronominal form for the unmarked term *śit*, and some speakers also have special pronominal forms *tămi* and *tumi*. No differences between adnominal

and pronominal forms were observed, besides syntactic differences. All three demonstratives may be used exophorically and anaphorically; no other uses were detected. Table 1 presents basic distinctions that Northern Khanty demonstratives reveal.

Table 1: The inventory of Northern Khanty demonstratives

Demonsrative	Deictic center	Spatial meaning
tăm	Speaker	Proximal
tum	Speaker&Addressee	Distal
śi	Speaker&Addressee	Unmarked

I also want to notice that unmarked term *śi* has grammaticalized into the emphatic particle (for instance, see ex. (10) below). I consider it to be a distinct lexeme synchronically and do not discuss it in this paper.

2.2 Exophoric uses

All three demonstratives in Northern Khanty may be used exophorically. Practically, speakers tend to volunteer *tăm* and *tum* wherever possible, and *śi* is more often accepted by the consultants than volunteered[1]. Proximal demonstrative *tăm* is usually used for the referents that speaker may reach with the arm, cf. examples (1) and (2) with reachable and unreachable referent respectively:

(1) tăm/#śi/#tum kinškaj-en ma λʉŋət-s-əm.
this/#DEM/#that book-POSS.2SG I read-PST-1SG
'I read this book.'(a book in front of the speaker, context 6 from Wilkins (1999b), further cited as WQ6)[2]

(2) #tăm/śi/OKtum kinškaj-en ma λʉŋət-s-əm.
#this/DEM/OKthat book-POSS.2SG I read-PST-1SG
'I read this book.' (WQ12: speaker and addressee next to each other, a book is to the side of the addressee and is unreachable for the speaker)

However, the scale of proximity is relative and may be expanded dependent on referent type and positions of the interlocutors. For example, *tăm* is used referring to large referents that are close enough but inaccessible with an arm (e.g. a house several meters away) and to small referents that are between speaker and addressee but unreachable for both of them.
Distal demonstrative *tum* is in complementary distribution with proximal term *tăm*. *Tum* is used for the referents outside the shared space[3] (see (2) above and (3) below):

(3) muj păta #tăm/OKśi/OKtum kinškaj-en kamən kɛrət'λ'ə-λ?
why #this/OKDEM/OKthat book-POSS.2SG outside lie-NPST[3SG]
'Why is that book lying outside?' (WQ21: a book is several meters away from speaker and addressee)

Unmarked demonstrative *śi* is not in complementary distribution with *tăm* and *tum*. In fact, it is allowed in all contexts where distal *tum* is used (e.g. (2) and (3) above). Regarding the

[1]In the examples I use no symbol for the volunteered variant and OK for the accepted one.
[2]In this example and in the majority of the following ones: the possessive marker of 2sg -en functions as a salient article, see Mikhailov (2024) for the discussion.
[3]Shared space is the space in-between the interlocutors.

variation of *tăm* and *śi*, *śi* is alternative to *tăm* in the situations when the referent is reachable for both interlocutors, but is located closer to the addressee:

(4) ^{OK}tăm/^{OK}śi/$^{\#}$tum kinškaj-en năŋ-en?
^{OK}this/^{OK}DEM/$^{\#}$that book-POSS.2SG you-DAT
'Is this book yours?' (WQ9: speaker and addressee sitting next to each other, a book is in front of the addressee)

When proximal and distal terms are unsuitable, *śi* arises as the only one possible option:

(5) $^{\#}$tăm/śi/$^{\#}$tum kinškaj-en χuλśa wʉ-s-ən?
$^{\#}$this/DEM/$^{\#}$that book-POSS.2SG from.where take-PST-2SG
'Where did you take that book from?' (WQ16: several meters between speaker and addressee, the referent is near the addressee)

In contrastive uses with two referents *tăm* and *tum* are volunteered regardless of the addressee's position; *śi* is rarely volunteered, but possible (ex. (6) below). This fact also speaks in favor of unmarkedness of *śi* (see Meira and Terrill (2005) for the discussion of unmarked demonstratives in contrastive contexts).

(6) tăm/$^{\#}$śi/$^{\#}$tum an-ən pun-man šaj, $^{\#}$tăm/^{OK}śi/tum an-ən pun-man jiŋk.
this/$^{\#}$DEM/$^{\#}$that cup-LOC put-CVB tea $^{\#}$this/^{OK}DEM/that cup-LOC put-cvb water
'This cup has tea in it, and that cup has water in it.' (Context based on Rostovtsev-Popel (2009), the first cup is closer to the speaker than the second cup)

Summing up, proximal *tăm* and distal *tum* encode clear distance distinctions, while *śi* may be used in a wide range of contexts and is semantically vague with respect to the distance.

2.3 Attention and salience

According to Diessel (2006), there are two main functions of demonstratives: indicating the location of a referent and coordinating the focus of interlocutors' joint attention. Since *śi* is unmarked in distance, I hypothesized that its functions are related to attention, i.e. it is favored for the referents in the joint focus of attention. This hypothesis was confirmed only partly.
In quasi-experiment, *śi* was not used for drawing addressee's attention to the referent at all. Instead, it was often used in the contexts of confirmation like (7), with the referents in the focus of joint attention:

(7) [There are two sweets that are not arranged yet. Participant looks at the photo and then points to the position near the previous placed object on the toy table]
Participant:
maw tiw pun-e.
sweet here.ILL put-IMP.SG>SG
'Put [some] sweet here.'
Assistant: [looks at two sweets that are not arranged yet and takes one of them]
Participant (confirming):
śi maw-en.
DEM sweet-POSS.2SG
'This sweet [is the one I intended].'

However, in other (non-confirming) contexts with referents in joint attention *śi* is not preferable: the choice of the demonstrative rather depends on distance to the referent, as in (8). It seems that in examples like (7) anaphoricity and not attentional focus matters (see the next subsection

for the discussion of anaphoric uses).

(8) [You visit me for the first time, and you do not know that I have a dog. We are sitting in a room and talking. Suddenly a dog runs into the room and starts barking. We look at it, then I say:]
aλ păλ-a, tăm/śit/tum ma amp-ɛm.
PROH be.afraid-IMP[SG] this/DEM/that I dog-POSS.1SG
'Don't be afraid, this/that is my dog.'

Thus, focus of attention itself turns out not to be a distinctive parameter for Northern Khanty demonstrative system. I discuss in Section 3.1 why in spontaneous speech *śi* is not chosen for drawing attention.

2.4 Anaphoric uses

Unmarked term *śi* is natural in anaphoric uses (in the sense of Ahn (2022)) (ex. (9) below), while *tăm* and *tum* rather require specific conditions. Anaphoric *tăm* and *tum* arise in the contexts with more than one salient referent as (10): *tăm* is used for a more salient referent (main character of the story), and *tum* for a less salient one (a secondary character).

(9) [At the table people are discussing a woman who lives in Novosibirsk and studies Khanty. You do not know her, but you heard that she came to Kazym a few years ago. You say:]
śi/#tăm/#tum imij-en Amńa woš-a juχət-s.
DEM#this/#that woman-POSS.2SG Amnya town-DAT come-PST[3SG]
'That woman visited Kazym.'

(10) [The speaker tells a folklore story. The main character could not sleep that night and saw a creature. It touched another man who woke up frightened.]
in tăm ike-λ pa śi nuχ oms-əmt-əs ińśəs-λ-əλλe: "χuti
now this man-POSS.3SG ADD EMPH up sit-PNCT-PST[3SG] ask-NPST-3SG>SG how
ji-s-ən?" in tum ik-en lup-ijəλ-λ: "pa muλsər
become-PST-LOC now that man-POSS.2SG say-FREQ-NPST[3SG] ADD what.INDEF
pun-əŋ jošəp ut-ən pəχəλ-mə-s-ij-əm."
hair-PROP hand-PROP.P thing-LOC poke-MOM-PST-PASS-1SG
'And this man [who told a story] sat and asked him: "What happened?" And that man [who woke up frightened] says: "Someone with a hairy hand poked me."'

3 Analysis

3.1 Current analysis

I combine features of the analyses of Reisinger and Huijsmans (2021) and Ahn (2022). Ahn argues that demonstratives introduce an additional syntactic argument R that can be occupied by a pointing gesture, an anaphoric index or a relative clause. In Northern Khanty R may be occupied by pointing gesture and anaphoric index only. I assume that this happens due to syntactic restrictions and do not discuss it further here. Following Reisinger&Huijsmans' analysis, I include spatial and salience components into the presuppositions of demonstratives. In the analysis I want to capture several crucial facts. Firstly, all three demonstratives may be used exophorically, *tăm* and *tum* are preferred in contrastive contexts. Secondly, *śi* is rarely volunteered in the contexts of drawing the addressee's attention to the referent. Thirdly, *śi* is preferred in anaphoric uses, and *tăm* and *tum* can be used anaphorically if there is more than one salient referent.

I propose to analyze *śi* as an unmarked demonstrative and postulate an underspecified presupposition for it. The referent must meet the descriptive content of N, and the R slot must be occupied, while no spatial or salience information is included:

(11) $[\![\text{śi}]\!]^c = \lambda N.\lambda R.\iota y.N(y) \wedge R(y)$

For *tăm* and *tum* I propose additional presuppositions of proximity and distality respectively (defined below):

(12) $[\![\text{tăm}]\!]^c = \lambda N.\lambda R.\iota y.N(y) \wedge R(y) \wedge PROX(y)$

(13) $[\![tum]\!]^c = \lambda N.\lambda R.\iota y.N(y) \wedge R(y) \wedge DIST(y)$

I propose that PROX and DIST are subject to allosemy: an appropriate interpretation for PROX/DIST is chosen during Spell-Out of the syntactic structure to Logical Form depending on the syntactic context of the demonstrative head. In our case, the conditioning factor is whether GestP, i.e., a gesture phrase, or IdxP, an index phrase, is in the R position:

(14) a. PROX = $\lambda x.\lambda N.$ x is close in space to l_c / ____GestP
b. PROX = $\lambda x.\lambda N.\ x \in Sal_c \wedge \exists y\,[y \in Sal_c \wedge N(y) \wedge y <_{sal} x] \wedge \forall z\ [z \in Sal_c \wedge N(z) \wedge z \neq x \rightarrow z <_{sal} x]$ / ____IdxP

(15) a. DIST = $\lambda x.\lambda N.$ x is far in space from l_c / ____GestP
b. DIST = $\lambda x.\lambda N.\ x \in Sal_c \wedge \exists y\ [y \in Sal_c \wedge N(y) \wedge x <_{sal} y]$ / ____IdxP

I use a contextual parameter l_c — the location of the utterance situation. I do not provide a more explicit definition for 'close' and 'far' since these concepts are to a large extent vague, as the data on exophoric uses show.

Informally, (14) means that if there is GestP in the R position, for *tăm* to be felicitous the referent x (denoted by *tăm*) must be located close to the location of the utterance situation l_c; if there is IdxP in the R position, for *tăm* to be felicitous the referent x must be the most salient in the context c (all other salient referents that meet the descriptive content of N are less salient), and there must be at least one referent that is less salient than x. Similarly, (15) means that if there is GestP in the R position, for *tum* to be felicitous the referent x (denoted by *tum*) must be located far from l_c; if there is IdxP in the R position, for *tum* to be felicitous the referent x (denoted by *tum*) must be salient in the context c, and there must be at least one referent that is more salient than x.

We can now appeal to *Maximize presupposition!* (Heim 1991) to explain the distribution in (1)-(6): if there are several contextually equivalent alternative elements, the element with the stronger presuppositions must be chosen. Applying *Maximize presupposition!* allows us to account for the fact that exophoric *śi* is rarely volunteered in the contexts of drawing attention (see Section 2.3). At the same time, wide distribution of *śi* in both exophoric and anaphoric contexts is predicted: a weaker presupposition of *śi* is less likely to fail, compared with presuppositions of *tăm* and *tum*.

In exophoric uses, *tăm* is predicted to be used for proximal referents and *tum* for distal ones, while *śi* must be preferred for referents that are neither really close to l_c nor quite far enough. In anaphoric uses, if there is more than one salient referent meeting the descriptive content of N, according to *Maximize presupposition! tăm* will be used for the most salient referent, and *tum* for the less salient one (ex. (10)). But if there is only one salient referent, the presuppositions of *tăm* and *tum* fail, so *śi* will be used instead (ex. (9)). Another correct prediction that the analysis makes is that *śi* is preferred for confirmation (ex. (7)): such contexts are treated as anaphoric with one salient referent.

3.2 Other predictions and challenges for the analysis

The proposed analysis with minor speculations lets us account for temporal deixis. Temporal uses of demonstratives may be analyzed as anaphoric. If the referent is a current time interval, it is the most salient referent among all salient referents of certain type, so the presupposition of anaphoric *tăm* is satisfied, as in (16). If the referent is some other time interval, it is not clear if it is more or less salient than the current time interval, so presuppositions of *tăm* and *tum* fail, and *śi* is used, as in (17).

(16) tăm/#śi/#tum jetn wɵn wot-a ji-λ.
this/#DEM/#that evening strong wind-DAT become-NPST[3SG]
'There will be a strong wind this evening.' (The referent is the most salient evening of all the evenings.)

(17) [Starting the story]
śi/#tăm/#tum jetn wotas wɵ-s.
DEM/#this/#that evening blizzard be-PST[3SG]
'There was a blizzard that evening.' (The evening under discussion is equally salient to current evening.)

However, there are some problems for the analysis. Firstly, it does not capture variability, as applying *Maximize presupposition!* implies having only one option in certain context. We may assume that variability is a problem of elicitation or that the speakers accommodate presuppositions, but these explanations are not always satisfactory. Secondly, the analysis does not capture for some contrasts detected in the new field data.
tăm is preferred in anaphoric uses when the referent is present in utterance situation and is located close to the speaker:

(18) [I am walking through Kazym with a new bicycle when I meet a friend. I say:]
want-e mujsər welosipet tăj-λ-əm! tăm/%%śi/#tum
look-IMP.SG>SG what bicycle have-NPST-1SG this/%%DEM/#that
welosiped-ɛm aŋk-ɛm-a mojλə-λ-ɛm.
bicycle-POSS.1SG mother-POSS.1SG-DAT gift-NPST-1SG>SG
'Look, what a bicycle I have! I will give this bicycle as a present to my mother'

Within our analysis, could we treat such uses as exophoric if *tum* behaved the same way. However, *tum* is prohibited in anaphoric contexts when the referent is present in the discourse situation and is far from the speaker; in such contexts *śi* is used. Additionally, there is some evidence that distal *tum* used exophorically requires visual landmarks or another salient referent in the context. Thus, further investigation is needed and the analysis presented here needs to be modified, although the main ideas are likely to remain the same.

4 Conclusions

The proposed analysis unifies two crucial uses of demonstratives, exophoric and anaphoric. It accounts for the main contrasts of the demonstrative system in Northern Khanty and proposes a new idea of limiting some demonstratives to contexts of salience-driven competition. However, current version of the analysis does not catch some data, so it needs modifications. It also does not focus on distance distinctions of exophoric demonstratives; a distinct analysis in the framework of Optimality Theory might be more applicable for the purposes of predicting the distribution in exophoric uses.

References

Ahn, Dorothy (2022). “Indirectly direct: An account of demonstratives and pointing”. In: *Linguistics and Philosophy* 45, pp. 1345–1393.

Diessel, Holger (2006). “Demonstratives, joint attention, and the emergence of grammar”. In: *Cognitive Linguistics* 17.4, pp. 463–489.

Heim, Irene (1991). “Artikel und Definitheit”. In: vol. 17. De Gruyter Mouton.

Kaksin, Andrey (2010). *Kazymskij dialekt khantyjskogo yazyka. (Russian) [Kazym dialect of Khanty language]*. Yugorskii gosudarstvennyi universitet.

Meira, Sérgio and Angela Terrill (2005). “Contrasting contrastive demonstratives in Tiriyó and Lavukaleve”. In: *Linguistics* 43.6, pp. 1131–1152.

Mikhailov, Stepan (2024). “Semantics of the Northern Khanty Salient Article: Definiteness, salience, and obviation”. In: *Voprosy jazykoznanija* 1, pp. 7–38.

Reisinger, Daniel and Marianne Huijsmans (2021). “Demonstratives in ʔayʔaǰuθəm”. In: UBCWPL.

Rostovtsev-Popel (2009). “Tipologiya demonstrativov: srednie dejktiki”. In: *Voprosy yazykoznaniya* 2, pp. 22–34.

Wilkins, David (1999a). “Eliciting contrastive use of demonstratives for objects within close personal space (all objects well within arm’s reach)”. In: Max Planck Institute for Psycholinguistics.

— (1999b). “The 1999 demonstrative questionnaire: “This” and “that” in comparative perspective”. In: Max Planck Institute for Psycholinguistics.

Neg-raising interacts with implicatures: The case of *doubt*[*]

Andrea Matticchio
UMass Amherst
Amherst MA, USA
amatticchio@umass.edu

Abstract

The verb *doubt* is a downward-monotonic attitude predicate. I argue that it should be analyzed as a negated existential modal, and that the inferences it comes with result from the interaction between neg-raising and a possibility implicature. In doing so, I provide a discussion of neg-raising and scalar implicatures arising from one and the same predicate.

1 Introduction

The verb *doubt* is an attitude predicate that creates a downward monotonic environment in its clausal complement; this licenses weak negative polarity items (NPIs) like *ever*.

(1) a. Abigail doubts that Taro eats meat.
∴ Abigail doubts that Taro eats chicken.
b. Abigail doubts that Taro will ever pass the exam.

This predicate has received some attention in the literature. It is generally considered a negative form of *believe* (see for instance Heim 1992, p. 190, Gajewski 2007, p. 297, Chierchia 2013, p. 200, or Abrusán 2016, p. 185), but some work focusing on this predicate has noticed peculiarities in the inferences it leads to and the types of clauses it can embed (Karttunen 1977; Fischer 2003; Biezma and Rawlins 2012; Anand and Hacquard 2013; Uegaki 2023). A shared intuition is that the meaning of *doubt* is more weakly negative than *not doubt* is positive. The examples in (2) show that Moorean contradictions arise with *not doubt*, but not with non-negated *doubt*.

(2) From Uegaki 2023, p.c. with F. Staniszewski
a. I doubt that Jo's sick, but she might be.
b. #I don't doubt that Jo's sick, but she might not be.

The main explanandum of this paper is why *not doubt* conveys a strong doxastic attitude towards the embedded proposition, whereas non-negated *doubt* has a weak meaning, such that the attitude holder can still entertain the possibility that the embedded proposition is true. Throughout the paper, I will make reference to English judgments by my informants as well as my own judgments on Italian, which align with English for all the examples considered.[1]

2 Not so weak

Uegaki (2023) proposes that *doubt* has a weak assertive meaning: it is a negated universal quantifier over the attitude holder's doxastic alternatives, in a Hintikkan semantics. When embedding a declarative CP (the case I'll focus on in this paper), it has the denotation in (3).

(3) $[\![\mathsf{doubt}]\!]^w = \lambda p_{\langle s,t\rangle}\,.\,\lambda x_e\colon \exists w' \in \textsc{dox}_x^w\ p(w')\,.\,\neg\forall w'' \in \textsc{dox}_x^w\ p(w'')$ (Uegaki 2023)

[*]I thank my advisor Vincent Homer and the audience of the MECORE 2024 closing workshop.
[1]Unlike English, Italian *dubitare* 'doubt' does not generally embed finite polar questions.

Under this account, a strong meaning with universal quantificational force is immediately returned when *doubt* is embedded under negation. This explains (2-b). To prevent that non-negated *doubt* has too weak a meaning, Uegaki proposes that an implicature is calculated on domain alternatives (following Bar-Lev 2021) in upward-entailing environments. However, the meaning of *doubt* cannot be strengthened to that of a negated existential quantifier: it is presupposed that at least one accessible world makes the embedded proposition true. For the presupposition of the assertion to be satisfied, some alternatives must be pruned, namely at least those that consider singleton domains. This minimal pruning moderates the strength of *doubt* to less than negated existential, thus guaranteeing that discourses like (2-a) are coherent.

A first concern with this analysis is the meaning that *doubt* has if only minimal pruning applies. Following Bar-Lev (2021), the exhaustification operator $\textsc{exh}^{\textsc{ie}+\textsc{ii}}$, which considers both exclusion and inclusion of alternatives (Bar-Lev and Fox 2020), asserts all propositions obtained by replacing the domain of quantification $\textsc{dox}_x^w$ with its subdomains D. Uegaki seems to implicitly assume that domain alternatives are not presuppositional. The resulting meaning in (4-a), where $|D| \geq 2$ because of the presupposition, is equivalent to the one in (4-b).

(4) Given (3), $[\![\textsc{exh}^{\textsc{ie}+\textsc{ii}}\ x\ \mathsf{doubts}_{\textsc{dox}_x^w}\ p]\!]^w = \mathtt{T}$ if and only if:
   a. $[\![x\ \mathsf{doubts}_{\textsc{dox}_x^w}\ p]\!]^w = \mathtt{T}$ and $[\![x\ \mathsf{doubts}_D\ p]\!]^w = \mathtt{T}$ for all $D \subseteq \textsc{dox}_x^w$ such $|D| \geq 2$
   b. Equivalently, $\exists! w' \in \textsc{dox}_x^w\ p(w')$

Proof. Assume (4-a). By the presupposition of (3), there is at least one $w' \in \textsc{dox}_x^w$ such that $p(w') = \mathtt{T}$. Suppose there is more than one: then, there is a subdomain D of two elements that both make p true. This contradicts (4-a): so, there is just one such w'. Conversely, assume (4-b). The presupposition and the assertion of (3) are satisfied, and no subdomain of two elements will contain more than one world that makes p true. So, all the asserted alternatives are true. □

A meaning that requires a proposition to be true in exactly one world of a domain is implausible. Moreover, Uegaki admits that pruning of more alternatives should be context dependent, but this returns an undesirable form of context dependency. First, it is possible to see, following the proof above, that if all alternatives on domains of cardinality less or equal to n are pruned, the meaning asserts that p is true of $n-1$ worlds in $\textsc{dox}_x^w$. If pruning is licensed by a question under discussion that makes certain alternatives irrelevant (following Bar-Lev 2021), a question that affects pruning for *doubt* must ask whether the number of worlds in $\textsc{dox}_x^w$ that make the embedded proposition true is higher or lower than a certain n. As argued by von Fintel and Iatridou (2008, p. 118), meanings that "count possible worlds" might be implausible.

A second reason to reject (3) is that entailments of sentences containing *doubt* are not captured. The inferences in (5) are intuitively valid, and (5-b) shows that (5-a) can be generalized to the conjunction of an indefinite number of proposition (here via universal quantification).

(5) a. I doubt that Taro has a red hat, and I doubt that he has a green hat.
   ∴ I doubt that Taro has a red or a green hat.
   b. For every student in the class, I doubt that s/he passed the exam.
   ∴ I doubt that any student in the class passed the exam.

In standard modal logic (K), these entailments are valid for a strong, but not for a weak negative modal, as in (6). Countermodels exist for hypothetical modals of intermediate strength as well: the symbol $\Box_{/2}$ below means that the proposition is true in at least half of the accessible worlds. Countermodels are notated with sets $R[w^0]$ of the worlds $w^1, \ldots, w^n$ accessible from the world of evaluation w^0, subscripts indicating what propositions are true at each world.

(6) a. $\neg\Diamond p \wedge \neg\Diamond q \models \neg\Diamond(p \vee q)$
   b. $\neg\Box p \wedge \neg\Box q \not\models \neg\Box(p \vee q)$ *countermodel*: $R[w^0] = \{w^1_{p,\neg q}, w^2_{\neg p,q}\}$

c. $\neg\Box_{/2}p \wedge \neg\Box_{/2}q \not\models \neg\Box_{/2}(p \vee q)$ *countermodel*: $R[w^0] = \{w^1_{p,\neg q}, w^2_{\neg p,\neg q}, w^3_{\neg p,q}\}$

Given (5), *doubt* should be analyzed as a negated existential quantifier on a modal domain. In Section 4, the intuitive weakness of *doubt* and its felicity in (2-b) are argued to be a function of the domain of quantification. Before moving on to that, consider that *not doubt* is strong, too: the relevant inferences are in (7), and (8) shows that universal quantificational force is necessary.

(7) I don't doubt that Taro has a red hat, and I don't doubt that he has a green hat.
∴ I don't doubt that Taro has a red and a green hat.

(8) a. $\Box p \wedge \Box q \models \Box(p \wedge q)$
b. $\Diamond p \wedge \Diamond q \not\models \Diamond(p \wedge q)$ *countermodel*: $R[w^0] = \{w^1_{p,\neg q}, w^2_{\neg p,q}\}$

3 A negative neg-raiser

If we model natural language attitude predicates and negation as modal operators and negation in modal logic, the observations above raise a symmetry problem. If ⌜doubt that p⌝ is translated as $\neg\Diamond p$, then ⌜not doubt that p⌝ should be $\Diamond p$; but this fails to capture (7). On the other hand, if the latter is translated as $\Box p$, the former should be $\neg\Box p$; but this fails to capture (5). A standard solution to this type of puzzle is to assume that one set of inferences is due to the lexical meaning of the predicate, whereas the other is the result of a meaning strengthening operation. I propose that the lexical meaning of *doubt* has strong quantificational force, as in (9).[2] Here, $\mathcal{B}^w_{\mathsf{d}}(x)$ returns the quantification domain of *doubt* for an agent x and world w.

(9) $[\![\mathsf{doubt}]\!]^w = \lambda p_{\langle s,t\rangle}\,.\,\lambda x_e\,.\,\neg\exists w' \in \mathcal{B}^w_{\mathsf{d}}(x)\ p(w')$

Instead, when *doubt* is embedded under negation, a meaning strengthening operation applies. This amounts to saying that *doubt* is a neg-raising predicate (Horn 1971; Horn 1978). Neg-raising alone does not explain why *not doubt* expresses certainty, but its role will become clear in Section 4. However, it crucially guarantees the validity of inferences like (7). Evidence that *doubt* indeed behaves like a neg-raiser comes from the disappearance of the strong reading when negation is introduced by a higher clause, as shown by the contrast in (10). Similarly, negating an utterance containing *doubt*, with an appropriate prosodic contour, can lead to a non-strengthened reading, as in (11). This resembles what has been labeled local accommodation under negation (Heim 1983), metalinguistic negation (Horn 1985), or contrastive negation (Bruno 2017), and the intonation profile is that of a contrastive topic (Büring 2003; Meyer 2016).

(10) a. #Abigail doesn't doubt that Taro eats meat; she thinks he might eat it.
b. It's not true that Abigail doubts that Taro eats meat; she thinks he might eat it.

(11) – I've heard that you doubt that Taro will pass the exam.
– I don't DOUBT that Taro will pass the exam. I'm merely agnostic.

Clause-mate negation anti-licenses an NPI in the complement of *doubt* (Homer 2021), as shown in (12). The effect seems to persist if negation comes from a superordinate *not believe*, but not *not want*, as shown in (13). If *doubt* is a neg-raiser, this contrast might be explainable because only *believe*, but not *want*, allows for cyclic neg-raising (Horn 1971; Gajewski 2007; Homer 2015).

(12) ??Abigail doesn't doubt that Taro will ever pass the exam.

(13) a. ??I don't believe Abigail doubts that Taro will ever pass the exam.
b. I don't want Abigail to doubt that Taro will ever pass the exam.

[2] This lexicalization, which gives *doubt* the meaning of a negated existential modal, is consistent with the generalization in Horn (1990) that natural language tends not to lexicalize negated universal quantifiers.

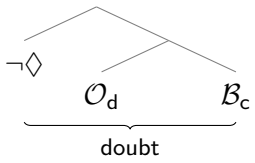


Figure 1: Decompositional analysis of the verbal head for *doubt*

4 A possibility implicature

Doubt being a neg-raiser is not enough to explain the difference in strength that arises under negation and the contrast in (2-b). My proposal builds on the following observation: the intuitive uncertainty that *doubt* carries can be reinforced or canceled like an implicature, as in (14).

(14) a. Abigail doubts that Taro will pass the exam, but she isn't certain that he won't.
b. Abigail doubts that Taro will pass the exam; in fact, she's certain that he won't.

I propose a way of deriving the possibility inference as an implicature. I assume that the modal domain of *doubt* competes and forms a scale with the modal domain of *be certain*; this scale is given by a subset relation between the former and the latter, and it is revealed by the asymmetric entailment between ⌜be certain that $\neg p$⌝ and ⌜doubt that p⌝ illustrated in (15).

(15) Abigail is certain that Taro doesn't eat meat.
∴ Abigail doubts that Taro eats meat.

I adopt a decompositional analysis of verbal heads for attitude predicates along the lines of Theiler, Roelofsen, and Aloni (2019), such that the quantificational component is detached from the domain of quantification in syntax. This reproduces the intuition that Kratzer (1977) had for modals. Unlike modals, attitude predicates in English are rigid in flavor, which might suggest that such a decomposition is unmotivated. However, recent research on other languages (see for example Bogal-Allbritten 2016; Močnik and Abramovitz 2019) showed that this difference is not cross-linguistically stable, and that attitude predicates can vary in their flavors just like modals.

The logical component of *doubt*, which I notate as ¬◊, provides negated existential quantificational force. The domain of quantification is derived by applying an ordering operator $\mathcal{O}_{\mathsf{d}}$ to the modal domain of quantification for the predicate *be certain*, notated as $\mathcal{B}_{\mathsf{c}}$. Denotations of terminals are given in (16), which compose as in Figure 1. For every world w and agent x, let $<_x^w$ be a strict partial ordering relation such that for all w', w'', it holds that $w' <_x^w w''$ if and only if w' is more plausible than w'' according to x in w. This is evocative of plausibility orderings from the literature on belief revision (Blumberg and Lederman 2021; Hansson 2022).

(16) a. $[\![\mathcal{B}_{\mathsf{c}}]\!]^w = \lambda x_e \,.\, \lambda w' \,.\, w'$ is compatible with what x is certain of at w
b. $[\![\mathcal{O}_{\mathsf{d}}]\!]^w = \lambda B_{\langle e,\langle s,t\rangle\rangle} \,.\, \lambda x_e \,.\, \lambda w' \,.\, w' \in B(x) \wedge \neg\exists w'' \in B(x)\ w'' <_x^w w'$
c. $[\![\neg\Diamond]\!]^w = \lambda B_{\langle e,\langle s,t\rangle\rangle} \,.\, \lambda p_{\langle s,t\rangle} \,.\, \lambda x_e \,.\, \neg\exists w' \in B(x)\ p(w')$

Given (16) and Figure 1, it is possible to check that the domain of quantification for *doubt* is a subset of that for *be certain*: for all x, w, it holds that $[\![\mathcal{O}_{\mathsf{d}}]\!]^w([\![\mathcal{B}_{\mathsf{c}}]\!]^w)(x) \subseteq [\![\mathcal{B}_{\mathsf{c}}]\!]^w(x)$. By adopting deletion alternatives from Fox and Katzir 2011, an alternative for a sentence containing *doubt* can be formally derived by deleting the ordering operator $\mathcal{O}_{\mathsf{d}}$ from the structure. If this strictly stronger alternative is negated by exhaustification, the result is that there is a world in the domain of quantification for *be certain* where the embedded proposition is true (see (20-b)).

The domain of *doubt* competes with the domain of *be certain*, but the reverse does not happen. Given that the latter corresponds to a syntactically simpler structure, it does not have an alternative where $\mathcal{O}_{\mathsf{d}}$ is inserted (Fox and Katzir 2011). This asymmetry in structural

complexity guarantees that the implicature in (17) is not generated.

(17) Abigail is not certain that Taro eats meat.
↛ Abigail doubts that Taro doesn't eat meat.

There is additional evidence that this inference has grammatical reality. The denotation proposed in (9) alone makes *doubt* an anti-additive function. Although theories of NPI licensing are very diverse, anti-additivity would generally lead to the expectation that not only weak, but also strong NPIs are licensed in the scope of *doubt* (Zwarts 1998; Gajewski 2005; Gajewski 2011; Chierchia 2013; Homer 2021). However, this is not the case for NPIs like *either*, puctual *until*, and the perfective modifier *in weeks*, or for the Italian *neppure* '(not) even', *fino a* 'until', and *affatto* 'at all'.[3] An example is given in (18). If *doubt* comes with a possibility implicature, the overall meaning is non-monotonic, and the degradedness of certain NPIs can be explained.[4]

(18) Will Taro call Abigail before Sunday?
a. I don't think that he will call her until Monday.
b. *I doubt that he will call her until Monday.

Going back to the original explanandum, it is possible to derive the strong meaning of *not doubt* as the interaction between the possibility implicature presented above and neg-raising. For concreteness, I will follow a proposal by Romoli (2013), who analyzes neg-raising as an implicature that arises from the negation of an excluded middle alternative. To make it compliant with the architecture in Fox and Katzir 2011, I introduce a modal quantifier, defined in (19), that competes as a scalemate item with the logical component of *doubt*, in (16-c).

(19) $[\![\neg\Diamond \vee \Box]\!]^w = \lambda B_{\langle e,\langle s,t\rangle\rangle} \,.\, \lambda p_{\langle s,t\rangle} \,.\, \lambda x_e \,.\, [\neg\exists w' \in B(x)\ p(w')] \vee [\forall w' \in B(x)\ p(w')]$

Alternative computation can replace $\neg\Diamond$ with $\neg\Diamond \vee \Box$. If multiple replacements are possible, as argued by Bar-Lev and Fox (2020), an alternative is available where both substitution of the logical component and deletion of $\mathcal{O}_{\mathsf{d}}$ apply. Implicature computation is informally summarized in (20), where the subscripts d and c specify whether the domain of quantification is that of *doubt* or *be certain*. I use the EXH operator defined in Fox 2007 and adopted in much subsequent literature; the set of alternatives is in (20-a). It is shown in (20-b) that in positive environments, alternatives with the excluded middle operator do not add to the meaning, and the implicature is as described above. Instead, when *doubt* is embedded under negation, negating the excluded middle alternatives strengthens the meaning as in (20-c). The result asserts that the embedded proposition is true in all worlds compatible with what the attitude holder is certain of.

(20) a. $Alt(\neg\Diamond_{\mathsf{d}}p) = \{\neg\Diamond_{\mathsf{d}}p, \neg\Diamond_{\mathsf{c}}p, (\neg\Diamond_{\mathsf{d}}p \vee \Box_{\mathsf{d}}p), (\neg\Diamond_{\mathsf{c}}p \vee \Box_{\mathsf{c}}p)\}$
b. $\text{EXH}(\neg\Diamond_{\mathsf{d}}p) = \neg\Diamond_{\mathsf{d}}p \wedge \Diamond_{\mathsf{c}}p \wedge \neg(\neg\Diamond_{\mathsf{c}}p \vee \Box_{\mathsf{c}}p) = \neg\Diamond_{\mathsf{d}}p \wedge \Diamond_{\mathsf{c}}p$
c. $\text{EXH}(\neg\neg\Diamond_{\mathsf{d}}p) = \neg\neg\Diamond_{\mathsf{d}}p \wedge \neg\neg(\neg\Diamond_{\mathsf{d}}p \vee \Box_{\mathsf{d}}p) \wedge \neg\neg(\neg\Diamond_{\mathsf{c}}p \vee \Box_{\mathsf{c}}p) = \Box_{\mathsf{c}}p$

5 Conclusions

The analysis provided above accounts for all the inferences presented in Sections 1 and 2: *doubt* validates inferences as a strong negative modal, but it conveys uncertainty; *not doubt*, which

[3] Certain strong NPIs that are licensed in questions, like *a single thing* (Homer 2021, p. 15) or Italian *nessuno* 'nobody', instead seem to be possible in the scope of *doubt*.

[4] A reviewer points out that as noted in Chierchia 2013, p. 201 (p.c. with Kai von Fintel and Sabine Iatridou), universal modals generally do not give rise to indirect scalar implicatures that disrupt (even strong) NPI licensing. Although it is not clear what expectations about *doubt* this observation should lead to, it is worth noticing that the same strong NPIs as above seem to be degraded in the scope of *not certain*.

behaves like a strong positive modal, conveys certainty and can lead to Moorean contradictions. This was achieved by making neg-raising and a possibility implicature interact, given the possibility of multiple replacements in the computation of alternatives (Bar-Lev and Fox 2020).

One crucial ingredient of the analysis is the decomposition of the attitude predicate in the syntax and the consequent possibility to generate scalar alternatives by either manipulating the domain of quantification (possibility implicature) or the logical component (neg-raising). This requires alternatives that are well formed in meaning and active in the computation of implicatures even if they do not have pronounceable trees: they are *conceptual alternatives* in the sense of Buccola, Križ, and Chemla (2021). A very similar analysis can be given of *be certain* under negation, which seems to give rise to an indirect scalar implicature (called a presupposition in Uegaki 2020) despite there not being any existential attitude predicate in the English lexicon.

(21) Abigail is not certain that Taro eats meat.
$\rightsquigarrow$ Abigail thinks it's possible that Taro eats meat.

This paper uses an implementation of Romoli's (2013) analysis of neg-raising as a scalar implicature. It is conceivable that another analysis of neg-raising can acheive the same results. The present account relies on the assumption that the property of being a neg-raiser is inert unless the predicate is in a negative environment (see Alxatib and Đurovic 2019 for possibly another case for the same intuition). Instead, analyzing neg-raising as a presupposition or as a scaleless implicature fails to achieve the same. If neg-raising comes from an excluded middle presupposition, and if alternatives are presuppositional, the alternative that manipulates the modal domain is never innocently excludable, regardless of polarity. A strong reading for *not doubt* can be achieved by innocently including the alternative, but the same move for $\ulcorner$doubt that $p\urcorner$ returns a meaning equivalent to $\ulcorner$be certain that not $p\urcorner$. The proofs are left to the reader.

The approach to neg-raising by Jeretic (2022) would not be suitable, either. That account, which derives neg-raising just like homogeneity in Bar-Lev 2021, uses exhaustification to assert subdomain alternatives. Given an attitude holder x and world w, if $\ulcorner x$ doesn't doubt that $p\urcorner$ is strengthened to an existential statement on each subset of $[\![\mathcal{B}_{\mathsf{c}}]\!]^w(x)$, a strong meaning for *not doubt* can be derived. However, the same mechanism also generates innocently excludable subdomain alternatives with non-negated *doubt*, namely those for $D \subseteq [\![\mathcal{B}_{\mathsf{c}}]\!]^w(x) \setminus [\![\mathcal{O}_{\mathsf{d}}]\!]^w([\![\mathcal{B}_{\mathsf{c}}]\!]^w)(x)$. The returned meaning asserts that the embedded proposition is true in all worlds in the domain of *be certain* that are not in the domain of *doubt*. Such a meaning is too strong and validates unacceptable inferences like (22). For reasons of space, the proofs are again left to the reader.

(22) I doubt that Taro bought the red hat, and I doubt that he bought the green hat.
$\not\rightsquigarrow$ I'm certain that he either bought both hats or neither one.

Finally, this paper adopts a method to distinguish modals of intuitively different strengths by operating on their domain of quantification instead of their quantificational force, as has been previously done by von Fintel and Iatridou (2008). In this way, it is possible to simultaneously capture inferences that are valid because of the quantificational strength of the modal, as (5) and (7), and contrasts that arise because of the domain of quantification, as (2). Moreover, different domains of quantification that are derived from one another can compete and form a scale, thus generating implicatures about the strength of the modal. This could be explored outside of the case of *doubt*, for example with weak necessity modals. If they quantify universally on a restricted domain, compared to strong necessity modals, and if this difference is represented structurally, strength implicatures can arise, and indeed, continuations like in (23) are felicitous.

(23) a. Taro should bring a drink, but he doesn't have to.
b. Taro should bring a drink; in fact, he has to.

References

Abrusán, Márta (2016). "Presupposition cancellation: Explaining the 'soft-hard' trigger distinction". In: *Natural Language Semantics* 24.2, pp. 165–202. DOI: 10.1007/s11050-016-9122-7.

Alxatib, Sam and Ivana Đurovic (2019). "Neg-raising *believe* and *Maximize Presupposition*". In: *Proceedings of the 22nd Amsterdam Colloquium*. Ed. by Julian J. Schlöder, Dean McHugh, and Floris Roelofsen, pp. 21–29.

Anand, Pranav and Valentine Hacquard (2013). "Epistemics and attitudes". In: *Semantics and Pragmatics* 6.8, pp. 1–59. DOI: 10.3765/sp.6.8.

Bar-Lev, Moshe E. (2021). "An implicature account of homogeneity and non-maximality". In: *Linguistics and Philosophy* 44.5, pp. 1045–1097. DOI: 10.1007/s10988-020-09308-5.

Bar-Lev, Moshe E. and Danny Fox (2020). "Free choice, simplification, and Innocent Inclusion". In: *Natural Language Semantics* 28.3, pp. 175–223. DOI: 10.1007/s11050-020-09162-y.

Biezma, María and Kyle Rawlins (2012). "Responding to alternative and polar questions". In: *Linguistics and Philosophy* 35.5, pp. 361–406. DOI: 10.1007/s10988-012-9123-z.

Blumberg, Kyle and Harvey Lederman (2021). "Revisionist reporting". In: *Philosophical Studies* 178.3, pp. 755–783. DOI: 10.1007/s11098-020-01457-4.

Bogal-Allbritten, Elizabeth (2016). "Building meaning in Navajo". PhD thesis. UMass Amherst.

Bruno, Chris (2017). "Contrastive negation and the theory of alternatives". In: *Proceedings of SALT 27*, pp. 59–74.

Buccola, Brian, Manuel Križ, and Emmanuel Chemla (2021). "Conceptual alternatives: Competition in language and beyond". In: *Linguistics and Philosophy* 45.2, pp. 265–291. DOI: 10.1007/s10988-021-09327-w.

Büring, Daniel (2003). "On D-trees, beans, and B-accents". In: *Linguistics and Philosophy* 26.5, pp. 511–545. DOI: 10.1023/a:1025887707652.

Chierchia, Gennaro (2013). *Logic in gramamar: Polarity, free choice, and intervention*. Oxford studies in semantics and pragmatics. Oxford: Oxford University Press.

von Fintel, Kai and Sabine Iatridou (2008). "How to say *ought* in foreign: The composition of weak necessity modals". In: *Studies in Natural Language and Linguistic Theory*. Springer Netherlands, pp. 115–141. DOI: 10.1007/978-1-4020-8354-9_6.

Fischer, Markus (2003). "Ein Zweifelsfall: *zweifeln* im Deutschen". In: *ZAS Papers in Linguistics* 30, pp. 1–40. DOI: 10.21248/zaspil.30.2003.179.

Fox, Danny (2007). "Free choice and the theory of scalar implicatures". In: *Presupposition and implicature in compositional semantics*. Ed. by Uli Sauerland and Penka Stateva. Palgrave Studies in Pragmatics, Language and Cognition. London: Palgrave Macmillan, pp. 71–120. DOI: 10.1057/9780230210752_4.

Fox, Danny and Roni Katzir (2011). "On the characterization of alternatives". In: *Natural Language Semantics* 19.1, pp. 87–107. DOI: 10.1007/s11050-010-9065-3.

Gajewski, Jon Robert (2005). "Neg-raising: Polarity and presupposition". PhD thesis. Cambridge, MA: MIT.

— (2007). "Neg-raising and polarity". In: *Linguistics and Philosophy* 30.3, pp. 289–328. DOI: 10.1007/s10988-007-9020-z.

— (2011). "Licensing strong NPIs". In: *Natural Language Semantics* 19.2, pp. 109–148. DOI: 10.1007/s11050-010-9067-1.

Hansson, Sven Ove (2022). "Logic of belief revision". In: *The Stanford encyclopedia of philosophy*. Ed. by Edward N. Zalta. Spring 2022. Metaphysics Research Lab, Stanford University. URL: https://plato.stanford.edu/archives/spr2022/entries/logic-belief-revision/.

Heim, Irene (1983). "On the projection problem for presuppositions". In: *Proceedings of the 2nd WCCFL*. Ed. by Michael Barlow, Daniel P. Flickinger, and Nancy Wiegand, pp. 114–125.

Heim, Irene (1992). "Presupposition projection and the semantics of attitude verbs". In: *Journal of Semantics* 9.3, pp. 183–221. DOI: 10.1093/jos/9.3.183.

Homer, Vincent (2015). "Neg-raising and positive polarity: The view from modals". In: *Semantics and Pragmatics* 8, pp. 1–88. DOI: 10.3765/sp.8.4.

— (2021). "Domains of polarity items". In: *Journal of Semantics* 38.1, pp. 1–48. DOI: 10.1093/jos/ffaa006.

Horn, Laurence R. (1971). "Negative transportation: Unsafe at any speed?" In: *Proceedings of CLS7*, pp. 120–133.

— (1978). "Remarks on neg-raising". In: *Syntax and Semantics*. Leiden: Brill, pp. 129–220. DOI: 10.1163/9789004368873_007.

— (1985). "Metalinguistic negation and pragmatic ambiguity". In: *Language* 61.1, pp. 121–174. DOI: 10.2307/413423.

— (1990). "Hamburgers and truth: Why Gricean explanation is Gricean". In: *Annual Meeting of the Berkeley Linguistics Society* 16.1, pp. 453–471. DOI: 10.3765/bls.v16i0.1693.

Jeretic, Paloma (2022). "Neg-raising with belief predicates as a scaleless implicature". In: *Semantics and Linguistic Theory* 1, pp. 891–914. DOI: 10.3765/salt.v1i0.5367.

Karttunen, Lauri (1977). "To doubt whether". In: *The CLS Book of Squibs*. Chicago, IL: Chicago Linguistic Society, pp. 58–59.

Kratzer, Angelika (1977). "What *must* and *can* must and can mean". In: *Linguistics and Philosophy* 1.3, pp. 337–355. DOI: 10.1007/bf00353453.

Meyer, Marie-Christine (2016). "Redundancy and embedded exhaustification". In: *Semantics and Linguistic Theory* 25, pp. 491–511. DOI: 10.3765/salt.v25i0.3486.

Močnik, Maša and Rafael Abramovitz (2019). "A variable-force variable-flavor attitude verb in Koryak". In: *Proceedings of the 22nd Amsterdam Colloquium*. Ed. by Julian J. Schloder, Dean McHugh, and Floris Roelofsen, pp. 494–503.

Romoli, Jacopo (2013). "A scalar implicature-based approach to neg-raising". In: *Linguistics and Philosophy* 36.4, pp. 291–353. DOI: 10.1007/s10988-013-9136-2.

Theiler, Nadine, Floris Roelofsen, and Maria Aloni (2019). "Picky predicates: Why *believe* doesn't like interrogative complements, and other puzzles". In: *Natural Language Semantics* 27.2, pp. 95–134. DOI: 10.1007/s11050-019-09152-9.

Uegaki, Wataru (2020). "The existential/uniqueness presupposition of *wh*-complements projects from the answers". In: *Linguistics and Philosophy* 44.4, pp. 911–951. DOI: 10.1007/s10988-020-09309-4.

— (2023). "The *doubt-whether* puzzle". In: *Non-interrogative subordinate* wh-*clauses*. Ed. by Łukasz Jędrzejowski and Carla Umbach. Oxford: Oxford Academic. Chap. 15, pp. 461–491. DOI: 10.1093/oso/9780192844620.003.0015.

Zwarts, Frans (1998). "Three Types of Polarity". In: *Plural quantification*. Ed. by Fritz Hamm and Erhard Hinrichs. Vol. 69. Studies in Linguistics and Philosophy. Dordrecht: Kluwer, pp. 177–238. DOI: 10.1007/978-94-017-2706-8_5.

Dutch Books, Indicative Conditionals, and Rational Updating

Calum McNamara
Yale University
New Haven, CT, USA
calum.mcnamara@yale.edu

Xueyin (Snow) Zhang
University of California, Berkeley
Berkeley, CA, USA
snowzhang@berkeley.edu

Abstract

Intuitive credence judgments in cases involving indicative conditionals sometimes seem to conflict with the claim that you should update your credences by conditionalization. For this very reason, however, these credence judgments appear Dutch-bookable. After all, a classic result of Lewis (1999) seems to show that, if you update your credences in any way that conflicts with conditionalization, then a cunning bookie can always lead you into a Dutch book situtation. In this paper, we argue that—despite appearances—the intuitive credence judgments in the relevant cases may not, in fact, be Dutch-bookable. The reason is that the Lewisian Dutch book argument for conditionalization turns out to rest on a hidden assumption. We drop that assumption, and give more general Dutch book results. Our results vindicate conditionalization only in special cases.

1 Introduction

Suppose we just rolled a six-sided die.[1] The only relevant information you have is that it's fair, and that it landed on one of its sides. Consider these sentences:

(1) The die landed on an even number. (Even)

(2) If the die landed on a low number (i.e., 1, 2, or 3), then it landed on 2. ($\mathsf{Low} \to 2$)

What are your credences in (1) and (2), respectively?

Answer: 1/2 and 1/3. In the first case, you know the die is fair, and you know that half of its sides are even. In the second case, you know there are three low numbers, each of them equally likely to come up, and only one of them is 2.

Next, suppose you learn that the die did, in fact, land even. Then, having learned this (and nothing stronger), what's your credence *now* in the sentence (2)?

Answer: 1. After all, if the die landed even, then it landed on 2, 4, or 6. So, *if* it also landed on a low number, that singles out one possibility in particular: it landed on 2.

These judgments are very natural. However, there appears to be a problem with them: they seem to leave you open to a Dutch book. To see why, suppose you value dollars linearly (with the obvious choice of units), and your credences match your fair betting quotients. Then, given the credence judgments we outlined above, it seems like the following bets should be acceptable to you, at a time before learning that the die landed even:

1. A \$25 on Even, which you'll sell for a price of \$11;

2. A \$30 bet on $\mathsf{Low} \to 2$, which you'll either buy or sell for a price of \$10.

Similarly, if you become certain of (2) after learning Even, then this bet should be acceptable to you as well:

[1] The case that follows is due to Ciardelli and Ommundsen 2022. For similar cases, see, e.g., Goldstein and Santorio 2021, Santorio 2022, and Khoo 2022.

3. A \$30 bet on $\mathsf{Low} \rightarrow 2$ that you'll buy for a price of \$25.

But if this is right, then it seems like a cunning bookie can Dutch book you, using one of the strategies below.

First, suppose the bookie knows your fair betting quotients. Then, suppose she asks you for your fair price for a \$30 bet on the conjunction $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$. If this price is above \$10, then she'll be willing to sell you that bet, at your fair price, as well as buy from you a \$30 bet on $\mathsf{Low} \rightarrow 2$ for a price \$10. By assumption, you're willing to accept both. But if you do, you'll be led into a *synchronic Dutch book*:[2]

	$\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$	$\mathsf{Even} \wedge \neg(\mathsf{Low} \rightarrow 2)$	$\neg\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$	$\neg\mathsf{Even} \wedge \neg(\mathsf{Low} \rightarrow 2)$
Bet 1	$30 - x$	$-x$	$-x$	$-x$
Bet 2	-20	10	-20	10
Net	$10 - x$	$10 - x$	$-20 - x$	$10 - x$

Table 1: Your net payoff if you sell the bet on $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$ for $\$x$ with $x > 10$

On the other hand—and more important for our purposes—if your fair price for the bet on $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$ is no more than \$10, then the bookie can buy that bet from you for a price \$10, and sell you a \$25 bet on Even for \$11. Again, *ex hypothesi*, both offers are fair or favorable, by your lights. So, you should be willing to accept them both. But if you do, then the bookie will wait to find out if the die landed even. And if it does, she'll offer to sell you a \$30 bet on $\mathsf{Low} \rightarrow 2$ for \$25, which is again an acceptable bet to you, by assumption. However, these transactions lead you into a *diachronic Dutch book*:

	$\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$	$\mathsf{Even} \wedge \neg(\mathsf{Low} \rightarrow 2)$	$\neg\mathsf{Even}$
Bet 1	14	14	-11
Bet 2	-20	10	10
Bet 3	5	-25	0
Net	-1	-1	-1

Table 2: Your net payoff if you sell the bet on $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$ for \$10 (or less)

In this case, the reason seems to be that your credence judgments after learning diverge from the prescriptions of *conditionalization*.[3] And a famous result of Lewis (1999) seems to show that, if you update your credences in any way that conflicts with conditionalization, then this makes you Dutch-bookable.

Thus, we seem to have a puzzle. On the one hand, the credence judgments which our die case elicits seem intuitively rational.[4] On the other hand, they also appear to license betting dispositions that are Dutch-bookable and thereby rationally defective. So: what should we say in response? (Note also that the case discussed here is not unique. Analogous problems seem to arise in *many* cases involving indicative conditionals—i.e., sentences like our (2).)

In this paper, we're going to offer a partial diagnosis of this puzzle. In contrast to the Dutch book arguments we've just spelled out, our view is that the credence judgments elicited

[2]The reason for this, of course, is that in this case, your credences wouldn't satisfy Kolmogorov's probability axioms. In particular, having a fair price of more than \$10 for $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$ would imply that you're more confident in a conjunction than you are in either conjunct.

[3]Recall that conditionalization says the following: if c is the probability function representing your credences at the current time, and c_A is your credence function after learning A (and nothing stronger), then, for any B, $c_A(B) = c(B \mid A) := \frac{c(A \wedge B)}{c(A)}$, provided that $c(A) > 0$. We spell out exactly how these judgents conflict with conditionalization in the last section, below.

[4]Indeed, several authors have defended the claim that they're intuitively rational—see, e.g., the references given in fn. 1.

by our die case may not be Dutch-bookable after all. To establish this, we'll show that the second Dutch book given above actually rests on a hidden assumption—namely, an assumption about how rational agents will evaluate *betting plans*. However, this assumption is controversial. And dropping it turns out to yield Dutch book results for a more general rule of rational updating—one that we call *conditional conditionalization*. This rule delivers the intuitive credence judgments in cases like our die case (at least given natural background assumptions). And it agrees with conditionalization when restricted to purely "factual" sentences (i.e., sentences that don't themselves contain conditionals).

2 Motivating the Proposal

Before we get to that, however, let's first consider a bad objection to the foregoing diachronic Dutch-book argument.[5] Thus, suppose someone complains: Table 2 is a bit misleading. That table shows that Bet 3 pays \$0 in cases where the die lands odd. But Bet 3 is an unconditional bet, and so has well-defined payoffs in those cases. In particular, you might win the bet in some cases where the die lands odd, and thereby collect a net *gain* of \$4. For this reason, then, we should reject the diachronic Dutch book argument against the intuitive credence judgments.

	$\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$	$\mathsf{Even} \wedge \neg(\mathsf{Low} \rightarrow 2)$	$\neg\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$	$\neg\mathsf{Even} \wedge \neg(\mathsf{Low} \rightarrow 2)$
Bet 1	14	14	-11	-11
Bet 2	-20	10	10	10
Bet 3	5	-25	5	-25
Net	-1	-1	4	-26

Table 3: A Better Table?

This, clearly, is a bad objection, because the payoffs of Bet 3 in cases where the die lands odd are, in some sense, *irrelevant* to you—Bet 3 won't be offered to you in those cases in the first place. True, you can dream about the \$5 that you *could* win, *if* the bet *were* offered to you in those situations. But the fact is that you won't be offered that bet in those cases, and you know it. So what's the upshot? It's that you should take into account the payoffs of Bet 3 only *if* that bet is offered to you.

However, reflecting on this idea raises another concern about the foregoing Dutch book argument. To see what it is, first consider this sentence:

(3) If Bet 3 is offered, then you will win that bet.

Now, should you be certain of (3) *before* you learn whether the die landed even? To us, the answer isn't obvious. Indeed, there seem to be two plausible replies:

- *First Answer.* No, it's not rational for you to be certain of (3) before you learn whether the die landed even. To see why, notice that, given your background knowledge, Bet 3 is offered if and only if Even is true, and you win Bet 3 if and only if $\mathsf{Low} \rightarrow 2$ is true. By your own lights, it's possible for Even to be true and yet $\mathsf{Low} \rightarrow 2$ to be false, since you think the former is more probable than the latter. Thus, it's *not* rational for you to be certain of sentence (3).

[5] Due to considerations of space, we'll focus on the diachronic Dutch book argument here, i.e., the one in which your fair price for the bet on $\mathsf{Even} \wedge (\mathsf{Low} \rightarrow 2)$ is no more than \$10. Note also that, in a longer version of this paper, we discuss several other objections to this Dutch book argument. For example, we discuss an objection which says that the Dutch book is "practically defective", because Bet 2 cannot be "settled". We also discuss *contextualist* and *expressivist* responses to the puzzle. Again, see McNamara and Zhang MS for further details.

- *Second Answer.* Yes, it *is* rational for you to be certain of (3) before you learn whether the die landed even. Again, given your background knowledge, Bet 3 is offered if and only if Even is true; and you win Bet 3 if and only if Low $\rightarrow$ 2 is true. The sentence (3) is true just in case 'If the die landed even, then it landed on 2 if it landed on a low number' is true. (In symbols: Even $\rightarrow$ (Low $\rightarrow$ 2).) And intuitively, it seems rational for you to be certain of *that*.

We think both answers here are plausible. However, the Dutch book argument outlined in Table 2 *assumes* the first answer to our question. More specifically, it assumes that, before you learn whether Even is true, it's a live possibility for you that you might lose \$25 if you take Bet 3.

However, suppose you endorse the second answer to our question, i.e., you think it's impossible for you to lose Bet 3 if it's offered and you take it. Then, it appears that you *won't* lose \$1 for sure by accepting Bets 1-3. For if the die lands even, then you'll get \$14 from Bet 1 and \$5 from Bet 3, and you *might* win Bet 2 and keep a \$10 premium from the bookie. As a result, there is a possibility in which you have a net profit of \$29. Thus, taking into account the payoffs of Bet 3 in cases where the die lands odd is a bit like taking into account the payoffs we mentioned in the bad objection outlined above, if you think *Second Answer* is right:

	Even $\wedge$ (Low $\rightarrow$ 2)	Even $\wedge$ $\neg$(Low $\rightarrow$ 2)	$\neg$Even $\wedge$ (Low $\rightarrow$ 2)	$\neg$Even $\wedge$ $\neg$(Low $\rightarrow$ 2)
Bet 1	14	14	-11	-11
Bet 2	-20	10	10	10
Bet 3	5	5	5 (✗)	-25 (✗)
Net	-1	29	-1	-1

Table 4: The Best Table?

We can make this point more general. According to Dutch book results like the one given by Lewis (1999), the core of a diachronic Dutch book argument is the observation that a *plan* for how to accept or decline certain bets is payoff-equivalent to certain compounds of unconditional bets. As Lewis describes it:

> Consider a conditional bet: that is, a bet that will be null and void unless its condition is met. We note, first, that the conditional bet is equivalent in its outcome, come what may, to a certain pair of unconditional bets. We note, second, that the conditional bet is also equivalent in its outcome, come what may, to a certain contingency plan whereby one's future betting transactions are made to depend on the arrival of new evidence. The first equivalence yields a well-known synchronic argument relating the prices of conditional and unconditional bets. The second equivalence yields a diachronic argument relating the present prices of conditional bets to the future prices, after various increments of evidence, of unconditional bets. (p. 404)

Thus, in our example, the outcome of taking Bet 3 if the die landed even is equivalent to the outcome of buying a \$30 bet on Even $\wedge$ (Low $\rightarrow$ 2) for \$17.5 and selling a \$25 bet on Even for \$12.5—or at least it is, according to Lewis. However, this compound of unconditional bets is unacceptable to you since, by assumption, \$12.5 is the minimum price at which you are willing to sell a \$25 bet on Even, and \$17.5 is above the maximum price at which you are willing to buy a \$30 bet on Even $\wedge$ (Low $\rightarrow$ 2). This is the sense in which you "hold two contradictory opinions about the expected value of the very same transaction", according to Lewis: you find the contingency plan acceptable, but the relevant pair of unconditional bets unacceptable, even though, *ex hypothesi*, they are the same.

What we are suggesting, however, is that whether the two betting arrangements really *are* the same depends on what you think will happen *if* the die ultimately lands on even. In turn, this depends on your interpretation of the conditional 'If Bet 3 is offered, then you will win that bet'. As we argued above, there's a reasonable interpretation of 'if' such that it's rational for you to be certain *ex ante* that this conditional is true. And if you *are* certain of this conditional, then, by your own lights, the plan of buying Bet 3 for \$25 if the die lands even will *not* be payoff-equivalent to buying a \$30 bet on $\mathsf{Even} \wedge (\mathsf{Low} \to 2)$ for \$17.5 and selling a \$25 bet on Even for \$12.5; the former won't lose you money, but the latter might. Thus, from this perspective, the Dutch book argument above doesn't show that you hold contradictory opinions about the expected value of the very same transaction. Instead, from your perspective, the two transactions are very different.

3 Conditional Conditionalization

In this section, we're going to try making that insight more precise. To do so, let's start by considering yet another sentence:

(4) If the die landed on an even number, then it landed on 2 if it landed on a low number. $(\mathsf{Even} \to (\mathsf{Low} \to 2))$

Now, what's your credence in sentence (4)?

Intuitively, the answer here is '1', and the reasoning is similar to the reasoning we gave for certainty in (2) after learning Even: the only even number that's low is 2. So, if the die landed even; and if, in addition, it landed on a low number; then it must have landed on 2—there are no other (relevant) possibilities.

If this is right, however, then it suggests the rationality of being certain of $\mathsf{Low} \to 2$ after learning Even is closely related to the rationality of being certain of $\mathsf{Even} \to (\mathsf{Low} \to 2)$ prior to this learning event.[6] With this in mind, then, we want to propose that the connection is captured by the following rule of rational learning:

> **Conditional Conditionalization**. After learning the truth of A (and nothing stronger), your new credence in any B should be equal your old conditional credence in the indicative conditional $A \to B$, conditional on A. Formally, if c is your prior credence function, and c_A is your posterior credence function, after learning A (and nothing stronger), then, if $c(A) > 0$:
>
> $$c_A(B) = c(A \to B \mid A) := \frac{c(A \wedge (A \to B))}{c(A)}.$$

Notice that, if $c(A \wedge (A \to B)) = c(A \wedge B)$, then this rule *entails* classical conditionalization (according to which $c_A(B) = c(B \mid A)$). On most theories of conditionals, moreover, this equality—known as *probabilistic modus ponens*—holds when both A and B are "factual" (i.e., they're sentences that do not themselves contain conditionals).[7] Thus, to this extent, classical conditionalization (restricted to factual sentences) can be viewed as a special case of conditional conditionalization.

On the other hand, the two rules come apart when extended to non-factual sentences. In particular, they disagree about the correct probability judgments in our die case. To see this,

[6] In turn, the rationality of being certain of $\mathsf{Even} \to (\mathsf{Low} \to 2)$ might depend on your preferred semantics for indicative conditionals. For example, on the "path semantics" for indicative conditionals given by Goldstein and Santorio 2021 and Santorio 2022, it's natural to think that certainty in this sentence is required.

[7] Note also that it holds for any semantic theory that validates plausible "centering" principles—even in the case in which A and B are non-factual.

suppose c is a classical probability function such that $c(\mathsf{Even} \wedge (\mathsf{Low} \to 2)) \leq c(\mathsf{Low} \to 2)$. Then, according to conditionalization, if $c(\mathsf{Even}) = 1/2$ and $c(\mathsf{Low} \to 2) = 1/3$, then:

$$c_{\mathsf{Even}}(\mathsf{Low} \to 2) = \frac{c(\mathsf{Even} \wedge (\mathsf{Low} \to 2))}{c(\mathsf{Even})} \leq \frac{1/3}{1/2} < 1.$$

On the other hand, conditional conditionalization vindicates the intuitive credence judgments the die case, provided that you are *ex ante* certain of the conditional $\mathsf{Even} \to (\mathsf{Low} \to 2)$. For, if you are certain of that conditional, then according to conditional conditionalization, your posterior credence in $\mathsf{Low} \to 2$ after learning Even should be:

$$\frac{c(\mathsf{Even} \wedge (\mathsf{Even} \to (\mathsf{Low} \to 2)))}{c(\mathsf{Even})}.$$

Then, since you are certain of $\mathsf{Even} \to (\mathsf{Low} \to 2)$, the numerator equals $c(\mathsf{Even})$, which means that the ratio equals 1. Crucially, moreover, this entailment holds for *any* positive credal assignments to Even and $\mathsf{Low} \to 2$.

Thus, to show that the intuitive credence judgments are not Dutch-bookable, it suffices for us to show that conditional conditionalization is not Dutch-bookable. As we hinted, however, in order to do this, we need to reject Lewis's assumption that the plan of taking a bet on B if A is true is payoff-equivalent to certain compounds of unconditional bets on $A \wedge B$ and A, respectively. The good news, however, is that, in the present case, this assumption seems questionable. After all, if you think you *won't* lose the bet on $\mathsf{Low} \to 2$ if you take it when Even is true—since you're *ex ante* certain of $\mathsf{Even} \to (\mathsf{Low} \to 2)$—then, by your lights, this plan is not payoff-equivalent to any compounds of unconditional bets on $\mathsf{Even} \wedge (\mathsf{Low} \to 2)$ and Even.

How, then, should you evaluate the payoffs of betting on $\mathsf{Low} \to 2$ if Even is true? Here is a natural thought: you should evaluate the payoffs of this plan in the same way you evaluate a conditional bet on $\mathsf{Even} \to (\mathsf{Low} \to 2)$, *conditional on the truth of Even*. More generally:

> **Plans-as-Conditional-Bets-on-Conditionals (PCBC)**. A plan of *taking a bet on B if A is true* is payoff-equivalent to a conditional bet on the conditional $A \to B$, conditional on A.

PCBC entails Lewis's assumption that betting plans are payoff-equivalent to conditional bets, under the assumption that a conditional bet on B given A is payoff-equivalent to a conditional bet on $A \to B$ given A. Again, this assumption is plausible if both A and B are factual, for then both bets are payoff-equivalent to a conditional bet on $A \wedge B$, given A. On the other hand, this assumption plausibly fails if B contains conditionals. And as we saw, these are also precisely cases where Lewis's assumption seems problematic. Thus, we have the following result:[8]

Theorem 3.1. *Suppose an agent is not synchronically Dutch-bookable and satisfies PCBC. Then she is not diachronically Dutch-bookable if and only if she updates by conditional conditionalization.*

Once again, this result entails classical conditionalization, given probabilistic *modus ponens*. Relatedly, our Dutch book theorem entails the Dutch book theorem of Lewis (1999) for classical conditionalization, if we assume a conditional bet on $A \to B$ given A is payoff-equialent to a conditional bet on $A \wedge B$ given A.[9] As we heard, however, it's not uncontroversial that this should be the case. And if we reject Lewis's assumption, then the intuitive credence judgments with which we started may not be Dutch-bookable after all.

[8] A proof of this result can be found in McNamara and Zhang MS.

[9] Incidentally, our main result also entails the Dutch book result of McGee (1989), given a different, and competing, assumption about the probabilities of conditionals—namely, that A and $A \to B$ are probabilistically independent. See McGee's paper for further discussion.

Acknowledgements. For helpful discussion and feedback, we are grateful to Fabrizio Cariani, Melissa Fusco, Wesley Holliday, Jim Joyce, Mikayla Kelley, Harvey Lederman, Matt Mandelkern, Sarah Moss, Eric Pacuit, Paolo Santorio, and audiences at Maryland, Michigan, and UT Austin.

References

Ciardelli, Ivano and Adrian Ommundsen (Nov. 2022). "Probabilities of conditionals: updating Adams". In: *Noûs*. DOI: 10.1111/nous.12437. URL: https://doi.org/10.1111%2Fnous.12437.

Goldstein, Simon and Paolo Santorio (2021). "Probability for Epistemic Modalities". In: *Philosophers' Imprint* 21.33.

Khoo, Justin (2022). *The Meaning of 'If'*. New York, USA: Oxford University Press.

Lewis, David (1999). "Why Conditionalize". In: *Philosophy of Probability: Contemporary Readings*. Ed. by Antony Eagle. Routledge, pp. 403–407.

McGee, Vann (Oct. 1989). "Conditional Probabilities and Compounds of Conditionals". In: *The Philosophical Review* 98.4, p. 485. DOI: 10.2307/2185116. URL: https://doi.org/10.2307%2F2185116.

McNamara, Calum and Snow Zhang (MS). "Why (Not) Conditionalize?" Unpublished Manuscript.

Santorio, Paolo (Jan. 2022). "Path Semantics for Indicative Conditionals". In: *Mind* 131.521, pp. 59–98. DOI: 10.1093/mind/fzaa101. URL: https://doi.org/10.1093%2Fmind%2Ffzaa101.

A localist approach to the semantics and distribution of X-marking*

Teruyuki Mizuno
Ochanomizu University
Tokyo, Japan
mizuno.teruyuki@ocha.ac.jp

Abstract

I propose an account that assigns a uniform semantics to each local occurrence of X-marking. The proposed account correctly captures the distribution of X-marking in conditionals and desire reports in English and Greek (and possibly many other languages).

1 Introduction

X-marking (von Fintel and Iatridou 2023; vFI) refers to grammatical items that imply suspension of some established facts, which is necessary, for instance, to discuss counterfactual situations.[1] Such items include Past (Imperfective) that occurs in conditionals (1) and desire reports (2).[2]

(1) [*An o archigos* ***pethene***$_{\text{X-MARKED}}$ *avrio*], [*tha ton* ***thavame***$_{\text{X-MARKED}}$ *eki*]. (Greek)
if the chief die.PAST.IMPF tomorrow FUT him buy.PAST.IMPF there
'If the chief **died**$_{\text{X-MARKED}}$ tomorrow, we **would**$_{\text{X-MARKED}}$ bury him there.' (vFI: 12)

(2) *Tha* ***ithela***$_{\text{X-MARKED}}$ [*na* ***erxotan***$_{\text{X-MARKED}}$ *avrio*]. (Greek)
FUT want.PAST.IMPF NA come.PAST.IMPF tomorrow
'I wish he **would**$_{\text{X-MARKED}}$ come tomorrow.'

There are multiple occurrences of X-marking in (1) and (2): in (1), X-marking appears in both the antecedent and the consequent; in (2), it appears both on the desire predicate and in the complement (at least in Greek). A challenge for any account of X-marking is to correctly figure out the contribution of *each local occurrence* of X-marking. This is left open in vFI. Many authors have similarly sidestepped this issue (e.g., Iatridou 2000; Mackay 2019), or have suggested that either one of them is the result of agreement with another (e.g., Schulz 2014; Grønn 2021).

I propose an account that assigns a uniform semantics to each local occurrence of X-marking in (1)-(2), thus dispensing with the assumption of agreement to explain the same facts.

2 Proposal

The idea is that each occurrence of X-marking presupposes that the local domain with which it is associated has undergone a *domain expansion* (i.e., an operation that suspends some facts/beliefs; see vFI). (3) is the proposed entry. $\mathcal{D}$ takes a world w and returns a set of worlds (a domain) for evaluating modal expressions with respect to w. Which world and domain function are default may be affected by contextual or grammatical factors; this will be clarified along the way.

(3) Let w^* be a default world and $\mathcal{D}^*$ be a default domain function. For any world w and domain function $\mathcal{D}$, $[\![\text{X } \phi]\!]^{w,\mathcal{D}}$ is defined only if $\mathcal{D}^*(w^*) \subset \mathcal{D}(w)$; if defined, $[\![\text{X } \phi]\!]^{w,\mathcal{D}} = [\![\phi]\!]^{w,\mathcal{D}}$.

*This work was supported by the NSF grant (#2116972 "Research on conditional and modal language"; PI: Magdalena Kaufmann; Co-PI: Stefan Kaufmann). I thank Thanos Iliadis for judgments on Greek data, Aarón Sánchez for judgments on Mexican Spanish data, and Si Kai Lee and Robert Lowe for judgments on English data.

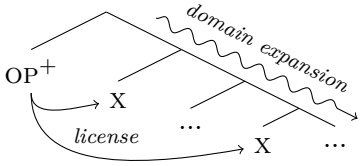
[1] This way of characterizing the purported contribution of X-marking, vFI note, is traced to Stalnaker 1975.

[2] Also known as "Fake Past" (see Iatridou 2000).

I assume that domain expansion is enforced by an operator OP^+, as defined in (4). Crucial about the proposed account is the system where the effect of domain expansion "trickles down" and licenses each local occurrence of X-marking one by one, as illustrated in (5).[3]

(4) $[\![\text{OP}^+\phi]\!]^{w,\mathcal{D}} = [\![\phi]\!]^{w,\mathcal{D}_+}$ (where $\mathcal{D}(w) \subset \mathcal{D}_+(w)$)

(5)



I assume that domains are determined by accessibility relations: for any $w, \mathcal{D}$, $\mathcal{D}(w)$ consists of worlds that are accessible from w through an accessibility relation R (i.e., $\mathcal{D}(w) = \{w' \mid wRw'\}$). I focus on those that are serial, transitive and euclidean, using "$\sim$" to refer to any such relations.[4] Domains created by $\sim$ may represent one's doxastic state (Stalnaker 2002; S. Kaufmann 2005) or context sets in non-defective cases (Stalnaker 2014; see also Yalcin 2024); my account generalizes to any of these.[5] An expanded domain $\mathcal{D}_+(w)$ is created from $\sim^+$, which is defined on $\sim$: $\sim^+$ is serial, transitive, euclidean, and such that $\{w' \mid w \sim w'\} \subset \{w' \mid w \sim^+ w'\}$.[6]

3 Conditionals

I assume that conditionals universally quantify over a domain that is restricted by the antecedent (Kratzer 1986, 1991), with the presupposition that the antecedent is compatible with the domain.

(6) a. $[\![\text{IF } \phi, \psi]\!]^{w,\mathcal{D}}$ is defined only if $[\![\phi]\!]^{w',\mathcal{D}} = 1$ for some $w' \in \mathcal{D}(w)$;
b. if defined, $[\![\text{IF } \phi, \psi]\!]^{w,\mathcal{D}} = 1$ iff $\forall w'' \in \{w' \in \mathcal{D}(w) \mid [\![\phi]\!]^{w',\mathcal{D}}\}$. $[\![\psi]\!]^{w'',\mathcal{D}}$

The conditionals in (1) are analyzed as involving the structure "$\text{OP}^+[\ \text{IF}\ [\text{X}\ \phi], [\text{X}\ \psi]]$". Let w and $\mathcal{D}$ be the default world and the default domain function: w is, say, the actual world, and $\mathcal{D}(w)$ the context set or the speaker's doxastic state at w, depending on utterance context.

The derivation goes as in (7). At (a), the compatibility presupposition is triggered for ϕ and $\mathcal{D}_+(w)$. At (b), in the scope of a universal quantifier, the two instances of X-marking are assumed to trigger a *universal presupposition* at their respective loci (i.e., for all $w' \in \mathcal{D}_+(w)$, $\mathcal{D}(w) \subset \mathcal{D}_+(w')$). This universal presupposition is satisfied trivially as $\mathcal{D}_+(w)$ forms an *equivalence class* of worlds (see Appendix A.1 for proofs). Since $\mathcal{D}_+(w)$ forms an equivalence class, it holds that for all $w' \in \mathcal{D}_+(w)$, $\mathcal{D}_+(w) = \mathcal{D}_+(w')$. This means that the universal presupposition is satisfied as long as $\mathcal{D}(w) \subset \mathcal{D}_+(w)$ holds, and it is indeed satisfied since $\mathcal{D}(w) \subset \mathcal{D}_+(w)$ holds trivially.

(7)
$[\![\ \text{OP}^+\ [\ \text{IF}\ [\text{X}\ \phi], [\text{X}\ \psi]]]\!]^{w,\mathcal{D}}$
$=$ $[\![\ \text{IF}\ [\text{X}\ \phi], [\text{X}\ \psi]]\!]^{w,\mathcal{D}_+}$ $\cdots(a)$ $[\![\phi]\!]^{w',\mathcal{D}_+} = 1$ for some $w' \in \mathcal{D}_+(w)$
$=$ 1 iff $\forall w'' \in \{w' \in \mathcal{D}_+(w) \mid [\![\text{X}\ \phi]\!]^{w',\mathcal{D}_+}\}$. $[\![\text{X}\ \psi]\!]^{w'',\mathcal{D}_+}$ $\cdots(b)$ $\forall w' \in \mathcal{D}_+(w)$. $\mathcal{D}(w) \subset \mathcal{D}_+(w')$ (Local)
$=$ 1 iff $\forall w'' \in \{w' \in \mathcal{D}_+(w) \mid [\![\phi]\!]^{w',\mathcal{D}_+}\}$. $[\![\psi]\!]^{w'',\mathcal{D}_+}$

The presupposition of each occurrence of X-marking is locally satisfied and filtered, and thus does not globally project; what globally projects in (7) is only the compatibility presupposition.

[3] My account thus takes a "reflective" view in the sense introduced by vFI; each occurrence of X-marking does not enforce domain expansion by itself but *reflects* the existence of an expansion operator higher in the structure.

[4] I adopt the symbol "$\sim$" from S. Kaufmann 2005. A relation R is *serial* iff for any x, there is y such that xRy; *transitive* iff for any x, y, z, if xRy and yRz, then xRz; *euclidean* iff for any x, y, z, if xRy and xRz, then yRz. Seriality implies consistency; transitivity and euclideanness imply positive and negative introspection respectively.

[5] Adding reflexivity results in "factive" domains. My account applies to these domains too.

[6] Expansion will be sensitive to, e.g., similarity ordering (Stalnaker 1968; Lewis 1973), which I gloss over here.

4 Desire reports

I assume that WANT universally quantifies over domains of worlds that are optimal with respect to the desirability ranking of an agent **a** (i.e., $\mathcal{D}^{\preccurlyeq_{\mathbf{a}}}(w)$; e.g., von Fintel 1999).[7] I also assume that WANT triggers unsettledness presupposition for its complement (Heim 1992; von Fintel 1999).

(8) a. $[\![\text{WANT}_a\ \phi]\!]^{w,\mathcal{D}}$ is defined only if the truth or falsity of ϕ is not settled in $\mathcal{D}(w)$;
b. if defined, $[\![\text{WANT}_a\ \phi]\!]^{w,\mathcal{D}} = 1$ iff $\forall w' \in \mathcal{D}^{\preccurlyeq_{\mathbf{a}}}(w).\ [\![\phi]\!]^{w',\mathcal{D}}$

The domain that WANT quantifies over is usually identified with the attitude holder's doxastic state (von Fintel 1999; see also Heim 1992).[8] I implement this idea by positing the covert belief operator as defined in (9) above the occurrence of WANT. (10) shows a sample derivation. $\Box_a^{\text{BEL}}$ introduces as a new domain the attitude holder's doxastic state. The outer layer of universal quantification at (*b*) is reduced, as the two quantifiers quantify over the same doxastic state.

(9) $[\![\Box_a^{\text{BEL}}\ \phi]\!]^{w,\mathcal{D}} = 1$ iff $\forall w' \in \mathcal{D}'(w).\ [\![\phi]\!]^{w',\mathcal{D}'}$ (where $\mathcal{D}' = \lambda w.\ \{w' \mid w \sim_{\mathbf{a}} w'\}$)

(10) $[\![\ \Box_a^{\text{BEL}}\ [\ \text{WANT}_a\ \phi\ ]\!]^{w,\mathcal{D}} = 1$
iff $\forall w' \in \mathcal{D}'(w).\ [\![\text{WANT}_a\ \phi]\!]^{w',\mathcal{D}'}$ $\cdots (a)\ \mathcal{D}' = \lambda w.\ \{w' \mid w \sim_{\mathbf{a}} w'\}$
iff $\forall w' \in \mathcal{D}'(w).\ \forall w'' \in \mathcal{D}'^{\preccurlyeq_{\mathbf{a}}}(w').\ [\![\phi]\!]^{w'',\mathcal{D}'}$ $\cdots (b)$ ϕ or $\neg\phi$ is unsettled in $\mathcal{D}'(w)$
iff $\forall w' \in \mathcal{D}'^{\preccurlyeq_{\mathbf{a}}}(w).\ [\![\phi]\!]^{w',\mathcal{D}'}$

With these in place, the Greek sentence in (2) above is analyzed as involving the structure "$\Box_a^{\text{BEL}}$ [OP$^+$ [X [WANT$_a$ [X ϕ]]]]" (I assume for simplicity that English *wish* is semantically the composite of X and WANT; it is undefined unless its domain is expanded). In the scope of $\Box_a^{\text{BEL}}$, the default domain function for X-marking is assumed to be identified with the attitude holder's doxastic state. The default world is assumed again to be the actual world, for which I use w.

The derivation goes as in (11). OP$^+$ applies below $\Box_a^{\text{BEL}}$, thus expanding the attitude holder's doxastic state. For proofs for the reduction from $(*)$ to (d), see Appendix A.2. As above, in the scope of universal quantifiers, the two occurrences of X-marking are assumed to trigger universal presuppositions at their respective positions (see (b) and (d)). The one triggered at (d) is trivially satisfied since $\mathcal{D}'(w) \subset \mathcal{D}'_+(w)$ and for all $w' \in \mathcal{D}'_+(w)$, $\mathcal{D}'_+(w) = \mathcal{D}'_+(w')$. Since $\mathcal{D}'(w) \subset \mathcal{D}'_+(w)$ and the restrictor of a universal quantifier is a downward-monotonic environment, the universal presupposition triggered at (b) is also satisfied. The two presuppositions are filtered locally.

(11) $[\![\ \Box_a^{\text{BEL}}\ [\ \text{OP}^+\ [\ \text{X}\ [\ \text{WANT}_a\ [\ \text{X}\ \phi\ ]]]]]\!]^{w,\mathcal{D}} = 1$
iff $\forall w' \in \mathcal{D}'(w).\ [\![\ \text{OP}^+\ [\ \text{X}\ [\ \text{WANT}_a\ [\ \text{X}\ \phi\ ]]]]\!]^{w',\mathcal{D}'}$ $\cdots (a)\ \mathcal{D}' = \lambda w.\ \{w' \mid w \sim_{\mathbf{a}} w'\}$
iff $\forall w' \in \mathcal{D}'(w).\ [\![\ \text{X}\ [\ \text{WANT}_a\ [\ \text{X}\ \phi\ ]]]\!]^{w',\mathcal{D}'_+}$ $\cdots (b)\ \forall w' \in \mathcal{D}'(w).\ \mathcal{D}'(w) \subset \mathcal{D}'_+(w')$ (Local)
iff $\forall w' \in \mathcal{D}'(w).\ [\![\ \text{WANT}_a\ [\ \text{X}\ \phi\ ]]\!]^{w',\mathcal{D}'_+}$ $\cdots (c)$ ϕ or $\neg\phi$ is unsettled in $\mathcal{D}'_+(w)$
iff $\forall w' \in \mathcal{D}'(w).\ \forall w'' \in \mathcal{D}'^{\preccurlyeq_{\mathbf{a}}}_+(w').\ [\![\text{X}\ \phi]\!]^{w'',\mathcal{D}'_+}$ $\cdots (*)$
iff $\forall w' \in \mathcal{D}'^{\preccurlyeq_{\mathbf{a}}}_+(w).\ [\![\text{X}\ \phi]\!]^{w',\mathcal{D}'_+}$ $\cdots (d)\ \forall w' \in \mathcal{D}'_+(w).\ \mathcal{D}'(w) \subset \mathcal{D}'_+(w')$ (Local)
iff $\forall w' \in \mathcal{D}'^{\preccurlyeq_{\mathbf{a}}}_+(w).\ [\![\phi]\!]^{w',\mathcal{D}'_+}$

5 Interaction with O-marking

My account can explain illicit combinations of X-marking and *O-marking*, such as Greek (12). The semantics of O-marking is defined as in (13): it presupposes that the local domain is identical

[7] Let $\preccurlyeq_{\mathbf{a}}$ be a (world-relative) desirability ranking attributed to an agent **a**. For any w and $\mathcal{D}$, $\mathcal{D}^{\preccurlyeq_{\mathbf{a}}}(w)$ is defined as follows: for all $w' \in \mathcal{D}(w)$, $w' \in \mathcal{D}^{\preccurlyeq_{\mathbf{a}}}(w)$ iff for all $w'' \in \mathcal{D}(w)$, if $w'' \preccurlyeq_{\mathbf{a},w} w'$, then $w' \preccurlyeq_{\mathbf{a},w} w''$.

[8] See Rubinstein 2012, 2017 for an alternative possibility. See also Villalta 2008; Anand and Hacquard 2013.

to the default domain (i.e., that no expansion has taken place).

(12) * ***Thelo***$_{\text{O-MARKED}}$ [*na* ***erxotan***$_{\text{X-MARKED}}$ *avrio*] (Greek)
want.NONPAST NA come.PAST.IMPF tomorrow
Intended: 'I wish he would come tomorrow.'

(13) Let w^* be a default world and $\mathcal{D}^*$ be a default domain function. For any $w, \mathcal{D}$, $[\![\text{O } \phi]\!]^{w,\mathcal{D}}$ is defined only if $\mathcal{D}^*(w^*) = \mathcal{D}(w)$; if defined, $[\![\text{O } \phi]\!]^{w,\mathcal{D}} = [\![\phi]\!]^{w,\mathcal{D}}$.

If (12) involves OP$^+$, it clashes with the presupposition of O on WANT (14a); if (12) does not involve OP$^+$, the presupposition of X in the complement cannot be satisfied (14b). Thus, the presupposition of either one *always* fails to be satisfied, hence the inescapable badness of (12).[9]

(14) a. OP$^+$ [O [WANT$_a$ [X ϕ]]] ⇝ fails to satisfy O's presupposition
b. O [WANT$_a$ [X ϕ]] ⇝ fails to satisfy X's presupposition

Assuming that *wish* lexically triggers the same presupposition as X-marking does, my account also predicts the contrast in (15): if the sentence involves OP$^+$, the presupposition of O is not satisfied (16a); if it does not involve OP$^+$, the presupposition of *wish* is not satisfied (16b).

(15) I **wish** it {***is**$_{\text{O-MARKED}}$ / **was**$_{\text{X-MARKED}}$} raining now. (Crowley 2024: 49-50)

(16) a. OP$^+$ [WISH$_a$ [O ϕ]] ⇝ fails to satisfy O's presupposition
b. WISH$_a$ [O ϕ] ⇝ fails to satisfy WISH's presupposition

My account also predicts the badness of the conditional in (17a) and the conditional in (17b): in either case, if the sentence involves OP$^+$, the presupposition of O fails to be satisfied (18a); if the sentence does not involve OP$^+$, the presupposition of X fails to be satisfied (18b).[10]

(17) a. *If John **were**$_{\text{X-MARKED}}$ here now, Mary {**is**$_{\text{O-MARKED}}$ / **will**$_{\text{O-MARKED}}$ be} here too.
b. *If John **is**$_{\text{O-MARKED}}$ here now, Mary **would**$_{\text{X-MARKED}}$ be here too.

(18) a. OP$^+$ [IF [X ϕ], [O ψ]] / OP$^+$ [IF [O ϕ], [X ψ]] ⇝ fails to satisfy O's presupposition
b. IF [X ϕ], [O ψ] / IF [O ϕ], [X ψ] ⇝ fails to satisfy X's presupposition

6 Comparison with Crowley 2024

Building on Leahy 2011, 2018, Crowley 2024 proposed a localist account that treats each instance of X-marking as semantically *vacuous*.[11] (19) is a rendition of his account in the current setup.

[9] As shown in (i), the opposite pattern (i.e., X-marking appears only on WANT) is acceptable in Greek (Iatridou 2000 noted that (i) expresses an indirect, still fulfillable desire). I note that, as shown in (ii), these two patterns of mixed combination are both unacceptable in Mexican Spanish (Aarón Sánchez, p.c.), confirming the prediction of the proposed account. The apparent cross-linguistic variation observed here must be left for future research.

(i) *Tha* ***ithela***$_{\text{X-MARKED}}$ [*na* ***ine*** *psilos*].
FUT want.PAST.IMPF NA be.NPST tall
'I would like him to be tall.' (Iatridou 2000: 33)

(ii) a. ***Quiero*** [*que* {I. ***venga*** / II. ****viniera***$_{\text{X-MARKED}}$} *mañana*].
want.PRES.IND.1SG that come.PRES.SUBJ.3SG come.PAST.IMPF.SUBJ.3SG tomorrow
(I) 'I want him to come tomorrow'. (II) Intended: 'I wish he would come tomorrow.'
b. ***Quisiera***$_{\text{X-MARKED}}$ [*que* {I. ****venga*** / II. ***viniera***$_{\text{X-MARKED}}$} *mañana*].
want.PAST.IMPF.SUBJ.1SG that come.PRES.SUBJ.3SG come.PAST.IMPF.SUBJ.3SG tomorrow
'I wish he would come tomorrow.'

[10] (17b) improves if focus is placed on *would* (Si Kai Lee, p.c.). I leave this for future investigation.
[11] See also Wimmer 2024.

(19) For any world w and domain function $\mathcal{D}$, $[\![\text{X}_{\text{CROWLEY}}\ \phi]\!]^{w,\mathcal{D}} = [\![\phi]\!]^{w,\mathcal{D}}$.

Crowley assumes that only O-marking is semantically non-vacuous. Recast in the present system, his analysis of O-marking triggers the presupposition that w in $[\![\cdot]\!]^{w,\mathcal{D}}$ is included in the context set of the factive common ground (conditionals) or the doxastic state of the attitude holder (desire reports); in the scope of universal quantifiers, it presupposes that the quantification domain is contained in the context set or the doxastic state. In his account, domain expansion associated with X-marking results from pragmatic competition with O-marking.

As he himself noted (fn.26), Crowley set aside cases like Greek (2) where X-marking appears both on WANT and in the complement, for which my proposal provides a straightforward account. It is in fact not clear how his account rules out the illicit combination of O/X in (12). (20) is the counterpart of (14b) that replaces my entry of X-marking with $\text{X}_{\text{CROWLEY}}$. Since it is semantically vacuous, $\text{X}_{\text{CROWLEY}}$ does not clash with O on WANT (even assuming his analysis of O-marking). His account thus does not seem to predict the badness of (12).[12]

(20) O [WANT_a [$\text{X}_{\text{CROWLEY}}$ ϕ]] ⇝ no semantic clash

In the same vein, Crowley's account does not predict the badness of the conditionals in (17): the $\text{X}_{\text{CROWLEY}}$-counterparts of (18b), as shown in (21), do not lead to semantic clash.[13]

(21) IF [$\text{X}_{\text{CROWLEY}}$ ϕ], [O ψ] / IF [O ϕ], [$\text{X}_{\text{CROWLEY}}$ ψ] ⇝ no semantic clash

7 Future outlook

I believe that the proposed account of X-marking explains similar data from many other languages beyond English and Greek, although I recognize that it alone does not capture the entirety of X-marking phenomena in natural language. Even in English, the empirical landscape is more complicated than what this paper has shown, as evidenced, e.g., by the contrast in (22) (Ippolito 2003: 41, 48). As Ippolito 2003 observed, when the existence presupposition of event participants is not satisfied at the utterance time, the conditional has to involve an *additional layer* of Past.[14]

(22) a. #Charlie is dead. If he **came** to the party tomorrow, he **would** meet Sally.
b. Charlie is dead. If he **had come** to the party tomorrow, he **would have** met Sally.

How does this fare with the present account? Recent literature has suggested that this additional layer of Past, unlike those appearing in (22a), contributes a real temporal meaning (Schulz 2017; S. Kaufmann 2023; see M. Kaufmann and Todorović 2023 for Serbian; see also Mizuno 2023 for cross-linguistic research on desire reports regarding this point). This layer of Past is assumed to make the same contribution as analyzed by the so-called "Past-as-Past" approach to Fake Past; it scopes over the conditional and expands the domain by enforcing temporal backshift.[15]

[12] Crowley proposed a way to derive the contrast in (15), but the derivation crucially hinges on his assumption that English *wish* lexically presupposes the falsity of its complement; it is less plausible that the same assumption holds for WANT+X in Greek (if so, one would have to assume that X-marking presupposes counterfactuality one way or another, which is not plausible given a range of non-counterfactual uses of X-marking; see, e.g., vFI). Note that my explanation of (15) does not hinge on whether *wish* presupposes the falsity of its complement.

[13] Crowley's account actually predicts that the two forms in (21) end up semantically equivalent to "IF [O ϕ], [O ψ]" (i.e., O-marked conditionals such as "*if John is here now, Mary is here too*"), because he assumes that O-marking, whether in the antecedent or in the consequent, triggers the same global presupposition (i.e., that the quantification domain is contained in the context set). More specifically, the two occurrences of O-marking in IF [O ϕ], [O ψ] are assumed to trigger an identical global presupposition at each locus in his account, but the net global presupposition is the same as those triggered in IF [$\text{X}_{\text{CROWLEY}}$ ϕ], [O ψ] and IF [O ϕ], [$\text{X}_{\text{CROWLEY}}$ ψ], because one occurrence of O-marking suffices to contribute the relevant presupposition. This I deem to be another undesirable prediction of his account.

[14] See also Ippolito 2013. See also S. Kaufmann 2023 for some discussion of this observation by Ippolito.

[15] I refer the reader to vFI for the debate between the Past-as-Past view and the "Past-as-Modal" view.

Leaving further investigation for future work, I speculate that conditionals like (22b) involve the additional layer of Past/Perfect in the position of OP$^+$ and thus license the occurrences of X.[16]

(23) PAST/PERFECT [IF [X ϕ], [X ψ]]

It also remains to be seen how my account extends to X-marking appearing in weak necessity modals (von Fintel and Iatridou 2008, 2023).[17] I leave all remaining issues for future research.

A Proofs

A.1 Domains created by $\sim$ and $\sim^+$ form equivalence classes

I first show as a lemma that $\sim$ and $\sim^+$ both have the properties in (24).

(24) An accessibility relation R is
   a. *shift-reflexive* iff for all w, w', if wRw', then $w'Rw'$;
   b. *shift-symmetric* iff for all w, w', w'', if wRw' and wRw'', then $w'Rw''$ and $w''Rw'$.

Lemma A.1. *$\sim$ and $\sim^+$ are both shift-reflexive and shift-symmetric.*

Proof. Both properties follow immediately from euclideanness of $\sim$ and $\sim^+$. □

Proposition A.2. *For any w, $\{w' \mid w \sim w'\}$ and $\{w' \mid w \sim^+ w'\}$ form equivalence classes.*

Proof. Transitivity follows from the definitions of $\sim$ and $\sim^+$. Reflexivity and symmetry within these sets follow from shift-reflexivity and shift-symmetry of $\sim$ and $\sim^+$. □

A.2 Reduction from $(*)$ to (d) in (11)

$(*)$ and (d) in (11) can be restated respectively as in (25a) and (25b) (ignoring the restriction of the domain through $\preccurlyeq_{\mathbf{a}}$, which I assume does not affect the reduction here). I first show as a lemma that the two underlined parts are equivalent. Equivalence between (25a) and (25b), and consequently the reduction from $(*)$ to (d), follow immediately from the lemma.

(25) a. For any w', w'' such that $\underline{w \sim w' \text{ and } w' \sim^+ w''}$, $[\![\text{X } \phi]\!]^{w'',\mathcal{D}_+} = 1$ $\cdots (*)$ in (11)
   b. For any w'' such that $\underline{w \sim^+ w''}$, $[\![\text{X } \phi]\!]^{w'',\mathcal{D}_+} = 1$ $\cdots (d)$ in (11)

Lemma A.3. *For any w, w', w'', the following holds: $w \sim w'$ and $w' \sim^+ w''$ iff $w \sim^+ w''$.*

Proof. ($\Rightarrow$) Suppose $w \sim w'$. Since $\{w' \mid w \sim w'\} \subset \{w' \mid w \sim^+ w'\}$, it follows that $w \sim^+ w'$. By this, $w' \sim^+ w''$ and transitivity of $\sim^+$, it follows that $w \sim^+ w''$.
($\Leftarrow$) Suppose $w \sim^+ w''$. By seriality of $\sim$, there exists w' such that $w \sim w'$. Take any w' such that $w \sim w'$. Since $\{w' \mid w \sim w'\} \subset \{w' \mid w \sim^+ w'\}$, it follows that $w \sim^+ w'$. By this, $w \sim^+ w''$ and euclideanness of $\sim^+$, it follows that $w' \sim^+ w''$. □

Proposition A.4. *(25a) and (25b) are equivalent.*

Proof. This follows immediately from Lemma A.3. □

[16] See Mizuno 2023 for a formal implementation of this idea for counterfactual desire reports across languages.

[17] Another remaining issue is occurrences of X-marking in deeply embedded positions, e.g., "*were*" in (ia). While (ia) requires *were* instead of *is* in this position (Si Kai Lee, Robert Lowe, p.c.), the opposite holds for (ib).

(i) a. [If I *knew*$_{\text{X}}$ [this ***were***$_{\text{X}}$ chocolate]], I *would*$_{\text{X}}$ eat it. (Iatridou 2000: 63a)
   b. If New York *were*$_{\text{X-MARKED}}$ the capital of the United States, John *would*$_{\text{X-MARKED}}$ know that it {✓***is***$_{\text{O-MARKED}}$ / ??***was***$_{\text{X-MARKED}}$ / #***were***$_{\text{X-MARKED}}$} (but Bill wouldn't).

References

Anand, Pranav and Valentine Hacquard (2013). "Epistemics and attitudes". In: *Semantics and Pragmatics* 6.8, pp. 1–59. DOI: 10.3765/sp.6.8.

Crowley, Paul (2024). "Modal tense: *if* and *wish*". In: *Linguistics and Philosophy* 47, pp. 401–443. DOI: 10.1007/s10988-023-09401-5.

Grønn, Atle (2021). "Justifying tense and mood morphology in counterfactuals". In: *Theoretical Linguistics* 47.3-4, pp. 233–250. DOI: 10.1515/tl-2021-2021.

Heim, Irene (1992). "Presupposition projection and the semantics of attitude verbs". In: *Journal of Semantics* 9.3, pp. 183–221. DOI: 10.1093/jos/9.3.183.

Iatridou, Sabine (2000). "The grammatical ingredients of counterfactuality". In: *Linguistic Inquiry* 31.2, pp. 231–270. DOI: 10.1162/002438900554352.

Ippolito, Michela (2003). "Presuppositions and implicatures in counterfactuals". In: *Natural Language Semantics* 11, pp. 145–186. DOI: 10.1023/A:1024411924818.

— (2013). *Subjunctive Conditionals: A Linguistic Analysis*. Cambridge, MA: MIT Press.

Kaufmann, Magdalena and Neda Todorović (2023). "The remote and the impossible in Serbian". Handout Presented at FDSL 16.

Kaufmann, Stefan (2005). "Conditional truth and future reference". In: *Journal of Semantics* 22.3, pp. 231–280. DOI: 10.1093/jos/ffh025.

— (2023). "How fake is fake past?" In: *Conditionals: Logic, Linguistics and Psychology*. Ed. by Stefan Kaufmann, David E. Over, and Ghanshyam Sharma. Palgrave Macmillan, pp. 389–423.

Kratzer, Angelika (1986). "Conditionals". In: *Papers from the Parasession on Pragmatics and Grammatical Theory at the Twenty-Second Regional Meeting of the Chicago Linguistic Society (CLS)*. Ed. by Anne M. Farley, Peter T. Farley, and Karl-Erik McCullough, pp. 1–15.

— (1991). "Conditionals". In: *Semantik: Ein Internationales Handbuch der Zeitgenössischen Forschung*. Ed. by Arnim von Stechow and Dieter Wunderlich. Berlin/New York: Walter de Gruyter, pp. 651–656. DOI: 10.1515/9783110126969.7.651.

Leahy, Brian (2011). "Presuppositions and antipresuppositions in conditionals". In: *Proceedings of SALT* 21, pp. 257–274. DOI: 10.3765/salt.v21i0.2613.

— (2018). "Counterfactual antecedent falsity and the epistemic sensitivity of counterfactuals". In: *Philosophical Studies* 175, pp. 45–69. DOI: 10.1007/s11098-017-0855-z.

Lewis, David (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Mackay, John (2019). "Modal interpretation of tense in subjunctive conditionals". In: *Semantics and Pragmatics* 12.2, pp. 1–29. DOI: 10.3765/sp.12.2.

Mizuno, Teruyuki (2023). "Counterfactual expressions: An investigation into their structures and meanings". PhD thesis. University of Connecticut, Storrs.

Rubinstein, Aynat (2012). "Roots of modality". PhD thesis. University of Massachusetts Amherst.

— (2017). "Straddling the line between attitude verbs and necessity modals". In: *Modality across Syntactic Categories*. Ed. by Ana Arregui, María Luisa Rivero, and Andrés Salanova. Oxford: Oxford University Press, pp. 109–131. DOI: 10.1093/acprof:oso/9780198718208.003.0007.

Schulz, Katrin (2014). "Fake tense in conditional sentences: A modal approach". In: *Natural Language Semantics* 22, pp. 117–144. DOI: 10.1007/s11050-013-9102-0.

— (2017). "Fake perfect in X-marked conditionals". In: *Proceedings of SALT* 27, pp. 547–570. DOI: 10.3765/salt.v27i0.4149.

Stalnaker, Robert (1968). "A theory of conditionals". In: *Studies in Logical Theory*. Ed. by Nicholas Rescher. Oxford: Blackwell, pp. 98–112.

— (1975). "Indicative conditionals". In: *Philosohpica* 5, pp. 269–286. DOI: 10.1007/BF02379021.

— (2002). "Common ground". In: *Linguistics and Philosophy* 25, pp. 701–721. DOI: 10.1023/A:1020867916902.

Stalnaker, Robert (2014). *Context*. Oxford: Oxford University Press.
Villalta, Elisabeth (2008). "Mood and gradability: an investigation of the subjunctive mood in Spanish". In: *Linguistics and Philosophy* 31, pp. 467–522. DOI: 10.1007/s10988-008-9046-x.
von Fintel, Kai (1999). "NPI licensing, Strawson entailment, and context dependency". In: *Journal of Semantics* 16.2, pp. 97–148. DOI: 10.1093/jos/16.2.97.
von Fintel, Kai and Sabine Iatridou (2008). "How to say *ought* in foreign: The composition of weak necessity modals". In: *Time and Modality*. Ed. by Jacqueline Guéron and Jacqueline Lecarme. Dordrecht: Springer, pp. 115–141. DOI: 10.1007/978-1-4020-8354-9_6.
— (2023). "Prolegomena to a theory of X-marking". In: *Linguistics and Philosophy* 46, pp. 1467–1510. DOI: 10.1007/s10988-023-09390-5.
Wimmer, Alexander (2024). "X- vs. O-marked want". In: *Journal of Semantics* 41.2, pp. 211–243. DOI: 10.1093/jos/ffae005.
Yalcin, Seth (2024). "Defining common ground". In: *Linguistics and Philosophy* 47, pp. 1045–1070. DOI: 10.1007/s10988-024-09415-7.

What Is Carnap's Problem and How Can We Solve It?

Julien Murzi
University of Salzburg
Salzburg, Austria
julien.murzi@plus.ac.at

Brett Topey
University of Salzburg
Salzburg, Austria
brett.topey@plus.ac.at

Abstract

On a plausible approach to logical metasemantics, our dispositions to treat a logical expression's I- and E-rules as valid determine its contribution to the truth conditions of sentences in which it appears. Carnap's so-called Categoricity Problem is that the rules in question don't seem to fix a unique interpretation of our logical vocabulary: there appear to be deviant interpretations of both the connectives and the quantifiers that are compatible with the validity of their rules. And although standard responses are available to Carnap's problem as it applies to propositional logic (by appeal, e.g., to bilateral rules or to a local notion of validity), the case of the quantifiers is more difficult. Here we develop a more precise account of how Carnap-style arguments work than has ever before been given, one that makes clear why certain such arguments succeed while others fail. In so doing, we demonstrate that despite recent criticisms, the account of the categoricity of the quantifiers we offered in previous work isn't threatened by any of the alleged deviant interpretations that have been discussed in the literature: each of these either is incompatible with the validity of the quantifier rules or else results in an illegitimate Carnap-style argument.

1 Introduction

Carnap's so-called Categoricity Problem is a metasemantic problem for moderate inferentialism: the plausible (by our lights) view that use determines meaning. More specifically, in the case of a logical operator \$, moderate inferentialists hold that our dispositions to treat \$'s I- and E-rules as valid categorically determine \$'s contribution to the truth conditions of the sentences in which it appears. Moderate inferentialism is inferentialist, in that it's a view on which the interpretation of a given expression is a product of the inference rules that govern its use, but it's also moderate, in that it's based on an entirely orthodox reference-based view of what interpretations are. It can also, unlike many philosophical views, be *tested*: since both the standard reference-based interpretation of our logical expressions and their I- and E-rules are relatively well understood, one can check, for any logical operator \$, whether the rules governing its use really do determine its interpretation. Carnap's problem, then, is that there seem to be deviant interpretations of both the connectives and the quantifiers that are nevertheless compatible with the validity of their I- and E-rules – i.e., that the validity those rules seemingly fails to categorically determine the interpretation of our logical expressions, in which case moderate inferentialism can't be correct. Consider, for instance, the trivial valuation: a valuation v^* that makes every sentence true, including 'A' and '$\neg A$'. This valuation is incompatible with the standard truth conditions for negation, and yet it appears to be compatible with the validity of all natural deduction rules, including the I- and E- rules for the negation sign.

Moderate inferentialism, though, is a highly unspecific doctrine: it entails that the validity of I- and E-rules fixes the interpretation of logical expressions, but it doesn't tell us how to formalise logic, nor does it tell us how to understand the notion of validity. And as we now know, Carnap's problem is highly sensitive to both the way logic is formalised and the way validity is understood. For instance, bilateral and multiple-conclusion formalisations of classical propositional logic (CPL) are categorical (see, e.g., Smiley 1996; Rumfitt 2000; Raatikainen 2008;

Murzi and Hjortland 2009), and so indeed can be standard single-conclusion formalisations, provided that validity is understood as *local* (see, e.g., Garson 2013; Murzi and Topey 2021). So there are plausible responses available to Carnap's problem for CPL.

Of course, natural language contains quantifiers in addition to propositional connectives, and so a full solution to Carnap's problem must generalize to first-order logic (FOL), and arguably also to second-order logic (SOL) and beyond. But here things are more complex. Bonnay and Westerståhl (2016) have argued that on certain semantic assumptions, the rules of FOL categorically determine the interpretation of the first-order quantifiers. Their claim, though, is highly problematic on at least two counts. First, the semantic assumptions in question are arguably illicit: the moderate inferentialist's claim was that our use of logical expressions determines their interpretation *on its own*, not that it does so when certain semantic facts are already in place. Second, as we show, Bonnay and Westerståhl's work depends on a mistaken understanding of what it is for a model to be deviant – a mistake that's inherited by del Valle-Inclán's (2024) recent critique of their proposal. Further complicating matters is the fact that we have argued in previous work – see our "Categoricity by Convention" (2021) – that the validity of the rules of FOL and SOL determines their standard, unrestricted interpretations, provided those rules are understood as *open-ended* – a conclusion that, according to Brîncuş (2024a) and Brîncuş (2024b), is disproved by deviant interpretations of the quantifiers given by Carnap (1943) and Garson (2013).

Here we offer a more precise understanding of Carnap-style objections to moderate inferentialism than has appeared in the literature before, an understanding that makes it clear why certain Carnap-style arguments succeed – in the sense that they genuinely show that certain ways of formalising logic, or certain notions of validity, are unavailable to moderate inferentialists – while others fail. In so doing, we demonstrate that despite recent criticisms, our own account of the categoricity of the quantifiers isn't threatened by the alleged deviant interpretations of the quantifiers offered by Carnap, Garson, Briîncuş, or del Valle-Inclán: each of these either is incompatible with the validity of the quantifier rules (as we formalise them) or else results in an illegitimate Carnap-style argument. The result is that moderate inferentialism, as we develop it, turns out to be tenable after all.

2 Background

We said that, on our view, the interpretation of a logical expression e is determined by our dispositions to treat the basic rules for e as valid. But what is a rule and what is our conception of validity? If e is a logical expression, we take e's basic rules to be its I- and E-rules in a *sequent-style single-conclusion natural deduction system*. For instance, we formalise the basic rules for $\rightarrow$ as follows:

$$\frac{\Gamma, \varphi \vdash \psi}{\Gamma \vdash \varphi \rightarrow \psi} \rightarrow\text{-I} \qquad \frac{\Gamma \vdash \varphi \rightarrow \psi \qquad \Delta \vdash \varphi}{\Gamma, \Delta \vdash \psi} \rightarrow\text{-E}$$

As for validity, there are three possible accounts:

(i) validity as preservation of truth;

(ii) validity as preservation of sequent satisfaction (local validity); and

(iii) validity as preservation of sequent validity (global validity).

Simple truth preservation doesn't apply to metainferences such as $\rightarrow$-I, $\neg$-I, $\vee$-E etc. So our choice is restricted to (ii) and (iii), which apply to (meta)rules whose inputs and outputs are sequents.

We begin by defining the notion of sequent satisfaction:

Definition 2.1 (Sequent satisfaction)**.** A valuation v satisfies a sequent $\Gamma \vdash \varphi$ iff either v makes some $\psi \in \Gamma$ false or v makes φ true.

We can then define global and local validity as follows:

Definition 2.2 (Global and local validity)**.** A metarule

$$\frac{\Gamma_1 \vdash \varphi_1 \quad \dots \quad \Gamma_n \vdash \varphi_n}{\Delta \vdash \psi}$$

- is locally valid with respect to a class of valuations V iff it preserves satisfaction – i.e. iff for all valuations $v \in V$, if v satisfies every $\Gamma_i \vdash \varphi_i$, it also satisfies $\Delta \vdash \psi$; and
- is globally valid with respect to V iff it preserves validity – i.e. iff, if all $v \in V$ satisfy every $\Gamma_i \vdash \varphi_i$, then all $v \in V$ satisfy $\Delta \vdash \psi$.

We take validity to be local validity, essentially because, in our setting, it is a generalisation of truth preservation. After all, if we let $\Gamma_1, \dots, \Gamma_n$ and Δ be empty, on the local account, a meta-rule

$$\frac{\Gamma_1 \vdash \varphi_1 \quad \dots \quad \Gamma_n \vdash \varphi_n}{\Delta \vdash \psi}$$

is valid if and only if the inference from $\varphi_1, \dots, \varphi_n$ to ψ preserves truth in v for all v's, as, we think, it should be.

Following Tennant (1999), we interpret '$\bot$' as a logical punctuation sign (formally, $\bot = \varnothing$). On this assumption, the local conception of validity gets us the determinacy of CPL, via the following theorem proved by Garson (2013):

Theorem 2.3 (Garson's Local Validity Theorem)**.** *The rules of* CPL *are locally valid with respect to a class of vlauations V only if all members of V obey the classical truth tables.*

Let us now move on to FOL. In general, establishing that our treating certain rules of inference as valid categorically determines particular truth-conditional meanings for our expressions involves showing that, among all the possible ways of assigning truth values to formulae – i.e., all *valuations* or *general models* – the *only* ones that are compatible with the validity of the rules are those that respect the truth-conditional meanings in question (where respecting those truth-conditional meanings amounts to satisfying the associated semantic clauses). But since we are interested here in the question of the categoricity of our *logical* expressions in particular, we can take for granted that nonlogical expressions have meanings of the usual sort, with each constant designating some member of a domain of objects, each predicate having as its extension some set of tuples of objects from the domain, and so on. So we can let the set of general models be the set of triples $\langle D, I, [\![]\!] \rangle$, where D and I are a domain and interpretation function of the usual sort and where $[\![]\!]$ is a function from pairs of formula and variable assignment to truth values that assigns truth values to atomic formulae in the usual way but may assign truth values to logically complex formulae in any way whatsoever. In this setting, solving Carnap's problem amounts to proving, just from the assumption that our logical operators' I- and E-rules are valid across some set of general models, that all of those models satisfy the usual truth-conditional semantic clauses for the operators in question.

We interpret natural deduction rules as being rules for formulae rather than for sentences. Thus, we adopt the following version of $\forall$-I:

$$\frac{\Gamma \vdash \varphi}{\Gamma \vdash \forall x \varphi} \forall\text{-I}$$

where x doesn't appear free in Γ. And since we now allow arbitrary formulae to occur in sequents, we must generalise our definitions of (sequent) satisfaction and validity.

Definition 2.4 (Sequent satisfaction relative to a variable assignment)**.** Let α be a variable assignment over v's domain. Then, α satisfies$_\alpha$ $\Gamma \vdash \varphi$ iff either for some $\psi \in \Gamma$, $[\![\psi]\!]_{v,\alpha} \neq 1$ or $[\![\varphi]\!]_{v,\alpha} = 1$.

Definition 2.5 (Sequent satisfaction, generalised)**.** A valuation v satisfies $\Gamma \vdash \varphi$ iff, for every variable assignment α over v's domain, α satisfies$_\alpha$ $\Gamma \vdash \varphi$.

(Note that this generalised definition gives the previous one as a special case when all the formulae in the sequent are sentences.)

Definition 2.6 (Local validity, generalised)**.** A meta-inference is locally valid with respect to a class of valuations V iff it preserves sequent satisfaction for all $v \in V$ – i.e. iff for all $v \in V$ either the premiss sequents are not all satisfied by v or the conclusion sequent is satisfied by v.

We accept, with McGee, that the first-order rules are '*open-ended*, meaning that the rules are valid... not only within the language... but they will remain valid however the language may enriched by the addition of new sentences' (2000, p. 66). On this assumption, it is possible to show that the local validity of ∀-I determines the standard, objectual interpretation of ∀:

(∀) $[\![\forall v\varphi]\!]_{\mathcal{M}} = 1$ iff $[\![\varphi[o/v]]\!]_{\mathcal{M}} = 1$ for all $o \in D$.

3 Three Types of Carnap-Style Argument

The various broadly Carnapian arguments for the non-categoricity of logic that have appeared in the literature can be understood as falling under one of three general types. Arguments of two of these types can, depending on their details, be successful, but arguments of the third type turn out to be entirely illegitimate. We proceed by distinguishing among these three types of non-categoricity argument and discussing them in turn.

3.1 Type I Arguments

In Type I arguments, a general model is presented that's compatible with the validity of the rules but incompatible with the standard semantic clauses for our logical operators. Insofar as the general model that's presented genuinely has these features, a Type I argument will certainly be successful: any such general model is straightforwardly a counterexample to the claim that the validity of the rules guarantees that the operators have their standard meanings.

The trivial valuation v^*, for example, is usually deployed in service of a Type I argument: insofar as that valuation (understood as a general model) is compatible with the validity of the rules – which it is, if the rules are the usual single-conclusion rules and their validity is understood by appeal to the usual consequence relation – the rules fail to pin down the standard semantic clauses for the negation and disjunction signs, since v^* is compatible *only* with deviant interpretations of those operators.

Of course, as suggested above, there are various ways to avoid this implication by building additional structure into either our formalisation of logic or our notion of validity. This is why our local-validity-based account of the categoricity of CPL, for instance, isn't threatened by this sort of Type I argument.

3.2 Type II Arguments

Type II arguments begin not from a general model but directly from a deviant set of semantic clauses for the logical operators; here what's shown is simply that the validity of the rules fails to rule out the deviant interpretation in question. To show this requires examining *all* general models compatible with that interpretation in order to determine whether the rules are valid across those general models. It also requires ensuring that the interpretation is genuinely deviant – i.e., that at least one general model compatible with that interpretation is *incompatible* with the standard interpretation. (An alternative semantic clause for the negation sign according to which '$\neg A$' is true in a model just in case 'A' is false *and* '$A \to A$' is true, for instance, is not deviant in any genuine sense.) Insofar as a Type II argument meets both constraints, it will be successful.

The argument from Carnap's deviant semantic clause for the universal quantifier, for example, is a Type II argument: all (propositionally standard) general models compatible with that deviant interpretation – an interpretation on which '$\forall xFx$' is true just in case every object in the domain is in the extension of 'F' *and* the object designated by 'b' is in the extension of 'G' – turn out also to be compatible with the validity of the rules (if those rules are given the sort of axiomatic formalisation that was common in Carnap's time).

Again, though, it's possible to avoid this result simply by building more structure into our formalisation of logic. On our own formalisation of the I- and E-rules for the universal quantifier, for instance, the categoricity of FOL isn't threatened by this sort of Type II argument.

3.3 Type III Arguments

Type I arguments work by holding fixed a general model and considering all possible interpretations of the logical operators that are compatible with it, and Type II arguments work by holding fixed an interpretation of the operators and considering all possible general models that are compatible with it. Both strategies are legitimate. In Type III arguments, though, elements of those strategies are illegitimately combined: restrictions are placed both on general models and on interpretations of the operators in such a way that no demonstration of non-categoricity is genuinely made available.

This is perhaps best clarified by example. So consider Briîncuş's claim that the universal introduction rule remains valid given Carnap's deviant interpretation of the universal quantifier if, in addition, "we take 'Gb' to be true" (2024, p. 349). This claim is intended to serve as an objection to various attempts to secure the categoricity of the quantifier, including our own. This, though, is a Type III argument: Briîncuş presents a deviant semantic clause for the quantifier while ruling out by fiat certain general models compatible with that interpretation. When *all* general models compatible with that interpretation are taken into account, as in a Type II argument, it's evident that the universal introduction rule, as we formalise it, isn't valid across those general models: some of the general models satisfy 'Fx' but fail to make true 'Gb', in which case '$\forall xFx$' turns out to be false on Carnap's deviant interpretation despite the fact that the validity of the rule would guarantee its truth.

4 More on Type III Arguments: A Case Study

Del Valle-Inclán's critique of Bonnay and Westerståhl's proposal goes via a Type III argument as well. In this case, though, the mistake isn't del Valle-Inclán's; it originates with Bonnay and Westerståhl themselves, who build too much structure into their notion of a general model and so are committed to describing as deviant certain valuations that are in fact compatible with the standard interpretation of the logical operators. In particular, they take a general model not

merely to assign truth values to logically complex formulae but to come with an interpretation of the universal quantifier built in. So del Valle-Inclán proceeds by presenting a general model that otherwise is compatible with the standard interpretation of the universal quantifier and then stipulating that the quantifier has a deviant interpretation. More precisely, he constructs a general model on which, for every predicate, either everything in the domain is in the extension of that predicate or nothing is, so that a universally quantified sentence (standardly interpreted) will be true just in case its existentially quantified counterpart (standardly interpreted) is, and then he stipulates that we "(mis)interpret $\forall$ as $\exists$" (2024, p. 7), in which case the general model is of course deviant.

Notice, though, that if we take a general model (as we should) merely to assign truth values to logically complex formulae, it's clear that del Valle-Inclán is here holding fixed *both* a general model *and* an interpretation of the universal quantifier. And if we hold fixed only the general model, as in a Type I argument, it's evident that the general model is in fact *not* deviant: since '$\forall$' satisfies the standard semantic clause for the existential quantifier, it will *also* satisfy the standard clause for the universal quantifier (since, again, the general model is designed in such a way that those two interpretations will correspond to the same assignment of truth values to quantified sentences).

5 Concluding Remarks

In short, it turns out that, by attending carefully to the differences between the three types of Carnap-style argument, we can see that the strategy for solving Carnap's problem we offered in "Categoricity by Convention" (2021) remains promising despite some recent suggestions to the contrary.

Acknowledgements. Funding for this research was provided by the Austrian Science Fund (Grant No. P33708), whose support we gratefully acknowledge. We're also grateful to audiences at uAnalytiCon-2022 at Ural Federal University, the Perspectives on Categoricity workshop at the University of Vienna, the Convention in Logic and Language conference at the University of Haifa, the 2022 FilMat Conference at the University of Pavia, the 1st EuPhiLo Conference at the University of Padua, the University of Salzburg, the 11th European Congress for Analytic Philosophy at the University of Vienna, the 10 Years of *Modal Logic as Metaphysics* conference at the University of Hamburg, and the (In)determinacy in Mathematics conference at the National University of Singapore for helpful discussion of the ideas presented here.

References

Bonnay, Denis and Dag Westerståhl (2016). "Compositionality solves Carnap's Problem". In: *Erkenntnis* 81.4, pp. 721–739. DOI: https://doi.org/10.1007/s10670-015-9764-8.

Brîncuş, Constantin C. (2024a). "Categorical quantification". In: *Bulletin of Symbolic Logic*. Advance online publication. DOI: https://doi.org/10.1017/bsl.2024.3.

— (2024b). "Inferential quantification and the ω-rule". In: *Perspectives on Deduction: Contemporary Studies in the Philosophy, History and Formal Theories of Deduction*. Ed. by Antonio Piccolomini d'Aragona. Springer, pp. 345–372. DOI: https://doi.org/10.1007/978-3-031-51406-7.

Carnap, Rudolf (1943). *Formalization of Logic*. Harvard University Press.

del Valle-Inclán, Pedro (2024). "Carnap's Problem, definability and compositionality". In: *Journal of Philosophical Logic*. Advance online publication. DOI: https://doi.org/10.1007/s10992-024-09767-2.

Garson, James (2013). *What Logics Mean: From Proof Theory to Model-Theoretic Semantics*. Cambridge University Press. DOI: https://doi.org/10.1017/CBO9781139856461.

McGee, Vann (2000). "'Everything'". In: *Between Logic and Intuition: Essays in Honor of Charles Parsons*. Ed. by Gila Sher and Richard Tieszen. Cambridge University Press, pp. 54–78. DOI: https://doi.org/10.1017/CBO9780511570681.

Murzi, Julien and Ole Thomassen Hjortland (2009). "Inferentialism and the categoricity problem: reply to Raatikainen". In: *Analysis* 69.3, pp. 480–488. DOI: https://doi.org/10.1093/analys/anp071.

Murzi, Julien and Brett Topey (2021). "Categoricity by convention". In: *Philosophical Studies* 178.10, pp. 3391–3420. DOI: https://doi.org/10.1007/s11098-021-01606-3.

Raatikainen, Panu (2008). "On rules of inference and the meanings of logical constants". In: *Analysis* 68.4, pp. 273–280. DOI: https://doi.org/10.1093/analys/68.4.282.

Rumfitt, Ian (2000). ""Yes" and "No"". In: *Mind* 109.436, pp. 781–823. DOI: https://doi.org/10.1093/mind/109.436.781.

Smiley, Timothy (1996). "Rejection". In: *Analysis* 56.1, pp. 1–9. DOI: https://doi.org/10.1093/analys/56.1.1.

Tennant, Neil (1999). "Negation, absurdity and contrariety". In: *What Is Negation?* Ed. by Dov M. Gabbay and Heinrich Wansing. Kluwer, pp. 199–222. DOI: https://doi.org/10.1007/978-94-015-9309-0.

Explaining universal semantic properties in the adjectival/adverbial domain

Irina Nemtcova
Center for Mind/Brain Sciences,
University of Trento
nemtzova.irina@gmail.com

Jakub Szymanik
Center for Mind/Brain Sciences,
Dep. of Information Engineering and Computer Science
University of Trento
jakub.szymanik@gmail.com

Abstract

Native speakers would generally agree that certain patterns in the language are preferable to others. The order of several adjectives modifying the same noun is the best example: various potential sequences of adjectives are all equivalent in meaning, yet speakers often prefer specific orders. Recently, with the availability of large quantities of data and significant development of processing tools and statistical methods, researchers have started to study word order preferences using quantitative approaches based on "efficiency-based" communication. In our work, we test three of these theories: pointwise mutual information (PMI) (Futrell, Qian, et al. 2019), integration complexity (IC) (Dyer 2017), and information gain (IG) (Futrell, Dyer, and Scontras 2020). We start by replicating the results obtained by (Futrell, Dyer, and Scontras 2020) on adjective order preferences in English. To verify the cross-linguistic consistency of the results, we analyze adjective order tendencies in Russian. Since the theories in the study provide insights into base word order, not limited to adjectives only, we continue by investigating in detail if the predictions made by these theories can generalize and be applied to explain the order of adverbs.

1 Introduction

Native speakers of a language often show a preference for the position of words in specific constructions. The order of several adjectives modifying the same noun is the best example: speakers consistently show preferences for a particular order of adjectives, and interestingly, these preferences are observed cross-linguistically. For instance, in a situation where English speakers would have to describe a chair that is both *small* and *red*, they would show a preference for *a small red chair* compared to *a red small chair* (Scontras, Degen, and Goodman 2017). The very same preferences are found not only in English but also in other languages where adjectives occupy the prenominal position, such as Greek, Polish, or Russian. The tendency for two or more consecutive adjectives to follow some order is also observed in postnominal languages, where adjectives come after the noun they modify. Various potential sequences of adjectives are all equivalent in meaning, yet specific orders are preferred to others.

Cross-linguistic similarities have also been found in the order of adverbs (Cinque 1999). Their distribution is challenging to study due to the high variation in adverbs' positioning. As long as there are no syntactic structure constraints, adverbs can be placed in various positions in the sentence, often without a change in meaning.

Many theories, primarily relying on descriptive approaches, have been proposed in an attempt to find a measure that governs the word order tendencies in the adjectival and adverbial domains (Danks and Glucksberg 1971; Scontras, Degen, and Goodman 2017; Cinque 1999.) However, these approaches often lack generalization and frequently describe correlation rather than actual cause.

Recent approaches that try to make more generic predictions that would hold cross-linguistically use the argument of "functionality." In other words, the communication process

is efficient-based and constrained by the brain: it tries to maximize information transfer while minimizing cognitive cost (Hawkins 2004). Efficiency can be achieved with simplicity by minimizing cognitive effort for both sides involved in communication: the sender and the receiver. It can, for instance, be quantified using the length of the transmitted message (Gibson et al. 2019). Thus, the reason we find patterns that are repeated across languages can be due to these constraints of cognition on communication.

Efficiency-based approaches have already been studied separately before on cross-linguistic data. However, except for a few studies, the results of their predictions have been analyzed in isolation and have been missing systematic comparisons. Following Futrell, Dyer, and Scontras 2020, our research aims to address this question and test the predictions from four efficiency-based theories: integration complexity (Dyer 2017), information gain, pointwise mutual information (Futrell, Dyer, and Scontras 2020), and subjectivity (Scontras, Degen, and Goodman 2017).

1.1 Predictors

Each theory in our analysis measures cognitive effort by examining a sentence's relationship between a headword and its dependents.

Integration complexity measurement is based on the co-occurrence of the dependent word and the syntactic category of its head. The theory predicts that a dependent with a narrow, predictable range of possible heads should be placed closest to the head since we are more likely to predict its head's syntactic category. On the other hand, a dependent with no clear tendency regarding the syntactic category of the head will be placed further. The range of possible heads of the depended is measured with Shannon's entropy using the probabilities of each possible outcome (x_i), as shown in Equation 1:

$$H(X) = -\sum_{i=1}^{n} P(x_i) * log_2 P(x_i). \tag{1}$$

PMI states that an efficient language will minimize the linear distance between elements with high pointwise mutual information (PMI) - an information-theoretic measure of how strongly two words predict each other. Thus, the dependent (x) that has high PMI with its head (y) will tend to be closer to the head:

$$PMI(x, y) = log \frac{p(x, y)}{p(x)p(y)}. \tag{2}$$

Information gain looks at a phrase as a decision tree for identifying a referent where each partitioning of a phrase is associated with some information gain. The proposal is that a word with smaller information gain will be placed earlier so that each word gradually narrows the set of referents. The information gain of a word is measured as a difference between the starting entropy of a headword $H[N]$ and the sum of the entropy of each partition of a dependent N_x (a subset of heads that take a given word as a dependent) and a negative evidence N_x^c (which in turn, do not have a given word as a dependent), as shown in the Equation 3:

$$IG(x) = H[N] - \left[\frac{N_x}{N}\ H[N_x] + \frac{N_x^c}{N}\ H[N_x^c]\right] \tag{3}$$

And lastly, **subjectivity** predicts that less subjective adjectives occur linearly closer to the nouns they modify (Scontras, Degen, and Goodman 2017). An empirical study of the English language supported the claim. Participants were shown a series of adjectives and had to indicate whether an adjective was objective or subjective using a sliding scale. After comparing subjectivity ratings with the sequences of several consecutive adjectives, it turned out that less subjective adjectives are placed closer to the noun they modify. The reason might be that the placement of adjectives from most to least subjective facilitates successful reference resolution.

2 Proposal

We propose to compare the accuracies of these theories and to find the predictor that would be the best in identifying the word order in phrases with two stacked adjectives/adverbs in pre-head position in typologically diverse languages: English and Russian.

First, we will calculate predictors based on the scoring function S(H, D) applied to a dependent (D) and a head (H). As we saw earlier, the estimated predictors come with theories describing their effect on word order. Thus, the difference in scoring functions between $S(D_1, H)$ and $S(D_2, H)$ can be enough to predict the order: dependents that have low PMI or low IG are placed father from the head; whereas dependents with low IC or subjectivity are placed closer to the head.

However, we can go further and fit a logistic regression classifier to predict a word order in a sequence D_1D_2H, given the difference between $\mathrm{S}(D_1, H)$ and $\mathrm{S}(D_2, H)$. The goal of the classifier is to predict the order from the unordered set of adjectives (A_1, A_2) and a noun; and adverbs (Adv_1, Adv_2) and a verb. Based on empirical results from previous studies, we expect that the fitted values will be negative for PMI and IG, and positive for IC and subjectivity. We can compare predictors in terms of accuracy by analyzing the classifier's results.

3 Data

We estimated predictors using an automatically annotated subsection of the Wikipedia corpus. First, the corpus was cleaned, and only sentences with alphanumeric characters and punctuation, with no less than 5 and no more than 35 items per sentence, were left. The limit on the sentence length was needed because of the particularities of the Wikipedia corpus to filter sentences that contained, for example, long enumeration lists, which did not have any meaningful information useful for the study. After cleaning the data, we got 22,534,266 tokens in English and 11,429,524 in Russian. The extracted text was then parsed with dependency annotations using Stanza parser (Qi et al. 2020).

From the corpus, we extracted:

– Russian, adjectives: For the Russian language, we worked with lemmas. First, we extracted adjective-noun (AN) pairs, a set of pairs $< A, N >$ where A is an adjective dependent on the noun N with dependency type *amod*. We extracted 895,485 AN pairs, out of which 345,675 unique pairs. AN pairs were used to calculate the values of corpus-based predictors and for clustering.

  Second, we extracted AAN triples, a set of triples $< A_1, A_2, N >$ where A_1 and A_2 are adjectives with relation type *amod*, dependent on a single noun head *N*. Additionally, A_1 and A_2 have no other dependents except *N*, must appear in order A_1A_2N with no intervening words in between. We extracted 56,801 unique triples, which we used to fit logistic regression from predictors to word order and evaluate the model.

– English, adjectives: We extracted 990,453 AN pairs in total, 430,698 unique ones, and 41,617 unique AAN triples following the above procedure. The analysis of the English language was done on word forms.

– A similar procedure was used for extracting adverb-verb pairs and adverb-adverb-verb triples, where the adverb is dependent on the verb and has a dependency type *advmod*. In English, we got 93,479 unique Adv-Verb pairs (279,509 total) and 3,320 Adv-Adv-Verb triples. In Russian, we got 71,073 unique Adv-Verb pairs (159,596 in total) and 1,172 Adv-Adv-Verb triples.

We estimated predictors on wordforms (for English) and lemmas (for Russian) as well as over clustering of words in embedding space.

To work with clusters, first, we substitute words with indexes of their respective clusters and then calculate the co-occurrences of these cluster indices (Figure 1). We use cluster analysis to deal with data sparsity problems (which might be present when working with the distribution of word forms). Additionally, this analysis can show us what information the predictors are sensitive to. For instance, clusters contain semantic information about the words that form the cluster.

$$\text{huge, big, massive} \longrightarrow Adj_{35} \ ; \ \text{house, building} \longrightarrow N_{21} \ ; \quad Adj_{35}N_{21}$$

Figure 1: Cluster analysis example: words are replaced with the respective number of their cluster, and then co-occurrences of these indices are calculated in pairs and triples.

We performed clustering with sklearn.cluster.KMeans (Pedregosa et al. 2011). The number of clusters was determined by choosing the largest k, which did not result in singleton clusters. For English, we used KMeans applied to a pre-trained set of 400k 300-dimension GloVe vectors generated from Wikipedia (Pennington, Socher, and Manning 2014). For Russian, we used Navec embeddings[1]. Adjectives were clustered into $k_A = 300$ and nouns $k_N = 1000$ clusters, both in English and Russian. Adverbs were clustered into $k_{Ad} = 50$ and verbs into $k_V = 150$ clusters in English. In Russian, we had $k_{Ad} = 50$ adverb and $k_V = 300$ verb clusters. The variation in the number of clusters for adverbs and verbs might reflect the inherent semantic differences between the two languages.

We calculated corpus-based predictors (PMI, IG, IC) using pairs of Adj-Noun and Adv-Verb. For English adjectives, we also include the subjectivity measure. The scores for 398 adjectives were provided by Futrell, Dyer, and Scontras 2020 following Scontras, Degen, and Goodman 2017 methodology: where 264 English-speaking participants indicated with the slider scale the subjectivity, by adjusting a slider (in a range from “completely objectivity" to “completely subjective"). Each out of 398 adjectives received an average of 20 ratings.

For comparative reasons, we keep triples where D_1 and D_2 have scores for all predictors; at the end, for testing a classifier, we have 8,752 AAN and 2,857 Adv-Adv-V triples in English, and 61,337 AAN and 1,085 Adv-Adv-V in Russian, which we use for fitting the model and evaluation.

4 Results

In both languages, the results differ based on whether the predictor is estimated on a wordform or a cluster, visualized in Figures 2 and 3. Still, we can point out that a specific pattern is present in both languages. When applied to adjective order, PMI is the best predictor when tested on word forms. However, when tested on clusters, its accuracy drops. In this case, integration complexity and information gain show better results. Interestingly, this pattern is present with adverbs too. This might be seen as a confirmation that PMI is good at tracing meaning closeness but may be less effective in capturing the general properties of the cluster to which a word belongs. Additionally, we can see that integration complexity shows better results when tested on clusters. The reason might be that IC is good at capturing general semantic properties of the cluster since it focuses not only on the distribution of words but also on the syntactic category of the heads. The same might be valid for information gain. This pattern aligns with the observations made by Futrell, Dyer, and Scontras 2020.

[1] https://github.com/natasha/navec

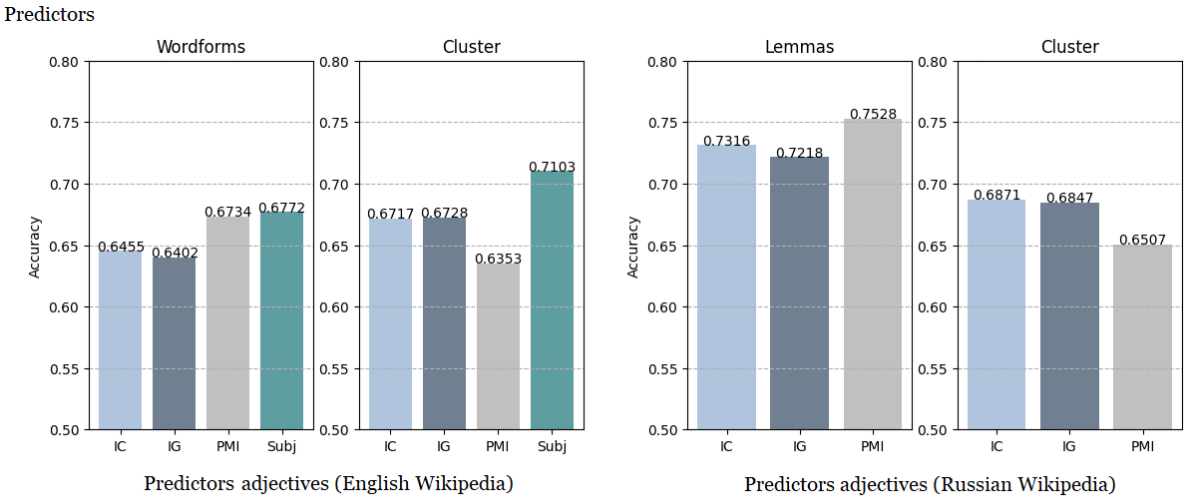


Figure 2: Test accuracy of logistic regression derived from 80:20 train/test split of 8,752 AAN triples extracted from a subpart of English Wikipedia and 61,337 AAN triples from subpart of Russian Wikipedia.

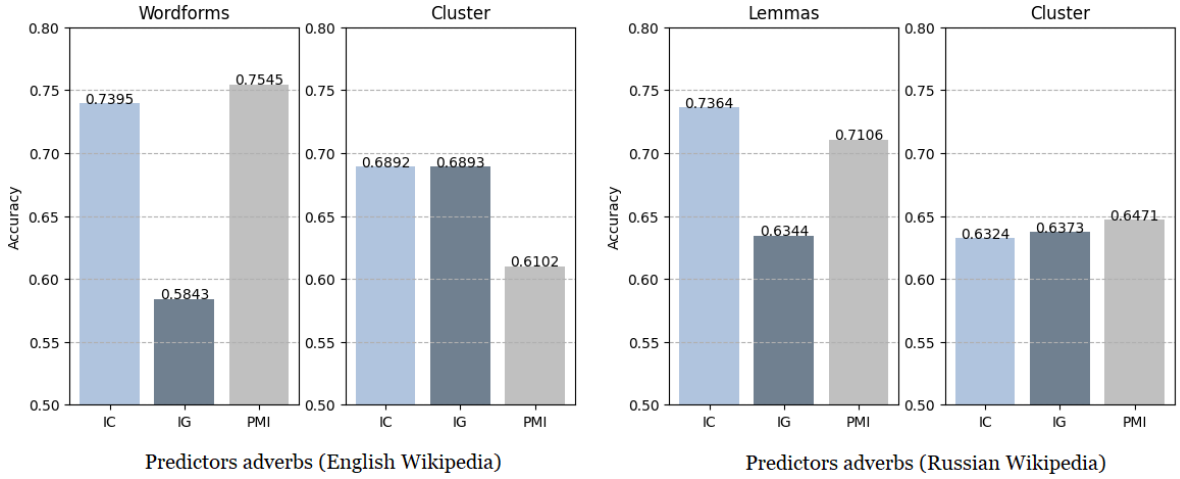


Figure 3: Test accuracy of logistic regression derived from 80:20 train/test split of 2,857 Adv-Adv-V triples extracted from a subpart of English and 1,085 Adv-Adv-V from subpart of Russian Wikipedia.

5 Discussion

The theories integrating syntax and processing efficiency can be tested to explain a broader linguistic phenomena. In our study, we started by replicating the results of Futrell, Dyer, and Scontras 2020 cross-linguistically (on two morphologically different languages) and across two different domains (adjectives and adverbs order). However, certain aspects should be taken into consideration when analyzing the results. First, the study's predictors are text-based, rely purely on syntactic annotations, and thus cannot differentiate between meanings. For future investigation one may want to focus more on the sematic aspect, for instance, by including more specific phrase annotation. Second, it was shown that subjectivity acts as the best predictor of adjective order in the prenominal position. However, collecting significant subjectivity ratings can not always be straightforward, as in the case of the English language. Future work for collecting subjectivity ratings is needed: either manually or including LLMs to get reliable scores

(Jumelet et al. 2024). Lastly, our study concentrates on the base order and shows tendencies rather than universals. We can see these tendencies as "soft biases" towards a base order that can be violated depending on the context and the meaning the speaker wants to convey.

6 Conclusion

In our work, we investigated whether processing efficiency theories can explain the tendencies in the attested order of adjectives and adverbs. Overall, the method of using the dependency structure of a sentence to study word order can be tested on other languages, provided they have a big enough corpus from which to extract pairs of heads and dependents, which are needed to calculate predictors. This research has many future directions: including more languages or examining the explanatory power coming from a combination of several predictors. Future work will be important for thoroughly understanding how information theoretic approaches explain word order.

References

Cinque, G. (1999). *Adverbs and Functional Heads: A Cross-Linguistic perspective.* Oxford University Press.

Danks, J. H. and S. Glucksberg (1971). "Psychological scaling of adjective orders." In: *Journal of Verbal Learning and Verbal Behavior, 10(1)*, pp. 63–67.

Dyer, W. (2017). "Minimizing Integration Cost: A General Theory of Constituent Order". PhD thesis. University of California, Davis.

Futrell, R., W. Dyer, and G. Scontras (2020). "What determines the order of adjectives in English? Comparing efficiency-based theories using dependency treebanks." In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 2003–2012.

Futrell, R., P. Qian, et al. (2019). "Syntactic dependencies correspond to word pairs with high mutual information". In: *Proceedings of the Fifth International Conference on Dependency Linguistics (Depling, SyntaxFest 2019).*

Gibson, E. et al. (2019). "How Efficiency Shapes Human Language". In: *Trends in Cognitive Sciences, 23(5)*, pp. 389–407. DOI: https://doi.org/10.1016/j.tics.2019.02.003.

Hawkins, J. A. (2004). *Efficiency and Complexity in Grammars.* Oxford, UK: Oxford University Press.

Jumelet, Jaap et al. (2024). *Black Big Boxes: Do Language Models Hide a Theory of Adjective Order?* arXiv: 2407.02136 [cs.CL]. URL: https://arxiv.org/abs/2407.02136.

Pedregosa, F. et al. (2011). "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research 12*, pp. 2825–2830.

Pennington, J., R. Socher, and C. D. Manning (2014). "GloVe: Global Vectors for Word Representation". In: *Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.

Qi, P. et al. (2020). "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages". In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations.*

Scontras, G., J. Degen, and N. D. Goodman (2017). "Subjectivity predicts adjective ordering preferences." In: *Open Mind: Discoveries in Cognitive Science, 1(1)*, pp. 53–65.

On the anti-exhaustive inference of *ya*

Andreea C. Nicolae
Leibniz-Zentrum Allgemeine
Sprachwissenschaft, Berlin, Germany
nicolae@leibniz-zas.de

Yasutada Sudo
University College London
London, UK
y.sudo@ucl.ac.uk

Muyi Yang
Osaka University
Osaka, Japan
muyi.yang.ym@gmail.com

Abstract

Sudo 2014 analyzes the connective *ya* in Japanese as an inclusive disjunction that triggers two scalar inferences, (i) a conjunctive inference and (ii) an anti-exhaustive inference, based on alternatives involving other connectives. Sauerland et al. 2017 propose a competing analysis of (i) that is based on domain alternatives, but they make no claim about (ii). We report on an experiment with an inference task whose results indicate that *ya* does indeed trigger (ii) but it is not as robust as (i). We argue that this is problematic for Sudo's (2014) analysis, but at the same time shows that Sauerland et al.'s (2017) analysis needs to be augmented with a mechanism for generating (ii). In order to account for the difference in robustness, we propose that (i) is a grammatically derived scalar inference, while (ii) is a pragmatically derived scalar inference.

1 Introduction

Japanese has a connective *ya*, which is often used to list examples (as Kuno 1973, p. 114 remarks). For instance, (1), is roughly translated with '... and ..., for example'.

(1) Taroo-wa amerika-ya furansu-ni itta.
Taro-TOP USA-YA France-to went
'Taro went to the USA and France, for example.'

Sudo 2014 proposes that the lexical meaning of *ya* is simply an inclusive disjunction but it is strengthened by two scalar inferences (SIs): (i) a conjunctive inference and (ii) an anti-exhaustive inference. Concretely, according to Sudo 2014, the semantics of (1) is as in (2), but it also involves two SIs, (2-a) and (2-b).

(2) Taro went to the USA $\vee$ Taro went to France
a. $\rightsquigarrow$ Taro went to both the USA and France. CONJUNCTIVE
b. $\rightsquigarrow$ Taro went to at least one more country. ANTI-EXHAUSTIVE

According to this idea, the underlying semantics of *ya* is disjunctive, despite the fact that the overall meaning entails conjunction, e.g., (1) entails that Taro went to both the USA and France. The strongest evidence for this comes from *ya* in quantified sentences like (3), which does not entail that every friend went to both countries, but only that every friend went to at least one of them (cf. Smith and Kobayashi 2018).

(3) dono tomodachi-mo amerika-ya furansu-ni itta.
which friend-MO USA-YA France-to went
'Every friend went to USA or France, for example.'

Importantly, the conjunctive counterpart of (3) formed with the conjunctive connective *to* in place of *ya*, as in (4), does entail that every friend went to both countries. This suggests that the universal quantifier *dono NP* is read distributively and the relevant reading of (3) is not a cumulative reading with respect to a conjunctive object.

(4) dono tomodachi-mo amerika-to furansu-ni itta.
which friend-MO USA-YA France-to went
'Every friend went to the USA and France, for example.'

There are two competing proposals about how the two SIs of *ya* are derived: Sudo 2014 analyzes both of them as (higher-order) SIs that arise in competition with the conjunctive and disjunctive alternatives. Sauerland et al. 2017, on the other hand, focus on the conjunctive inference and derive it with domain alternatives (cf. Bowler 2014; Singh et al. 2016).

In this paper, we report on an experiment with an inference task that compared the two SIs of *ya* between them as well as with the exhaustive inference of the conjunctive connective *to* and the anti-conjunctive, alternatively, the exclusive inference of the disjunctive connective *ka*. The latter two inferences are illustrated by (5) and (6), respectively.

(5) Taroo-wa amerika-to furansu-ni itta.
Taro-TOP USA-TO France-to went
'Taro went to the USA and France.'
⇝ Taro did not go to other countries EXHAUSTIVE

(6) Taroo-wa amerika-ka furansu-ni itta.
Taro-TOP USA-KA France-to went
'Taro went to the USA or France.'
⇝ Taro did not go to both countries ANTI-CONJUNCTIVE

The results indicate that *ya* does trigger an anti-exhaustive inference but it is not as robust as the other three inferences we tested, which were all very robust. We argue that these experimental findings are problematic for Sudo's (2014) analysis, which derives the two SIs of *ya* by the same mechanism, whereby failing to capture their difference in robustness. The results also show that Sauerland et al.'s (2017) analysis needs to be augmented with a mechanism for generating the anti-exhaustive inference. To explain the difference in robustness, we propose a hybrid account where the conjunctive inference of *ya* is a *grammatically* derived SI, as proposed by these previous studies, but its anti-exhaustive inference is a *pragmatically* derived SI (see Marty et al. 2021; Marty et al. 2024c for similar ideas).

2 Experiment

The goals of our experiment were to show that *ya* has an anti-exhaustive inference and to compare its robustness with its conjunctive inference, as well as with other inferences that are related to these two inferences, namely the exhaustive inference of the cojunctive connective *to* and the anti-conjuctive inference of the disjunctive connective *ka* (The latter also typically triggers an exhaustive inference but we did not test it).

2.1 Material

Our experiment employed a version of the inference task. The inference task has not only been extensively used to investigate various types of SIs, but also is known to yield more robust results than other experimental tasks (Geurts and Pouscoulous 2009; Ronai and Xiang 2021; Marty et al. 2024b), therefore making it the ideal task for purposes of showing the existence of an SI.

In order to further boost the robustness of the target inferences, following Ronai and Xiang's (2021) findings, we included a *wh*-question in each trial and presented the test sentence as an answer to it, as illustrated by the following example stimulus. (Note that nouns in Japanese are generally number-neutral.)

(7) Shota: Takuya-wa donna petto-o katteru no?
Takuya-TOP what.kind pet-ACC have Q
'Shota: What kind of pet does Takuya have?'

Misaki: neko-ya kame-da yo.
cat-YA turtle-COP PRT
'Misaki: A cat(s) and a turtle(s)'.

In addition to the existence of the anti-exhaustive inference, we are also interested in its quality. Intuitively, the anti-exhaustive inference of *ya* feels weaker than the conjunctive inference, as previously suggested by Sudo 2014, fn. 2, and we hypothesized that this is because the anti-exhaustive inference is by default perceived as a primary implicature, rather than a secondary implicature, in the sense of Sauerland 2004. For example, in the case of (1), the anti-exhaustive inference is '$\neg B_s$(Taro went to only USA and France)', rather than 'B_s(¬Taro went to only USA and France)'. In order to test this hypothesis, we prepared two types of probe questions, PRIMARY and SECONDARY. For (7), for example, the English translations of the two probe questions are as in (8). (We used the epistemic possibility modal *kamoshirenai* to phrase the PRIMARY question.)

(8) a. Can you conclude the following from this?: Misaki thinks that Takuya might have a pet(s) different from a cat(s) or a turtle(s). PRIMARY
b. Can you conclude the following from this?: Misaki thinks that Takuya has a pet(s) different from a cat(s) or a turtle(s). SECONDARY

Both questions were answered by pressing a button labeled *Hai* 'Yes' or *Iie* 'No'.

One caveat is in order: if an inference is perceived as a primary implicature, the PRIMARY question should be accepted, while the SECONDARY question rejected. If an inference is perceived as a secondary implicature, the SECONDARY question should be accepted, but whether the PRIMARY question will be rejected depends on whether the probe question itself is read with an implicature implying ignorance.

2.2 Procedure

The experiment consisted of 4 types of INFERENCE (ANTI-EXHAUSTIVE, CONJUNCTIVE, EXHAUSTIVE, ANTI-CONJUNCTIVE) × 2 types of PROBE (PRIMARY, SECONDARY) = 8 experimental conditions. Both factors were tested within participants. Each experimental condition had 4 items, meaning there were 32 target items in total. They were presented in a random order together with 32 filler items on Gorilla Experiment Builder (`www.gorilla.sc`; Anwyl-Irvine et al. 2020). The experiment was pre-registered on the Open Science Framework at `https://osf.io/mtu7z`.

We employed 39 self-claimed Japanese native speakers on CrowdWorks (`crowdworks.jp`) (see Majima et al. 2017 for comparisons with other services for the purposes of behavioral research). They were all paid 350–385 yen, including the platform fee. Two of them were excluded from analysis as they failed to reach a pre-established threshold accuracy of 80% for filler items (Their accuracy rates were 56.2% and 68.8%, suggesting that their answers were largely random). For the remaining 37 participants, the mean accuracy rate was 93.3%.

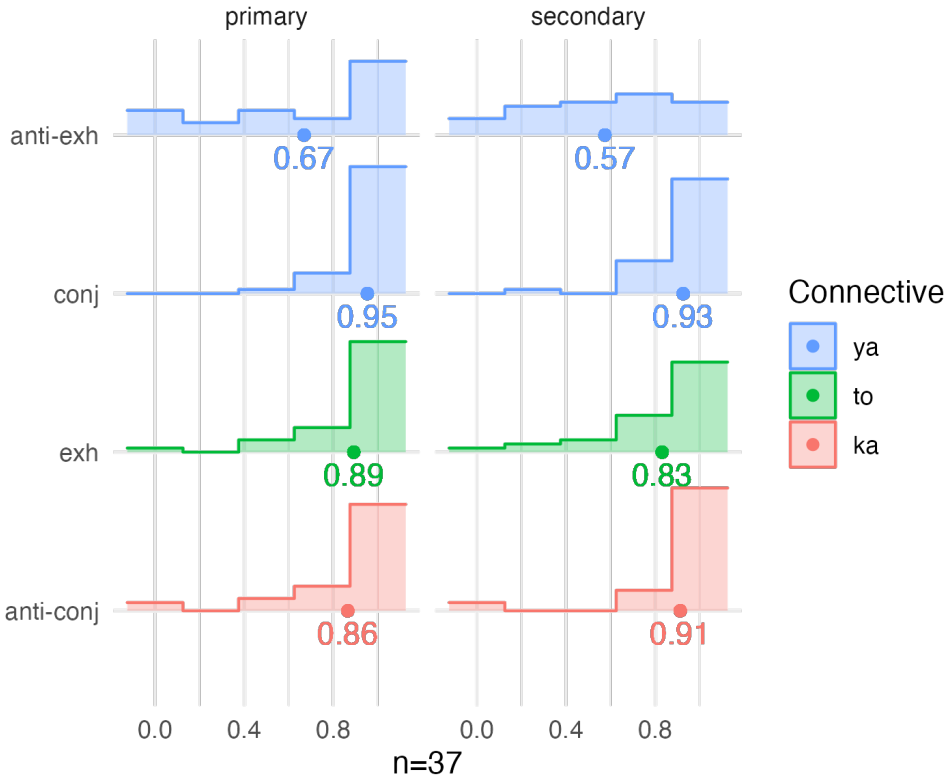


Figure 1: Proportions of *Yes* answers in experimental conditions. The histograms represent by-participant proportions across 4 items in the respective experimental conditions.

2.3 Results

The results are summarized in Figure 1. We fitted a mixed-effects logistic regression model using the `lme4` package (Bates et al. 2015) for `R` (R Core Team 2024) with INFERENCE and PROBE as fixed effect factors, with both factors treatment coded with ANTI-EXHAUSTIVITE and PRIMARY as reference levels, respectively. The model revealed a significant simple effect of INFERENCE with respect to each non-reference level, a significant simple effect of PROBE, and a significant interaction with respect to the anti-conjunctive inference of *ka* (see Table 1 for details).

2.4 Discussion

The experimental results show that the anti-exhaustive inference of *ya* is much less robust than the other three inference types, and is characterized more aptly as a primary implicature. The significant interaction effect with respect to the anti-conjunctive inference of *ka* suggests that the latter is more likely to be perceived as a secondary implicature, while the non-significant interaction effects with the other inference types could well be due to the ceiling effect.

We take these results as evidence for the anti-exhautive inference, because were it not for it, the simple effect of PROBE and the significant interaction effect would not be explained. This in turn means that Sauerland et al.'s (2017) account is incomplete, as it is entirely silent about this inference.

At the same time, the low robustness of the anti-exhaustive inference poses a serious challenge for Sudo's (2014) account, according to which it is an SI of the same kind as the other three inferences tested in our experiment. In particular, it proposes to derive the anti-exhaustive and conjunctive inferences of *ya* with a single operator.

Fixed effects	β	s.e.	z	p
Intercept	1.22	0.47	2.57	0.010*
$\text{INFERENCE}_{\text{CONJ}}$	3.41	1.13	3.01	0.003**
$\text{INFERENCE}_{\text{EXH}}$	1.89	0.77	2.44	0.014*
$\text{INFERENCE}_{\text{ANTICONJ}}$	4.39	1.87	2.35	0.019*
PROBE	−0.74	0.34	−2.19	0.028*
$\text{INFERENCE}_{\text{CONJ}}\times\text{PROBE}$	0.20	0.68	0.30	0.767
$\text{INFERENCE}_{\text{EXH}}\times\text{PROBE}$	0.07	0.53	0.13	0.900
$\text{INFERENCE}_{\text{ANTICONJ}}\times\text{PROBE}$	1.83	0.73	2.51	0.012*

Random effects	σ^2	σ	ρ		
By-participant					
Intercept	4.25	2.06			
$\text{INFERENCE}_{\text{CONJ}}$	6.57	2.56	−0.77		
$\text{INFERENCE}_{\text{EXH}}$	9.15	3.03	−0.84	0.80	
$\text{INFERENCE}_{\text{ANTICONJ}}$	24.29	4.93	−0.29	0.19	0.41
By-item					
α	1.07	1.03			
$\text{INFERENCE}_{\text{CONJ}}$	2.12	1.46	−0.46		
$\text{INFERENCE}_{\text{EXH}}$	1.86	1.36	−0.97	0.63	
$\text{INFERENCE}_{\text{ANTICONJ}}$	1.20	1.10	−0.63	−0.04	0.59

Table 1: Parameter estimates of the mixed-effects logistic regression model.

While the experimental design has some room for improvement, e.g., the effect of *Probe* came out as relatively small, overall, it is a reasonable conclusion that a novel account of the anti-exhaustive inference is called for. To this end, we will sketch a novel 'hybrid' account in what follows.

3 A hybrid account

Firstly, it should be noticed that our experimental results are compatible with both Sudo's (2014) and Sauerland et al.'s (2017) account of the conjunctive inference of *ya* as a grammatically derived SI. We therefore remain neutral as to which analysis of the conjunctive inference is on the right track, but we crucially follow them in assuming that the conjunctive inference is a grammatically derived SI.

Turning now to the anti-exhaustive inference, as we discussed above, we reject Sudo's (2014) idea that the anti-exhaustive inference is also an SI of the same kind. Instead, we propose that it is a *pragmatically* derived SI.

Marty et al. (2021, 2024) recently suggested a similar hybrid account for SIs generated by *or* vs. negated *and*. In their experiments with truth-value judgment tasks, they observed that the SIs of sentences with *or* occurring in various modal constructions are generally much more robust than those of sentences containing negated *and* instead, despite the fact that the contents of the sentences are quite similar, if the SIs are taken into consideration. To account for this, they suggest that *or* generates a scalar inference with some grammatical mechanism (e.g., EXH), while a scalar inference for negated *and* is due to pragmatic reasoning in the spirit of (neo-)Gricean Pragmatics. If pragmatic reasoning is contingent on various contextual factors, as is commonly assumed, it is not surprising that pragmatically derived SIs come out not so

robustly in experimental settings. Assuming that grammatically derived scalar inferences are more reliably present (although in principle also optional), we can account for the observed difference in robustness.

Applying this idea to the two scalar inferences of *ya*, we analyze its conjunctive inference as a grammatically derived SI, and its anti-exhaustive inference as a pragmatically derived SI. This analysis not only accounts for the difference in robustness in the same way as Marty et al.'s account of *or* vs. negated *and*, but it also naturally follows from it that the anti-exhaustive inference is more likely to be perceived as a primary implicature, on the assumption that pragmatically derived SIs are born as primary implicatures and only optionally strengthen to secondary implicatures with further reasoning (see, e.g., Gazdar 1979; Soames 1982; Sauerland 2004).

In order for this hybrid account to explain the full set of observations from our experiment, some auxiliary assumptions are necessary. Firstly, it is crucial that the conjunctive inference is analyzed as a *grammatically* derived SI. Secondly, we have to assume that the exhaustive inference of *to* and the anti-conjunctive inference of *ka* are likewise both grammatically derived SIs, as they were both much more robust than the anti-exhaustive inference of *ya* in our results. Thirdly, the pragmatic mechanism that derives the anti-exhaustive inference should (i) reason about an alternative expression which has an exhaustive inference and (ii) derive the inference that the speaker couldn't utter it while respecting the conversational maxims. This crucial alternative can be the conjunctive alternative with its exhaustive inference (which is a grammatically derived SI in our account) for *ya* in positive contexts and the disjunctive alternative with its exhaustive inference for *ya* in negative contexts. This account of the anti-exhaustive inference is very similar to Sudo's (2014), except that the core mechanism that 'negates' the alternative is pragmatic reasoning, rather than a grammatical operator.

4 Outlook

While variation in robustness among different scalar expressions ('scalar diversity') has recently garnered a lot of attention in the experimental literature (e.g., Tiel et al. 2016; Ronai and Xiang 2021; Marty et al. 2024a; Nicolae et al. 2024a; Nicolae et al. 2024b), the theoretical understanding of the variation is arguably still rudimentary. We certainly do not think everything about scalar diversity, which is a continuum, is to be accounted for in terms of the dichotomy of grammatical vs. pragmatic inferences, but if we (and Marty et al. 2021; Marty et al. 2024c) are right, at least part of the diversity should be.

Furthermore, our hybrid account makes experimentally testable predictions that go beyond robustness alone, e.g., that pragmatic inferences are more likely to be perceived as primary implicatures and are more sensitive to the speaker's epistemic state, etc., which are left for future research.

Acknowledgements. A.N. received support from the DFG grant NI-1850/2-1 and the ERC Synergy Grant 856421. The work has also been partly supported by JSPS Core-to-Core Program (#JPJSCCA20210001; PI: Yoichi Miyamoto).

References

Anwyl-Irvine, Alexander L. et al. (2020). "Gorilla in our midst: An online behavioral experiment builder". In: *Behavior Research Methods* 52, pp. 388–407. DOI: 10.3758/s13428-019-01237-x.

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Bowler, Margit (2014). "Conjunction and disjunction in a language without 'and'". In: *Proceedings of SALT 24*. Ed. by Todd Snider, Sarah D'Antonio, and Mia Weigand. LSA and CLC, pp. 137–155. DOI: 10.3765/salt.v24i0.2422.

Gazdar, Gerald (1979). *Pragmatics: Implicature, Presupposition, and Logical Form*. Academic Press.

Geurts, Bart and Nausicaa Pouscoulous (2009). "Embedded Implicatures?!?" In: *Semantics and Pragmatics* 2.4, pp. 1–34. DOI: 10.3765/sp.2.4.

Kuno, Susumu (1973). *The Structure of the Japanese Language*. Cambridge, MA: MIT Press.

Majima, Yoshimasa et al. (2017). "Conducting online behavioral research using crowdsourcing services in Japan". In: *Frontiers in Psychology* 8.378. DOI: 10.3389/fpsyg.2017.00378.

Marty, Paul et al. (2021). "Negative Free Choice". In: *Semantics and Pragmatics* 14.13, pp. 1–54. DOI: 10.3765/sp.14.13.

— (2024a). "On the salience of linguistic alternatives in the inference task for scalar implicature". Talk at Experiments in Linguistic Meaning (ELM) 3, University of Pennsylvania.

— (2024b). "Probing the probe: why inference tasks may inflate response rates for Scalar Implicature". Ms.

— (2024c). "What makes linguistic inferences robust?" In: *Journal of Semantics* 41.1, pp. 1–52. DOI: 10.1093/jos/ffad010.

Nicolae, Andreea C. et al. (2024a). "Do languages have exclusive disjunctions?" In: *Open Mind*.

— (2024b). "Exclusivity and exhaustivity of disjunction(s): A cross-linguistic study". To appear in *Sinn und Bedeutung (SuB) 28*.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Ronai, Eszter and Ming Xiang (2021). "Exploring the connection between Question Under Discussion and scalar diversity". In: *Proceedings of the Linguistic Society of America* 6.1, pp. 649–662. DOI: 10.3765/plsa.v6i1.5001.

Sauerland, Uli (2004). "Scalar implicatures in complex sentences". In: *Linguistics and Philosophy* 27.3, pp. 367–391. DOI: 10.1023/B:LING.0000023378.71748.db.

Sauerland, Uli et al. (2017). "Tracking down disjunction". In: *New Frontiers in Artificial Intelligence: JSAI-isAI 2015 Workshops, LENLS, JURISIN, AAA, HAT-MASH, TSDAA, ASD-HR, and SKL, Kanagawa, Japan, November 16–18, 2015, revised selected papers*. Ed. by Mihoko Otake et al. Springer, pp. 109–121. DOI: 10.1007/978-3-319-50953-29.

Singh, Raj et al. (2016). "Children interpret disjunction as conjunction: Consequences for theories of implicature and child development". In: *Natural Language Semantics* 24, pp. 305–352. DOI: 10.1007/s11050-016-9126-3.

Smith, Ryan Walter and Ryoichiro Kobayashi (2018). "Alternating conj/disjunctions: the case of Japanese *-toka* and *tari*". In: *Proceedings of Sinn und Bedeutung 22*. Ed. by Uli Sauerland and Stephanie Solt. ZAS. Berlin: ZASPiL, pp. 393–406.

Soames, Scott (1982). "How presuppositions are inheritied: A solution to the projection problem". In: *Linguistic Inquiry* 13.3, pp. 483–545.

Sudo, Yasutada (2014). "An anti-exhaustive, polarity sensitive connective in Japanese and higher-order scalar implicatures". Talk at TEAL-9.

Tiel, Bob van et al. (2016). "Scalar diversity". In: *Journal of Semantics* 33.1, pp. 137–175. DOI: 10.1093/jos/ffu017.

Generic Explanations for Generics

Tom Ralston
St John's College, University of Oxford
Oxford, United Kingdom
tomgjralston@gmail.com

Abstract

It is widely accepted that generics express law-like generalisations, on the basis of evidence that their truth does not depend only on how things are in the actual world at one moment in time. In this paper, I propose a temporal truthmaker semantics for generics and argue that this theory can account for the intuition that generics express law-like generalisations. The law-likeness of generics therefore does not motivate the introduction of either the unarticulated constituent 'GEN' or a domain of abstract kinds. Instead, it can be explained by the nature of the states which make generics true, which may be either generic facts about a category, or particular facts about the instances of a category.

1 Introduction

While there are many points of disagreement regarding the meaning of generic sentences, one article of faith is that they express nomic, law-like or non-accidental generalisations. According to this orthodoxy, generics say not only how things are, but how things would be even if things were different in many respects, holding fixed some set of rules or laws.

(1) American Presidents are male.

If generics simply 'stated the facts', we should expect (1) to be true, since at the time of writing every actual American President has been male. Yet (1) is at best misleading and at worst false, since it suggests that there is some rule or requirement in force that states that in order to be an American President, a person must be male. The dominant attitude in the literature is that generics have a modal component to their semantics, usually contributed by the phonologically null operator, 'GEN' (e.g. Krifka et al., 1995; Nickel, 2016; Kirkpatrick, 2023). Simplifying a little, on this view, the logical form of (1) is:

(2) GEN(x) : American-President(x)[male(x)]

The semantic contribution of 'GEN' is something like a modal quantifier over objects in possible worlds or situations. Since these situations may be non-actual, the truth-conditions of generics do not depend only on what is the case in the actual world. The main alternative is the kind-reference theory (e.g. Carlson, 1977; Liebesman, 2011; Teichman, 2023); according to the simplest versions, the logical form of (1) is:

(3) male(american-president-kind)

On this view, generically interpreted bare plural noun phrases (henceforth generic BPs), are referential and refer to abstract objects: kinds. The intuition that generics are nomic is accounted for in what it is for a kind to have a property: kinds are more than just their actual members.

My objective in this paper is not to show that modal and kind-reference theories of generics are not viable. There are, however, reasons to be dissatisfied with the dominant modal theories which attribute a significant component of the meaning of generics to a constituent, 'GEN', which is not articulated in any known language. Instead, my aim is programmatic: to show that a viable alternative to both theories exists, which can explain why generics are nomic. To this end, I will give a truthmaker semantics for generics, which treats generic BPs as referential expressions whose reference extends across time, but not across possible worlds or situations.

2 Preliminaries

2.1 Nomicity

I begin with some brief remarks on what I take nomicity to be. I will use 'nomic' as a synonym for 'non-accidental' and 'law-like' (I choose the label for reasons of style alone). To begin with, we can distinguish between nomic *sentences* and nomic *regularities*. The first category is linguistic and includes sentences which are made true not only by how the world happens to be, but how things would be if things in the world were different, holding fixed some set of rules of laws.

Nomic regularities are not linguistic: they are regularities in the world. To label a regularity as nomic is not to say that it reflects a law of nature; laws of nature, if there are any, apply everywhere with no exceptions, yet many nomic regularities (especially those which are associated with generics) admit exceptions. I don't want to take on any metaphysical commitments about the nature of nomic regularities. My objective is to explain why generics are nomic in the linguistic sense. For now I will just note two characteristics of nomic regularities which will become relevant in the course of this paper:

Counterfactuals They support counterfactuals (Goodman, 1983). Suppose there is a law which underpins the regularity that all copper conducts electricity. Then, if I were to connect this particular strip of copper wire to a power source in the right way, it would conduct electricity. That is true, even if I never in fact do such a thing.

Unobserved Instances They support inferences about unobserved instances. Nomic regularities may be projected on to instances of a category which have not yet been observed, including those in the future (Jackson and Pargetter, 1980).

2.2 Generics and temporary generalisations

There is good evidence that generics are not about only the way things are at a particular moment in time. For instance, (4) would not be true even if every black raven died, leaving behind only white ravens (Nickel, 2016, p. 16).

(4) Ravens are white.

This is also nicely illustrated by Cohen's famous example:

(5) (a) *Situation:* By chance, every current Supreme Court judge has a prime Social Security number.
(b) Supreme Court judges have a prime Social Security number (Cohen, 1999, p. 35).

As Cohen argues, the sentence sounds false in this situation, since it is insufficient that every Supreme Court judge just happens to have a prime Social Security number for it to be true. More generally, if we construct contexts in which a generalisation holds of all instances of a category but only at a particular moment in time, a generic will sound false in that context.

This rules out a very simple referential theory according to which generic BPs refer to actual objects at the reference time of an utterance. But it does not rule out a temporal referential theory according to which generic BPs refer to actual objects at more than one time. This would explain, for example, why (4) is false: the BP refers not only to the ravens that there are now, but the ravens that there have been and will be, so the temporary fact that all *current* ravens are white does not suffice to make (4). Similarly, (5b) sounds false in Cohen's context because we assume that Supreme Court judges in the past and future have non-prime Social Security numbers. In what follows, I will expand on this initial sketch of a theory.

2.3 Branching time

I will assume a branching metaphysics of time (Thomason, 1970). On this picture, the past is fixed, but the future is open. Call a pair of times and an order on times, $\langle T, \prec \rangle$, a temporal frame. $\prec$, the precedence relation, has the following three properties:

Transitive $\forall t, t', t'' \in T(t \prec t' \wedge t' \prec t'' \rightarrow t \prec t'')$

Irreflexive $\forall t \in T(\neg t \prec t)$

Forward Branching $\forall t, t', t'' \in T(t' \prec t \wedge t'' \prec t \rightarrow (t' \neq t'' \rightarrow (t' \prec t'' \vee t'' \prec t')))$

The ordering is linear with respect to the past, but open with respect to the future, so for any time t it has at most one past but may have more than one future. A *history* is a maximal subset of T that is linearly ordered by $\prec$, i.e. it is a complete branch which cannot be extended (note that a history extends into the future).

This metaphysical picture raises the question: what is it for a sentence about the future to be true? If time branches into the future, there is no one history which contains *the* future, so we can't simply say that for a sentence about the future to be true at a time t is for there to be some time t' such that $t \prec t'$ at which the sentence is true, since there may be many other future times at which it is false. Following the terminology of Kaufmann et al. (2006), I will call a sentence *settled* at time t iff it is true in every history which includes t.

3 Truthmaker semantics for generics

I now turn to the semantics of generics. The proposal is that generics can be *made true* in different ways. In truthmaker semantics, propositions are identified with sets of states which make a sentence true (or *verify* it) (Fine, 2017). To a temporal frame, add a domain of states to be a set S which is ordered with respect to a partial ordering, $\sqsubseteq$, and a corresponding fusion operation, $\sqcup$, such that $a \sqsubseteq b \leftrightarrow a \sqcup b = b$: $\langle T, \prec, S, \sqsubseteq \rangle$. In a truthmaker semantic framework, whether a sentence is true depends on what states exist. The notion of verification ($\Vdash$) is primitive and holds between states and sentences. The idea that I will pursue here is that for a state s to verify a sentence ϕ is for there to be no need to look beyond the information contained in s to explain ϕ (Linnebo, 2022, p. 357).[1] This notion of verification naturally includes the relation of settling: if a state settles how things will be in the future, then it contains enough information to explain why things will be that way.

3.1 Generic and instance-based explanations

Fine informally states his semantics for the universal quantifier as follows: "A state verifies $\forall x \phi(x)$ if it is the fusion of verifiers of its instances $\phi(a_1), \phi(a_2), \ldots$" (Fine, 2017, p. 567). But when the domain of the universal quantifier is indeterminate, a truthmaker of this sort will not generally be available. This would be the case when the domain of the quantifier extends into the future, which I have assumed to be indeterminate. So how can a universal generalisation be true when it is not determinate which of its instances exist? Linnebo (2022) addresses this question by arguing that the truth of a universal generalisation can in principle be explained either by the nature of its instances or the nature of the categories that it describes (see also Fine, 2017, pp. 568–569). Linnebo distinguishes between two ways that a sentence might be made true:

[1] Note that here, unlike in Fine's truthmaker semantics, verification is monotonic, i.e. if $s \Vdash \phi$ and $s \sqsubseteq t$, $t \Vdash \phi$.

> An explanation is **generic** insofar as it does not proceed via individual instances but is based on general facts about the properties or operations involved in the claim that is generalized (Linnebo, 2022, p. 349).

An explanation is **instance-based** to the extent that it is not generic. Recall that we are considering *metaphysical* explanations here, so the question is: what currently existing state settles a sentence of the form ⌜All Fs are G⌝ when it is not determinate what things are Fs? No instance-based explanation will be possible in such a situation, but a generic explanation will be, provided that it contains enough information about what it is to be an F that it guarantees that any F is G.

Carrying over this distinction to the meaning of generics like (1), the problem is that it is not yet determinate who will be the President in the future. So no instance-based explanation will be possible. For there to be a generic explanation for (1), there would need to be a state which now settles for any given American President, their sex. If, for example, there were a rule in force which stated that the President must be male, then the existence of that rule would settle (1). That state would not itself need to settle who becomes the President in the future, provided that it settles whether they are male or female. But since there is no such state, (1) is not true.

3.2 Intrinsic truthmaker semantics

To make this intuition precise, I will augment Linnebo's truthmaker semantics for first-order sentences, which are reproduced in full below. Since I am introducing a temporal dimension to the semantics, I will index predicates to times; if P is a predicate, then $P_{\mathcal{T}}$ is a predicate true of everything which is P at some time in a set of times $\mathcal{T}$. $\mathbb{1}$ is the inconsistent state; $\bot$ is a sentential constant which is false on every interpretation; where P is an n-place predicate, $[\![P]\!]_s$ is the set of n-tuples which satisfy P in a state s; $D(s)$ is the domain of objects whose existence is verified by a state s.

$(\Vdash \mathbb{1})$	$\mathbb{1} \Vdash \phi$	if	$s \Vdash \bot$, then $s = \mathbb{1}$
$(\Vdash \wedge)$	$s \Vdash \phi \wedge \psi$	iff	$s \Vdash \phi$ and $s \Vdash \psi$
$(\Vdash \vee)$	$s \Vdash \phi \vee \psi$	iff	$s \Vdash \phi$ or $s \Vdash \psi$
$(\Vdash \rightarrow)$	$s \Vdash \phi \rightarrow \psi$	iff	for every s', if $s' \Vdash \phi$, then $s \sqcup s' \Vdash \psi$
$(\Vdash \neg)$	$s \Vdash \neg\phi$	iff	for every s', if $s' \Vdash \phi$ then $s \sqcup s' \Vdash \bot$
$(\Vdash At)$	$s \Vdash P(a_1, \ldots, a_n)$	iff	$\langle a_1, \ldots a_n \rangle \in [\![P]\!]_s$
$(\Vdash \exists)$	$s \Vdash \exists x \phi(x)$	iff	$s \Vdash \phi(a)$ for some $a \in D(s)$
$(\Vdash \forall)$	$s \Vdash \forall x \phi(x)$	iff	$s \sqcup s' \Vdash \phi(a)$ for each s' and each $a \in D(s')$

The intuition behind $(\Vdash \forall)$ is that a state can verify a universal generalisation without verifying the existence of each instance, provided that the state would verify the generalisation when fused with states which verify each instance. Extending this intuition to generics, I will use '$\sigma x {}^*P(x)$' to translate the bare plural 'Ps', where ${}^*P$ is the algebraic closure of P; $\leq$ is the mereological parthood relation; '$\sigma x P(x)$' denotes the least upper bound of P (Link, 1983). It is a term which denotes the things each of which satisfies P. Then we have the following clause:

$(\Vdash \sigma)$ $\quad s \Vdash Q(\sigma x {}^*P(x)) \quad$ iff $\quad$ $s \sqcup s' \Vdash Q(a)$ for every s' and every a such that for each $x \in D(s'), x \leq a$ and $s' \Vdash {}^*P(a)$.

The intuition behind $(\Vdash \sigma)$ is that for a state to verify $Q(\sigma x {}^*P(x))$, it is sufficient that, given any plurality of objects which are Ps, it verifies that those objects are Q. Since I claim that generic BPs refer to objects which satisfy a description at more than one time, their representation must be indexed to some set of times. Taking the temporal index of 'whale' to be the set of all times T, the translation of (6a) will be as follows:

(6) (a) Whales are mammals.
(b) ${}^*\text{mammal}(\sigma x {}^*\text{whale}_T(x))$

Since T includes times from more than one possible future, there is no fact of the matter which plurality of whales the σ-term refers to. Despite this indeterminacy, (6b) may still be verified by a state s, however, if any way of adding whales to s results in a state which verifies that those whales are mammals. This is what enables a generic to have a generic explanation: a state may verify a generic without proceeding via facts about individuals, but rather facts about a category in general. Let's now turn to a true generic which admits exceptions:

(7) (a) Ravens are black.
(b) $^*\text{black}(\sigma x^*\text{raven}_T(x))$

Since 'are black' is a distributive predicate, a purely instance-based explanation of this generic would be the fusion of all of the states which each verify sentences of the form $\ulcorner\text{black}(a)\urcorner$ where a is a raven. But, as I argued, that will need to include verifiers for sentences about not only the current ravens in the world, but the plurality of things which are ravens at any time in T. If those states don't yet exist, or it is not determinate which of those states will exist, an instance-based explanation of (7a) won't be available. So we should look for a verifier which is, to at least some extent, generic. The catch here, of course, is that not all ravens are black, so it looks like the clause in ($\Vdash \sigma$) will be too strong, since it requires that any state which verifies that somethings are ravens also verifies that those things are black.

I will follow Schwarzschild (1994) and Križ and Spector (2021), among many others, who propose that plural predications (a category which I claim includes BP generics) may be underspecified with respect to the proportion of the argument which must satisfy the predicate for the sentence to be true. Adding 'all' to the sentence eliminates the underspecification. On this view, there is no fact of the matter as to how many ravens must be black in order for (7a) to be true. But the context of utterance and the purposes of conversation can help to narrow down the intended meaning of a sentence, by filtering out unintended interpretations. On Križ and Spector's theory, this is because the meaning of the plural predicate itself is underspecified. So even for distributive predicates such as 'are black', it is not the case that if $a \in [\![\text{black}]\!]_s$, then for every $x \leq a$, $x \in [\![\text{black}]\!]_s$. That is, it can be true of a plurality a that they are black, even though it is not the case that for each thing among a, it is black.

Returning to the question of the nature of the verifiers of (7a), a general fact about the genotype of ravens and its relation to the phenotype might verify this sentence, without containing any information about particular ravens. And it may still be the case that a state satisfies ($\Vdash \sigma$) with respect to (7a), even though not every individual raven is black in every state which is fused with that state is black.

4 Explicit temporal restrictions

In the previous two examples, I assumed that the temporal index of the predicate was the set of all times, T. But that isn't always the case, since generics can be explicitly restricted by temporal adverbials:

(8) (a) Superpowers are referred to by abbreviations the first letter of which is 'U' (Dahl, 1975).
(b) In the 1970s, superpowers were referred to by abbreviations the first letter of which is 'U'.

As Dahl observed of (8a), the temporally unrestricted generic suggests that being referred to by an abbreviation the first letter of which is 'U' was a prerequisite for being a superpower (Dahl, 1975, p. 109). Yet the temporally restricted (8b) has a much more prominent 'accidental' interpretation, according to which it was simply true of the superpowers in the 1970s that they were referred to by such an abbreviation. Clearly, an instance-based explanation is possible for (8b), since there were just two things which were superpowers during the 1970s and the reference

of the BP is limited to that period of time (i.e. the temporal index $\mathcal{T} = \{t : t$ occurred during the 1970s$\}$). The instance-based explanation would be the state $s \sqcup t$, where:

(i) $s \Vdash$ the United States of America was referred to by an abbreviation the first letter of which is 'U';

(ii) $t \Vdash$ the Union of Soviet Socialist Republics was referred to by an abbreviation the first letter of which is 'U'.

On the other hand, it is consistent with there being a completely instance-based explanation for (8b) that there is *also* a generic explanation for it, so (8b) can be interpreted either as a nomic or a descriptive generalisation. In contrast, (8a) has no such temporal restriction and hence no instance-based explanation is possible. When a speaker utters (8a), in order to interpret the sentence as true, a hearer must then assume that there is a (partially) generic explanation for (8a). It is difficult to imagine what a generic explanation for (8a) could be – hence its oddness – but if there were a generic explanation for (8a), that would support counterfactuals – if China were a superpower, it would be referred to by an abbreviation the first letter of which is 'U' – and inferences about unobserved instances, since a generic explanation would be a state which contains information about what it is to be a superpower, irrespective of the identity of the particular object which is a superpower.

5 Implicit temporal restrictions

Finally, purely descriptive or statistical interpretations of unmodified generics must be accounted for. For instance, we can imagine both a nomic and a descriptive interpretation of (9):

(9) Firefighters wear protective clothing.

On its nomic interpretation, (9) says that in order to be a firefighter, a person must wear protective clothing (in appropriate situations). On its descriptive interpretation, it says that actual firefighters (typically) wear protective clothing. On the present theory, an instance-based explanation corresponds to the descriptive interpretation and a generic explanation – e.g. a rule requiring firefighters to wear protective clothing – corresponds to the nomic interpretation. More speculatively, it seems that only generics which are plausibly supported by generic explanations are felicitous as indefinite singulars (see Leslie et al., 2009; Lawler, 1973):

(10) (a) A firefighter wears protective clothing.

(b) ?A superpower is referred to by an abbreviation the first letter of which is 'U'

For an instance-based explanation of (9) to be available, the temporal index on 'firefighters' must not include times in the future, otherwise it will be indeterminate which objects are firefighters. When the temporal index includes only the present and the past, descriptive interpretations are possible, even when there is no explicit temporal restriction. That is, descriptive generics are statements about how things have been up until now.

Given that the value of the temporal index need not be made explicit, this represents a source of confusion in the meaning of generics, since a speaker may intend to communicate a descriptive generalisation about past/present instances, yet be interpreted as communicating a generalisation which holds of future instances too, supported by a nomic generalisation. The possibility of this kind of miscommunication means that hearers may "slip" from descriptive interpretations of generics to nomic ones (Langton et al., 2012). For instance, I might conclude that (9) is true on the basis observing a large number of firefighters wearing protective clothing, yet when I utter (9), a hearer may assume that I possess knowledge of a generic explanation for (9), which can then be extended to new firefighters who I have not observed.

Acknowledgements. Thanks to Ian Rumfitt and Paul Elbourne for extensive comments.

References

Carlson, G. N. (1977). "Reference to Kinds in English". PhD thesis. Massachusetts, USA: University of Massachusetts Amherst.

Cohen, A. (1999). *Think Generic!: The Meaning and Use of Generic Sentences*. Stanford, California: Center for the Study of Language and Information.

Dahl, Ö. (1975). "On Generics". In: *Formal Semantics of Natural Language*. Ed. by E. L. Keenan. Cambridge: Cambridge University Press, pp. 99–111. DOI: 10.1017/CBO9780511897696.009.

Fine, K. (2017). "Truthmaker Semantics". In: *A Companion to the Philosophy of Language*. John Wiley & Sons, Ltd, pp. 556–577. DOI: 10.1002/9781118972090.ch22.

Goodman, N. (1983). *Fact, Fiction, & Forecast*. Fourth. London: Athlone Press.

Jackson, F. and Pargetter, R. (1980). "Confirmation and the Nomological". In: *Canadian Journal of Philosophy* 10.3, pp. 415–428. DOI: 10.1080/00455091.1980.10715734.

Kaufmann, S., Condoravdi, C., and Harizanov, V. (2006). "Formal Approaches to Modality". In: *The expression of modality* 1. Ed. by W. Frawley, pp. 71–105.

Kirkpatrick, J. R. (2023). "Generic Conjunctivitis". In: *Linguistics and Philosophy* 46, pp. 379–428. DOI: 10.1007/s10988-022-09371-0.

Krifka, M., Pelletier, F. J., Carlson, G. N., ter Meulen, A., Chierchia, G., and Link, G. (1995). "Genericity: An Introduction". In: *The Generic Book*. Ed. by G. N. Carlson and F. J. Pelletier. Chicago ; London: University of Chicago Press, pp. 1–124.

Križ, M. and Spector, B. (2021). "Interpreting Plural Predication: Homogeneity and Non-Maximality". In: *Linguistics and Philosophy* 44.5, pp. 1131–1178. DOI: 10.1007/s10988-020-09311-w.

Langton, R., Haslanger, S., and Anderson, L. (2012). "Language and Race". In: *Routledge Companion to Philosophy of Language*. Ed. by G. Russell and D. G. Fara. London: Taylor & Francis Group.

Lawler, J. M. (1973). "Studies in English Generics". PhD thesis. United States – Michigan: University of Michigan.

Leslie, S.-J., Khemlani, S., Prasada, S., and Glucksberg, S. (2009). "Conceptual and Linguistic Distinctions between Singular and Plural Generics". In: *Proceedings of the Annual Meeting of the Cognitive Science Society* 31.31.

Liebesman, D. (2011). "Simple Generics". In: *Noûs* 45.3, pp. 409–442. DOI: 10.1111/j.1468-0068.2010.00774.x.

Link, G. (1983). "The Logical Analysis of Plurals and Mass Terms: A Lattice-theoretical Approach". In: *Meaning, Use, and Interpretation of Language*. Ed. by R. Bäuerle, C. Schwarze, and A. von Stechow. De Gruyter, pp. 302–323. DOI: 10.1515/9783110852820.302.

Linnebo, Ø. (2022). "Generality Explained". In: *The Journal of Philosophy* 119.7, pp. 349–379. DOI: 10.5840/jphil2022119725.

Nickel, B. (2016). *Between Logic and the World: An Integrated Theory of Generics*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199640003.001.0001.

Schwarzschild, R. (1994). "Plurals, Presuppositions and the Sources of Distributivity". In: *Natural Language Semantics* 2.3, pp. 201–248. DOI: 10.1007/BF01256743.

Teichman, M. (2023). "The Sophisticated Kind Theory". In: *Inquiry* 66.9, pp. 1613–1654. DOI: 10.1080/0020174X.2016.1267407.

Thomason, R. H. (1970). "Indeterminist Time and Truth-Value Gaps". In: *Theoria* 36.3, pp. 264–281. DOI: 10.1111/j.1755-2567.1970.tb00427.x.

Dynamic Effects of Modalized Questions

Valentin D. Richard
ILLC, University of Amsterdam
Amsterdam, The Netherlands
Université de Lorraine, CNRS, Inria, LORIA, F-54000
Nancy, France
valentin.richard@loria.fr

Abstract

A dynamic analysis of modalized questions faces several difficulties. When an existential modality scopes over the wh-phrase, two kinds of phenomena are observed. First, the external staticity of the modal operator prevents the wh-phrase from being referred back to in indicative follow-up sentences. Anaphoric reference to the wh-phrase should yet be allowed through modal subordination. Second, exhaustivity and global presuppositions are weakened in modalized *which* questions, leading to mention-some readings and local presuppositions. We design a model capturing these two desiderata, based on dynamic inquisitive semantics and stack-based semantics for local contexts. Our model provides a uniform analysis of modalized and conditional questions using the Kratzerian theory of modality. It also correctly predicts discrepancies in presupposition projection between modalized and conditional questions.

1 Modalized Questions

1.1 Possibility Weakens Questions

Singular *which* questions have a uniqueness presupposition (Higginbotham and May 1981; Dayal 1996). For example, (1-b) carries the presupposition that a single letter is hidden in the actual world. Either the true answer is *R* (for FORM), or it is *A* (for FOAM).

(1) a. SITUATION: *Alice and Bob are playing a game. Alice wrote a mystery English word of 4 or 5 letters on the board, but her hand hides a part of it. The skeleton is FO__M.*
b. ALICE: Which letter is hidden here?
c. ALICE: Which letter could be hidden here?

Hirsch and Schwarz 2019 observed that this presupposition is weakened with an existential modality operator. Modalized question (1-c) does not presuppose that there exists a single letter that is actually hidden. But it still presupposes something, which could be paraphrased as: *For each possible answer, a single letter is hidden.* Roelofsen and Dotlačil 2023 call this a local presupposition.

Moreover, as opposed to (1-b), where only one true answer exists in every world, several true answers are possible for (1-c) in the actual world. We could answer (1-c) by saying *A*, *R*, or a distributive conjunction or disjunction *A and/or R*. Ultimately, question (1-c) requires some information about the content of the English lexicon, not about the actual mystery word. In particular, the hidden letters might actually be *RU* (for FORUM). But (1-c) presupposes that the English lexicon does not *only* have *RU* as possible completion.

Plural *which* questions are strongly exhaustive (Groenendijk and Stokhof 1984). For example, in (2-b), Mary requires the whole set of eczema patients.

(2) a. SITUATION: *Mary would like to experiment with two or three patients with eczema. She asks Ann, who knows the medical files of all patients with skin conditions in the hospital, a question.*

b. Which patients have eczema?
c. Which patients could I invite for my experiment?

However, in modalized question (2-c), Mary is open to a non-exhaustive set of eczema patients. Modalized plural *which* questions allow mention-some readings. Xiang 2016 observed this phenomenon, and Xiang and Cremers 2017 confirmed it through quantitative data.

We follow the analysis of Hirsch and Schwarz 2019 claiming that these phenomena are explained by the modality operator scoping over the *which* phrase.

1.2 Modal Subordination with Modalized Questions

In declarative sentences, a modal scoping over an indefinite blocks its anaphoric potential (Karttunen 1969). However, modally subordinated sentences have access to the embedded antecedent (Roberts 1989). We observe the same with modalized questions.

Under the reading of (3-b) where Hansel is not asking *What is the specific animal which might eat us?*, but where he focuses on imagining a scenario, response (3-c) feels off. The discourse referent u raised by *which*^{u} under the scope of *might* cannot be accessed in this indicative sentence. However, in sentence (3-d), u can be referred to by *it*$_u$ because of modal subordination.

(3) a. SITUATION: *Hansel and Gretel are walking in the woods. They saw scary shadows and heard noises. They believe that child-eating animals are hiding in these woods.*
b. HANSEL: Which^u animal might eat us? (*might* > *which* reading)
c. GRETEL: ?I don't know, but it$_u$ is very large.
d. GRETEL: I don't know, but it$_u$ would eat you first.

2 Previous works

Our goal is to define an existential modal operator in dynamic semantics that captures all these properties. It has (1.) to be externally static while (2.) still allowing modal subordination. It has (3.) to obviate the global uniqueness presupposition while still providing a local one. Finally, it has (4.) to be able to turn a mention-all question into a mention-some one.

Previous dynamic semantic models do not account for all four desired properties. The dynamic model of questions and modalities of Groenendijk 1998; Groenendijk, Stokhof, and Veltman 1996 satisfies (1.) without allowing (2.). van Rooij 1998 manages to encode modal subordination (2.), but the DPPL substrate (from van den Berg 1996) does not structurally provide external staticity (1.). Similarly, Brasoveanu's (2010) account of modal subordination fails to model in which worlds individual discourse referents are *not* accessible.

In appendix A, I briefly discuss another issue of previous definitions of modals in dynamic semantics, which is solved by using Kratzer 1991 theory.

2.1 Stack-based Treatment of Local Contexts

Implementing externally static modal operators prompts us to be able to retrieve modal local contexts somehow (Stalnaker 1981). Many authors resolve local context as anaphora, using world referents (Brasoveanu 2010; Hofmann 2019) or context referents (Frank 1997; Geurts 1999). Kibble 1998 and Asher and McCready 2007 claim that, unlike individual anaphora, modal local contexts are more restricted and have no real resolution ambiguity. Here, we follow their arguments and opt for a different system.

We adopt a stack-based semantics (Kaufmann 2000). Sentences are denoted by functions updating macro-contexts τ. A stack τ contains local contexts. The bottom element c_0 corresponds to the common ground. Modal operators push on top of τ a new local context, which can contain

more discourse referents than c_0. A follow-up modal sentence is evaluated in that topmost local context. Indicative sentences presuppose that the topmost element is c_0. We assume here that popping elements out of τ is performed by pragmatics (e.g. via discourse relations (Asher and Lascarides 2003)).

Isaacs and Rawlins 2008 use Kaufmann's (2000) system to model modal subordination of conditional questions within partition semantics (Groenendijk and Stokhof 1984). Hara and Sano 2017 revise their model to provide better predictions for conditional questions, using inquisitive semantics (Ciardelli, Roelofsen, and Theiler 2017; Ciardelli, Groenendijk, and Roelofsen 2018). We extend their approach by (i.) adding referents in the model, by upgrading to dynamic inquisitive semantics $\mathsf{Inq}_\mathsf{B}^\mathsf{D}$ (Dotlačil and Roelofsen 2019; Dotlačil and Roelofsen 2021); (ii.) defining the denotation of modals in $\mathsf{Inq}_\mathsf{B}^\mathsf{D}$ based on Kratzerian theory; and (iii.) specifying their presuppositions and actions on stacks.

The choice of $\mathsf{Inq}_\mathsf{B}^\mathsf{D}$ is motivated by another factor. Socolof, Hirsch, and Schwarz 2020 showed that exhaustivity and presupposition weakening also happen with disjunction. $\mathsf{Inq}_\mathsf{B}^\mathsf{D}$ was proven to be able to model desiderata (3.) and (4.) for disjunction (Roelofsen and Dotlačil 2023).

3 Modal Dynamic Inquisitive Semantics

3.1 Inquisitive Possibility

In Kratzerian theory, every world is equipped with a modal base (a set of classical propositions) and an ordering source. Due to the lack of space, we only focus on the modal base f here. For each world $w \in W$, call $\mu(w) = \bigcap f(w)$ the modal set at w. The intensional denotation of $\Diamond$ is (1), with R_μ being the accessibility relation associated with μ.

$$\begin{aligned} [\![\Diamond\varphi]\!]_{\mathsf{Int}} = &\quad \{w \in W \mid \mu(w) \cap [\![\varphi]\!]_{\mathsf{Int}} \neq \varnothing\} \quad = \mathsf{R}_\mu^{-1}([\![\varphi]\!]) \\ \text{where } \mathsf{R}_\mu^{-1}(s) := &\quad \{w \in W \mid \mu(w) \cap s \neq \varnothing\} \end{aligned} \tag{1}$$

The inquisitive existential modality (2) proposed by Ciardelli 2016, § 7.1 is sensitive to the inquisitiveness of its complement but produces non-inquisitive propositions. Yet, a modalized question is questioning. Therefore, we propose another definition (3). By inverting quantifiers $\forall$ and $\exists$, we let inquisitiveness project from φ to $\Diamond\varphi$. In this different modal inquisitive semantics (MInqB), a set of worlds s resolves $\Diamond\varphi$ iff s resolves some $\Diamond\psi$, where ψ is a proposition resolving φ.

$$[\![\Diamond\varphi]\!]_{\mathsf{InqBK}} = \{s \subseteq W \mid \forall w \in s.\ \exists s' \in [\![\varphi]\!]_{\mathsf{InqBK}}.\ \mu(w) \cap s' \neq \varnothing\} \tag{2}$$

$$[\![\Diamond\varphi]\!]_{\mathsf{MInqB}} := \{s \subseteq W \mid \exists s' \in [\![\varphi]\!]_{\mathsf{MInqB}}.\ \forall w \in s.\ \mu(w) \cap s' \neq \varnothing\} \tag{3}$$

$$= \{s \subseteq W \mid \exists s' \in [\![\varphi]\!]_{\mathsf{MInqB}}.\ s \subseteq \mathsf{R}_\mu^{-1}(s')\} \tag{4}$$

To illustrate this definition consider example (4).

(4) a. SITUATION: *Mary supervised an exam for students a and b. Mary and John know that exactly one student cheated. John knows that Mary saw one or both of them suspiciously looking at their pencil case during the exam. This counts as a cheat suspicion. He asks her about that.*

b. JOHN: Which student may have cheated?

Let's model this situation with four worlds. The subscript x of world w_x^z is the real cheater: a or b. The superscript z is the set of students Mary saw suspiciously looking at their pencil case: a, b or ab (i.e. a and b). World w_x^z is accessible from $w_{x'}^{z'}$ iff. $z = z'$. The semantics of (4-b) in this model is the issue on the right in Fig. 1.



Figure 1: Illustration of the semantics of (4-b) in MInqB. Arrows represent the accessibility relation R_μ.

The alternatives of the prejacent φ are projected to $\Diamond\varphi$ by mediation of R_μ^{-1}. If the accessibility relation R_μ is an equivalence relation, $\mathsf{R}_\mu^{-1}(s)$ contains s. Hence, output resolving proposition s is wider than prejacent resolving proposition s'. This widening is what weakens the question.[1]

3.2 Dynamic Inquisitive Possibility

In GSV, the dynamic semantics of Groenendijk, Stokhof, and Veltman 1996, information states s are sets of pairs $\langle w, g\rangle$ of a world w and an assignment function g. A Kratzerian externally static denotation of $\Diamond$ in GSV would be (5).[2] It restricts the input information states s to the possibilities whose world content is in relation R_μ with the world content of some possibility in $[\![\varphi]\!]_{\mathsf{MGSV}}$.

$$[\![\Diamond\varphi]\!]_{\mathsf{MGSV}} := s \mapsto \{\langle w, g\rangle \in s \mid \exists\langle w', g'\rangle \in [\![\varphi]\!]_{\mathsf{MGSV}}.\ w' \in \mu(w)\} \tag{5}$$

$$= s \mapsto \{\langle w, g\rangle \mid w \in \mathsf{R}_\mu^{-1}(\mathsf{WC}([\![\varphi]\!]_{\mathsf{MGSV}}))\} \tag{6}$$

where the world content of an information state is defined as $\mathsf{WC}(s) = \{w \mid \exists\langle w, g\rangle \in s\}$.

Putting all together, we obtain a denotation of $\Diamond$ in Modal Dynamic Inquisitive Semantics $\mathsf{MInq}_\mathsf{B}^\mathsf{D}$ as (7), where a context c is a downward-closed nonempty set of information states.[3]

$$\Diamond\mathcal{U} := c \mapsto \{s \in c \mid \exists s' \in \mathcal{U}(c).\ \forall\langle w, g\rangle \in s.\ \exists\langle w', g'\rangle \in s'.\ w' \in \mu(w)\} \tag{7}$$

$$= c \mapsto \{s \in c \mid \exists s' \in \mathcal{U}(c).\ \mathsf{WC}(s) \subseteq \mathsf{R}_\mu^{-1}(\mathsf{WC}(s'))\} \tag{8}$$

Formula (8) has the same form as the definition of $\Diamond$ in MInqB (4). The only difference is the presence of the world content projection WC, so that the inclusion is an inclusion between sets of worlds.

3.3 Pushing and Percolating

The syntactic language of our model $\mathsf{MInq}_\mathsf{B}^\mathsf{D}$ has two formula sorts. The first formula sort is type $k \to k$, where k is the type of contexts. Formulas of dynamic inquisitive semantics $\mathsf{Inq}_\mathsf{B}^\mathsf{D}$ have this type. By taking ℓ the type of stack of contexts $\tau = \langle c_0, ..., c_n\rangle$, the second sort of formula is of

[1] Information state widening was also advocated by Giannakidou and Mari 2019 for questions with an overt possibility modal.

[2] In formula (5), quantified assignment g' is independent from g. This allows us to solve the modal identity problem (Beaver 2001, § 8.3.1). For example, in the following discourse, interpreting both it_u in (i-b) and (i-c) with the same assignment g leads to a contradiction because $g(u)$ cannot equal A and R. The modality operator must thus make it possible to evaluate it_u in (i-b) with another (accessible) assignment. This is unnecessary for universal modality (13).

(i) a. A^u letter is hidden.
b. It$_u$ might be R.
c. But it$_u$ is A.

[3] Here, for simplicity, we do not take assignment matrices as in Dotlačil and Roelofsen 2021, but simple assignments. Note that it would work the same.

type $\ell \to \ell$, i.e. macro-context update functions. Like Kaufmann 2000, we take modals to be operators of type $(k \to k) \to (\ell \to \ell)$. After pushing a new local context, modals percolate the information (but not referents) to lower contexts, viz. (9).

$$\begin{aligned} [\![\mathsf{if}]\!]_{\mathsf{MInq}^{\mathsf{D}}_{\mathsf{B}}} &:= \quad \mathcal{U} \mapsto \textsc{push}\ \mathcal{U} \\ [\![\mathsf{then}]\!]_{\mathsf{MInq}^{\mathsf{D}}_{\mathsf{B}}} = [\![\mathsf{would}]\!]_{\mathsf{MInq}^{\mathsf{D}}_{\mathsf{B}}} &:= \quad \mathcal{U} \mapsto \textsc{push}\ \mathcal{U}; \textsc{perc}\ \mathcal{U} \\ [\![\mathsf{might}]\!]_{\mathsf{MInq}^{\mathsf{D}}_{\mathsf{B}}} = [\![\mathsf{could}]\!]_{\mathsf{MInq}^{\mathsf{D}}_{\mathsf{B}}} &:= \quad \mathcal{U} \mapsto \textsc{push}\ \mathcal{U}; \textsc{perc}\ \Diamond\mathcal{U} \end{aligned} \tag{9}$$

where $\ell \to \ell$ conjunction is defined as $\mathcal{T}; \mathcal{S} := \tau \mapsto \mathcal{S}(\mathcal{T}(\tau))$.

The analyses of Roberts 1989 and Gillies 2004 for conditionals and modal subordination are given in (10).

$$\begin{aligned} \textit{if } \varphi\textit{, then } \psi &\rightsquigarrow \Box(\varphi \to \psi) \\ \textit{if } \varphi\textit{, then might } \psi &\rightsquigarrow \Box(\varphi \to \Diamond\psi) \\ \textit{might } \varphi\textit{, would } \psi &\rightsquigarrow \Diamond\varphi \wedge \Box(\varphi \to \psi) \end{aligned} \tag{10}$$

To obtain these predictions, we define the percolation of a context update function $\mathcal{U}$ as follows. $\textsc{perc}\ \mathcal{U}$ applies $\mathcal{U}$ on the penultimate context c_{n-1} and modal conditionalization wrt. c_{n-1} and $\mathcal{U}$ on lower ones (11). Intuitively, $c[c' \vdash \mathcal{U}]$ is context c after learning that *if* c' *then* $\mathcal{U}$.

$$\begin{aligned} \textsc{push}\ \mathcal{U} &:= \lambda\langle c_0, ..., c_n\rangle.\ \langle c_0, ..., c_n, \mathcal{U}(c_n)\rangle \\ \textsc{perc}\ \mathcal{U} &:= \lambda\langle c_0, ..., c_{n-2}, c_{n-1}, c_n\rangle.\ \langle c_0[c_{n-1} \vdash \mathcal{U}], ..., c_{n-2}[c_{n-1} \vdash \mathcal{U}], \mathcal{U}(c_{n-1}), c_n\rangle \end{aligned} \tag{11}$$

with $c[c' \vdash \mathcal{U}]$ defined as $(\Box(c' \to \mathcal{U}))(c)$. Universal modality and implication (13) are consistent with $\mathsf{Inq}^{\mathsf{D}}_{\mathsf{B}}$ and (7).[4]

$$\begin{aligned} c' \to \mathcal{U} &:= \quad c \mapsto \{s \in c \mid \forall t \subseteq s. \forall t' \in c'.\ t \sqsubseteq t' \to t' \Subset \mathcal{U}(c')\} \\ \Box\mathcal{U} &:= \quad c \mapsto \{s \in c \mid \exists s' \in \mathcal{U}(c).\ \forall\langle w, g\rangle \in s.\ \forall w'.\ w' \in \mu(w) \to \langle w', g\rangle \Subset s'\} \end{aligned} \tag{13}$$

3.4 Meeting the Requirements

As an illustration, let's interpret sentence (5-a) in our model. Like Roelofsen and Dotlačil 2023, I assume functional heads. Focus head Foc_u provides the witness request operator $?u$. Interrogative clause type head Int ensures non-informativeness, by providing the presuppositional closure operator † (Roelofsen 2015). To weaken the uniqueness presupposition, we require the modality to raise above Foc_u but below Int, hence logical form (5-b).

(5) a. Which^u letter could be hidden in FO__M?
b. Int (*could* (Foc_u (*which*^{u} *letter hidden*)))
c. $\mathcal{T}$ = †($\textsc{push}\ \mathcal{V}; \textsc{perc}\ \Diamond\mathcal{V}$), with $\mathcal{V} = [u]; \mathbf{atom}\{u\}; \mathbf{letter}\{u\}; \mathbf{hidden}\{u\}; \mathbf{max}\{u\}; ?u$

The denotation of formula (5-c) is illustrated in Fig. 2.

Referent u introduced by $\mathcal{U}$ is not projected to the common ground c_0, ensuring external staticity (1.). A subsequent modal operator has access to the topmost local context c_1, containing assignments defined on u, thus ensuring modal subordination (2.).

In $\mathsf{Inq}^{\mathsf{D}}_{\mathsf{B}}$, mention-all reading is triggered by the interaction between the exhaustivity operator **max**, provided by *which*, and the witness request operator $?u$, creating the alternatives. Context $c_1 = \mathcal{V}(c_0)$ contains the alternatives "*d is hidden*" for every single letter d. Crucially, $\Diamond$ projects this inquisitiveness onto c'_0. In c'_0, the alternatives "*A can be hidden*" and "*R can be hidden*" overlap on $\{w^*_A, w^*_{RU}, w^*_R\}$, creating a mention-some reading (4.).

[4] Extension $\leqslant$ and subsistence are $\Subset, \sqsubseteq$ are defined as follows. Notation $\Subset, \sqsubseteq$ is taken from Dekker 1992.

$$\begin{array}{rcllllcll} \langle w, g\rangle & \leqslant & \langle w', g'\rangle & \text{if} & w = w' \wedge g \subseteq g' & \quad i & \Subset & s' & \text{if} \quad \exists j \in s'.\ i \leqslant j \\ s & \leqslant & s' & \text{if} & \forall j \in s', \exists i \in s.\ i \leqslant j & \quad s & \sqsubseteq & s' & \text{if} \quad s \leqslant s' \wedge (\forall i \in s.\ i \Subset s') \\ & & & & & \quad s & \Subset & c' & \text{if} \quad \exists s' \in c'.\ s \sqsubseteq s' \end{array} \tag{12}$$

World w	w_A^A	w_A^*	w_{RU}^*	w_R^*	w_R^R	w_{RU}^{RU}
Words with skeleton FO__M in English	FOAM	FOAM, FORM, FORUM	FOAM, FORM, FORUM	FOAM, FORM, FORUM	FORM	FORUM
Actual word on the board	FOAM	FOAM	FORUM	FORM	FORM	FORUM
Modal set $\mu(w)$	$\{w_A^A\}$	$\{w_A^*, w_R^*, w_{RU}^*\}$	$\{w_A^*, w_R^*, w_{RU}^*\}$	$\{w_A^*, w_R^*, w_{RU}^*\}$	$\{w_R^R\}$	$\{w_{RU}^{RU}\}$

Table 1: Illustration model $\mathcal{M}$ made of the set of worlds $W = \{w_A^A, w_A^*, w_{RU}^*, w_R^*, w_R^R, w_{RU}^{RU}\}$.

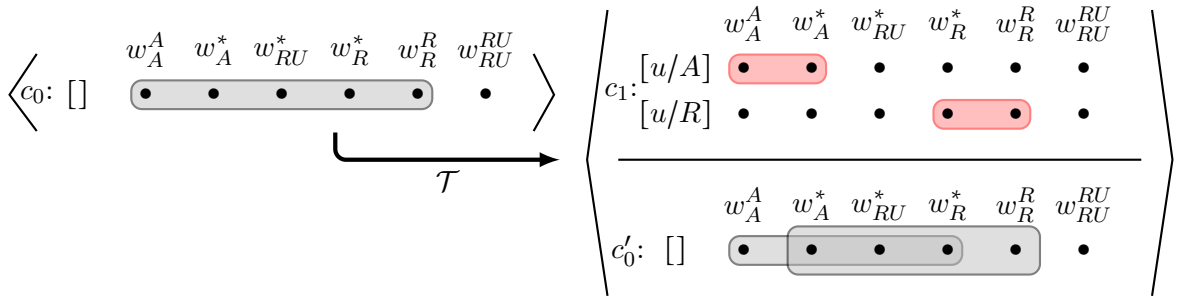


Figure 2: Diagram of the actions of the semantics of question (5-a) in model $\mathcal{M}$ on initial stack $\tau_0 = \langle c_0 \rangle$. We only represent alternatives of contexts, as sets of dots. A dot represents a possibility, i.e. a pair of a possible word (in abscissa) and an individual assignment function (in ordinate).

World w_{RU}^* is included in the alternatives of c_0'. Thus, global presupposition is obviated. However, if we take $c_0'' = \wp(\{\langle w, [] \rangle \mid w \in W\})$[5], $\mathcal{T}$ is undefined on stack $\langle c_0'' \rangle$, because of world w_{RU}^{RU}, where no single letter can fit in skeleton FO__M. The † operator (14), by applying on every element τ_i of stack τ, ensures local uniqueness presupposition (3.).

$$\dagger\mathcal{S} := \tau \mapsto \begin{cases} \mathcal{S}(\tau) & \text{if } \forall i < |\tau|.\ \bigcup \tau_i \sqsubseteq \bigcup \mathcal{S}(\tau)_i \\ \text{undefined} & \text{otherwise} \end{cases} \tag{14}$$

Finally, our model allows us to compare (5-a) with (6). The conditional antecedent introduces a context where w_{RU}^{RU} is excluded. Thus, †, provided in the consequent, acts vacuously. It correctly predicts that both (5-a) and (6) have the same effect, except for one point: (6) has no presupposition.

(6) If a^v single letter is hidden, which^u letter is it$_v$?

4 Conclusion

We designed a modal dynamic inquisitive semantics which accounts for exhaustivity and presupposition weakening of modalized questions. It also captures modal subordination with an externally static existential modal. To do so, we used stack-based semantics, allowing us to treat conditional and modalized questions uniformly.

Acknowledgements. I am grateful to Floris Roelofsen, Reinhard Muskens, and Philippe de Groote for precious feedback about this work.

[5] [] is the empty assignment, and $\wp$ the powerset operator.

A Non-Distributive Test Semantics for Modals

A side issue of previous dynamic semantics of modals concerns their non-distributive test interpretation. The definition advocated by Veltman 1996, and used by Dekker 1992; Groenendijk, Stokhof, and Veltman 1996; van Rooij 1998; Kaufmann 2000; Beaver 2001; Asher and McCready 2007, considers that epistemic possibility $\Diamond\varphi$ is a test: if φ is true at some world $w \in s$, then $[\![\Diamond\varphi]\!](s) = s$, otherwise $[\![\Diamond\varphi]\!](s) = \varnothing$. This has several drawbacks.

First, it fails to model other modal flavours, like circumstantial or deontic modalities. Second, it fails to model knowledge variations because it assumes that all worlds are accessible from any world. Finally, it implies that a modalized declarative either brings no at-issue information or leads to a contradiction. As a consequence, Veltman's (1996) existential modality cannot account for weakenings (3.) and (4.). Worse, extending this definition to inquisitive semantics would prevent an interrogative formula ψ from raising alternatives because $[\![\psi]\!]$ would either output the input context or the contradiction state.

We argue against modelling modals as tests, for both declaratives and interrogatives. Modalities can safely remove possible worlds from the global context. For example, sentence S in (7) does provide some new (non-contradictory) at-issue information about the content of the English lexicon. Suppose we are in an initial information state s which contains the actual world plus a world w_R^R (see Tab. 1) where, in English, the only word of the form FO__M is FORM. Then, updating s with $[\![S]\!]$ should remove w_R^R from s. With a test, on the contrary, updating s with $[\![S]\!]$ would not remove w_R^R. This is unwanted.

(7) A vowel can be inserted in FO__M to make an English word.

We can solve these problems by adopting a Kripkean accessibility semantics for modals (Kripke 1959), and more generally, Kratzerian theory of modals (Kratzer 1977; Kratzer 1991). Our account is both eliminative and distributive.[6]

References

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge University Press.

Asher, Nicholas and Elin McCready (May 1, 2007). "Were, Would, Might and a Compositional Account of Counterfactuals". In: *Journal of Semantics* 24.2, pp. 93–129. ISSN: 0167-5133. DOI: 10.1093/jos/ffl013. URL: https://doi.org/10.1093/jos/ffl013 (visited on 04/23/2024).

Beaver, David I. (June 2001). *Presupposition and Assertion in Dynamic Semantics: A Critical Review of Linguistic Theories of Presupposition and a Dynamic Alternative*. Studies in Logic, Language, and Information. Center for the Study of Language and Information. 250 pp. ISBN: 978-1-57586-120-3. URL: https://press.uchicago.edu/ucp/books/book/distributed/P/bo3639443.html (visited on 11/04/2024).

Brasoveanu, Adrian (Nov. 1, 2010). "Decomposing Modal Quantification". In: *Journal of Semantics* 27.4, pp. 437–527. ISSN: 0167-5133. DOI: 10.1093/jos/ffq008. URL: https://doi.org/10.1093/jos/ffq008 (visited on 07/11/2023).

Ciardelli, Ivano (2016). "Questions in Logic". PhD thesis. Amsterdam: University of Amsterdam. 398 pp.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (Nov. 27, 2018). *Inquisitive Semantics*. Oxford Surveys in Semantics and Pragmatics. Oxford, New York: Oxford University Press. 240 pp. ISBN: 978-0-19-881478-8.

[6] The epistemic behaviour which motivated Veltman 1996 can be retrieved by requiring the accessibility relation to be an equivalence relation.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler (Feb. 1, 2017). "Composing Alternatives". In: *Linguistics and Philosophy* 40.1, pp. 1–36. ISSN: 1573-0549. DOI: 10.1007/s10988-016-9195-2. URL: https://doi.org/10.1007/s10988-016-9195-2 (visited on 03/15/2021).

Dayal, Veneeta (1996). *Locality in WH Quantification*. Red. by Gennaro Chierchia, Pauline Jacobson, and Francis J. Pelletier. Vol. 62. Studies in Linguistics and Philosophy. Dordrecht: Springer Netherlands. ISBN: 978-0-7923-5478-9. DOI: 10.1007/978-94-011-4808-5. URL: http://link.springer.com/10.1007/978-94-011-4808-5 (visited on 07/11/2023).

Dekker, Paul (Jan. 1, 1992). *An Update Semantics for Dynamic Predicate Logic*. Report. URL: https://eprints.illc.uva.nl/id/eprint/1177/ (visited on 11/07/2024).

Dotlačil, Jakub and Floris Roelofsen (2019). "Dynamic Inquisitive Semantics: Anaphora and Questions". In: *Proceedings of Sinn Und Bedeutung 23*. Universitat Autònoma de Barcelona. DOI: 10.18148/sub/2019.v23i1.538. URL: https://dare.uva.nl/search?identifier=e767de8b-ac55-4848-af4f-7bf63dfbfb95 (visited on 06/25/2021).

— (Mar. 2, 2021). "A Dynamic Semantics of Single-Wh and Multiple-Wh Questions". In: *Semantics and Linguistic Theory* 30.0, pp. 376–395. ISSN: 2163-5951. DOI: 10.3765/salt.v30i0.4839. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.376 (visited on 08/09/2021).

Frank, Anette (1997). "Context Dependence in Modal Constructions". PhD thesis. Stuttgart: Universität Stuttgart. 411 pp.

Geurts, Bart (1999). *Presuppositions and Pronouns*. Current Research in the Semantics / Pragmatics Interface 3. Leiden, The Netherlands: Brill. ISBN: 978-0-08-043592-3.

Giannakidou, Anastasia and Alda Mari (2019). *Modalization and Bias in Questions*. Unpublished manuscript. University of Chicago and Institut Jean Nicod Paris.

Gillies, Anthony S. (2004). "Epistemic Conditionals and Conditional Epistemics". In: *Noûs* 38.4, pp. 585–616. DOI: 10.1111/j.0029-4624.2004.00485.x.

Groenendijk, Jeroen (1998). "Questions in Update Semantics". In: *Formal Semantics and Pragmatics of Dialogue: Proceedings of the Thirteenth Twente Workshop on Language Technology (Twendial '98)*. Twente: Universiteit Twente, Faculteit Informatica, pp. 125–137. URL: https://dare.uva.nl/search?identifier=e7f44061-4fc2-4924-8811-848a0c9630ed (visited on 07/24/2023).

Groenendijk, Jeroen and Martin Stokhof (Jan. 1, 1984). "Studies on the Semantics of Questions and the Pragmatics of Answers". PhD thesis. Amsterdam, The Netherlands: Institute for Logic, Language and Computation (ILLC), University of Amsterdam.

Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman (1996). "Coreference and Modality". In: *The Handbook of Contemporary Semantic Theory*. Ed. by Shalom Lappin. Cambridge, Mass., USA: Blackwell Refernce, pp. 179–216. ISBN: 978-0-631-18752-3. URL: https://dare.uva.nl/search?identifier=c655089e-9fae-4842-90da-30e46f37825b (visited on 01/31/2024).

Hara, Yurie and Katsuhiko Sano (2017). "Conditional Questions Revisited". In: InqBnB2: Inquisitiveness Below and Beyond the Sentence Boundary. Amsterdam. URL: https://yuriehara.github.io/website/download/CQR_NINJALpdf.pdf.

Higginbotham, James and Robert May (Jan. 1, 1981). "Questions, Quantifiers and Crossing". In: *The Linguistic Review* 1.1, pp. 41–80. ISSN: 1613-3676. DOI: 10.1515/tlir.1981.1.1.41. URL: https://www.degruyter.com/document/doi/10.1515/tlir.1981.1.1.41/html (visited on 07/27/2023).

Hirsch, Aron and Bernhard Schwarz (2019). "Singular Which, Mention-Some, and Variable Scope Uniqueness". In: *Semantics and Linguistic Theory* 29.0, pp. 748–767. ISSN: 2163-5951. DOI: 10.3765/salt.v29i0.4637. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/29.748 (visited on 09/05/2023).

Hofmann, Lisa (Dec. 20, 2019). "The Anaphoric Potential of Indefinites under Negation and Disjunction". In: *Proceedings of the 22nd Amsterdam Colloquium*. Amsterdam, pp. 181–190.

Isaacs, James and Kyle Rawlins (Aug. 1, 2008). "Conditional Questions". In: *Journal of Semantics* 25.3, pp. 269–319. ISSN: 0167-5133. DOI: 10.1093/jos/ffn003. URL: https://doi.org/10.1093/jos/ffn003 (visited on 04/18/2024).

Karttunen, Lauri (1969). "Discourse Referents". In: *Proceedings of the 1969 Conference on Computational Linguistics (COLING '69)*. Sång-Säby, Sweden: Association for Computational Linguistics, pp. 1–38. DOI: 10.3115/990403.990490. URL: http://portal.acm.org/citation.cfm?doid=990403.990490 (visited on 07/07/2023).

Kaufmann, Stefan (2000). "Dynamic Context Management". In: Faller, Martina, Stefan Kaufmann, and Marc Pauly. *Formalizing the Dynamics of Information*. CSLI Lecture Notes 92. Stanford, CA: CSLI Publications, pp. 171–181. ISBN: 1-68400-007-6. URL: https://scholar.google.com/citations?view_op=view_citation&hl=fr&user=gffJf60AAAAJ&citation_for_view=gffJf60AAAAJ:aqlVkmm33-oC (visited on 04/18/2024).

Kibble, Rodger (1998). "Modal Subordination, Focus, and Complement Anaphora". In: *The Tblisi Symposium on Logic, Language and Computation: Selected Papers*. Ed. by Jonathan Ginzburg et al. Stanford, USA: CSLI Publications, pp. 73–86. ISBN: 978-1-57586-098-5.

Kratzer, Angelika (Jan. 1, 1977). "What 'Must' and 'Can' Must and Can Mean". In: *Linguistics and Philosophy* 1.3, pp. 337–355. ISSN: 1573-0549. DOI: 10.1007/BF00353453. URL: https://doi.org/10.1007/BF00353453 (visited on 01/26/2024).

— (1991). "Modality". In: *Semantik / Semantics: Ein internationales Handbuch zeitgenössischer Forschung*. Ed. by Arnim van Stechow and Dieter Wunderlich, pp. 639–650. DOI: 10.1515/9783110126969.7.639.

Kripke, Saul A. (1959). "A Completeness Theorem in Modal Logic". In: *The Journal of Symbolic Logic* 24.1, pp. 1–14. ISSN: 0022-4812. DOI: 10.2307/2964568. JSTOR: 2964568. URL: https://www.jstor.org/stable/2964568 (visited on 05/21/2021).

Roberts, Craige (Dec. 1, 1989). "Modal Subordination and Pronominal Anaphora in Discourse". In: *Linguistics and Philosophy* 12.6, pp. 683–721. ISSN: 1573-0549. DOI: 10.1007/BF00632602. URL: https://doi.org/10.1007/BF00632602 (visited on 07/26/2023).

Roelofsen, Floris (2015). "The Semantics of Declarative and Interrogative Lists". Report. URL: http://semanticsarchive.net/Archive/jdiNjljO/paper.pdf.

Roelofsen, Floris and Jakub Dotlačil (June 1, 2023). "Wh-Questions in Dynamic Inquisitive Semantics". In: *Theoretical Linguistics* 49.1-2, pp. 1–91. ISSN: 1613-4060. DOI: 10.1515/tl-2023-2001. URL: https://www.degruyter.com/document/doi/10.1515/tl-2023-2001/html (visited on 12/13/2023).

Socolof, Michaela, Aron Hirsch, and Bernhard Schwarz (2020). "Which-Questions, Uniqueness, and Answerhood: Evidence from Disjunction". In: *Proceedings of NELS*. Vol. 50.

Stalnaker, Robert C. (1981). "A Theory of Conditionals". In: *IFS: Conditionals, Belief, Decision, Chance and Time*. Ed. by William L. Harper, Robert Stalnaker, and Glenn Pearce. Dordrecht: Springer Netherlands, pp. 41–55. ISBN: 978-94-009-9117-0. DOI: 10.1007/978-94-009-9117-0_2. URL: https://doi.org/10.1007/978-94-009-9117-0_2 (visited on 06/06/2024).

Van Rooij, Robert (1998). "Modal Subordination in Questions". In: *Formal Semantics and Pragmatics of Dialogue. Proceedings of Twendial '98*. Ed. by J. Hulstijn and A. Nijholt. Enschede: University of Twente, pp. 237–247. URL: https://dare.uva.nl/search?identifier=31017f54-cf94-4cfd-b5c1-a4aa9f433b55 (visited on 07/07/2023).

Van den Berg, Martin H. (Jan. 1, 1996). "Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora". PhD thesis. Amsterdam: University of Amsterdam. 202 pp. URL: https://eprints.illc.uva.nl/id/eprint/1996/ (visited on 07/24/2023).

Veltman, Frank (1996). "Defaults in Update Semantics". In: *Journal of Philosophical Logic* 25.3, pp. 221–261. ISSN: 0022-3611. JSTOR: 30226572. URL: https://www.jstor.org/stable/30226572 (visited on 05/07/2024).

Xiang, Yimei (May 17, 2016). "Interpreting Questions with Non-Exhaustive Answers". PhD thesis. Cambridge, Massachusetts, USA: Harvard University, Graduate School of Arts & Sciences. 300 pp. URL: https://www.semanticscholar.org/paper/Interpreting-Questions-with-Non-Exhaustive-Answers-Xiang/3f3183c8150447d1cfdb5ab0faf6bc14bb1bf316 (visited on 09/05/2023).

Xiang, Yimei and Alexandre Cremers (2017). "Mention-Some Readings of Plural-Marked Questions: Experimental Evidence". In: *Proceedings of the Forty-Seventh Annual Meeting of the North East Linguistic Society (NELS 47)*. Ed. by Andrew Lamont and Katerina Tetzloff. Vol. 3. University of Massachusetts, Amherst, USA: GLSA (Graduate Linguistics Student Association), pp. 261–274. ISBN: 978-1-976344-60-2. URL: https://www.semanticscholar.org/paper/Mention-some-readings-of-plural-marked-questions%3A-Xiang-Cremers/f29b8de6ff99142533c255fc4c831ac1fc4237bc (visited on 07/10/2023).

Just-asking questions

Tom Roberts
Institute for Language Sciences, Utrecht University
Utrecht, The Netherlands
t.d.h.roberts@uu.nl

Abstract

While the 'canonical' function of questions—utterances of interrogative clauses—is to solicit *answers*, this is not a definitive property of questions, since some varieties have an interrogative form but a non-inquisitive function, such as rhetorical questions. I examine a new type of such questions in English which I dub *controversy questions*, whose answer is unknown to conversational participants but nevertheless steers the discourse towards answering them. I propose that such questions motivated a bleached pragmatics for question-asking: they set a conversational agenda of exploring a Question Under Discussion (Roberts 1996/2012) and nothing further. I show that this view can derive a range of attested possible discourse effects of questions given independent features of contexts of utterance. Moreover, I argue that this perspective also explains how question-asking can be weaponized as a rhetorical strategy to sow doubt under the guise of 'just asking questions' in contexts like newspaper headlines.

1 What does it mean to ask a question?

Asking a question—that is, uttering a sentence with interrogative form—is canonically equivocated with requesting *information*: an ignorant asker seeks an answer from a possibly knowledgeable addressee (Searle 1969; Dayal 2016, a.m.o.). While intuitive, the notion that information-seeking is part and parcel of question-asking has long been challenged by the existence of 'noncanonical' species of questions, which seem to generate very different discourse effects despite having the (morphosyntactic and prosodic) appearance of typical questions. Two famous cases are rhetorical questions, whose answer is obvious and are used to highlight the obviousness of some other proposition (1), and exam questions (2) which assess the addressee's knowledge for some discourse-external goal rather than rectify the speaker's ignorance (Sadock 1971; Han 2002; Caponigro and Sprouse 2007; Biezma and Rawlins 2017; Farkas 2022; Farkas 2024, a.m.o.).

(1) A: Did you go to Gertrude's party?
B: Is the Pope Catholic?
⇝ B went to the party and A should have known that

(2) (On a geography quiz)
Is Bratislava the capital of Slovakia?

These exceptional types of questions present an apparent classic mismatch of form and function. By appearances, the questions in (1)-(2) are instantiated by ordinary interrogatives, but they neither request information nor necessitate speaker ignorance. These observations indicate that neither of these functions are inherent to question-asking itself, motivating an alternative view of questions not as *intrinsic* requests for information, but as setting an immediate goal for the discourse of mutually resolving the issue raised by the question, e.g. in Question Under Discussion- or Table-based frameworks (Roberts 1996/2012; Ginzburg 1996; Farkas and Bruce 2010). Uttering an interrogative sentence, then, which might serve different functions depending on features of the context. However, such theories typically also assume that information-seeking

questions are the default flavor, which is descriptively plausible if theoretically arbitrary, and often stipulate in the requirement that the addressee answer the question into the conventional discourse effects of uttering interrogative clauses.

In this paper, I examine a variety of question I dub **controversy questions** (ConQs) in English, which have received little previous attention: questions which do not necessarily impose a requirement of answering, but rather communicate that the issue raised is somehow up for debate or worthy of asking. I propose that these questions support a bare-bones dynamics for question-asking inspired by Rawlins (2024): asking a question updates the QUD—opening an issue for eventual resolution—but in and of itself makes no demands on the addressee. Rather, the default obligation of question answering on the addressee comes from assumptions about how QUDs should be handled given different configurations of *authority* of discourse participants, which do not hold in the case of ConQs. I further show how this open-endedness of ConQs can serve as a rhetorical tool to sow doubt and misinformation, particularly in headlines.

2 What controversy questions are

Controversy questions are questions which aim to set a topic for discussion, but with no expectation of obtaining a full answer or even placing a demand on the addressee to provide one While ConQs can in principle be either polar or constituent questions, I restrict my attention to polar ConQs in this paper for reasons of space. ConQs are commonly used in headlines (3) to introduce articles which explore the possible answers to a question, often without settling on a particular answer (in the sense of Groenendijk and Stokhof 1984). In monologic contexts, such as giving presentations (4), and in open-ended dialogues (5), ConQs can be used to start a sincere conversation about the issue, with the tacet acknowledgment that a full resolution may not be reached.

(3) a. Can Barcelona survive mass tourism? (*New York Times*, 08/2024)
b. Is your home security system really secure? (*Fox News*, 08/2024)

(4) *First slide of astrobiology talk:*
Could humans survive on Mars?

(5) A: Is Kamala Harris going to win the election(, do you think)?
B: I don't know. It's a toss-up.

ConQs resemble rhetorical questions, whose primary function is to make a 'statement' by posing a question whose answer is obvious or unknowable (Rohde 2006; Biezma and Rawlins 2017; Farkas 2024). Like their rhetorical cousins, ConQs have the appearance of standard information-seeking questions, but functionally impose no obligation to resolve an issue on the addressee (and in the case of headlines, an answer from the reader is not even possible), though they do require a follow-up which is nevertheless *relevant* to the question (cf. (3)). But unlike rhetorical questions, ConQs do not have the 'feel' of an assertion: they seem to sincerely express a desire for answers, even if the addressee can't immediately provide one.

One additional key difference is that rhetorical questions seem to be infelicitous out of the blue–asking *Is the Pope Catholic?* with no prior discourse context can only be a genuine inquiry from a very uninformed person. ConQs, by contrast, are not only fine discourse-initially, but fit there most naturally. In this sense, they resemble the 'non-intrusive questions' of Farkas (2022). However, ConQs crucially lack the explicit grammatical marking which Farkas argues is a hallmark of NIQs, suggesting we should interrogate what contextual and pragmatic factors lead to ConQs having the interpretation they do. In short, ConQs are apparently grammatical unmarked questions whose answer is assumed to be not known to any participants in the discourse, but contribute a function of setting a topic for downstream exploration.

3 Asking questions in the Table model

I adopt the common assumption that polar questions denote the set containing a proposition p and its complement $\neg p$, à la Hamblin (1973) and Karttunen (1977), and couch my analysis in a slightly simplified version of the Table model of discourse contexts (Farkas and Bruce 2010; Farkas and Roelofsen 2017), which is well suited to describe the relationship between the denotation of a sentence and how uttering that sentence updates a discourse context.

In the basic version of this framework, a discourse context c consists of three components: (a) the *Table*, a stack of issues (sets of propositions) to be resolved in the current discourse, (b) a set of discourse commitments DC_x for each participant in the discourse, consisting of propositions to which that participant is committed, from which a common ground CG can be derived, and (c) a projected set of possible future common grounds as a result of the most recent discourse move. Moment-to-moment, discourse participants are primarily driven by the desire to remove issues from the Table by resolving them (or, in rarer cases, rejecting them); the overarching goal of a conversation is to shrink the context set cs ($= \bigcap CG$), or the set of worlds compatible with joint commitments of all participants.

I assume, as is standard in this framework, that the basic function of uttering an unmarked polar interrogative is to set the conversational agenda of resolving the issue that utterance places atop the Table. I follow Farkas and Bruce (2010) in assuming this element of the Table corresponds to the QUD when it is a non-singleton issue, but I crucially diverge from them in that I assume that this is **all** uttering an interrogative clause does, whereas F&B (and many subsequent authors) assume that uttering p? adds two elements to the projected set: one of CG enhanced with p, and one enhanced with $\neg p$.

Removing this step might seem like a pure technicality, since the ps can be derived from the Table, but it is frequently interpreted as a "menu" of licit follow-up discourse moves (see Rudin 2022 for a formalization of this idea). What I am proposing is that asking a question merely establishes a QUD, presenting it to the addressee and nothing further. What they do with that QUD depends squarely on key aspects of the the context of utterance.

3.1 What we can do with questions and why

Rather than being guided by the projected set, I treat discourse participants as having as their lodestar *addressing* the QUD, if not ultimately resolving it. In this section, I show how the act of introducing a QUD is interpreted depends on the addressee's beliefs about the discourse context, particularly with respect to the notion of AUTHORITY on an issue I. Drawing inspiration from Northrup (2014) and Heritage and Raymond (2005) for authority on propositions, I conceive here of authority with respect to a question as 'how good' a particular agent is as a source of information with respect to settling that question. Formally, I define authority with respect to an issue I in terms of an agent's probability of knowing the true answer to I:

(6) AUTHORITY: Given an information state s, the authority of agent x with respect to issue I in s is $p_s(K_x(I))$.

The way the addressee reasons about the speaker's utterance of question q depends on their beliefs about the speaker's beliefs. The speaker, in turn, must make some assumptions about these beliefs of the addressee in order for their question to have the desired effect. I will refer to the information state corresponding to what the speaker takes the addressee's doxastic state to be as s. What is crucial is that the interlocutors' relative authority in s can lead the very some utterance to have dramatically different interpretation. For example, if the addressee takes the speaker to already believe that the true answer to q is common ground, the utterance of q is unlikely to be a sincere information-seeking question.

To see how this works, I will sketch how a few possible configurations of authority relative to s can explain why adding q to the QUD can have the effects it does.

(7) **Sincere question**: $p_s(K_{Sp}(q)) < p_s(K_{Ad}(q))$

With sincere information-seeking questions, the addressee believes that they are taken to be a greater authority on q than the speaker. Since the utterance of q sets resolving q as the immediate goal of the conversation, the addressee can reason that they are being invited to provide that resolution thanks to their greater authority.

(8) **Exam question**: $p_s(K_{sp}(q)) = 1 \wedge p_s(K_{Ad}(q)) < 1$

In situations in which the speaker is assumed to know the true answer to q and the addressee may or may not, we need to reason about why the speaker is raising the QUD of q if not for their own information. Given the relative authority of the speaker over the addressee, it is natural to interpret such moves as a test of the addressee's knowledge. The addressee is still expected to provide an answer, but to different ends then informing the speaker.

(9) **Rhetorical question**: $p_s(K_{Sp}(q)) = p_s(K_{Ad}(q)) \wedge p_s(K_{Sp}(q)) \in \{0, 1\}$

Rhetorical questions are, as Farkas (2024) puts it, 'closed': they can't be resolved in the current discourse context, either because they are already answered by the common ground, or couldn't possibly be. The apparent futility of such a move indicates the speaker's goal is not to seek an answer to q per se, but rather to draw a parallel between the obviousness/uselessness of q and the analogous obviousness/uselessness of a preceding discourse move. For a more fleshed-out version of this idea, see Farkas (2024).

(10) **Controversy question**: $p_s(K_{Sp}(q)) = \text{low} \wedge p_s(K_{Ad}(q)) = \text{low}$

Turning at last to controversy questions, they are born of in contexts where the speaker is taken to believe that both they and the addressee have relatively low authority on q, where 'low' is contextually specific. Since both speaker and addressee have more than 0 authority on q, the sincere goal of addressing the QUD can still be the guiding hand of the discourse. However, the acknowledgment that neither speaker nor addressee are any great authority on p lifts the burden of providing a full answer on the addressee. If the addressee thinks they aren't assumed to be an authority on q, they should still respond something relevant to it since the speaker wants to know the answer, but the addressee cannot be reasonably expected to fully provide that answer.

3.2 Summary

When the utterance of a question q has less then full authority on q, it has an inquisitive function: since uttering q establishes it as the QUD, putting it on top of the conversational agenda, the addressee should answer it directly if they can. On the other hand, asking a question whose answer is known to the speaker, as in the case of rhetorical or exam questions, indicates that the speaker's goal cannot possibly be to gain information about whether q for themselves.

Note that this raising of 'insincere' QUDs is not a special feature of questions; for instance, sarcastic declarative sentences involve making utterances that obviously don't intend to inform addressees of the truth of their content:

(11) A: I just passed my geography test about Slovakia!
B: And I'm the Queen of Canada.
$\rightsquigarrow$ B doesn't believe that A passed[1]

[1]Example modified from a conversation with Donka Farkas, p.c.

The addressee's plausible follow-ups are dictated by reasoning about how the QUD should be addressed. the addressee is best equipped to address the QUD raised by a sincere question by answering it straightforwardly, since they are the local authority. On the other hand, ConQs don't put the addressee on the hook for an answer because they (definitionally) occur in contexts in which it is mutually understood that neither the speaker nor addressee know the answer, but they still set an unanswerable QUD as a topic of discussion, inviting exploration of possibilities. I leave a full-fledged explanation of how the precise inferences associated with different question types and exploration of the possible typology of question interpretations for future work.

4 Betteridge's law and questions as tools of misinformation

I've proposed that questions don't inherently require answers, but rather that making utterances with the kinds of semantic content that questions have can prompt an addressee for answers given the right contexts. The general pragmatic rules at play can also shed light on how questions can be used to all kinds of non-inquisitive rhetorical ends. I focus here on a single case study: The tongue-in-cheek journalistic adage Betteridge's Law, named for tech journalist Ian Betteridge, which describes an intuition about questions as headlines:

(12) **Betteridge's Law of Headlines** (Betteridge 2009)
Any headline that ends in a question mark can be answered by the word *no*.

Of course, this is not true of all headline questions (see (3)). But the intuition behind this 'law' is that a headline of polar question p? is a backdoor technique to put forward weak-evidenced claim that p–or establish p as a controversy–without explicitly committing to it. As Betteridge puts it, 'The reason why journalists use that style of headline is that they know the story is probably bullshit, and don't actually have the sources and facts to back it up, but still want to run it.' Because newspapers are supposed to be knowledgeable, opting for an interrogative headline as opposed to an assertive one generates the inference that p is up for debate.

This line of reasoning can be described more formally as a result of two independently-motivated general pragmatic considerations. The first is AUTHORITY. The second is Rudin's (2022) pragmatic maxim of VIABILITY, which militates against making utterances that raise issues containing alternatives that will never enter the common ground, such as questions to which the addressee has already committed to an answer. The idea here is that all else being equal, speakers aim to avoid asking questions with potential answers that are useless because someone in the conversation already thinks they are false. (Note that Rudin's version is specified in terms of the projected set, which I have altered here to align with the proposal that uttering interrogatives adds nothing to the projected set.)

(13) VIABILITY: Only raise issues for which all alternatives are compatible with what you believe are each interlocutor's commitments and private beliefs. (adapted from Rudin 2022: ex. 50)

Betteridge's law can be understood as an interaction between AUTHORITY and VIABILITY. Uttering a question p? (denoting $\{p, \neg p\}$) only avoids violating VIABILITY if the speaker takes both p and $\neg p$ to be possible given what they know about their own and the addressee's beliefs: that is, they don't know p and they don't know $\neg p$; if the speaker was committed to one or the other, they could avoid violating VIABILITY by uttering that proposition outright.

But news sources are also very special kinds of discourse participants. They are meant to be reliable informers, in the sense that they should give as much information as is (publicly) known as possible about the topics they cover. I refer to an agent that is supposed to be maximally authoritative about an issues as an ABSOLUTE AUTHORITY (14):

(14) An ABSOLUTE AUTHORITY on issue I relative to information state s is an agent y s.t. $\forall z \neq y : p_s(K_y(I)) \geq p_s(K_z(I))$.

If a speaker is an ABSOLUTE AUTHORITY on p? relative to the addressee's doxastic state, the addressee should infer that a speaker who utters p? indicates that *no one* knows whether p or $\neg p$. This is the crucial element for headline questions: news sources are assumed to be reliable information sources. I assume readers, who function as addressees in the sort of one-sided dialogue with a news source, treat news sources as absolute authorities. It follows that a headline p? appearing in a news source suggests that no one knows whether or not p.

Getting readers to draw the inference that the true answer to p? is currently unknown can be weaponized for questions for which one alternative is outlandish or inflammatory. For example, ConQs in highly partisan or controversial publications (15) can serve two purposes: making readers believe the issue is up for debate *and* raising the possibility of a dubious claim which otherwise saying outright might result in negative consequences:

(15) a. Is Amazon sabotaging first book about Trump assassination attempt? (*Infowars*, 09/2024)
b. Are abortion funds running out of money now? (*WorldNetDaily*, 09/2024).
c. Will Trump turbocharge your finances? (*Daily Mail*, 11/2024)

This tactic provides an author with the plausible deniability of 'just asking questions'; they can plant the seed of controversy while evading responsibility for its flowering, no matter how outlandish. The intentional use of implicature to communicate ideas without being on the hook for committing to them is nothing new (see Grice 1975, among many others), but what sets the inference generated by headline ConQs apart from more familiar implicatures is that their discourse effects do not involve making commitments at all.

Finally, there are other reasons why an author (or publisher) might opt for questions in headlines other than sowing doubt or propagating misinformation. In the current online news media ecosystem, revenue is generally proportional to volume of web traffic, which incentivizes headlines that drive clicks. While a potential reader could learn the main point of an article like *President X wins re-election* without actually reading the body, *Did President X win re-election?* suggests uncertainty, inviting the reader to click through to learn more. Subsequent research may reveal how exactly questions are leveraged in headlines, and to what ends.

5 Conclusion

Conversations are more than a back-and-forth volley of information; often they can just be the exploration of an idea without any particular resolution. Uttering a question, in virtue of simply setting the QUD, directs our attention to a particular topic and delineate the boundaries of said conversation, but doesn't always demand an answer, in ways that flow from how we understand the context of a conversation.

Much work remains to be done in this domain. One interesting direction, particularly in reasoning about inferences derived from the very fact of speakers making a particular utterance, could be connecting the pragmatics of questioning with logics of attention (e.g. Jager 2009; Westera 2017). Empirically, documenting the range of non-canonical questions cross-linguistically is also essential. Here I have explored only one corner of English, but it's not at all obvious that the story I have sketched here should apply to any other language, let alone all of them. Are ConQs a linguistic universal?

Acknowledgements. Thanks to Donka Farkas, Kyle Rawlins, Floris Roelofsen, and anonymous AC reviewers for useful ideas and comments. Errors were supplied in-house.

References

Betteridge, Ian (2009). *TechCrunch: Irresponsible journalism*. Technovia.co.uk. Archived at https://web.archive.org/web/20090226202006/http://www.technovia.co.uk/2009/02/techcrunch-irresponsible-journalism.html.

Biezma, Maria and Kyle Rawlins (2017). "Rhetorical questions: Severing asking from questioning". In: *Proceedings of Semantics and Linguistic Theory*. Ed. by Dan Burgdorf et al. Vol. 27, pp. 302–322. DOI: 10.3765/salt.v27i0.4155.

Caponigro, Ivano and Jon Sprouse (2007). "Rhetorical questions as questions". In: *Proceedings of Sinn und Bedeutung*. Vol. 11, pp. 121–133. DOI: 10.18148/sub/2007.v11i0.635.

Dayal, Veneeta (2016). *Questions*. Oxford University Press.

Farkas, Donka (2022). "Non-intrusive questions as a special type of non-canonical questions". In: *Journal of Semantics* 39.2, pp. 295–337. DOI: 10.1093/jos/ffac001.

— (2024). "Rhetorical questions revisited". In: *Proceedings of the Chicago Linguistics Society 60*.

Farkas, Donka and Kim Bruce (2010). "On Reacting to Assertions and Polar Questions". In: *Journal of Semantics* 27.1, pp. 81–118. DOI: 10.1093/jos/ffp010.

Farkas, Donka and Floris Roelofsen (2017). "Division of Labor in the Interpretation of Declaratives and Interrogatives". In: *Journal of Semantics* 34.2, pp. 237–289. DOI: 10.1093/jos/ffw012.

Ginzburg, Jonathan (1996). "Dynamics and the semantics of dialog". In: *Language, Logic, and Computation, Volume 1*. Ed. by J. Seligman. Stanford: CSLI.

Grice, Paul (1975). "Logic and Conversation". In: *Syntax and Semantics 3: Speech Acts*. Ed. by P. Cole and J. Morgan. Leiden, The Netherlands: Brill. DOI: doi.org/10.1163/9789004368811_003.

Groenendijk, Jeroen and Martin Stokhof (1984). "Studies on the Semantics of Questions and the Pragmatics of Answers". PhD thesis. University of Amsterdam.

Hamblin, Charles L. (1973). "Questions in Montague English". In: *Foundations of Language* 10.1, pp. 41–53.

Han, Chung-hye (2002). "Interpreting interrogatives as rhetorical questions". In: *Lingua* 112.3, pp. 201–229. DOI: doi.org/10.1016/S0024-3841(01)00044-4.

Heritage, John and Geoffrey Raymond (2005). "The Terms of Agreement: Indexing Epistemic Authority and Subordination in Talk-in-Interaction". In: *Social Psychology Quarterly* 68.1, pp. 15–38.

Jager, Tikitu de (2009). ""Now that you mention it, I wonder...": Awareness, Attention, Assumption". PhD thesis. University of Amsterdam.

Karttunen, Lauri (1977). "Syntax and semantics of questions". In: *Linguistics and Philosophy* 1.1, pp. 3–44. DOI: 10.1007/BF00351935.

Northrup, Oliver (2014). "Grounds for Commitment". PhD thesis. UC Santa Cruz.

Rawlins, Kyle (2024). *Asking questions*. Presentation at Semantics and Linguistic Theory 32.

Roberts, Craige (1996/2012). "Information structure in discourse". In: *OSU Working Papers in Linguistics* 49, pp. 91–136.

Rohde, Hannah (2006). "Rheotrical questions as redundant interrogatives". In: *San Diego Linguistics Papers*.

Rudin, Deniz (2022). "Intonational commitments". In: *Journal of Semantics* 39.2, pp. 339–383. DOI: 10.1093/jos/ffac002.

Sadock, Jerry M. (1971). "Queclaratives". In: *Seventh Regional Meeting of the Chicago Linguistic Society*, pp. 223–232.

Searle, John (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

Westera, Matthijs (2017). “Exhaustivity and intonation: A unified theory”. PhD thesis. University of Amsterdam.

Bounds for metaphysical modality

Paolo Santorio
University of Maryland, College Park
paolosantorio@gmail.com

Abstract

Epistemic modals exhibit a kind of nonclassical behavior, exemplified by epistemic contradictions like (1). Conversely, metaphysical modals seemingly conform to the predictions of semantics based on classical modal logics, as witnessed by the consistency of (2).

(1) # Al is tall and he might not be.

(2) Al is tall and he might not have been.

I argue that this difference is illusory. Metaphysical modals display nonclassical behavior mirroring that of epistemics. This nonclassicality is manifested in puzzles concerning the interaction of *would* and *might*. I show how these facts be captured, while still preserving the contrast between (1) and (2), building on Mandelkern's recent work on bounds.

1 Introduction

Epistemic modals notoriously exhibit a kind of nonclassical behavior, which is hard to capture on semantic theories based on Kripkean accessibility relations (Kratzer 1981, 2012 a.o.). This nonclassicality is typically exemplified by so-called epistemic contradictions, i.e. sentences of the form $\ulcorner A \wedge \Diamond \neg A \urcorner$ (Veltman 1985; Yalcin 2007; Willer 2013; Mandelkern 2019; Mandelkern 2024 a.o.). These sentences are consistent by the lights of classical theories, but they are infelicitous. For a simple example, consider (1).

(1) #Ada is tall and she might not be.

Epistemic modality is taken to be unique in this behavior. Seemingly, other modal flavors are better behaved and conform to the predictions of a modal semantics based on standard modal logics. In particular, the flavor of modality that is often labeled 'metaphysical', and that includes both *would*-conditionals and contrary-to-fact *might*-claims, is usually considered a paradigm of classicality. At first sight, this is supported by empirical evidence. The metaphysical counterparts of epistemic contradictions, such as (2), are perfectly felicitous.

(2) Ada is tall and she might not have been.

This short paper argues that, contrary to first appearances, the divergence between epistemic and metaphysical modality is largely illusory. Metaphysical modals display a kind of nonclassical behavior that mirrors that of epistemic modals. This kind of nonclassicality is manifested in puzzles concerning the interaction of *would* and *might*, some well-known and some new. Moreover, this nonclassicality can be captured in a theory that generalizes current theories of epistemic modals. For current purposes, I build on Mandelkern's (2019, 2024) recent proposal about epistemic modality, which ties the domain of epistemic modals to the notion of local context. I state the motivating data in §2 and the positive account in §3. §4 concludes by discussing some general theoretical morals.

2 Nonclassicality for metaphysical modals: two puzzles

By 'metaphysical modality' I mean the modal flavor expressed by *would*-conditionals, non-epistemic *might*-conditionals, and non-epistemic *might* claims. Some examples are in (3).

(3) a. If the die had landed odd, it would not have landed on two.
b. If the die had landed odd, it might have landed on three.
c. The die might have landed on three.

As often noticed, this flavor of modality involve characteristic tense and aspectual marking, which has come to be called 'X-marking' (Iatridou 2000, von Fintel and Iatridou 2023, a.o.).[1]

Given restrictions of space, here I cannot make a full case for the claim that metaphysical modals are nonclassical. But I present two puzzles. Taken together, they are suggestive of the idea that there is an analogy between epistemic and metaphysical modals.

2.1 Puzzle #1: *would-might* interactions

The first puzzle I discuss is a variant on a puzzle recently presented by Greenberg 2021.[2] To start, consider sentence (4), as uttered in the following scenario.

> A die was tossed and landed 1. You don't know anything about whether the die was fair or biased in any way. I say:

(4) #The die might have landed on 2, and if it had landed even, it would not have landed on 2.

(4) sounds infelicitous. Notice also that this infelicity is generated by the fact that the two conjuncts clash with each other. Given the information available in the scenario, each of the two conjuncts of (4) might be true, and in any case could be uttered felicitously. But, once we conjoin them, we obtain something that is clearly infelicitous.

The first conjunct of (4) is equivalent to *The die might have landed even and 2*. Hence (4) suggests that conjunctions of the form $\ulcorner \Diamond(\mathsf{A} \wedge \mathsf{B}) \wedge (\mathsf{A} \boxrightarrow \neg \mathsf{B}) \vDash \urcorner$ are, in some sense, inconsistent.

Unconditional Might. (UM) $\Diamond(\mathsf{A} \wedge \mathsf{B}) \wedge (\mathsf{A} \boxrightarrow \neg \mathsf{B}) \vDash \bot$

To support this, notice that, like (4), all instances of $\ulcorner \Diamond(\mathsf{A} \wedge \mathsf{B}) \wedge (\mathsf{A} \boxrightarrow \neg \mathsf{B}) \urcorner$ are infelicitous.

(5) a. #Ada and Bashir might have both come to the party, and, if Ada had come, Bashir would not have come.
b. #Coins A and B might have both landed heads, and, if coin A had landed heads, coin B would not have landed heads.
c. #Xintong and Yimei might have both been in DC, and, if Xintong had been in DC, Yimei would not have been in DC.

Interestingly, the infelicity of (4) is unexpected on many standard theories. Assume a simple Lewis-style semantics for counterfactuals with limit assumption.[3]

(6) $[\![\mathsf{A} \boxrightarrow \mathsf{B}]\!]^{w,g} = \text{true iff } \forall w' \in \text{BEST}_{w,g}\ ([\![\mathsf{A}]\!]_g),\ [\![\mathsf{B}]\!]^{w',g} = \text{true}$

Assume also a simple semantics for metaphysical *might*: *might* is an existential quantifier over a contextually determined domain of metaphysically accessible worlds. Then the infelicity of (4) is

[1] For current purposes, I set aside *could* and *could have*. While they are obviously related to *would* and *might*, it is unclear that they express unambigously metaphysical modality. *can*, which is the non-X-marked counterpart of *could*, is notoriously polysemous, and has a habitual, cirumstantial, epistemic, and ability interpretations.

[2] Greenberg's original puzzle involves patterns of entailment between the conjuncts of sentences like (4). Using conjunctions makes things clearer for my purposes, and actually improves the quality of the data.

[3] For discussion of the limit assumption see, among many, Lewis 1973, Stalnaker 1981, Kaufmann 2017. I assume the limit assumption throughout my discussion. Also, a word about notation: 'g' is a variable for orderings over worlds. I use the notation '$[\![\mathsf{A}]\!]_g$)' to denote the set of worlds $\{w : [\![\mathsf{A}]\!]^{w,g}\}$, i.e. essentially the possible worlds proposition expressed by A, relative to ordering g.

unexpected. It can be that the closest A-worlds in the domain are non-B-worlds, but there are some A-and-B worlds further off.

To be sure, some versions of classical theories do predict the infelicity of (4). Assume that metaphysical *might* invariably quantifies over a set of 'maximally similar' worlds.[4] Then the first conjunct of (4) says that one of the most similar worlds to the actual world is a die-landing-two worlds. This is indeed incompatible with standard variably strict truth-conditions. But, as Greenberg points out, making *might* strong in this way creates trouble elsewhere. Sentences of the form $\ulcorner\neg\Diamond(\mathsf{A}\wedge\mathsf{B})\wedge(\mathsf{A}\boxrightarrow\mathsf{B})\urcorner$, such as (7), are wrongly predicted to be consistent.

(7) # The die mightn't have landed 2, and if it had landed even, it would have landed 2.

2.2 Puzzle #2: CEM and Conditional Might

The second puzzle is familiar from the classical debate on counterfactuals (Stalnaker 1968, Stalnaker 1981, Lewis 1973). Counterfactuals seem to validate two plausible principles.

Conditional Excluded Middle. (CEM) $\vDash(\mathsf{A}\boxrightarrow\mathsf{B})\vee(\mathsf{A}\boxrightarrow\neg\mathsf{B})$

Conditional Might. (CM) $(\mathsf{A}\boxrightarrow\mathsf{B})\wedge(\mathsf{A}\diamondrightarrow\neg\mathsf{B})\vDash\bot$

CEM has been subject to extensive empirical discussion; at this point, the case for it is very strong (see, a.o., Klinedinst 2011, Marty, Romoli, and Santorio 2021, Ramotowska et al. 2024). The core evidence for CEM concerns interactions between counterfactuals and other logical operators. In particular, semantics that conform to CEM predict that counterfactuals can commute with a number of logical operators, including negation, without affecting meaning. This prediction is borne out. For example, assuming that *lose* is equivalent to *not win*, the sentences in (8) sound equivalent.

(8) a. Each of these tickets would have lost, if it had been bought. $\forall x(Buy(x)\boxrightarrow\neg Win(x))$
b. None of the tickets would have won, if it had been bought. $\forall x\neg(Buy(x)\boxrightarrow Win(x))$

As for **CM**, the evidence for it is simple: conjunctions like (9) sound straight-up contradictory.

(9) #If the die had landed even, it would have landed 2; and if it had landed even, it mightn't have landed 2.

Unfortunately, **CEM** and **CM** are jointly untenable against a classical background, as they entail that *would* and *might*-counterfactuals are equivalent ($\mathsf{A}\boxrightarrow\mathsf{B}\dashv\vDash\mathsf{A}\diamondrightarrow\mathsf{B}$). And in fact, classical theories for counterfactuals choose one between **CEM** and **CM**, and reject the other. Famously, Stalnaker validates **CEM** and rejects **CM** (Stalnaker 1968, Stalnaker 1981), while Lewis takes the opposite route (Lewis 1973).

3 Bounded semantics for metaphysical modality

I suggest that the puzzles of §2 show that metaphysical modals behave nonclassically. Moreover, the relevant kind of nonclassicality parallels the nonclassical behavior that motivates dynamic theories of epistemic modals. Consider again epistemic contradictions like (1):

(1) #Ada is tall and she might not be.

[4] On some interpretations, this is the view defended by Kratzer (1981, 2012, a.o.). It is unclear to me whether Kratzer would want to endorse the assumption that all occurrences of *might* quantify over closest worlds, without allowing context to relax this constraint.

Theories in the dynamic tradition vary widely, but they share a basic explanation for the infelicity of (1). Roughly, in (1) the right conjunct is evaluated against a set of worlds that has been updated with the other conjunct. As a result, the possibility modal quantifies over a set of worlds where Ada is tall, and the possibility claim is false. This despite the fact that there is a clear sense in which the conjuncts of (1) are consistent. It is consistent that Ada is tall, and that the set of worlds representing speakers' information includes a world where she is not.

I suggest that something structurally analogous happens for metaphysical modality. Consider the problem conjunction (4):

(4) #The die might have landed on 2, and if it had landed even, it would not have landed on 2.

I suggest that the basic truth conditions of the two conjuncts of (4) are simple. The first says that there is a metaphysically accessible world where the die landed on 2. The second says that, in the closest world where the die lands even, it lands on 2. Hence, in a sense, they are consistent. But, in the conjunction, the quantificational domains of the modals are constrained in a way that makes the sentence infelicitous. In particular, the first conjunct is evaluated against a set of worlds that is updated with the second conjunct, i.e. the counterfactual $\ulcorner \textit{even} \boxdot\!\!\rightarrow \neg \textit{two} \urcorner$. Given basic logical facts about counterfactuals[5], this has the effect of removing all even-and-two-worlds from the set. Hence the possibility modal quantifies over a set of worlds where the die did not land on 2, and is guaranteed to be false.

The remainder of the paper sketches a formal implementation of the idea. The system builds on the 'bounded theory' developed by Mandelkern (2019, 2024), *modulo* some simplifications.[6] Aside from the technical analogies, the proposal involves the introduction of a new parameter, i.e. what I call 'metaphysical domain'. I first discuss metaphysical domains, and then present the formal system.

3.1 Metaphysical domains

The literature on presupposition introduces the notion of a local context (Stalnaker 1970, Karttunen 1974, Schlenker 2009 a.o.). Roughly, a local context is a set of epistemic possibilities that are available in the compositional computation, and are partly determined by surrounding sentential material. For example, the local context of the second conjunct in (10) is determined by the initial context set, updated with the proposition that France has a king.

(10) France has a king, and the king of France is a snob.

Mandelkern (2019, 2024) has recently shown that the notion of local context is linked in interesting ways to the semantics of epistemic modality.[7] In particular, the domain of quantification of epistemic modals appears to be constrained by the local context.

I propose that the interpretation of non-epistemic modality is similarly constrained. I introduce a non-epistemic counterpart of local contexts, which I call 'metaphysical domain' (MD for short). The notion of MD can be linked to the notion of openness, which is employed in the semantics for historical modals (see Copley 2009, Cariani and Santorio 2018 a.o.). At every point in time, we can individuate a set of worlds that represent 'open possibilities': roughly, ways that the future course of history might evolve. The set of open worlds at a context, $\textsc{open}_c$,

[5] In particular, the fact that counterfactuals (at least, non-nested counterfactuals) validate Weak Centering, i.e. they entail the corresponding material conditional ($\mathsf{A} \boxdot\!\!\rightarrow \mathsf{B} \vDash \mathsf{A} \supset \mathsf{B}$).

[6] I use Mandelkern's system simply because it is currently the best architecture we have to capture the behavior of epistemic contradictions and related phenomena. In principle, the idea can be implemented in other frameworks.

[7] Strictly speaking, Mandelkern's bounded theory does not exploit a notion of local context, but rather a closely related notion of a bound. Conceptually, the two notions are very closely related.

provides the starting value for the MD parameter at c. This parameter is then shifted by various operator in the course of the compositional computation.

Crucially, some items can expand the metaphysical domain. In particular, several authors have noticed that X-marked conditionals exploit a broader domain of quantification than non-X-marked ones (see e.g. Stalnaker 1975, Stalnaker 1988, von Fintel and Iatridou 2023). I assume that this expansion operation is tied to PAST, and represent its effects via a dedicated operator '↑'. ↑ D is the domain we obtain by expanding D.[8]

The expansion operation introduces a complication. To track both whether the domain has been expanded and what information we have from the surrounding bits of the sentences, we cannot track the MD as a single set of worlds. Rather, we need to use a pair $\langle \text{D}, \text{I} \rangle$. The first member, D (the 'background domain'), tracks whether expansion has happened. The second member, I, represents the information added during composition. The MD is determined by the intersection of the two sets. For a simple example, consider:

(11) The die will land odd and the coin will land even.

I assume that the MD of each conjunct is updated with the information provided by the other conjunct (see below). Hence the MD of the right conjunct is the pair consisting of D, and I augmented with the proposition **odd**. Assuming that the intial values of D and I are, respectively, OPEN_c and the set of all worlds, the MD of the right conjunct is $\text{OPEN}_c \cap \textbf{odd}$, i.e. the set of open worlds at the context where the die will land odd.

3.2 Formal system

The formal system assigns three layers of meaning to each sentence: (i) classical truth-conditionals; (ii) definedness conditions; (iii) an update function for the MD parameter.

Logical Forms. I assume that, at LF, both *would*- and *might*-claims involve an expansion operator '↑'. (Arguably, the expansion operator is realized by 'PAST'; see Santorio 2024 for a theory of X-marking along these lines). For current purposes, I assume that the LFs of (12-a) and (13-a) are in (12-b) and (13-b). I will also use the shorthand in (12-c) and (13-c).[9]

(12) a. If die had landed even, it would have landed 2.
b. ↑ [the die_1 land even $\boxdot\!\!\rightarrow$ it_1 land 2]
c. the die_1 land even $\boxdot\!\!\rightarrow_{\uparrow}$ it_1 land 2

(13) a. The die might have landed 2.
b. ↑ [◇ (the die land 2)]
c. $\Diamond_{\uparrow}$ (the die land 2)

Layer 1: truth-conditions. At the basic layer, the semantics exploits just basic truth conditions. (For simplicity, I ignore the truth-conditional role of tense.)

Counterfactuals are selectional (Stalnaker 1968): a selection function s selects an antecedent-world within a modal base $f(w)$.

(14) $[\![\mathsf{A} \boxdot\!\!\rightarrow_{\uparrow} \mathsf{B}]\!]^{c,s,f,w,\text{D}}$ = true iff (i) it is defined and (ii) $[\![\mathsf{B}]\!]^{c,s,f,s([\![\mathsf{A}]\!] \cap f(w),w),\text{D}}$ = true

I assume that the selection function obeys Centering: whenever $w \in [\![\mathsf{A}]\!]$, $s([\![\mathsf{A}]\!], w) = w$.

Possibility modals work simply as a existential quantifiers over a modal base $f(w)$.

(15) $[\![\Diamond_{\uparrow}\mathsf{A}]\!]^{c,s,f,w,\text{D}}$ = true iff (i) it is defined and (ii) $\exists w' \in f(w)$ s.t. $[\![\mathsf{B}]\!]^{c,s,f,w',\text{D}}$ = true

[8] For current purposes, it doesn't matter exactly what the expanded set is. For convenience, we can just take it to be the set of all metaphysically possible worlds. (This assumption entails that iterating the expansion operator doesn't lead to further expansions of the domain. So far as I can see, this seems a desirable consequence.)

[9] For the case of conditionals, this is a simplification: there is evidence that sentences like (12-a) involve two semantically active occurrences of PAST, which are interpreted locally in the antecedent and the consequent (see Santorio 2024). But this is irrelevant here.

Layer 2: definedness conditions. The second layer assigns recursively stated definedness conditions to sentences. (Definedness conditions can, but need not, be equated with presuppositions.) Following Mandelkern's proposal about epistemic modals (Mandelkern 2019, Mandelkern 2024), the definedness conditions for modals requires that the modal base be a subset of the MD. The expansion operator maps $\textsc{d}$ to an expanded domain $\uparrow\textsc{d}$.

Connectives $\ulcorner\neg A\urcorner$ is defined at $\langle \textsc{d}, \textsc{i}\rangle$ iff A is defined at $\langle \textsc{d}, \textsc{i}\rangle$
$\ulcorner A \wedge B\urcorner$ is defined at $\langle \textsc{d}, \textsc{i}\rangle$ iff A is defined at $\langle \textsc{d}, \textsc{i}\rangle + B$ and B is defined at $\langle \textsc{d}, \textsc{i}\rangle + A$

Exp. op. $\ulcorner\uparrow(A)\urcorner$ is defined at $\langle \textsc{d}, \textsc{i}\rangle$ iff A is defined at $\langle \uparrow\textsc{d}, \textsc{i}\rangle$

Modals $\ulcorner\Diamond(A)\urcorner$ is defined at $\langle \textsc{d}, \textsc{i}\rangle$ iff, for all $w \in \textsc{d} \cap \textsc{i}$, $f(w) \subseteq \textsc{d} \cap \textsc{i}$, and A is defined at $\langle \textsc{d}, \textsc{i}\rangle$
$\ulcorner A \boxdot\!\!\rightarrow B\urcorner$ is defined at $\langle \textsc{d}, \textsc{i}\rangle$ iff (i) for all $w \in \textsc{d} \cap \textsc{i}$, $f(w) \subseteq \textsc{d} \cap \textsc{i}$; (ii) for all $w \in \textsc{d} \cap \textsc{i}$, $s(\llbracket A \rrbracket, w) \in \textsc{d} \cap \textsc{i}$; (iii) A and B are defined at $\langle \textsc{d}, \textsc{i}\rangle$.

Layer 3: updates. Let $\llbracket A \rrbracket_c$ be the proposition expressed by A at c. We have:

(i) For any A that is not of the form '$\uparrow A$': $\langle \textsc{d}, \textsc{i}\rangle + A = \langle \textsc{d}, \textsc{i} \cap (\overline{\textsc{d}} \cup \llbracket A \rrbracket_c)\rangle$

(ii) $\langle \textsc{d}, \textsc{i}\rangle + \uparrow A = \langle \uparrow\textsc{d}, \textsc{i}\rangle + A$

Update with regular sentences is essentially intersective update, with a minor twist.[10] Update with a sentence involving an expansion operator triggers expansion of the background domain.

Consequence. I define consequence as preservation of definedness and truth.

(16) $A_1, \ldots, A_n \vDash B$ iff, for all c s.t. $A_1, \ldots, A_n$ are def. and true at c, B is def. and true at c.

Predictions. Let us consider now some predictions.

(i) Instances of **Unconditional Might**, i.e. sentences of the form $\ulcorner\Diamond_{\uparrow}(A \wedge B) \wedge (A \boxdot\!\!\rightarrow_{\uparrow} \neg B)\urcorner$ are never both defined and true. This explains the infelicity of the sentences in (4) and (5).

(ii) Conversely, sentences of the form $\ulcorner A \wedge \Diamond_{\uparrow}\neg A\urcorner$ like (2) can be true and defined. The key reason (and the key difference with the case of epistemic modals) is that the expansion operator widens the metaphysical domain beyond the initial set $\textsc{open}_c$.[11]

(iii) **Conditional Excluded Middle** and **CM** both come out valid on this system (all instances of **CEM** are true, while instances of **CM** are always true if defined). This vindicates the motivating data in §2.2.

4 Conclusion

Nonclassical accounts of epistemic modals seem to require a major bifurcation in modal semantics: modals of different flavors work in very different ways. The current theory shows that we can recover the unity of modal semantics by generalizing mechanisms used to account for epistemic contradictions. Independently of whether the current system is ultimately correct, the fact that this can be done offers an important proof of concept, and enriches the space of options for our semantics for modality.

[10] The twist is that we update the information parameter with the proposition expressed by A disjoined with the complement of $\textsc{d}$. This is to prevent the update from affecting non-$\textsc{d}$ worlds when the domain is expanded.

[11] The computation of the MD of $\ulcorner A \wedge \Diamond_{\uparrow}\neg A\urcorner$ works as follows. (i) The initial MD of the sentence is $\textsc{open}_c$. (ii) Via the conjunction clause, the MD of the right conjunct is $\langle \textsc{d}, \textsc{i}\rangle + A$, i.e. (via the definition of update) $\langle \textsc{d}, \textsc{i} \cap (\overline{\textsc{d}} \cup \llbracket A \rrbracket_c)\rangle$. At this point, the MD consists of only A-worlds. (iii) But the expansion operator in the second conjunct expands the MD to $\langle \uparrow\textsc{d}, \textsc{i} \cap (\overline{\textsc{d}} \cup \llbracket A \rrbracket_c)\rangle$, hence to $\uparrow\textsc{d} \cap (\textsc{i} \cap (\overline{\textsc{d}} \cup \llbracket A \rrbracket_c))$. This set can include non-A worlds.

References

Cariani, Fabrizio and Paolo Santorio (2018). "Will Done Better: Selection Semantics, Future Credence, and Indeterminacy". In: *Mind* 127.505, pp. 129–165.

Copley, Bridget (2009). *The Semantics of the Future*. New York: Routledge.

von Fintel, Kai and Sabine Iatridou (2023). "Prolegomena to a theory of X-marking". In: *Linguistics and Philosophy* 46.6, pp. 1467–1510.

Greenberg, Gabriel (2021). "Counterfactuals and modality". In: *Linguistics and Philosophy* 44.6, pp. 1255–1280.

Iatridou, Sabine (2000). "The grammatical ingredients of counterfactuality". In: *Linguistic inquiry* 31.2, pp. 231–270.

Karttunen, Lauri (1974). In: *Theoretical Linguistics* 1.1-3, pp. 181–194. DOI: doi:10.1515/thli.1974.1.1-3.181. URL: https://doi.org/10.1515/thli.1974.1.1-3.181.

Kaufmann, Stefan (2017). "The Limit Assumption". In: *Semantics and Pragmatics* 10.18. DOI: 10.3765/sp.10.18.

Klinedinst, Nathan (2011). "Quantified Conditionals and Conditional Excluded Middle". In: *Journal of Semantics* 28.1, pp. 149–170. DOI: 10.1093/jos/ffq015.

Kratzer, Angelika (1981). "Partition and Revision: The Semantics of Counterfactuals". In: *Journal of Philosophical Logic* 10.2, pp. 201–216.

— (2012). *Modals and Conditionals: New and Revised Perspectives*. Vol. 36. Oxford University Press.

Lewis, David K. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Mandelkern, Matthew (2019). "Bounded Modality". In: *Philosophical Review* 128.1, pp. 1–61. DOI: 10.1215/00318108-7213001.

— (2024). *Bounded Meaning: The Dynamics of Interpretation*. Oxford University Press.

Marty, Paul, Jacopo Romoli, and Paolo Santorio (2021). "Counterfactuals and Undefinedness: Homogeneity vs. Supervaluations". In: *Proceedings of SALT*. Vol. 30, pp. 603–623.

Ramotowska, Sonia et al. (2024). "Undefined Counterfactuals: Experimental Evidence for Selection Semantics". draft.

Santorio, Paolo (2024). "Modalizing the Past: A Theory of C-Marking". draft, University of Maryland, College Park.

Schlenker, Philippe (July 2009). "Local Contexts". In: *Semantics and Pragmatics* 2.3, pp. 1–78. DOI: 10.3765/sp.2.3.

Stalnaker, Robert (1968). "A Theory of Conditionals". In: *Studies in Logical Theory*. Ed. by N. Recher. Oxford.

— (1970). "Pragmatics". In: *Synthese* 22.1-2, pp. 272–289. DOI: 10.1007/bf00413603.

— (1975). "Indicative Conditionals". In: *Philosophia* 5.

— (1981). "A Defense of Conditional Excluded Middle". In: *Ifs*. Ed. by William Harper, Robert C. Stalnaker, and Glenn Pearce. Reidel, pp. 87–104.

— (1988). "Belief Attribution and Context". In: *Contents of Thought*. Ed. by Robert H. Grimm and Daniel Davy Merrill. Tucson, pp. 140–156.

Veltman, Frank (1985). "Logic for Conditionals". PhD thesis. University of Amsterdam.

Willer, Malte (2013). "Dynamics of Epistemic Modality". In: *Philosophical Review* 122.1, pp. 45–92.

Yalcin, Seth (2007). "Epistemic Modals". In: *Mind* 116.464, pp. 983–1026.

Age stratification in collective readings of Present Day English *every*-DPs

Fabian Schlotterbeck
University of Tübingen
Tübingen, Germany
fabian.schlotterbeck@uni-tuebingen.de

Polina Berezovskaya
University of Tübingen
Tübingen, Germany
polina.berezovskaya@uni-tuebingen.de

Sigrid Beck
University of Tübingen
Tübingen, Germany
sigrid.beck@uni-tuebingen.de

Abstract

We present an experimental study on English determiner phrases (DPs) containing *every* in combination with collective sentence predicates. The experiment is based on the hypothesis that *every*-DPs are undergoing semantic change towards a referential plural interpretation. The results support this hypothesis. In particular, change in progress is visible from variable acceptability depending on the type of predicate, and an age gradation effect. In addition, the results indicate that there is an unexpected asymmetry between collective interpretations with *every*-DPs in subject vs. object position.

1 Introduction

English DPs with the determiner *every* can be used in sentences like (1). This is unexpected given the standard Generalized Quantifier semantics (cf. Barwise and Cooper 1981) of *everyone* in (2-a), predicting that the DP cannot be combined with a collective predicate like *gather*, which requires a plural subject, cf. (3). Our study tests the hypothesis that Present Day English *every* is in the process of acquiring a referential plural reading, (2-b), where *everyone* denotes the group containing all (relevant) persons (cf. Champollion 2010).

(1) a. Everyone gathered in the hallway.

(2) a. $[\![everyone]\!] = \lambda Q.\forall x[person(x) \rightarrow Q(x)]$ b. $[\![everyone]\!] = max(^*\lambda x.person(x))$

(3) a. The students / Amy, Bill and Cyril gathered in the hallway.
b. *The student / Amy gathered in the hallway.

Collective predicates can be subdivided into the classes in (4) (Brisson 2003; Dowty 1987). The corpus study and the informant judgments reported in Beck (in prep.) led us to believe that the use of *every*-DPs in sentence contexts like (1) is an innovation, and that the type of collective predicate matters for the acceptability of *every*-DPs. Beck compared a historical corpus with data up until 1914 to a modern corpus and found that plural *every*-DPs are more common today than they used to be. They occur in the modern corpus with subdistributive predicates. We conjectured that plain collective predicates, as in (4-b/c), are not (yet) acceptable with *every*-DPs today, cf. (5), which contrasts with the subdistributive predicate in (1) above. We further conjectured that the prospective change is more advanced in younger speakers than in older speakers (the age gradation or Apparent Time Hypothesis according to which such effects may be indicative of change in progress; see Labov 1963; 1966; 1994; see also e.g. D'Arcy 2005 for a more recent example study and discussion). These conjectures are visualized in Fig. 1. Finally, we test if the change is more advanced in *everyone*/*everybody* than in *every*+NP, compare (5-c).

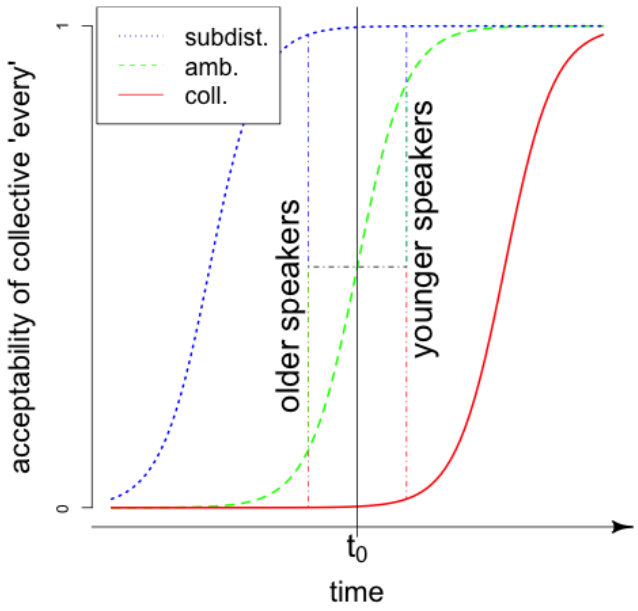


Figure 1: Inconic sketch of hypothesized stepwise change in progress for different types of predicates. coll.: plain collective; amb.: ambiguous collective; subdist.: subdistributive predicates; t_0 stands for the present day.

(4) a. The students *gathered*. (subdistributive)
b. The students are *numerous*. (plain collective (subject))
c. The workstations *outnumber* the programmers. (plain collective (sub., obj.))
d. Irene *weighed* the kittens. (ambiguous (obj.))
e. The children *bought* the house. (ambiguous (sub.))

(5) a. *Every student is numerous.
b. *The workstations outnumber every programmer.
c. ?Every programmer gathered in the hallway.

In summary the hypotheses we aimed to test are (i) that Present Day English *every* is in the process of acquiring a referential plural reading; (ii) that the type of collective predicate matters for the acceptability of *every*-DPs; and (iii) that the prospective change is more advanced in younger speakers than in older speakers.

2 Methods

To test our hypotheses, we conducted a web-based experiment in which participants read contexts and then judged the acceptability of a phrase that was targeting a crucial (in some conditions disambiguating) bit of the context (e.g. the weighing event in the example item in (6) below). Acceptability was operationalized as a binary judgment that participants provided by pressing one of two response keys. Response key assignments were counterbalanced between participants. We tested the different types of predicates in (4). In particular, we tested three predicates of each predicate type (subdistributive: *be spread out*, *gather in the hallway*, *crowd around the monitor*; collective subject: *be numerous*, *be a good team*, *be politically homogeneous*; collective object: *the workstations outnumber* _, *the committee is formed by* _, *Alex compared* _ ; collective subject of ambiguous predicate: *carry the piano upstairs*, *own the house*, *write the play*; collective object of ambiguous predicate: *I played against* _, *I weighed* _, *the lead ball outweighs* _). We embedded the 15 predicates in sentence constructions in which a collective interpretation of *everyone* (as

in (6-ii)) and *every*+NP (as in (6-iii)) was made salient by the context.[1] These constructions were compared to a clearly acceptable control with a definite plural subject (e.g., *weigh the guests* in (6-i)) and a clearly unacceptable control (e.g. *weigh the explosion* in (6-iv)). The controls served the purpose of providing a baseline of clearly acceptable and clearly unacceptable verb-argument combinations against which responses in the target conditions could be compared. We generated 15 items (3 per PREDICATE TYPE) and collected data from 199 participants in five AGE GROUPS ($18-30:40$; participants $31-40:39$; $41-50:40$; $51-60:40$; $61-75:40$) via the platform `prolific.co`. We had originally aimed for 40 participants per AGE GROUP, but one submission from the $31-40$ range was not returned. Altogether this yielded a mixed $4 \times 5 \times 5$ design in which we manipulated the three factors CONSTRUCTION (cf. (6), manipulated within items; 4 levels: *acceptable control*, *everyone*, *every+NP*, *unacceptable control*), PREDICATE TYPE (cf. (4) manipulated between items; 5 levels: collective (amb, obj), collective (amb, subj), subdistributive, collective (obj) and collective (subj)) and AGE GROUP (manipulated between participants; see above). Using a Latin square, the items were distributed across four lists in such a way that each participant saw each item (i.e. each predicate; e.g. *weigh*) in only one CONSTRUCTION. Each participant thus saw one fourth of the entire set of the materials. A set of 17 filler items from an unrelated sub-experiment were added to each list. Lists were randomly assigned to participants.

(6) Example item (with *collect. amb. obj* predicate)

Context: You own a small boat on which you take people fishing. A family of three has hired you and your boat by phone for the afternoon. When you see them at the harbour, you get worried because they are big people. Your boat can carry 320kg and you weigh 60kg. Fortunately, the harbour owns a scale. You put the three people on the scale and read off the weight. The scale says 250kg. You decide to go ahead with the trip. As you board, your cousin Erica passes by. She looks worried and asks you if you aren't taking on a bit too much.

Target: Can you say the following?

(i)	I weighed the guests.	*acceptable control*
(ii)	I weighed everyone.	*everyone*
(iii)	I weighed every guest.	*every+NP*
(iv)	I weighed the explosion.	*unacceptable control*

To analyze the binary acceptability judgments statistically and specifically test for age effects, we used a form of *penalized logistic regression* (cf. Firth 1993). In particular, we computed so-called Bayesian generalized linear mixed-effects models as implemented in the `blme` package (Chung et al. 2013). This type of analysis is similar to analyses based on logit mixed-effects models, but, in addition, allows the specification of prior distributions for Bayesian maximum a posteriori inference on fixed effects. These prior distributions can be used to penalize large absolute parameter estimates that may result from separation in the data (e.g. if there is no variation in responses in some cells of the experimental design; see Clark et al. 2023 for discussion of linguistic applications in studies involving quasi-complete separation). We included fixed effects of the factors PREDICATE TYPE, AGE GROUP and CONSTRUCTION and their interactions as well as random intercepts of items and participants (given their inclusion allowed for convergence). The fixed effect of PREDICATE TYPE was dummy coded with *subdist* as reference level, i.e. all other PREDICATE TYPES were compared to *subdist*. For the other two factors we specified the following contrasts. For CONSTRUCTION, we compared *unacceptable controls* to the mean ratings of the other CONSTRUCTIONS (contrast *c1*), the *acceptable controls* to *everyone* (contrast *c2*), and the *acceptable controls* to *every*+NP (contrast *c3*). For AGE GROUP, we compared mean ratings in the $18-30$ and $31-40$ groups to the other groups (contrast *ag1*), the $18-30$ group

[1] Materials, data and analyses are made available at https://osf.io/gnjqt.

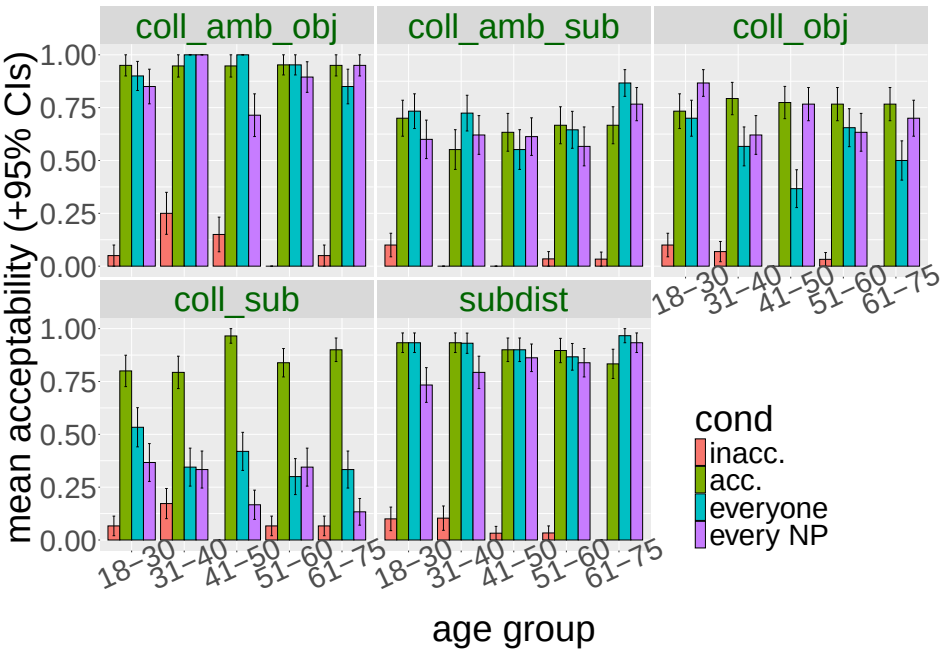
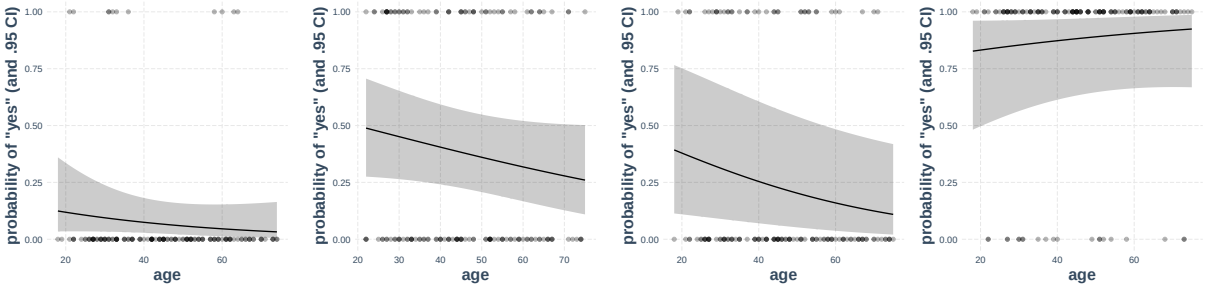


Figure 2: Mean ratings and 95% CIs by PREDICATE, AGE GROUP and CONDITION. CIs were approximated using the `summarySE` function in the `Rmisc` package (Hope 2022). acc: acceptable; inacc: inacceptable.

to the $31-40$ group (contrast *ag2*), the $41-50$ and $51-60$ groups to the $61-75$ group (contrast *ag3*) and the $41-50$ to the $51-60$ group (contrast *ag4*). In addition, we conducted two post-hoc analyses. The first post-hoc analysis focussed on the one PREDICATE TYPE were we found an indication of the age gradation effect (see results below). In this PREDICATE TYPE, we analyzed the four CONSTRUCTIONS separately and regressed response probabilities against the AGE of each individual participant (except for one who did not provide the information) as a (quasi) continuous predictor. The purpose was to further test for age gradation in acceptability. The second post-hoc analysis compared *every*+NP to *everyone* across PREDICATE TYPEs and AGE GROUPS. For this analysis, we do not report p-values because they are not independent of the contrasts tested in the main analysis.

3 Results

Mean ratings are shown in Fig. 2. There was substantial variation between predicates. The *every*-conditions patterned with the *acceptable controls* in subdistributive (88% vs. 90% on average, resp.) and ambiguous predicates (*coll amb obj.*: 91% vs. 95%; *coll amb sub.*: 67% vs. 64%;), but led to substantially lower ratings than the controls in collective predicates (*coll obj.*: 64% vs. 77%; *coll sub.*: 33% vs. 86%;). The statistical analysis revealed a marginally significant three-way interaction (model comparison: $\chi^2(48) = 61.57, p = .09$) which was driven by significant interaction terms in the fitted model that involved AGE GROUPs and CONSTRUCTIONs wrt. collective predicates (e.g. (4-b) and (4-c)). To resolve the three-way interaction, we analyzed the different PREDICATE TYPES separately. We found that *unacceptable controls* received significantly lower ratings than the other CONSTRUCTIONs across PREDICATE TYPEs (all $|z| > 7$). In addition, we observed differences between the *acceptable controls* and the universal conditions in all PREDICATEs except *coll. amb. sub.* In particular, the condition with *everyone* received lower ratings than the *acceptable controls* in *coll. sub.* ($c2 : z = 4.20, p < .001$) and *coll. obj.* ($z = 4.07, p < .001$). The condition with *every*+NP received lower ratings than



(a) *unaccaptable control* (b) *everyone* (c) *every*+NP (d) *acceptable control*

Figure 3: Results of analysis regressing response probabilities against AGE in the four CONSTRUCTIONs involving *coll. sub.* PREDICATE TYPEs. The lines show model predictions. The shaded area represents 95% confidence intervals around these predictions, which were calculated based on the fixed effects in the logit mixed effects models. The dots above and below the regression line represent the given responses.

.

the *acceptable control* in *coll. sub.* ($c3 : z = 6.88, p < .001$), *coll. amb. obj* ($z = 1.98, p = .048$) and *subdist.* ($z = 2.42, p = .015$). Moreover, there were effects of AGE GROUP in the two ambiguous PREDICATE TYPES, with relatively high ratings across constructions for the AGE GROUP $31 - 40$ in *coll. amb. obj.* ($z = 2.56, p = .011$) and for the AGE GROUP $61 - 75$ in *coll. amb. sub.* ($z = -2.64, p = .008$). Crucially, the separate analysis of PREDICATE TYPEs revealed an interaction between AGE GROUP and CONSTRUCTION but only in the *coll. sub.* PREDICATE TYPE ($\chi^2(12) = 22.95, p = .028$), where we found an increase in ratings of the *every*+NP condition as compared to the *acceptable control* for the youngest two AGE GROUPs ($c3 \times ag1 : z = -1.92, p = .055$) and an increase for the $41 - 50$ vs. the 51-60 AGE GROUP ($c3 \times ag4 : z = -2.31, p = .021$).

The first post-hoc analysis, regressing response probabilities against AGE in the *coll. sub.* PREDICATEs, revealed that the probability of accepting the target sentence decreased significantly with AGE in the *every*+NP CONSTRUCTION (one-tailed test: $\beta = -.029, z = -2.16, p = .015$; see Fig. 3c) and marginally in the *everyone* CONSTRUCTION ($\beta = -.019, z = -1.56, p = .059$; see Fig. 3b). In the control conditions, no significant association with AGE was found (*unacceptable*: $\beta = -0.026, z = -1.2, p = .116$; see Fig. 3a; *acceptable*: $\beta = 0.016, z = 1.04, p = .15$; see Fig. 3d). The second post-hoc analysis comparing *every*+NP and *everyone* directly across AGE GROUPs and PREDICATE TYPEs revealed an indication that the difference between *every*+NP and *everyone* depended on the PREDICATE TYPE ($\chi^2(4) = 21.63$) but not on AGE GROUP ($\chi^2(4) = 0.96$). Separate analyses of the PREDICATE TYPEs showed that, in *subdist* and *coll. sub.* PREDICATE TYPEs, *everyone* was judged acceptable more often than *every*+NP ($z = -2.13$ and $z = -2.4$). There was no indication of significant differences in any of the other PREDICATE TYPEs, but a numerical trend in the opposite direction in *coll. obj.* ($z = 1.74$).

4 Discussion

We observed relatively high acceptance of collective readings of *every*-DPs across the board. This indicates that the change in progress that our study anticipated is relatively advanced; i.e. in the iconic Fig. 1, the solid black bar, i.e. t_0, should be shifted towards the right in our case study. The high acceptability was, however, accompanied by substantial variation between predicate types. The observation that not all sentence contexts show acceptability of the new meaning of *every*-DPs to the same degree supports our hypothesis of change in progress. In particular, we derived from the semantic analysis developed in (Beck in prep.) that the acceptability would be higher or equal in subdistributive vs. ambiguous and ambiguous vs. plain collective predicate

types (all relative to the corresponding control conditions). The change in progress hypothesis is further substantiated by the age gradient in the acceptability of *every*+NP in the *coll. sub.* predicate type (cf. (5-b)). We predicted the existence of age gradients based on the Apparent Time Hypothesis (cf. the classical work on language change by Labov 1963; 1966; 1994). Given the assumed succession of the change (cf. Fig. 1), it was not surprising to us that we found an age gradient only in collective, but not ambiguous or subdistributive predicate types. The idea is that the change is already too advanced in the latter two types of predicates for an age gradient to be detectable. In the following paragraph, we provide a brief description of key assumptions by Beck (in prep.) that led us to expect variation between predicates. Afterwards, we address an explorative aspect of our study and provide an explanation for the observed asymmetry between plural denoting DPs in subject vs. object position. What we did not find, despite small numerical differences, was reliable evidence for further development of *everyone* vs. *every*+NP.

At the outset of the change, the new meaning of *every*-DPs requires a distributive sentence context. With Beck (in prep.), we implement this as an uninterpretable feature [u∀] on the DP. A distributive operator DIST (e.g. Link 1983) carries the required [i∀] feature. The resulting meaning is indistinguishable from a Generalized Quantifier interpretation, as shown in (7). This requirement is dropped in a stepwise fashion. Subdistributive predicates contain a distributive subcomponent. Following Brisson (2003), we analyze them as containing DIST in a subordinate position, see (8). The semantic change becomes visible when such a subordinate DIST becomes an available licensing context for *everyone*[u∀]. Speakers generally accept the plural interpretation in (8) with this weakened requirement today. Some speakers, predominantly the youngest ones, accept *every*-DPs without any distributivity in the sentence context, i.e. with plain collective predicates. In those speakers, the change is the most advanced. There is no longer any distributivity requirement, i.e. the feature [u∀] is dropped and LFs without plural operators like DIST, cf. (9), become grammatical.

(7) a. Everyone smiled.
b. everyone [u∀] [DIST[i∀] smiled]
c. $\forall x[x \leq max(*person) \rightarrow smiled(x)]$

(8) a. Everyone gathered.
b. [IP everyone [u∀] [1 [[DIST[i∀] t_1 DO] [PRO_1 be-gathered]]]]
c. $\forall x[x \leq max(*person) \rightarrow x$ did something that lead to $max(*person)$ being together.]

(9) a. ??Everyone is numerous.
b. [IP everyone [VP is numerous]]

Apart from variation between predicates, we also found differences between cases where the collective *every*-DP appears in subject vs. object position, i.e. we observe an asymmetry between subjects and objects. In particular, collective *every*-DPs seem more acceptable in object vs. subject position. This finding was not predicted. We think an explanation may be provided based on syntactic agreement. The new meaning of *every*-DP leads to a mismatch between syntax and semantics. The NP sister of *every* has to be semantically plural on the new interpretation of *every* as *max*, but it is syntactically singular. When the *every*-DP is in subject position, because of subject-verb agreement, the finite verb should also be marked as plural, but it is syntactically singular. When the *every*-DP is in object position, this second mismatch does not arise. We conjecture that the overall higher acceptability of *every*-DPs in collective object positions is due to this difference between subject and object in English. In sum, our paper presents an experimental study that applies classical methods from diachronic linguistics to a semantic phenomenon and presents a formal semantic analysis of the change in progress. Strikingly, the change targets one of the archetypal Aristotelian quantifiers in natural language, English *every*.

Acknowledgements. We thank Robin Edds for help with the English experimental items.

References

Barwise, John and Robin Cooper (1981). "Generalized quantifiers and natural language". In: *Linguistics and Philosophy* 4.2, pp. 159–219. DOI: 10.1007/bf00350139.

Beck, Sigrid (in prep.). *The Universal Semantic Cycle*. Ms., University of Tübingen.

Brisson, Christine (2003). "Plurals, all, and the non-uniformity of collective predication". In: *Linguistics and Philosophy* 26, pp. 129–184. DOI: 10.1023/A:1022771705575.

Champollion, Lucas (2010). "Cumulative readings of *every* do not provide evidence for events and thematic roles". In: *Proceedings of the 17th Amsterdam Colloquium*. Vol. 4. 2. Springer, pp. 213–222. DOI: 10.1007/978-3-642-14287-1_22.

Chung, Yeojin et al. (2013). "A nondegenerate penalized likelihood estimator for variance parameters in multilevel models". In: *Psychometrika* 78.4, pp. 685–709. DOI: 10.1007/s11336-013-9328-2.

Clark, Robert G. et al. (2023). "Dealing with complete separation and quasi-complete separation in logistic regression for linguistic data". In: *Research Methods in Applied Linguistics* 2.1, pp. 1–11. DOI: 10.1016/j.rmal.2023.100044.

D'Arcy, Alexandra (2005). "Like: syntax and development". PhD thesis. University of Toronto.

Dowty, David (1987). "Collective predicates, distributive predicates, and all". In: *Proceedings of the 3rd ESCOL*. Ed. by F. Marshall. Vol. 4. Ohio State University.

Firth, David (1993). "Bias reduction of maximum likelihood estimates". In: *Biometrika* 80.1, pp. 27–38. DOI: 10.2307/2336755.

Hope, Ryan M. (2022). *Rmisc: Ryan Miscellaneous*. R package version 1.5.1. URL: https://CRAN.R-project.org/package=Rmisc.

Labov, William (1963). "The social motivation of a sound change". In: *Word* 19.3, pp. 273–309. DOI: 10.1080/00437956.1963.11659799.

— (1966). *The Social Stratification of English in New York City*. 2nd ed. Cambridge University Press. ISBN: 978-0-521-82122-3.

— (1994). *Principles of Linguistic Change*. Oxford, UK: Blackwell.

Link, Godehard (1983). "The logical analysis of plurals and mass terms: a lattice-theoretical approach". In: *Meaning, Use, and Interpretation of Language*. Ed. by Rainer Bäuerle, Christoph Schwarze, and Arnim von Stechow. Berlin, Boston: De Gruyter, pp. 302–323. DOI: doi:10.1515/9783110852820.302.

What appositives can tell us about names and definite descriptions

Eli Sharf
University of California, Santa Cruz
Santa Cruz, United States
ebsharf@ucsc.edu

Abstract

Previous research has shown that while hearer-new unique definites are felicitous on their own, hearer-new names and incomplete definites are typically infelicitous (Sommers 1982; Roberts 2003). This paper presents two new observations regarding how appositives interact with these referring expressions. First, the infelicity of both hearer-new names and incomplete definites can be ameliorated by appositives that, pre-theoretically, clarify the reference of their anchor. Second, hearer-new incomplete definites are more constrained than names in terms of which appositives can effectively clarify their reference. I account for these observations by assuming that definites—but not names—conventionally encode uniqueness. These semantics are integrated into the system of Beaver and Coppock (2015). In this system, appositives disambiguate the values of indices attached to referring expressions, in some cases circumventing a pragmatic principle that adjudicates between competing logical forms.

1 Introduction

It has long been observed that using a hearer-new name out of the blue is infelicitous[1], while a hearer-new description is felicitous (Sommers 1982; Maier 2009). This contrast is shown below:

(1) **Coffee Shop**: *Ana and Bella just met for the first time at a coffee shop, and are becoming fast friends. Bella says she's considering getting a degree in psychotherapy. Ana's son goes to a school that employs a therapist named Mia, but she hasn't shared this information with Bella yet. Ana intends to put Bella in touch with Mia.*

a. #You should talk to Mia about this.
b. You should talk to the therapist at my kid's school about this.

This paper is concerned with the ability of certain appositives to ameliorate this infelicity:

(2) **Coffee Shop**: You should talk to Mia, the therapist at my kid's school, about this.

Pre-theoretically, the appositive is used to clarify the reference of its **anchor** (the modified NP).

I identify novel observations about which appositives can succesfully ameliorate the infelicitous use of hearer-new names and definites. With names, I show that appositives have to (i) connect the name to some individual in the context (in a specific sense) or (ii) communicate that the intended referent of the name is not in the context. Then, I show that incomplete definites are infelicitous in parallel contexts to names, but the range of appositives that can be used to clarify their reference in these contexts is much more restricted.

I present an analysis of these observations in the system of Beaver and Coppock (2015). The core observations will be accounted for by an analysis where definite descriptions, but not names, conventionally encode uniqueness, and where appositives disambiguate the value of indices attached to referring expressions, in some cases circumventing Beaver and Coppock (2015)'s comprehension-based variant of Maximize Presupposition.

[1]The judgment marked by # is that knowing the facts at hand, it would be entirely uncooperative to use the name. Note that there are certain intonational contours, and other contextual strategies, that alleviate the infelicity of (1-a). This judgment only applies to (1-a) used as a canonical assertion with unmarked intonation.

2 Familiar and Novel, Old and New

In order to explain constraints on hearer-new names, I need to define "hearer-new". The definition here piggybacks on a theory of discourse reference outlined in Roberts (2003) (**weak familiarity**), which treats discourse referents (drefs) as individuated contextual representations linked to specific existence entailments in the common ground. These entailments can arise through linguistic mention, joint perceptual experience, or common knowledge. I will use this conception of discourse referents to analyze the data in the next section, as well as in the proposal.

Definition 2.1. Hearer old/new. The use of a referring expression E in context C is **hearer-old** if the speaker intends E to refer to a weakly familiar dref in C that is presupposed to satisfy the informational content of E in C. Otherwise it is **hearer-new**.

I assume the informational content of a name is bearing the name and the informational content of a description is the property in the restriction of *the*. In (1), the uses of *Mia* and *the therapist at my kid's school* are hearer-new, since the speaker intends to refer to a **novel dref**. But uses of referring expressions can be hearer-new even if the speaker intends to use the expression to refer to a (weakly) familiar dref, so long as that dref is not presupposed to bear the name or meet the property in the description. Hearer-new names are also infelicitous in this latter case:

(3) **Party**: *Adam takes Ben, a new friend, to a party. They are greeted at the door by Mia, who does not introduce herself, so Adam knows that Ben does not know her name. Later on, Ben tells Adam that he is considering getting a degree in psychotherapy.*
Adam: #[Without pointing.] You should talk to Mia about this.

Example (3) demonstrates that the issue with the hearer-new name in (1-a) is not about the intended referent of the name being novel, it's that there isn't a way to link the name to any familiar dref because no familiar dref is presupposed to bear the name.

3 Which appositives can clarify the reference of their anchor?

3.1 Names

One group of appositives that facilitate the use of hearer-new names are appositives that specify a weakly familiar dref. Many definite nominal appositives are in this group, including those used anaphorically (4), deictically (5), and those that pick out an entity in the common ground (6).

(4) I'm excited for the dinner party tonight. A fireman is coming I want to talk to. Also, the dishes will be so good. I'm making chicken and Joe #(, the fireman,) is making a salad.

(5) *The speaker and hearer are looking at a number of women.*
Mia #([*pointing*], the girl with blue hair,) is coming over later.

(6) *The hearer knows that the UK has a prime minister but doesn't know his name.*
Keir Starmer #(, the prime minister of the UK,) is currently visiting the United States.

Appositive relative clauses can also serve this purpose. Consider the sentences below:

(7) **Party** (*see (3) above for context*)
a. You should talk to Mia, who greeted us on the way in, about this.
b. #You should talk to Mia, who knows a lot about therapy, about this.

The appositive *who greeted us on the way in* specifies a referent that is weakly familiar in virtue of Adam and Ben's joint experience of her earlier in the party. The appositive *who knows a lot about therapy* doesn't specify any weakly familiar dref, and the infelicity remains.

The other group of appositives that facilitate the use of a hearer-new name are appositives that communicate that the intended referent of the name is not weakly familiar, i.e., novel. One such case was in (2): since being *the therapist at Ana's kid's school* provides a source for Ana's, but not Bella's, acquaintance with the referent, the appositive signals that the referent is novel to Bella (and thus to the common ground). This group also includes certain indefinite nominal appositives, such as (8-a) but not (8-b), used in a variation on the **Coffee Shop** scenario:

(8) **Coffee Shop-Friend**: *Ana and Bella just met for the first time at a coffee shop, and are becoming fast friends. Bella says she's considering getting a degree in psychotherapy. Ana has a friend named Mia who's a therapist, but she hasn't shared this information with Bella yet. Ana intends to put Bella in touch with Mia.*

a. You should talk to Mia, a friend of mine, about this.
b. #You should talk to Mia, a nice woman, about this.

The property in (8-a) provides an exclusive basis for acquaintance in the same manner as (2), and in so doing communicates that the referent is novel to the context. The property in (8-b) does provide such a basis, and the infelicity remains.

Notice that with both groups of appositives above, the speaker communicates that they are aware that there is no familiar dref that is presupposed to bear the name – i.e., they are aware that the name is hearer-new. With the first group, this is accomplished by (i) attributing the name to some familiar dref, in conjunction with (ii) the requirement that appositives be non-trivial (Potts 2004). With the second group, this is accomplished by attributing the name to a novel dref. In the analysis in section **4**, I will argue that this implication allows the speaker to circumvent the default application of a pragmatic rule that legislates between logical forms.

3.2 Definites

I have already shown that definite descriptions can be hearer new in (1-b). This is only possible if the description is **semantically unique** – if there is only one satisfier in the world (Roberts 2003; Beaver and Coppock 2015). Indeed, what makes the description in (1-b) felicitous is the fact that it is somewhat regular for schools to employ a single therapist. It's infelicitous to use a hearer new description in this context for which such a uniqueness inference is unavailable (#*You should talk to the teacher at my kid's school.*).

Incomplete definites – definites that are not semantically unique – cannot be hearer-new (Roberts 2003; Beaver and Coppock 2015). We can see this in the **Coffee Shop** scenario:

(9) **Coffee Shop**: #You should talk to the therapist about this.

So long as Bella does not know about the therapist at Ana's kid's school, it is infelicitous for Ana to refer to her using an incomplete definite.

Crucially, however, the use of hearer-new incomplete definites cannot be facilitated by the same appositives as names:

(10) **Coffee Shop-Friend** (*see (8) above for context*)

a. (8-a) You should talk to Mia, a friend of mine, about this.
b. #You should talk to the therapist, a friend of mine, about this.

Communicating that the referent is novel is not enough to ameliorate its infelicity.

Despite this contrast, there are appositives that can clarify the reference of an incomplete definite. In this paper I will only discuss the example below, used in a variation on the **Party** scenario in (3). In this example, a hearer-new incomplete definite is accompanied by subsequent verbal instructions for how to pick out the referent of the description in the context:

(11) *Adam takes Ben, a new friend, to a party. Adam knows that there's a single therapist at the party, and that Ben doesn't know this. This therapist is standing across the room from them. Ben tells Adam that he is considering getting a degree in psychotherapy.*
You should talk to the therapist, [*pointing*] who is over there in the red hat, about this.

To summarize the data above: hearer-new names are infelicitous (1-a), but can be made felicitous by any appositive that links the name to a specific (weakly) familiar or novel dref. While hearer-new unique definites are felicitous (1-b), incomplete definites are not (9). Unlike names, the reference of an incomplete definite cannot be clarified by linking the name to a novel dref (10-b), but can be clarified by a deictically used appositive (11).

4 Deriving the observations

Beaver and Coppock (2015) propose a semantics for definite descriptions intended to capture, among other things, the inverse relationship between uniqueness and familiarity in definite descriptions exemplified by the contrast between (1-b) and (9). Below, I walk through their theory, and then show how, with a few additional assumptions, it can derive the data above.

Beaver and Coppock (2015)'s theory is built off their previous work in Coppock and Beaver (2015), which proposes that the uniqueness of definite descriptions is two-fold: first, there is **definiteness** – a conventional presupposition encoded by the definite determiner *the* that *at most one* (≤ 1) entity in the context satisfies its informational content. Second, there is **determinacy** – a presupposition of uniqueness + existence, or *exactly one* ($= 1$), contributed via a type-shifter **Iota** that applies to noun phrases in argument positions. This division is intended to account for argumental uses of definites that presuppose uniqueness but not existence (e.g., in *Mary didn't score the only goal*, there is not just one goal). This is derived in the following way: in this theory, two different type-shifters are available: **Iota** (the determinate reading) and **Ex** (the indeterminate reading); the latter is an existential generalized quantifier. A general principle prefers **Iota** over **Ex**, but this principle can be overridden so long as there is explicit evidence in the message that the presupposition of **Iota** cannot be met. In the case of *Mary didn't score the only goal.*, this evidence is focus on *only*, which, in conjunction with negation, signals that the focus alternative – *multiple goals* – is satisfied. This means that the sentence explicitly denies the existence of an *only goal*, overriding the default use of **Iota** over **Ex**. The general principle that regulates the use of **Iota** and **Ex** is defined in Beaver and Coppock (2015) as below:

(12) **Maximize presupposition (comprehension principle)**: Suppose two LFs for a sentence are identical except for one item which could take values A or B differing only in that A has stronger presuppositions than B. Then, in the absence of evidence to the contrary [as to what the speaker presupposes], prefer the interpretation with A, if necessary [accommodating in order to meet] that interpretation.

This principle is distinct from the classic principle in Heim (1991), as it legislates between competing logical forms rather than expressions like *the* and *a*.

Beaver and Coppock (2015) embed **Iota** and **Ex**, as well as the conventional meaning of *the* (≤ 1), into a system called *Partial File Logic* (*PFL*), where formulas are interpreted relative to input and output sequences. This semantic system is wrapped in an update mechanism, where formulas are evaluated on each world-sequence $< w, f >$ pair in a "file" F representing the common ground; what's important for our purposes is that if any $< w, f > \in F$ does not meet a presupposition of a given formula, the file F as a whole will not **accept** the formula.[2]

[2]See Beaver and Coppock (2015) for the full formal theory. This section will only define additions and modulations of the system.

In this system, noun phrases in argument positions are labeled with indices. Labeling is done by a predicate **Labeled**$(i)(x)$, which returns true so long as the output sequence assigns i to x:

(13) **Labeled** $:= \lambda i.\lambda x.\lambda f.\lambda g.\ g(i) = x \wedge g \geq_i f$

The subformula $g \geq_i f$ is defined such that if i is familiar, $g(i) = f(i)$, while if i is novel, $g(i)$ can be assigned to any individual. This means that relative to an input sequence f, only one value of x will satisfy **Labeled**$(i)(x)$ if i is familiar, but every value of x will satisfy **Labeled**$(i)(x)$ if i is novel. Crucially, this predicate is involved in meeting the uniqueness presuppositions of **Iota** and *the*, so any i that is familiar will meet uniqueness in virtue of this predicate, while any i that is novel will have to use other material to meet uniqueness.

Let's see how their system derives the felicitous use of the unique description in (1-b). The description is translated in (14), where **TherapistAtKidsSchool**$_i$ is a predicate conjoining the dynamic property *x is a therapist at my kid's school* with the dynamic property **Labeled**$(i)(x)$.

(14) *the therapist at my kid's school* $\rightsquigarrow$ **The**(**TherapistAtKidsSchool**$_i$)

In the **Coffee Shop** scenario, there is no explicit evidence against choosing the determinate reading, so the sentence is interpreted with **Iota**. Here, the file F will not accept any formula that labels the description with a familiar index i, because the presupposition is not met (you can always find some $< w, f >$ pair where $f(i)$ is not a therapist at Ana's kid's school). Despite having to choose a novel index, the conjoined property **TherapistAtKidsSchool**$_i$ will still meet **Iota**'s presupposition in virtue of the overt property in the description, so there's no infelicity.

I adopt a semantics for names where they are (i) always labeled with some index i, and (ii) presuppose the existence of an individual with the name. This is essentially the proposal of Geurts (1997), although I do not follow him in analyzing names as definite descriptions, as the latter in this proposal conventionally encode uniqueness. Below, I define a dynamic predicate representing the property of being called *Mia* (below, **And** represents dynamic conjunction):

(15) **Mia** $:= \lambda x.\lambda f.\lambda g.\ f = g \wedge x$ is called Mia

(16) **Mia**$_i$ $:= \lambda x.$ **Labeled**$(i)(x)$ **And Mia**(x)

For argumental names like *Mia* in (1-a), I propose this predicate is embedded under a covert determiner **REF** that adds the existence presupposition:

(17) **Exists** $:= \lambda\mathcal{P}.\lambda f.\lambda g.\ f = g \wedge \exists h.\exists x.f[\mathcal{P}(x)]h$

(18) **REF** $:= \lambda\mathcal{P}.\lambda x.\ \partial_d$(**Exists**$(\mathcal{P})$) **And** $\mathcal{P}(x)$

(19) *Mia* $\rightsquigarrow$ **REF**(**Mia**$_i$)

Above, $\mathcal{P}$ represents a variable over dynamic properties, and ∂_d represents the dynamic presupposition operator, which eventually returns the null truth value $\#_t$ if its prejacent is false.

Now we can derive the infelicity of hearer-new names, and how certain appositives ameliorate this infelicity. First, consider the use of the name *Mia* in the **Coffee Shop** scenario in (1-a). Since there is no evidence to the contrary, **Iota** is chosen. Moreover, like the description in (1-b), the file F will not accept any familiar index i, as you can always find some $< w, f >$ pair where $f(i)$ is not called Mia. Unlike this description, uniqueness is not able to be met in virtue of the dynamic property **Mia** either, since many individuals are or could be called Mia (Geurts 1997). Lastly, I assume accommodating there is one individual called Mia in the world is ruled out by **consistency**, a constraint on accommodation (Van der Sandt 1992; Beaver 2001). So there is no way to meet the presupposition of **Iota**, even by accommodation, and the name is infelicitous.

As discussed in section **3**, appositives that communicate that the name is hearer-new are able to ameliorate this infelicity. Now, these appositives are exactly the ones that would be able

to provide evidence about the speaker's presuppositions that can circumvent the comprehension variant of **Maximize Presupposition**, as they require the index on the name to be novel, and novel indices on names can never meet uniqueness.[3] By circumventing this principle, these appositives allow the indeterminate reading to surface. Of course, **REF** still has to be met in this reading, but this only requires that the existence of some individual with the name *Mia* is accommodated, which I assume is straightforward. There is thus no apparent infelicity.

Before moving to incomplete definites, I should note one essential part of this proposal: how the classic version of Maximize Presupposition in Heim (1991) interacts with names and definites. This latter principle is used in Beaver and Coppock (2015) to derive why **Ex** is always chosen over **Iota** in the case of indefinites: since *the* has strictly stronger presuppositions than *a*, Heim (1991)'s MP requires *the* to be used if the determinate reading is intended. This reasoning crucially does not apply to names in our proposal as the presupposition of *the* (≤ 1) is not strictly stronger than the presupposition of **REF** (≥ 1).

The infelicitous use of a hearer-new incomplete definite (9) is derived in the same way as names. Moreover, our assumption that *the* encodes uniqueness derives the infelicity of incomplete definites with indefinite appositives (#*the therapist, a friend of mine*). In this case, even if **Ex** is used, the presupposition of *the* still has to be met, and the appositive *a friend of mine* is not able to help as it (i) lies outside the scope of *the* (Potts 2004) and, even more, (ii) doesn't contribute a property that is unique.

In fact, with a few more assumptions, our analysis can also explain how the deictic appositive in (11) *who is over there in the red hat*, which is also outside the scope of *the*, can ameliorate the infelicity of *the therapist*. Here, I propose that the appositive disambiguates the value of a **domain restriction index** that *is* in the scope of *the* (Stanley and Szabó 2000; Elbourne 2002). Below, I assume that indices can be typed as both individuals and properties, and that *the* can be labeled with a property index that helps meet its uniqueness requirement:

(20) $C_d = \lambda x.\lambda f.\lambda g.\exists P.f[\mathbf{Labeled}(C)(P)]g \wedge P(x)$

(21) $\mathbf{The}_C := \lambda\mathcal{P}.\lambda x.\ \partial_d(\mathbf{Unique}(\lambda x.\mathcal{P}(x)\ \mathbf{And}\ C_d(x)))\ \mathbf{And}\ \mathcal{P}(x)\ \mathbf{And}\ C_d(x)$

(22) *the therapist* $\rightsquigarrow \mathbf{The}_C(\mathbf{Therapist}_i)$

Although this domain restriction index is available in (9) and (10-b), it cannot be disambiguated to a specific value that helps meet uniqueness in these examples (it is not clear whether the speaker intends to talk about *the therapist that she goes to*, or *the therapist at this party*, etc.). The deictic appositive in (11), in contrast, requires the referent of the description to be at the party, and thus can disambiguate the domain restriction to something like $\lambda x.$ *x is at this party.*

5 Conclusion

A number of recent theories have argued that names are covert definite descriptions (Geurts 1997; Elbourne 2002; Matushansky 2008; Fara 2015). I demonstrated in this paper that names and descriptions behave quite differently in similar contexts, and that their differences can be accounted for under the assumption that names lack a determiner with the semantics of *the*. Under an analysis where names embed *the*, these differences would have to be traced to some other property unique to names. In this way, the constraints on clarificatory appositives described in this paper provide a novel empirical argument for distinguishing the semantics of names and definites. Lastly, these appositives provide novel empirical terrain for investigating the semantics and pragmatics of referring expressions.

[3] With the group of appositives that connect the name to some familiar dref, the *initial* index on the name is required to be novel in order for the appositive to be non-trivial (Potts 2004). Of course, after the appositive update, this index becomes conversationally equivalent to the familiar index described by the appositive.

Acknowledgements Thanks so much to Maziar Toosarvandani and Pranav Anand for their help with this project.

References

Beaver, David (2001). Presupposition and assertion in dynamic semantics. CSLI Publications.
Beaver, David and Elizabeth Coppock (2015). "Novelty and Familiarity for Free". In: Proceedings of the 2015 Amsterdam Colloquium.
Coppock, Elizabeth and David Beaver (2015). "Definiteness and determinacy". In: Linguistics and philosophy.
Elbourne, Paul (2002). "Situations and Individuals." PhD thesis. Massachusetts Institute of Technology.
Fara, Delia G. (2015). "Names are predicates." In: Philosophical Review.
Geurts, Bart (1997). "Good news about the description theory of names." In: Journal of Semantics.
Heim, Irene (1991). "Artikel und definitheit". In: Handbuch der Semantik. De Gruyter.
Maier, Emar (2009). "Proper names and indexicals trigger rigid presuppositions." In: Journal of Semantics.
Matushansky, Ora (2008). "On the linguistic complexity of proper names." In: Linguistics and Philosophy.
Potts, Chris (2004). The logic of conventional implicatures. Oxford University Press.
Roberts, Craige (2003). "Uniqueness in definite noun phrases." In: Linguistics and philosophy.
Sommers, Fred (1982). The logic of natural language. Clarendon Press.
Stanley, Jason and Zoltán G. Szabó (2000). "On quantifier domain restriction". In: Mind Language.
Van der Sandt, Rob (1992). "Presupposition projection as anaphora resolution". In: Journal of Semantics.

Generic Modals in Mandarin

Chun-Hung Shih
University of Pennsylvania
Philadelphia, United States of America
chshih@sas.upenn.edu

Abstract

Generics express events occurring with some regularity or general stable properties (Krifka et al. 1995). While generics are mostly realized in the present tense without a dedicated morpheme in English, the morpheme *hui* is used in Mandarin in certain cases, which has been claimed to be a generic modal and related to its epistemic/future and ability usages (Chang 2001, Tsai 2015, Wu 2020). However, this generic modal *hui* is not present in all generic sentences. With an investigation of the distribution and licensing of *hui*, this paper shows (1) that this morpheme occurs with episodic/stage-level predicates but not stative/individual-level ones, and (2) that adverbs, negators, and polarity operators can license this morpheme in stative/individual-level predicates (see also Wu 2020). This paper claims that Mandarin has two generic operators, providing independent evidence for Chierchia's (1995) theory of a distinction between two generic operators. One type is realized by *hui* and situated in Spec,AspP with restrictions on the conditions under which the relevant event typically occurs, while the other is realized covertly and situated in Spec,VP with a restriction on the location where the eventuality holds. This paper further proposes a stacking of two generic operators, based on the empirical facts in Mandarin.

1 Introduction

Generics, also known as generic sentences or characterizing sentences, capture generalizations that we observe about the world. Instead of particular events or properties, generics express the regularities of events (1a-b) and the general properties (1c) (Krifka et al. 1995).

(1) a. John **smokes** a cigar after dinner. (Krifka et al. 1995:3)
b. John **used to** smoke a pipe. (ibid:7)
c. A potato contains vitamin C, amino acids, protein, and thiamine. (ibid:3)

In terms of morphology, generics are realized differently across languages, e.g., generics in English are realized as the present tense (1a) or with an auxiliary construction for past events (1b), while in Swahili, they are realized as an affix (2).

(2) Wanawake **hu**-fanya kazi ya kochokoa pwesa.
women **HABIT**-do work of catching squid
'The women (generally) do the work of catching squid.' (ibid:8)

In Mandarin, generics can be expressed by a morpheme *hui* (3), which is categorized as a modal on a par with its other usages as future and ability (Chang 2001, Tsai 2015, Wu 2020). However, this generic modal in Mandarin does not appear in all constructions of genericity (4).

(3) a. Xiong **hui** dongmian.
bear **HUI** hibernate
'Bears hibernate.'
b. Xiaoming chi-bao fan hou **hui** chouyan.
Xiaoming eat-full rice after **HUI** smoke
'Xiaoming smokes after eating.'

(4) a. Gou (#**hui**) you si tiao tui.
dog (#**HUI**) have four CL leg
'Dogs have four legs.'

b. Haitun (#**hui**) hen congming.
dolphin (#**HUI**) very smart
'Dolphins are very smart.'

In this paper, I use diagnoses of generics (Krifka et al. 1995) to confirm the status of *hui* as a generic modal and also present the distribution and licensing of this morpheme in different contexts. Following Chierchia's (1995) theory, I propose two generic operators in Mandarin, *hui* and Ø, and argue that they can be stacked in Mandarin.

2 Diagnoses, Distribution, and Licensing

2.1 Diagnoses

The morpheme *hui* found in generic sentences as in (3) is associated with genericity not only by intuitive interpretation but also by the empirical diagnoses as defined in Krifka et al. (1995).

First, the morpheme *hui* can co-occur with the adverbs *usually* and *basically*, which both characterize genericity (5) despite slight differences in meanings.

(5) a. Xiong **tongchang** / **jibenshang** hui dongmian.
bear **usually** / **basically** HUI hibernate
'Bears usually / basically hibernate.'

b. Xiaoming chi-bao fan hou **tongchang** / **jibenshang** hui chouyan.
Xiaoming eat-full rice after **usually** / **basically** HUI smoke
'Xiaoming usually / basically smokes after eating.'

Second, it is infelicitous for *hui* to co-occur with progressive aspect (6b-c), given that generics are typically stative and that statives are incompatible with progressives as seen in (6a).

(6) a. Wo **(#zhengzai)** xihaun ta.
1SG **(#PROG)** like 3SG
'I like / #am liking him.'

b. #Xiong hui **zhengzai** dongmian.
bear HUI **PROG** hibernate
'Bears hibernate.'

c. #Xiaoming chi-bao fan hou tongchang hui **zhengzai** chouyan.
Xiaoming eat-full rice after usually HUI **PROG** smoke
'Xiaoming smokes after eating.'

Third, *hui* is incompatible with accidental properties, namely, the properties that are not characteristic of the subjects in question. In (7a), 'getting hurt' is not a characteristic property of bears but an accidental event, thus causing the infelicity of the presence of the generic *hui*.

(7) a. ??Xiong hui (rongyi) shou-shang.
bear HUI (easily) get-hurt
'Bears (easily) get hurt.'

b. ??Xiaoming chi-bao fan hou hui diedao.
Xiaoming eat-full rice after HUI fall.over
'Xiaoming (usually) falls over after eating.'

From the above diagnoses, the morpheme *hui* is apparently associated with genericity, and thus can be categorized as a generic modal in line with the assumption in the previous literature (Chang 2001, Lin 2012, Tsai 2015).

2.2 Distribution

Though the generic modal *hui* helps express generics in Mandarin, not all generic sentences require this morpheme. The distribution is dependent on the type of predicate.

The generic modal *hui* is obligatorily present with episodic/stage-level verbs (8).

(8) a. Xiong #(hui) dongmian.
bear #(HUI) hibernate
'Bears hibernate.'
b. Xiaoming chi-bao fan hou #(hui) chouyan.
Xiaoming eat-full rice after #(HUI) smoke
'Xiaoming smokes after eating.'

However, the generic modal *hui* is absent with individual-level predicates, such as stative verbs (9) and individual-level adjective predicates (10).

(9) a. Gou (#hui) you si tiao tui.
dog (#HUI) have four CL leg
'Dogs have four legs.'
b. Juzi (#hui) fuhan weitamin C.
tangerine (#HUI) contain vitamin C
'Tangerines contain vitamin C.'

(10) a. Haitun (#hui) hen congming.
dolphin (#HUI) very smart
'Dolphins are very smart.'
b. Xifang ren (#hui) hen gao.
western person (#HUI) very tall
'Westerners are tall.'

In brief, episodic/stage-level predicates but not stative/individual-level ones can have the generic modal *hui*. However, though this generalization holds water for the present data, we will see some contexts can license the presence of *hui* with stative/individual-level predicates.

2.3 Licensing

While the generic modal *hui* is not found with individual-level predicates (11a), adverbs, negators, and polarity operators however can help license the presence of *hui* (11b-d) (see also Wu 2020).

(11) a. Shizi (#**hui**) you zongmao.
lion (#**HUI**) have mane
'Lions have a mane.'
b. Shizi tongchang / jibenshang (**hui**) you zongmao.
lion usually / basically (**HUI**) have mane
'Lions usually / basically have a mane.'
c. Mu-shizi bu **hui** you zongmao.
female-lion NEG **HUI** have mane
'Lionesses don't have a mane.'
d. Mu-shizi (**hui**) you zongmao ma?
female-lion (**HUI**) have mane Q
'Do lionesses have a mane?'

Although the morpheme *hui* may seem optional,[1] the meanings with and without it are

[1]Strictly speaking, the morpheme *hui* in negation (11c) is not optional. The sentence without *hui* requires

slightly different. Without the modal *hui*, the generic sentences are merely a plain description, whereas with the modal *hui*, the generic sentences appear to be judged according to normal conditions and (im)possibilities as relevant to our world knowledge. For example, (11b) with *hui* is uttered in a context where we may be seeing a lion without a mane but still confirm this regularity based on normal conditions. Similarly, (11c) and (11d) with *hui* respectively negate and doubt the possibility of lionesses with a mane according to normal conditions.

3 Two Generic Operators

A dyadic approach to generics is adopted here, where the generic operator relates to two propositions. Namely, the semantic structure of generic sentences consists of a matrix, which introduces the proposition of generalizations, and a restrictor, which restricts the condition for generalizations parallel to *when-* and *if*-clauses (Carlson 1989, Schubert and Pelletier 1989, Krifka et al. 1995). Thus, the generic sentence in (12a) can be thought of as having the semantic structure as (12b).

(12) a. John smokes (when he comes home).
b. Gen s [$_{\text{Matrix}}$ John smokes in s] ([$_{\text{Restrictor}}$ when he comes home in s]).

Building on the dyadic analysis, Chierchia 1995 distinguishes an inherent generic operator from a general generic operator,[2] given the pair below. The sentence (13a) has a stage-level predicate and involves with a general generic operator. This operator is situated in Spec,AspP as related to the habitual aspect and has a condition-relation (**C**) in the restrictor, which requires the smoking event to take place in certain conditions such as when Fred feels like smoking but not when he is asleep. On the other hand, the sentence (13b) has an individual-level predicate and occurs with an inherent generic operator. This operator is in Spec,VP as inherent to the predicate itself and has a locative relation (**in**) in the restrictor. Namely, when Fred is in the target worlds, the general property holds true in all circumstances, including when he falls asleep.

(13) a. Fred smokes. → Gen s [**C**(f,s)] [smoke(f,s)]
= 1 iff in all the worlds maximally similar to ours where the felicity conditions for Fred's smoking is met, Fred does the action of smoking.
b. Fred is a smoker. → Gen s [**in**(f,s)] [smoker(f,s)]
= 1 iff in all the worlds maximally similar to ours where Fred is located, Fred has the property of being a smoker.

Based on this distinction, I propose that these two generic operators have different morphological realizations in Mandarin (14). The morpheme *hui*, as occurring with episodic/stage-level predicates, realizes the general generic operator in the Spec,AspP position. This generic modal *hui* has the condition-relation in the restrictor, requiring certain conditions to be met for the relevant action to take place (15a). On the other hand, stative/individual-level predicates occur with a covert inherent generic operator in the Spec,VP position with a locative relation in the restrictor (15b). This explains why *hui* is not present in (14b) and (9-10).

(14) a. Xiong **hui** dongmian.
bear **GEN** hibernate
'Bears hibernate.'

another negation word *mei*, as will be seen in (18).

[2]The term general generic operator is coined by me in order to differentiate it from the inherent one.

b. Gou **Ø** you si tiao tui.
dog **GEN** have four CL leg
'Dogs have four legs.'

(15) a. Gen-*hui* s [bear(x) ∧ **C**(x,s)] [hibernate(x,s)]
= 1 iff in all the worlds maximally similar to ours where the felicity conditions for bears' hibernating are met, bears hibernate.
b. Gen-Ø s [dog(x) ∧ **in**(x,s)] [have-four-legs(x,s)]
= 1 iff in all the worlds maximally similar to ours where dogs are located, dogs have four legs.

A difference emerges between these generic operators, with respect to quantification inside the predicate. An inherent generic operator can easily have quantification in predication without any overt restrictor (16c), as the predicate describes the inherent property. On the other hand, a general generic operator needs an overt restrictor in order to have quantification in predication (16a-b), as the quantification is not inherent and needs to be licensed relative to a certain condition (see Krifka et al. 1995 for English).

(16) a. Xiong **hui** chi (#wushi tiao) yu.
bear **GEN** eat (#fifty CL) fish
'Bears eat (#fifty) fish.' (context-less, without QUD)
b. Xiong [zai dongmian hou] **hui** chi (wushi tiao) yu.
bear [LOC hibernate after] **GEN** eat (fifty CL) fish
'Bears eat (fifty) fish after hibernation.'
c. Gou **Ø** you (si tiao) tui.
dog **GEN** have (four CL) leg
'Dogs have (four) legs.'

However, as seen above, adverbs, negators, and polarity operators can license the presence of *hui* with stative/individual-level predicates and force the reading where speakers moreover consider normal conditions (11b-d). I thus propose that the generic modal *hui* in Spec,AspP can be stacked over the covert inherent generic operator in Spec,VP. In this configuration, the reading of normal conditions comes from the semantics of *hui*, as it has a condition-relation in the restrictor (where conditions are contentful and salient (Chierchia 1995)). On the other hand, the general stable property is still captured by the inherent generic operator Ø. In other words, the operator *hui* introduces normal conditions, in which the general stable property holds true. The semantic structure is shown in (17): when the condition is met as licensed by the general generic operator, then an entity has its inherent property as licensed by the inherent generic operator.

(17) Gen-*hui* s [C(x,s)] [Gen-Ø s [in(x,s)] [V(x,s)]]
= 1 iff in all the worlds maximally similar to ours where the felicity condition for x is met, when x is located in them, x has the property V.

With this proposed structure, we can thus capture the intuition in (18). The proposition (18a) expresses a plain description of negating that whales have ears, while the proposition (18b) indicates that speakers are considering normal conditions similar to the actual world according to world knowledge. Hence, in addition to negating that whales have ears like (18a), (18b) appears more of an opinion based on the speaker's thinking about conditions. In other words, with (18b), speakers may have an opinion that it is impossible for whales to have ears, as having ears appears weird and not suitable according to whales' body structures and their living environments in ocean (where speakers think of ears as protruding ones as humans).

(18) a. Jingyu mei you erduo.
whale NEG have ear
'Whales don't have ears.'

b. Jingyu bu **hui** you erduo.
whale NEG **HUI** have ear
'Whales don't have ears.'

This proposal can be further evidenced by the interactions between the properties of subject noun phrases and predicate types. As noted by Chierchia 1995, in addition to actions and predicates, the featural restrictions of subject nouns may also serve as a condition for a generic statement to be true. Hence, for the generic statement in (19a) to hold true, the conditions concern both the flying action and the bird types (19b), e.g., excluding penguins or ostriches.

(19) a. A bird flies.
b. Gen s [bird(x) ∧ C(x,s)] [fly(x,s)] where C = flying conditions; birds that can fly

The presence of *hui* shows a gradient effect in this respect. The more restricting conditions come with the subject noun phrase, the more likely the morpheme *hui* is licensed with a stative/individual-level predicate (20a-c). In brief, the generic modal *hui* is licensed because of its condition-relation in the restrictor. It has to be noted that this gradient effect can be seen with accidental generalizations as in (20) but not for law-like generalizations (21). In (21), despite more featural restrictions on the subject noun, the presence of *hui* does not improve the example. The observation is presented here, but a full account of how restricting conditions interact with different types of generalizations is left for future research.

(20) a. nanren (#**hui**) hen huaxin.
man (#**HUI**) very unfaithful
'Men are unfaithful.'
b. Niuyue de nanren (?**hui**) hen huaxin.
New.York de man (?**HUI**) very unfaithful
'Men from New York are unfaithful.'
c. You zhe zhong mianxiang de nanren **hui** hen huaxin.
have this kind face DE man **HUI** very unfaithful
'Men with this kind of face is unfaithful.' (in contexts of physiognomy or face-reading)

(21) a. Gou (#**hui**) you si tiao tui.
dog (#**HUI**) have four CL leg
'Dogs have four legs.'
b. Heise de gou (#**hui**) you si tiao tui.
black DE dog (#**HUI**) have four CL leg
'Black dogs have four legs.'

4 Conclusion

Two generic operators were proposed in Mandarin to capture the distribution of *hui*: a general generic operator in AspP (Gen with **C**ondition) realized as *hui* and an inherent one in VP (Gen with **in**-location) realized as Ø. Furthermore, it was demonstrated that the two generic operators can co-occur, which suggested their respective presence, as seen in the presence of *hui* with stative/individual-level predicates. It takes place when the conditions in restrictors are activated, either by adverbs, negators, and polarity operators or by featural restrictions of subject noun phrases. Hence, this paper provided independent support in Mandarin for Chierchia's (1995) theory.

Acknowledgements. I'd like to express my gratitude to Florian Schwarz for advice, to Julie Legate and Paloma Jeretič for comments, to Yun Chuang, Mingyang Bian, Xuanzhu Jin, and Aini Li for judgments/discussions, and to the audience in Semantics II and Penn Meaning Lab.

References

Carlson, Gregory Norman (1989). "On the semantic composition of English generic sentences". In: *Properties, types and meaning: Volume ii: Semantic issues*. Springer, pp. 167–192.

Chang, Ching-Hsiu (2001). "Semantic Analysis on the Modal Verb "HUI" in Mandarin Chinese". Master's thesis. National Chengchi University.

Chierchia, Gennaro (1995). "Individual-level predicates as inherent generics". In: *The generic book*. University of Chicago Press, pp. 176–223.

Krifka, Manfred et al. (1995). "Genericity: An Introduction". In: *The generic book*. University of Chicago Press, pp. 1–124.

Lin, Tzong-Hong Jonah (2012). "Multiple-modal constructions in Mandarin Chinese and their finiteness properties". In: *Journal of Linguistics* 48.1, pp. 151–186.

Schubert, Lenhart K and Francis Jeffry Pelletier (1989). "Generically speaking, or, using discourse representation theory to interpret generics". In: *Properties, types and meaning* 2, pp. 193–268.

Tsai, Wei-Tien Dylan (2015). "On the topography of Chinese modals". In: *Beyond functional sequence* 10, pp. 275–294.

Wu, Jiun-Shiung (2020). "Ambiguity of epistemic Hui in mandarin Chinese revisited: toward a unified semantics". In: *From Minimal Contrast to Meaning Construct: Corpus-based, Near Synonym Driven Approaches to Chinese Lexical Semantics*, pp. 51–62.

The interpretation of logophoric and ordinary pronouns in Ewe: an experimental study

Silvia Silleresi
University of Milano-Bicocca
Milan, Italy
silvia.silleresi@unimib.it

Itai Bassi
Leibniz-Centre General Linguistics (ZAS)
Berlin, Germany
& Ben Gurion University, Israel
bassi@leibniz-zas.de

Abigail Anne Bimpeh
Humboldt University of Berlin
Berlin, Germany
abigail.anne.bimpeh@hu-berlin.de

Imke Driemel
University of York
York, UK
imke.driemel@york.ac.uk

Anastasia Nuworsu
Ho Technical University
Ho, Ghana
anuworsu@htu.edu.gh

Maria Teresa Guasti
University of Milano-Bicocca
Milan, Italy
mariateresa.guasti@unimib.it

Abstract

This study examines the interpretation of logophoric and ordinary pronouns in Ewe. While the logophor *yè* is predicted to obligatorily refer to the attitude holder in indirect speech reports, the ordinary pronoun's reference pattern remains debated. Thirty-eight Ewe-speaking adults completed an acceptability judgment task evaluating pronoun usage in self and anti-self contexts for communicative (*say*) and mental (*think*) verbs. The findings align with the principle of Maximize Presupposition, assuming *yè* encodes a presupposition: *yè* is preferred when its presupposition is met, blocking the occurrence of the ordinary pronoun *é*. In turn, *é* is favored only where the presupposition of *yè* fails, though less robustly than *yè*.

1 Introduction

Several West-African languages are known to employ logophoric pronouns in indirect speech reports (e.g., clausal complements of verbs like *say*, *think*, a.o.) as special anaphoric elements which are used to denote the author of the speech or thought event referred to in the matrix clause. A typical case of logophoricity is that of Ewe (a Niger-Congo language of the Kwa group), in which the logophoric pronoun (henceforth LogP) *yè* occurs in the context of an attitude predicate and unambiguously refers to the attitude holder, as illustrated in (1a) (Clements 1975; Pearson 2015; Bimpeh 2023; Bimpeh et al. 2024, a.o.). Besides the logophoric pronoun, Ewe also has an ordinary pronoun (henceforth OrdP) *é*, which can also occur in – but it's not limited to – the context of an attitude predicate. However, there is a debate regarding the type of co-reference that OrdP admits. Some researchers observed that *é* must not refer to the attitude holder (Kofi in (1b)), but to some contextually salient individual other than the attitude holder (Clements 1975; Bimpeh 2023; Bimpeh et al. 2024). Conversely, Pearson (2015) reports on the possibility of OrdP referring to the attitude holder. This interpretation was noted by only two (out of five) speakers of the Mina dialect – an Ewe dialect spoken in Togo – who participated as consultants in Pearson's study.

(1) a. *Kofi*$_1$ *gblɔ be* ***yè***$_{1/*2}$ *dzo.*
Kofi say COMP **LogP** left
'Kofi said that he left.'

b. *Kofi*$_1$ *gblɔ be* ***é***$_{\%1/2}$ *dzo.*
Kofi say COMP **OrdP** left
'Kofi said that he left.'

In sum, while the co-reference possibilities of LogP are consistently reported to be restricted to the attitude holder, diverging observations have been made for OrdP. In this study, we aim to

clarify the disagreements using quantitative data. To our knowledge, this is the first experimental study investigating the interpretation of both LogP and OrdP in Ewe. In particular, we sought to determine (i) whether the co-referent possibilities of the LogP, *yè*, in Ewe are restricted to the attitude holder, and (ii) whether the OrdP, *é*, can or cannot co-refer with the attitude holder, as debated in the literature (Pearson 2015 vs. Clements 1975; Bimpeh 2023). If it turns out that (i) LogP obligatorily co-refers with the attitude holder and (ii) there is a disjointness effect between OrdP and the attitude holder, we can derive the pattern by the principle of Maximize Presupposition! (Heim 1991; Sauerland 2008).

(2) Maximize Presupposition! (MP): (adapted from Sauerland 2003)
Of two lexical items of the same complexity, where one has stronger presuppositions than the other but which lead to the same truth conditions in all contexts where both of their presuppositions are satisfied, the presuppositionally stronger item must be used whenever its presuppositions are contextually satisfied.

Following Bimpeh et al. (2024), we assume that LogP introduces a presupposition deriving the co-reference requirement with the attitude holder introduced by the matrix predicate. OrdP, however, acts like a standard 3rd person pronoun, thus it introduces no presupposition. Given (2), LogP is chosen over OrdP in contexts where its presupposition is met. In other words, Ewe speakers should generally prefer *yè* to refer to the attitude holder, while disprefer *é* in such contexts. With nothing else being said, OrdP *é* should, however, be able to refer to a contextually salient individual other than the attitude holder, as it is a pronoun after all. LogP *yè*, in contrast, cannot do so due to its presuppositional requirement (more details in section 4).

2 The study

2.1 Participants

Thirty-eight Ewe-speaking adults (21 females, 16 males, 1 neutral) participated in the study.[1] Adults' age ranged from 20 to 55 years (M= 29;5, SD= 11.2). Participants were recruited from the Volta region in Ghana. Following Ameka's (1991) division of Ewe into Northern (e.g., Ewedome) and Southern (e.g., Anlo, Tonu) dialects, our participants were distributed as follows: 15 spoke Northern dialects and 23 spoke Southern dialects. No participant was diagnosed with certified language and/or hearing difficulties at the moment of the testing. Additionally, we collected demographic information such as years of education and level of education via a questionnaire participants had to fill after the experiment.[2]

2.2 Materials and procedure

We conducted an acceptability judgment task (Marty, Chemla, and Sprouse 2020), presenting our participants with minimal pairs of test sentences (joint presentation) differing only in that one sentence had LogP and the other OrdP (4), and we asked them to judge such pairs against two types of contexts: self context and anti-self context. Abstractly, in a self context the attitude holder refers to themselves with the pronoun; in anti-self contexts, they refer to someone else. All contexts were presented in Ewe (but see (3) for English version).

[1] A total of 53 participants were recruited for the experiment. However, 15 of them were excluded for the following reasons: 8 participants declared to be not fully proficient in Ewe, 6 participants performed below chance on the filler items and 1 participant performed the task twice. All participants were Ewe-English bilinguals, as reported per the bilingual questionnaire they had to fill in after the experimental session. The study has been pre-registered at this link.

[2] Levels of education are based on the UNESCO International Standard Classification of Education (ISCED). All demographic information, including gender, age and spoken-dialect were collected via the questionnaire.

(3) **CONTEXTS:**
Common incipit: Sefa and Fafali are at home and decide to have a singing competition. In order to decide the winner they decide to record their voices. At the end of the competition Sefa hears one of the recordings. Then Sefa says:

a. **Self**: "What a beautiful voice! It must be Fafali's voice. In fact Fafali is very good at singing!" Then Sefa realizes that it is her own voice. So Sefa says "Oh no wait! But this is my voice! So I'm very good at singing, not Fafali!".

b. **Anti-self**: "What a beautiful voice! It must be my voice. In fact I'm very good at singing!" Then Sefa realizes that it is Fafali's voice. So Sefa says "Oh no wait! But this is Fafali's voice! So Fafali is very good at singing, not me!".

(4) **TEST SENTENCES:**

a. *Mɔ̀ɔ̀èbá la, Sefa gblɔ be* ***yè****-nyé hà-dzì-lá nyuie.*
In.end DEF, Sefa say COMP **LogP**-is song-sing-one.who great
'In the end, Sefa said that she is a great singer.'

b. *Mɔ̀ɔ̀èbá la, Sefa gblɔ be* ***é****-nyé hà-dzì-lá nyuie.*
In.end DEF, Sefa say COMP **OrdP**-is song-sing-one.who great
'In the end, Sefa said that she is a great singer.'

At the beginning of each context (as in (3) common incipit), two characters – the attitude holder (e.g., Sefa) and another character (e.g., Fafali) – were introduced. The story then depicted a scenario where both characters could plausibly be the referent of an action described in the narrative. Ultimately, only one of them was revealed as the true referent (this contextual information was encoded by the use of 'in the end' at the beginning of the test sentences). In the self attitude context (3a), the attitude holder – and crucially not the other character – turns out to be the only possible referent for the pronoun used in the scope of the attitude predicate in (4), while in the anti-self attitude context (3b) the second character – and crucially not the attitude holder – was the sole possible referent for the pronoun used in the scope of the attitude predicate in (4). Our contexts were either adapted from existing literature on attitude reports (Perry 1979; Kiemtoré 2022; Bimpeh 2023) or developed from scratch. Participants were asked to read each context and then rate both test sentences in (4) on a Likert scale (1-7 points), indicating their level of acceptance or rejection relative to the context. Half of the test sentences contained a verb entailing a communicative attitude – *gblɔ* 'say' – and the other half a verb entailing a mental attitude – *súsú* 'think'. Each participant saw a total of 18 items: 6 self contexts (3 with *say*, 3 with *think*), 6 anti-self contexts (3 with *say*, 3 with *think*) and 6 fillers. All items were presented in a randomized order, including the order of the test sentences. All contexts and test sentences were presented in Ewe. Two experimental lists were created: each list contained either the self or anti-self version of each context (e.g., (3a) was in list 1, (3b) was in list 2). Eighteen participants completed list 1, and twenty completed list 2. Each participant completed the task with an on-line version of the experiment. Before starting the experimental phase, participants went through a training phase to familiarize themselves with the task.

3 Results

To prevent scale biases, each participant's raw scores were transformed into z-scores prior to analysis (Sprouse and Aronoff 2013; Schütze 2016). Z-scores are normally distributed with a mean of 0 and a standard deviation of ±1: a positive score would indicate, at least generally,

that the sentence was acceptable, whereas a negative score would indicate unacceptability. Linear mixed-effects models were performed for each test sentence type (LogP and OrdP) using the z-scores as the dependent variable, contexts, verbs and lists as fixed effects, items and participants as random effects. An effect of contexts emerged for both LogP (Figure 1) and OrdP (Figure 2): logophoric pronouns were significantly preferred in self contexts ($X^2(1) = 47.34, p < .001$), while ordinary pronouns were slightly preferred in anti-self contexts ($X^2(1) = 4.12, p = .042$). Additionally, there was a marginally significant verb-by-contexts interaction in LogP ($X^2(1) = 3.76, p = .052$). This suggests that the judgments were slightly more polarized (acceptance of LogP in self contexts and rejection of LogP in anti-self contexts) with communicative verbs (*say*) than with mental verbs (*think*). No such effect was observed in OrdP ($X^2(1) = 0.53, p = .463$).

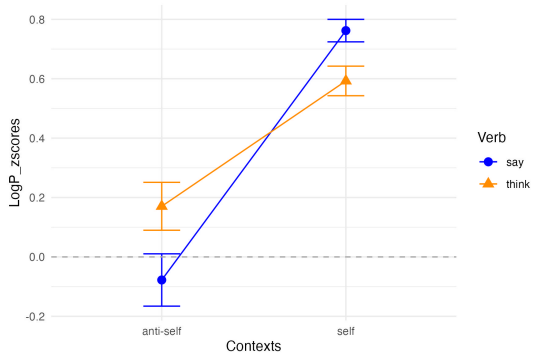


Figure 1: Z-scores of LogP in self and anti-self contexts for *say* and *think* verbs.

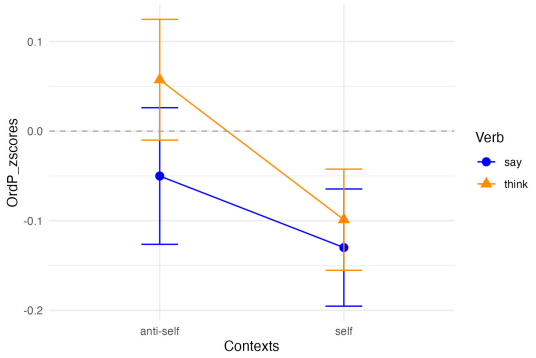


Figure 2: Z-scores of OrdP in self and anti-self contexts for *say* and *think* verbs.

When considering the potential impact of education level and years of education, a significant interaction was found between education level and contexts in LogP ($X^2(1) = 17.93, p < .001$). For OrdP a significant main effect of education level emerged ($X^2(1) = 6.51, p = .010$), together with a significant interaction between years of education and contexts ($X^2(1) = 5.92, p = .014$). These results indicate that (i) the higher the level of education, the more participants rejected LogP in the anti-self contexts and (ii) the higher the level and the number of years of education, the more participants accepted OrdP in the anti-self contexts and rejected them in the self ones. Finally, we checked for possible dialectal variations, but no effect emerged in our models, neither for LogP ($X^2(1) = 0.67, p = .412$) nor for OrdP ($X^2(1) = 0.01, p = .920$).

4 Discussion

Our results show that Ewe speakers display a strong preference for using the logophoric pronoun *yè* to indicate the attitude holder in self contexts, in line with previous literature (Clements 1975; Pearson 2015; Bimpeh et al. 2024; Bimpeh 2023). Concerning the interpretation of OrdP, our results confirmed the data by Clements (1975) and Bimpeh et al. (2024), in that *é* was dispreferred in self contexts (contra Pearson 2015).

These findings are consistent with the MP theory of LogPs (Bimpeh et al. 2024). We assume that the obligatory self reading of LogP is encoded as a presupposition, which in turn also derives the *de se* nature of logophoric pronouns which takes up the slot of a person feature. This presupposition is introduced by a pronominal feature Log, akin to person features. Formally, LogP consists of two elements in the syntax, one is a pronoun that denotes a variable, and the other is a morpho-syntactic feature Log: **LogP ≡ [Log [*pro*$_i$]]**. Bimpeh et al. (2024) implement this idea using the centered-world theory of Lewis (1979), according to which attitude ascriptions involve quantification over world-individual pairs. The Lewisian analysis of *de se* makes use

of the notion of a world's center, which is the individual who the attitude holder locates as themselves in the relevant worlds. Essentially, the LOG feature introduces a presupposition that makes the logophor as a whole denote the attitude holder's center.

The syntax of ORDP, on the other hand, is just like that of LOGP, except that LOG is replaced by a semantically-vacuous 3rd person feature[3]: **ORDP ≡ [3RD [pro_i]]**. If LOG is absent, so is its semantic contribution. In other words, ORDP is semantically unconstrained (apart from contextual recoverability). Because there are no semantic constraints on its denotation, ORDP is predicted to be compatible with a co-reference reading with the attitude holder. However, as we saw from our results, Ewe speakers systematically rejected ORDP in self contexts. If we couple the hypothesis that logophors introduce a *de se* co-reference presupposition with the competition idea as a case of MP (2), we can explain why ORDPs in Ewe are blocked in self contexts. In such contexts, both LOGP and ORDP are in principle possible candidates for the test sentences to yield a felicitous continuation; they are of the same complexity, lead to the same truth conditions, and thereby compete via MP. The test sentence including LOGP, however, is presuppositionally stronger, thereby blocking the occurrence of ORDP. Consequently, (2) indirectly predicts a disjointness effect of ORDPs in self contexts by an 'anti-presupposition' (Percus 2006). We should stress that 'complexity' in (2) refers to the number of nodes, not the number of (decomposed) person features. So while LOGP and ORDP are of the same syntactic complexity, the number of person subfeatures encoded in the morpho-syntactic component of the pronoun may vary. In other words, our MP based account of the disjointness effect of ORDP in self contexts has no bearing on the nature of the underlying person feature system (e.g., privative vs. binary, feature geometries, LOG features, etc.).

Although the general trend of our results confirms our main expectations, we also encountered some unexpected results, particularly regarding anti-self contexts. First, we need to account for the relative acceptability of LOGP in anti-self contexts, as shown in Figure 1. We speculate that this effect was not due to a failure in (the parsing of) MP, but rather as a consequence of the impact of participants' educational level. In fact, our results showed that a higher level of education was negatively related to the acceptance of LOGP in the anti-self contexts. We interpret this result as indicating that educational level could influence participants' judgments potentially making them more attuned to the complexity of the contexts presented. Consistent with previous literature, we propose that higher educational levels contribute to increased metalinguistic awareness, which may serve as a prerequisite for accurately interpreting linguistic contexts and assign linguistic judgments (Bialystok and Ryan 1985; Schütze 2016).

Second, the expected ORDP preference in anti-self contexts was attested, but not as pronounced as that observed in self contexts for LOGP. Again, our results showed how these findings were related to educational levels and years of education of the participants, in that the more education the speakers received, the more they tended to accept the ordinary pronoun as the referent of the attitude predicate in anti-self contexts. Nonetheless, level of education cannot explain why there was a relatively lower acceptance rate for ORDP overall compared to LOGP. One possible explanation is that ORDP, being the anti-presupposed alternative to LOGP, is more challenging to process in a joint presentation design. The processing of the 3rd person pronoun arises indirectly through a two-step process: first, participants consider the uniqueness presupposition of LOGP, and then subsequently they negate it. This approach aligns closely with the experimental data by Schneider et al. (2019) for definite and indefinite determiners. Additionally, this trend may be attributable to the fact that attitude reports with ORDP in embedded subject position are generally less frequent in the input than those with LOGP.[4]

Moreover, our analysis revealed a marginal effect related to attitude verbs in the context of LOGP: sentences with speech verbs (*say*) exhibited a clearer pattern than those with thought

[3]That is, it denotes the identity function.

[4]Abigail Anne Bimpeh and Anastasia Nuworsu, p.c.

verbs (*think*). While these results should be interpreted with caution due to the marginality of the statistical effects, it is still important to consider possible explanations for this pattern. We try to account for this finding on the basis of one (or a combination) of these two hypothesis: (i) Culy's (1994) hierarchy of attitude verbs, (ii) the diachronic nature of the complementizer in Ewe. Regarding (i) Culy suggests a hierarchy among attitude verbs based on the frequency of verbs allowing logophoric marking across languages, where speech verbs are most frequent, and perception verbs are the least frequent (speech > thought > knowledge > direct perception) (see also Bimpeh 2023 for a revised hierarchy of logophoric predicates in Ewe). This hierarchy reflects the perceived reliability of information, with speech verbs considered more objective and factual than thought or perception verbs (see also Koopman and Sportiche 1989 on this point). Speech can be heard or seen, while thought (of others) can only be inferred indirectly. Our data appears to support this hypothesis, indicating that information conveyed by *say* verbs is viewed as more objective and factual, making it easier to judge for our participants.

Concerning (ii), attitude verbs like *say* and *think* make use of the logophoric complementizer *be* in Ewe, which retains the semantic properties of *say* – even though it is often used (and generally glossed) as COMP, as shown in (4) (see Clements 1975; Lord 1993; Bimpeh 2023 on the diachronic interpretation of *be* in Ewe). In other words, *be* serves double duty: as a complementizer and as an element that still carries the force of the verb *say*.[5] This results in a reduplication of meaning in the *say* construction (*gblɔ* + *be* = 'say' + 'say'), unlike what we find in the *think* construction (*súsú* + *be* = 'think' + 'say'). We speculate that this makes the *say* construction in Ewe more transparent, and therefore easier to judge.[6]

Finally, no dialectal variations (Northern vs. Southern dialects of Ewe in Ghana's Volta region) were observed in our dataset, suggesting that Pearson's (2015) report of ORDP possibly referring to the attitude holder might be explained by the dialectal variety of Ewe spoken in Togo, which may not necessarily extend to the varieties of Ewe spoken in Ghana.

5 Conclusion

This study sheds light on the interpretation of logophoric and ordinary pronouns in Ewe, providing empirical support for theoretical claims about their reference patterns. Logophoric pronouns are strongly associated with the attitude holder in self contexts, while ordinary pronouns are dispreferred in these scenarios, consistent with predictions from the principle of Maximize Presupposition. The observed effects of educational background underscore the role of metalinguistic awareness in shaping linguistic judgments. Our findings also suggests a processing challenge associated with with anti-presuppositions triggered by ordinary pronouns. These insights contribute to a deeper understanding of pronoun interpretation in Ewe and its broader implications for the theoretical discussions about *de se* – and *de re* readings – in 'mistaken identity' scenarios.

Acknowledgements. We thank all the participants who took part in the study. The study was evaluated by the local commission for minimal-risk studies of the Psychology Department, which operates under the mandate of the Ethics Committee of the University of Milano-Bicocca (RM-2020-229). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (LeibnizDream, grant agreement No 856421).

[5]Notice that *be* can serve as the main verb in *say* in sentences like (1a), in which case Clements (1975) argues that *gblɔ* is deleted by an optional rule.

[6]This aligns with findings from the acquisition literature showing that children often struggle to understand sentences with *think* well into their fourth year, while they acquire *say* verbs early on (De Villiers and Pyers 2002; Papafragou, Cassidy, and Gleitman 2007, a.m.o.).

References

Ameka, Felix K (1991). "Ewe: Its grammatical constructions and illocutionary devices". PhD thesis. Australian National University Canberra.

Bialystok, Ellen and Ellen Bouchard Ryan (1985). "A metacognitive framework for the development of first and second language skills". In: *Metacognition, cognition, and human performance*. Ed. by Donna-Lynn Forrest-Pressley, GE MacKinnon, and T Gary Waller. Academic Press Orlando, FL, pp. 207–252.

Bimpeh, Abigail Anne (2023). "Logophoricity: an empirical-semantic assessment of yè in Ewe". PhD thesis. Universitätsbibliothek Johann Christian Senckenberg.

Bimpeh, Abigail Anne et al. (2024). "Obligatory de se logophors in Ewe, Yoruba and Igbo: Variation and competition". In: *Proceedings of the 40th West Coast Conference on Formal Linguistics*. Ed. by Jiayi Lu et al. Somerville, MA: Cascadilla Press, pp. 1–10. URL: `https://www.lingref.com/cpp/wccfl/40/paper3692.pdf`.

Clements, George (1975). "The Logophoric Pronoun in Ewe: Its Role in Discourse". In: *Journal of West African Languages* 10, pp. 141–177.

Culy, Christopher (1994). "Aspects of logophoric marking". In: *Linguistics*, pp. 1055–1094.

De Villiers, Jill G and Jennie E Pyers (2002). "Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding". In: *Cognitive development* 17.1, pp. 1037–1060.

Heim, Irene (1991). "Artikel und Definitheit". In: *Semantik / Semantics: Ein internationales Handbuch zeitgenössischer Forschung*. Ed. by Arnim von Stechow and Dieter Wunderlich. De Gruyter Mouton, pp. 487–535.

Kiemtoré, Alassane (2022). "Issues in Jula complementation: Structures, relations and matters of interpretation". PhD thesis. Univeristy of Stuttgart.

Koopman, Hilda and Dominique Sportiche (1989). "Pronouns, logical variables, and logophoricity in Abe". In: *Linguistic Inquiry* 20, pp. 555–588.

Lewis, David (1979). "Attitudes de dicto and de se". In: *The philosophical review* 88.4, pp. 513–543.

Lord, Carol (1993). *Historical Change in Serial Verb Constructions*. Amsterdam: Benjamins.

Marty, Paul, Emmanuel Chemla, and Jon Sprouse (2020). "The effect of three basic task features on the sensitivity of acceptability judgment tasks". In: *Glossa: a journal of general linguistics (2016-2021)* 5.1, p. 72.

Papafragou, Anna, Kimberly Cassidy, and Lila Gleitman (2007). "When we think about thinking: The acquisition of belief verbs". In: *Cognition* 105.1, pp. 125–165.

Pearson, Hazel Anne (2015). "The interpretation of the logophoric pronoun in Ewe". In: *Natural Language Semantics* 23.2, pp. 77–118.

Percus, Orin (2006). "Antipresuppositions". In: *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science*, pp. 52–73.

Perry, John (1979). "The problem of the essential indexical". In: *Noûs*, pp. 3–21.

Sauerland, Uli (2003). "A New Semantics for Number". In: *Proceedings of SALT 13*. Ed. by R. Young and Y. Zhou. Ithaca, NY: Cornell University, pp. 258–275.

— (2008). "Implicated presuppositions". In: *The discourse potential of underspecified structures* 8, pp. 581–600.

Schneider, Cosima et al. (2019). "Pragmatic processing: An investigation of the (anti-) presuppositions of determiners using mouse-tracking". In: *Cognition* 193, p. 104024.

Schütze, Carson (2016). *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Language Science Press.

Sprouse, Jon and M Aronoff (2013). *Acceptability judgments*. Oxford University Press Oxford.

Causal abductive reasoning in discourse processing: Evidence from eyetracking

Torgrim Solstad
Bielefeld University / CRC 1646
Bielefeld, Germany
torgrim.solstad@uni-bielefeld.de

Oliver Bott
Bielefeld University / CRC 1646
Bielefeld, Germany
oliver.bott@uni-bielefeld.de

Abstract

Psycholinguistic research has provided abundant evidence that both the nature of a causal relation as well as world knowledge have an immediate influence on the processing of explanatory discourse (e.g., Canestrelli, Mak, and Sanders 2013; Köhne-Fuetterer et al. 2021; Xiang and Kuperberg 2015). However, little is known about which particular reasoning processes are immediately triggered as we integrate new information. The present study provides first evidence that two types of abductive reasoning (Aliseda 2006) can be observed during processing, involving partial explanations that would either contradict world knowledge (abductive anomaly) or not be predicted by such knowledge (abductive novelty). In an experiment applying eyetracking during reading we identified two distinct processing signatures for the two types of abduction.

1 Introduction

Research on the comprehension of causal discourse has provided ample evidence that contingency relations trigger immediate inferences that may also take world knowledge into account (Traxler, Bybee, and Pickering 1997; Canestrelli, Mak, and Sanders 2013; Xiang and Kuperberg 2015; Noordman et al. 2015). However, the precise nature of the inference patterns involved, their interaction with world knowledge, and how they relate to behavioural measures is still understudied. The present study provides evidence that two distinct patterns of abductive reasoning (Aliseda 2006; Aliseda 2017) display different processing signatures in eyetracking during reading, showing that subtle causal distinctions are evaluated incrementally in comprehension.

The following mini-discourses illustrate the two types of abductive reasoning we manipulated:

(1) a. **Abductive anomaly:** $\Theta \not\Rightarrow \phi$**,** $\Theta \Rightarrow \neg\phi$
Peter recently installed a new dishwasher. Because [he followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

b. **Abductive novelty:** $\Theta \not\Rightarrow \phi$**,** $\Theta \not\Rightarrow \neg\phi$
Every morning, Marcus takes the eight o'clock train to Stuttgart. Because [he overslept by an hour today]$_{\text{cause2}}$, [he arrived safely]$_{\text{effect}}$.

The two scenarios correspond to Aliseda's (2006) notions of *abductive anomaly* (1a) and *novelty* (1b), respectively. For abductive anomaly (1a), the particular cause in combination with our world knowledge, Θ, *cause2*, cannot explain the effect ϕ in question, while they would readily explain the "opposite" effect, $\neg\phi$. Thus, the effect is of the right type, but its particular realization contradicts world knowledge. For *abductive novelty*, cause and world knowledge combined offer no support for the effect in question: Neither the presented effect ϕ, nor its negation $\neg\phi$ would seem explainable given the cause presented. A possible remedy could involve the abductive step of assuming an additional cause to be present, Θ, *cause1*, *cause2* $\Rightarrow \phi$.

We predicted the two reasoning types to be reflected in distinct processing signatures. Testing discourses involving these relations in eyetracking during reading, we expected *anomaly* to trigger

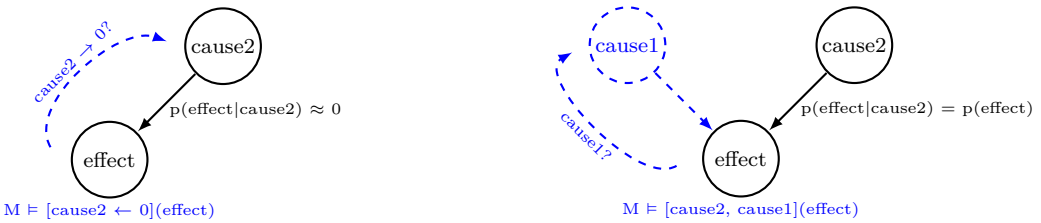


Figure 1: Anomalous (left) and novel (right) explanatory relation and their hypothesized repair strategies

more backtracking of preceding discourse since negating the cause, *not* having followed the instructions, could remedy the explanation. For *novelty*, proceeding to subsequent discourse would seem more promising, searching for an additional fact that might help explain the effect.

These abductive patterns were modeled using structural equation modelling (Figure 1, see Pearl 2009; Halpern 2016). For abductive anomaly and novelty alike, some eventuality *cause2* is explicitly claimed ("because") to be causally connected to an *effect*. The two instances differ in one important respect, though. In cases of abductive anomaly (left-hand side), the effect has a (stipulated) probability close to zero given cause2 ($p(\text{effect}|\text{cause2}) \approx 0$). If *cause2* were to be false, however, *effect would* be explainable ($M \models [\text{cause2} \leftarrow 0](\text{effect})$). Given these facts, we assume backtracking to constitute a plausible strategy, as indicated by the dashed blue arrow in the left-hand figure, checking whether a negation may have been overlooked.

For abductive novelty (right-hand side of Figure 1), we assume this strategy to be less promising, since the probability of the *effect* occurring seems to be independent of *cause2* ($p(\text{effect}|\text{cause2}) = p(\text{effect})$). In this case, a more plausible strategy would seem to assume that there must be some additional eventuality influencing the *effect* ($M \models [\text{cause2}, \text{cause1}](\text{effect})$). Thus, while the *effect* is surprising, it would seem more promising to continue reading to check whether such a possible cause (cause1) would be presented in subsequent discourse.

In order to investigate the different reasoning patterns illustrated in Figure 1, we constructed experimental items in German realizing the patterns of abductive anomaly and novelty as in (1) (between-item factor CAUSAL RELATION). For both types, we additionally included a second cause – *cause1* – that would make the effect follow naturally. We manipulated whether this effect was presented before or after *cause2, effect* as discussed for (1) (within-item factor CONSTRUCTION), see the following example item for anomaly (English translation only):

(2) a. **control (cause1 and cause2 conjoined under *because*):**
Because [the assembly instructions contained an error]$_{\text{cause1}}$ and [he followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

b. **left dislocation (cause1 before *because* clause):**
[The assembly instructions contained an error]$_{\text{cause1}}$. Because [he followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

c. **right dislocation (cause1 after *cause2, effect* sequence):**
Because [he followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$. You see, [the assembly instructions contained an error.]$_{\text{cause1}}$.

Of note, the sequence *cause2, effect* is the same across conditions. The control and left dislocation conditions in (2a) and (2b) both present two causes that are jointly suited to explain the effect in question, cancelling the need for abduction. The left dislocation condition was included to assess whether introducing one cause before the causal connective leads to processing difficulties as compared to a situation where the explanation is maximally explicit (Hagoort and van Berkum 2007). The right dislocation condition in (2c) crucially postpones revealing the necessary cause1 until after the partial explanation in the sequence *because cause2, effect*, calling for abduction.

2 Experiment 1: Contingency Ratings

The first experiment investigated *anomalous* and *novel* causal relations in three conditions, testing for the assumed logical properties of the two reasoning patterns (Aliseda 2006; Aliseda 2017) and the appropriateness of the complex scenarios in (2).

2.1 Methods

24 native German speakers (mean age 24.9 years, 19 female, 5 male) took part in the study for monetary compensation. 15 anomaly and 15 novelty relations were constructed in three conditions each in a 3 (CONSTRUCTION) × 2 (CAUSAL RELATION) design. The *complex cause* scenario (3a) served as control, testing whether two causes *cause1* and *cause2* were jointly appropriate to explain an *effect*. This condition was compared to a *simple cause* condition and a *negated simple cause condition*.

(3) a. **complex cause (anomaly):**
Because [the assembly instructions contained an error]$_{\text{cause1}}$ and [Peter followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

b. **simple cause (anomaly):**
Because [Peter followed the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

c. **negated simple cause (anomaly):**
Because [Peter *didn't follow* the instructions]$_{\text{cause2}}$, [the device broke down]$_{\text{effect}}$.

Participants' task was to rate these conditions on a seven-point bidirectional causality rating scale (Neunaber and Wassermann 1986; Ng, Lee, and Lovibond 2024), ranging from −3 ("completely reversed causal relation") to +3 ("completely normal causal relation") with a mid value of 0 explicitly indicating a "complete lack of causal connection". The *complex cause* condition should be rated equally well for anomaly and novelty. The two relation types should diverge with respect to the other two conditions, however: For novelty, neither the positive simple cause nor the negated simple cause condition should be causally related. For anomaly, however, the positive simple cause should be rated as "perfectly reversed", whereas the negated simple cause condition should be rated as a "perfectly normal causal relation".

The judgments were analyzed in R (R Core Team 2021) in a *cumulative link mixed effect* regression analysis with the R package *ordinal* (Christensen 2020). The models included the fixed effects of CONDITION (complex causes vs. simple cause vs. negated simple cause; treatment coded with complex causes as base category), CAUSAL RELATION (centered) and their interactions.

2.2 Results and Discussion

The mean contingency ratings are shown in Figure 2. As desired, the *complex cause* scenarios were clearly perceived as perfectly causally related. The *simple cause* and the *negated simple cause scenarios* received ratings fully in line with our expectations for *abductive novelty* and *anomaly*, respectively. For *novelty*, the simple cause and the negated simple cause condition were equally perceived as without any causal connection with mean ratings close to 0. For *anomaly*, however, the simple cause condition was perceived as contradicting normality while the negated simple cause condition was perceived to be as natural as the complex causes control condition.

This was corroborated by the statistical analysis giving rise to significant interaction effects for both CONDITION contrasts × CAUSAL RELATION ($\chi^2(2) = 269.32; p < .01$). Planned simple comparisons revealed no significance difference in perceived causal strength between *novelty* and *anomaly* in the complex causes control condition ($\hat{\beta} = -.05, z = -.27, p = .79$), but clear differences in the simple cause conditions ($\hat{\beta} = -1.44, z = -6.75, p < .01$) and in the negated causes conditions ($\hat{\beta} = 1.92, z = 8.27, p < .01$).

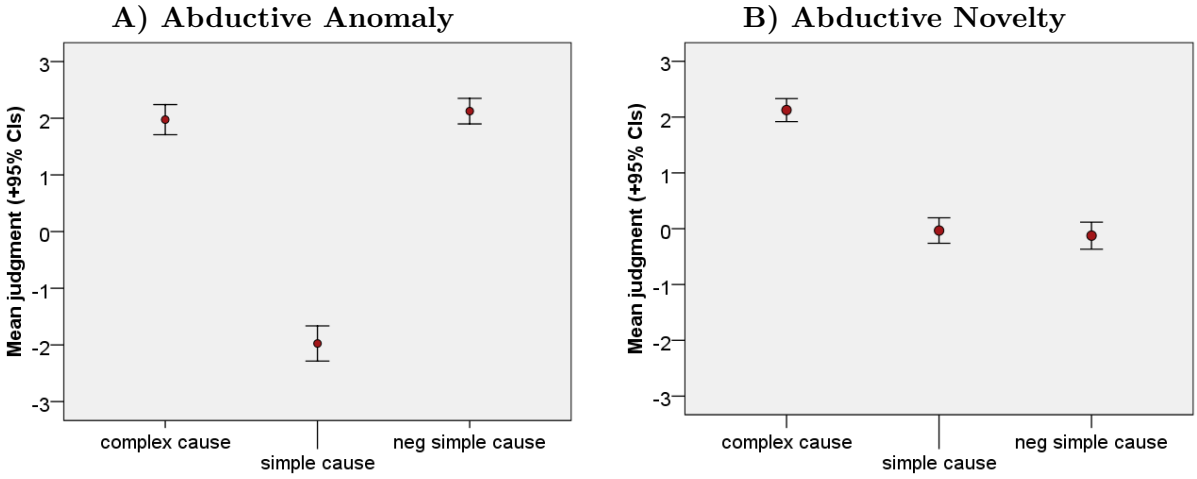


Figure 2: Mean contingency ratings and 95% confidence intervals in Exp. 1 for *anomaly* (A) and *novelty* (B). Positive values indicate "normal causal relations" and negative values indicate causal relations that are "opposite to normal". A value of 0 indicates no causal connection. 95% CIs estimated based on standard errors of the means.

3 Experiment 2: Eyetracking during reading

Our second experiment investigated the real-time comprehension of discourse with conjunctive causal scenarios triggering abductive *novelty* and *anomaly* in the three conditions in (2), respectively. The two reasoning patterns are hypothesized to lead to two very different processing strategies in real-time comprehension. *Novelty* should lead to integration difficulty of the surprising causal outcome given the lack of a causal connection to the preceding discourse context. For novel cause-effect relations, we thus expected longer first-pass reading time in right dislocation relative to both control and left dislocation, but no increase in the amount of regressions from the *effect* ROI to the preceding, causally unrelated *cause2* region. For anomalous cause-effect effect relations, causal integration during first-pass is expected to be easy an effect of this type (ϕ) is expected. However, since it is actually the reversed effect that would be predicted, first integration should be followed by backtracking to the preceding context in order to check for encoding errors of the preceding context. We operationalized such backtracking as a higher proportion of regressions out of the *effect* region in the right-dislocation condition. Finally, we expected longer second-pass times of the *cause2* region for *anomaly* than for *novelty*.

3.1 Methods

Twenty-seven students from Tübingen University (mean age: 25.9 years of age, 20 female, 7 male) were included in the analysis of the eyetracking experiment. All participants were native speakers of German and had normal or corrected-to-normal vision.

15 *novelty* and 15 *anomaly* items were constructed according to the three conditions in (2) and embedded in further context as follows: A short introductory context of two to three sentences set the scene, followed by the discourses displayed in (2). It may be noted that all *effect* regions were followed by a spillover region that we won't discuss here. All conditions were wrapped up with a concluding sentence, intended to increase overall coherence. The items in the three conditions were distributed to three lists in a Latin Square design. The same set of 30 filler discourses was added to the lists. In addition, 30 comprehension questions with two answer alternatives were formulated for 15 experimental items and 15 fillers.

A desktop-mounted Eyelink 1000 Plus eyetracker monitored the gaze location of participants' dominant eye. Participants viewed the stimuli binocularly on a 21″ monitor on a distance of 70cm from their eyes. Stories were displayed on five to eight lines with automatic line breaks in

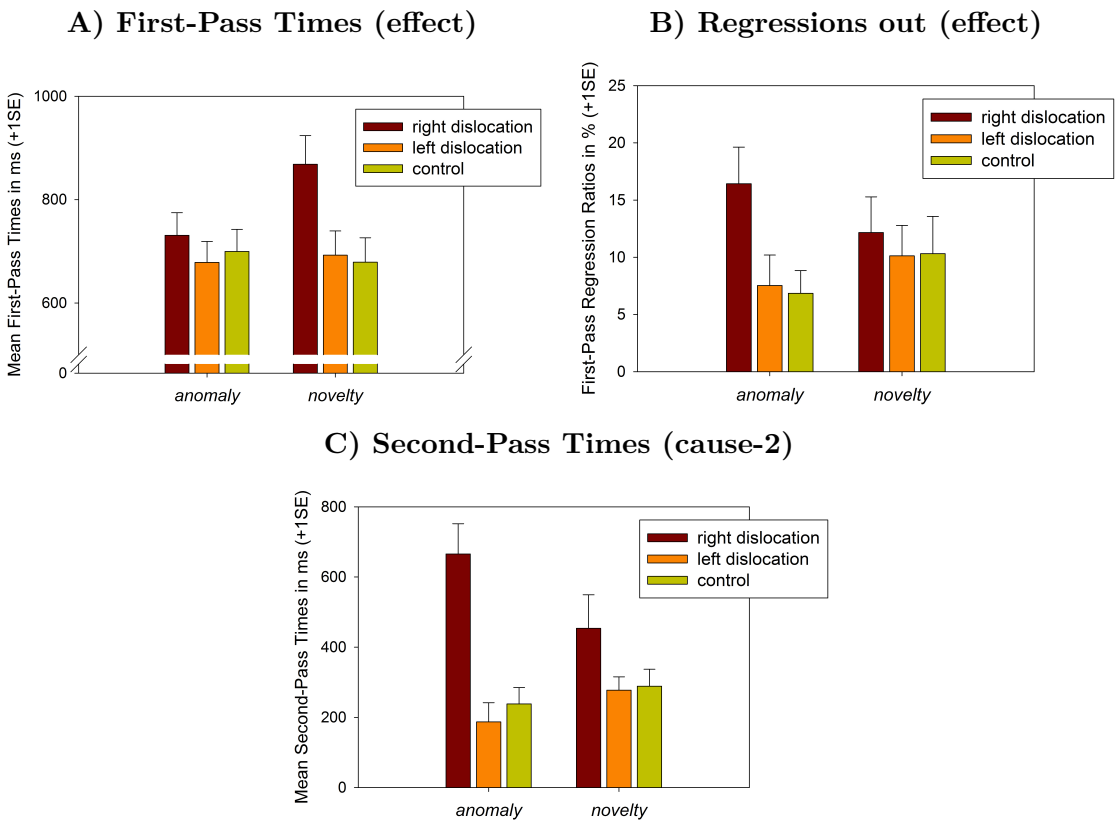


Figure 3: Mean first-pass times (top left) and first-pass regression ratios in the EFFECT region (top right) as well as second-pass times of the CAUSE-2 region (bottom row) in the eyetracking during reading experiment (Exp. 2). Error bars correspond to one standard error of the mean.

a single paragraph. After reading a story, participants were presented with a question or the subsequent trial. There was no time limit for answering the questions.

Three eyetracking measures were analyzed (cf. Clifton, Staub, and Rayner 2007): First-pass times (the sum of all fixations during first-pass readings), first-pass regression ratios (whether a regression was launched during first-pass reading) and second-pass times (the sum of all fixations in rereading a region). For statistical analysis of first-pass times and second-pass times linear mixed-effects regression models (Bates et al. 2015) were fit to the RTs of *cause2* and *effect* ROIs. The first-pass regression analysis employed logit mixed-effects analyses (Jäger 2008), with Bonferroni correction compensating multiple comparisons.

3.2 Results and Discussion

Comprehension questions were answered 95.8% correct on average showing that participants read attentively. Every participant scored at least 80% correct.

First-Pass Times The mean first-pass times for the *effect* ROI are shown in the top left of Figure 3. There were no differences between left dislocation and control conditions (saturated vs. intercept only model: $\chi^2(3) = .75., p = .86$). Thus, the effect sentences could be integrated equally easily irrespective of whether CAUSE1 appeared outside or in the scope of *weil* 'because'. For right dislocation, the predicted interaction was observed ($\hat{\beta} = -79.42, SE = 27.18, \chi^2(1) = 8.83, p < .01\overline{6}$). Breaking down the interaction effect in subset analyses, pairwise comparisons revealed that first-pass times were significantly higher for right dislocation relative to control in the NOVELTY scenarios ($\hat{\beta} = 190.44, SE = 40.40, \chi^2(1) = 21.28, p < .0\overline{3}$), but that there was no reliable difference in first-pass time of the ANOMALY scenarios ($\chi^2(1) = .79., p = .37$).

Thus, integration of the EFFECT ROI during first-pass reading in fact caused difficulty in causal NOVELTY but not in ANOMALY relations. This is what we predicted from the fact that the effect type as such is expected for *anomaly*, but not for *novelty*.

First-Pass Regression Ratios The analysis of first-pass regression ratios complemented the first-pass time analysis. The first-pass regression ratios are shown in the top right of Figure 3.

The analysis of the EFFECT ROI, revealed more regressions out in the right dislocation conditions than in the control and left dislocation conditions. ANOMALY displayed an average of 16% regressions after first pass reading (relative to 8 and 7% regressions in the left dislocation and control conditions, respectively), whereas NOVELTY received on average 13% regression in the right dislocation condition (compared to 10 and 9% regressions in left dislocation and control). In the GLMER analysis this pattern was reflected in a main effect of right-dislocation ($\hat{\beta} = .69, SE = .29, \chi^2(1) = 5.92, p < .01\overline{6}$) in the absence of an interaction ($\chi^2(1) = 1.09, p = .30$). Testing for the predictions outlined above, the right-dislocation effects were analyzed separately for the two causal relations in subset analyses of the data. These pairwise comparisons revealed a reliable right dislocation effect in ANOMALY ($\chi^2(1) = 5.24, p < .0\overline{3}$), but not in NOVELTY ($\chi^2(1) = .95, p = .33$). This pattern is consistent with the prediction that the likelihood to regress to the preceding discourse should be higher for causal ANOMALY than for NOVELTY.

Second-Pass Times The last analysis investigated the time spent in rereading the right-dislocation conditions. We hypothesized that the CAUSE2 region should lead to stronger second-pass effects in right dislocation of abductive anomaly relative to right dislocation in abductive novelty. The mean second-pass times of the *cause2* ROI are shown in panel C of Figure 3. The GLMER analyses also revealed no left-dislocation effects.

The analysis of second-pass times revealed a significant interaction (likelihood ratio test of the RIGHT DISLOCATION × CAUSAL RELATION interaction $\hat{\beta} = -258.67, SE = 119.13, \chi^2(1) = 4.69, p < .03$), not present in first-pass times (not discussed here). Thus, readers in fact spent more time rereading the CAUSE2 region in right-dislocated discourses involving abductive *anomaly* than for *novelty*. In addition to this interaction, a main right-dislocation main effect was observed ($\hat{\beta} = 431.15, SE = 83.94, \chi^2(1) = 25.13, p < .01$). On average, right-dislocation led to a 302 ms increase in second-pass time. However, the size of the right-dislocation effects differed quite dramatically between *anomaly* and *novelty* with a difference of approximately 260 ms. In summary, the need for rechecking the CAUSE ROI in right-dislocated discourses was especially strong after having encountered a causal relation displaying abductive anomaly. Again, the GLMER analyses revealed no left-dislocation effects.

4 General Discussion

The present paper presented an eyetracking study on the processing of causal relations in discourse. In particular, the study investigated the reading behaviour of causal scenarios in which two causes are required to explain an effect. The crucial comparison involved scenarios displaying abductive anomaly relative to ones exhibiting abductive novelty (Aliseda 2006).

While Experiment 1 used a rating task to confirm that the discourse materials displayed the desired reasoning patterns, Experiment 2 provided evidence that abductive anomaly and novelty are associated with different discourse processing strategies. *Anomaly* led to an increase in regressions and substantially longer rereading of the CAUSE2 region whereas *novelty* led to longer reading times during first-pass reading but no increase in regressions to earlier text.

The present study thus presents online processing evidence for sub-types of causal reasoning patterns in discourse comprehension not – as far as we are concerned – shown before. Future research should establish whether the dissociation of effects observed for the two abductive reasoning patterns yields distinct processing signatures beyond eyetracking during reading.

Acknowledgements We would like to thank Panagiota Rassia for assisting us with the statistical analysis. Furthermore, we gratefully acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) to subproject B01, CRC 1646, project number 512393437.

References

Aliseda, Atocha (2006). *Abductive reasoning: Logical Investigations into Discovery and Explanations*. Dordrecht: Springer.

— (2017). "The Logic of Abduction: An Introduction". In: *Springer Handbook of Model-Based Science*. Ed. by Lorenzo Magnani and Tommaso Bertolotti. Dordrecht, Heidelberg, London, New York: Springer, pp. 219–230.

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Canestrelli, Anneloes R., Willem M. Mak, and Ted J. M. Sanders (2013). "Causal connectives in discourse processing: How differences in subjectivity are reflected in eye movements". In: *Language and Cognitive Processes* 28.9, pp. 1394–1413. DOI: 10.1080/01690965.2012.685885.

Christensen, Rune H. B. (2020). *ordinal: Regression Models for Ordinal Data*. R package version 2018.4-19. URL: https://CRAN.R-project.org/package=ordinal.

Clifton, Chuck Jr., Adrian Staub, and Keith Rayner (2007). "Eye movements in reading words and sentences". In: *Eye movements: A window on mind and brain*. Ed. by Lorenzo Magnani and Tommaso Bertolotti. Amsterdam: Elsevier, pp. 341–372.

Hagoort, Peter and Joost J. A. van Berkum (2007). "Beyond the sentence given". In: *Philosophical transactions of the Royal Society of London* 362, pp. 801–811.

Halpern, Joseph Y. (2016). *Actual Causality*. Cambridge, Massachusetts/London, England: MIT Press.

Jäger, T.F. (2008). "Categorical data analysis: away from ANOVAs (transformation or not) and towards Logit Mixed Models". In: *Journal of Memory and Language* 59, pp. 434–446.

Köhne-Fuetterer, Judith et al. (2021). "The online processing of causal and concessive discourse connectives". In: *Linguistics* 59.2, pp. 417–448.

Neunaber, D. J. and E. A. Wassermann (1986). "The Effects of Unidirectional versus Bidirectional Rating Procedures on College Students' Judgments of Response-Outcome Contingency". In: *Learning and Motivation* 17 (2), pp. 162–179. DOI: 10.1016/0023-9690(86)90008-1.

Ng, Dawid W., Jessica C. Lee, and Peter F. Lovibond (2024). "Unidirectional rating scales overestimate the illusory causation phenomenon". In: *Quarterly Journal of Experimental Psychology* 77 (3), pp. 551–562. DOI: 10.1177/17470218231175003.

Noordman, Leo G. M. et al. (2015). "Causal inferences and world knowledge". In: *Inferences during reading*. Ed. by Edward J. O'Brien, Anne E. Cook, and Robert F. Jr Lorch. Cambridge: Cambridge University Press, pp. 260–289.

Pearl, Judea (2009). *Causality*. Cambridge: Cambridge University Press.

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Traxler, Matthew J., Michael D. Bybee, and Martin J. Pickering (1997). "Influence of connectives on language comprehension: Eye tracking evidence for incremental interpretation". In: *The Quarterly Journal of Experimental Psychology: Section A* 50.3, pp. 481–497.

Xiang, Ming and Gina R. Kuperberg (2015). "Reversing expectations during discourse comprehension". In: *Language, Cognition and Neuroscience* 30.6, pp. 648–672. DOI: 10.1080/23273798.2014.995679.

Expressive inferences & mighty obstacles

Andrés Soria-Ruiz
LOGOS, BIAP, University of Barcelona
Barcelona, Spain
asoriaruiz@ub.edu

Abstract

Predicates of personal taste (PPTs, i.e. *tasty*) carry so-called ACQUAINTANCE INFERENCES (AI), to the effect that the speaker has first-hand experience. In addition, Willer and Kennedy (2020) have observed a parallel phenomenon in the moral domain: moral predicates (MPs, i.e. *okay*) carry what they call PRACTICAL INFERENCES (PI), to the effect that the speaker has a practical stance towards whatever they are judging. My purpose is two-fold. First, whereas extant accounts have characterize modals and other operators as *obviating* AI/PI, the right generalization is stronger, namely that these inferences "flip" in those environments. Secondly, available theories provide either a principled account of AI that can hardly be extended to PI (Anand and Korotkova 2018; Ninan 2024), or a uniform but unexplanatory account of both inferences (Willer and Kennedy 2020). I propose a minimal amendment of Ninan 2024 so that it can be extended to PI in a principled way.

1 Data

1.1 The acquaintance inference of PPTs

Predicates of personal taste (PPTs, i.e. *tasty*) carry so-called ACQUAINTANCE INFERENCES (AI), to the effect that the speaker has first-hand experience (Ninan 2014; Pearson 2013; Bylinina 2017).[1] Compare *tasty* vs *vegetarian*:

(1) I have never tried Kaassoufflé.
   a. # ...It's tasty.
   b. ✓ ...It's vegetarian.

AI has peculiar properties. First, it cannot be cancelled (as just shown). Secondly, the AI has a peculiar projective pattern: whereas it projects under negation,

(2) I have never tried Kaassoufflé. # It's not tasty.

the AI does not project under modals and certain operators. But crucially, the AI does not just disappear in such environments; it *flips*. That is, the following sentences imply that the speaker lacks acquaintance (call this ¬AI):

(3) I have tried Kaassoufflé.
   a. # ...It might/must be tasty.
   b. # ...Probably, it is tasty.
   c. # ...If it is tasty, then...
   d. # ...Apparently, it is tasty.

[1] I set aside *exocentric* interpretations of PPTs, which do not carry this requirement (Stephenson 2007; Lasersohn 2005). In what follows, all PPTs are intended to be interpreted *autocentrically*. I similarly set aside PPTs with overt judges (*tasty to me*, see Anand and Korotkova 2018).

Whereas all theorists correctly observe that AI *vanishes* under the operators in (3) (Pearson 2013; Ninan 2014; Franzén 2018; Anand and Korotkova 2018; Willer and Kennedy 2020; Ninan 2024), they hardly make the stronger observation that AI *flips* in such environments (although see Anand and Korotkova 2018, p.13, ex. (39)). But it AI was merely obviated in these environments, the pattern in (3) would be surprising, as embedding PPTs under these operators should be compatible, not just with lacking acquaintance, but with having it as well.

1.2 The practical inference of MPs

Willer and Kennedy 2020 have uncovered a broadly parallel phenomenon in the moral domain: moral predicates (MPs, i.e. *okay*) carry what they call PRACTICAL INFERENCES (PI), to the effect that the speaker has a practical stance towards whatever they are judging. PIs are broadly similar to AIs, in that they project under negation and flip under operators:[2]

(4) I have no opinion about whether to commit tax fraud.

a. # ...It's okay. b. # ...It's not okay.

(5) I have an opinion about whether to commit tax fraud.

a. # ...It might/must be okay.
b. # ...Probably, it's okay.
c. # ...If it's okay, then...
d. # ...Apparently, it's okay.

2 Previous accounts

2.1 The implicature view: Ninan 2014

Ninan accounts for AI as a quality implicature, arising from a general norm of assertion (*assert what you know*, Williamson 1996), combined with a specific *Acquaintance Principle* for PPTs: If you know whether a is tasty, then you have tried a. This view rightly predicts the data about bare assertions and negation (the reader can check this). But it only makes the weaker prediction that in (3) the AI goes away. The explanation is that conversational implicatures similarly vanish in these environments.[3]

2.2 Presuppositional views: Anand and Korotkova 2018

Pearson 2013; Anand and Korotkova 2018 argue that PPTs carry an acquaintance presupposition. Anand and Korotkova 2018 (hf. A&K) encode this as a presupposition that the judge has direct evidence concerning the object. This direct-evidence requirement is captured using the notion of a KERNEL (von Fintel and Gillies 2010).

Definition 2.1 (Kernel)**.** A (K)ernel is a set of propositions that represent direct knowledge. K settles whether p iff $\exists q \in K[q \subseteq p \vee q \subseteq \neg p]$

Kernels are an additional parameter of evaluation, together with world and judges. PPTs presuppose that the judge's kernel settles the proposition:

[2]There's much more data on the behavior of AI/PI in different environments, such as attitude verbs, connectives, and quantifiers (see especially Ninan 2024; Willer and Kennedy 2020). Here, my focus is on AI/PI under modals.

[3]One might complain that any CI-view wrongly predicts that AI is cancellable. But as Ninan notes (2014, p. 297), this possibly makes AI a non-standard kind of conversational implicature.

(6) $[\![a \text{ is tasty}]\!]^{\langle j,w,K\rangle}$ = defined only if $K_{\langle j,w\rangle}$ settles whether a is tasty. If so,
$[\![a \text{ is tasty}]\!]^{\langle j,w,K\rangle} = 1$ iff a is tasty to j at w.

A&K then assign a meaning to modals where the direct evidence requirement of PPTs is overwritten, by shifting the kernel parameter in the prejacent from K to $\{\bigcap K\}$. $\bigcap K$ is meant as a vanilla doxastic state, comprising direct and indirect knowledge (Anand and Korotkova 2018, p. 10). However, as Willer and Kennedy 2020, p. 847 observe, $\{\bigcap K\}$ settles a proposition only if K settles it too, and thus the AI actually projects over 'might'. Thus, Ninan 2024, p. 8 suggests substituting $\{\bigcap K\}$ for $\mathcal{P}(W)$, the set of all propositions:

(7) $[\![\text{might } p]\!]^{\langle j,w,K\rangle}$ = defined only if $[\![p]\!]^{\langle j,w,\mathcal{P}(W)\rangle}$ = defined. If so,...

The problem now is that (7) merely *obviates* the acquaintance requirement. Thus, it does not predict the flip, i.e. (3a). To fix this, we should overwrite the kernel requirement, but so that K does *not* settle the prejacent (this, in fact, is von Fintel & Gillies' (2010) requirement). Here is the full entry for 'might':

(8) $[\![\text{might } p]\!]^{\langle j,w,K\rangle}$ = defined only if $K_{\langle j,w\rangle}$ does not settle whether p. If so,
$[\![\text{might } p]\!]^{\langle j,w,K\rangle} = 1$ iff $\exists w' \in \bigcap K_{\langle j,w\rangle} : [\![p]\!]^{\langle j,w',K\rangle} = 1$.

This provides a straightforward account of the contrast between (1a) and (2) vs. (3). However, there are 2 difficulties. First, note that 'a might be tasty' is defined only if 'a is tasty' is undefined. This implies that the former cannot be true. For if it is defined, then its prejacent will be undefined, and so there can be no epistemically accessible world at which it is true (this might just be a technicality, which can be solved by letting the kernel parameter shift in the assertive component of 'might' as well):

(9) $[\![a \text{ might be tasty}]\!]^{\langle j,w,K\rangle}$ = defined only if $[\![a \text{ is tasty}]\!]^{\langle j,w,K\rangle} \neq$ defined. If so,
$[\![a \text{ might be tasty}]\!]^{\langle j,w,K\rangle} = 1$ iff $\exists w' \in \bigcap K_{\langle j,w\rangle} : [\![a \text{ is tasty}]\!]^{\langle j,w',K\rangle} = 1$.

Second, it is not trivial to extend the presuppositional view to MPs/PI. A natural thought would be that MPs carry a "practical opinionated" presupposition (parallel to the acquaintance presupposition of PPTs). But given (8), embedding MPs under modals will not interact with such presupposition. Further tweaks in the semantics of modals would be necessary.[4]

2.3 Expressivist views

Expressivists treat AI as a *sui generis* inference, explained as follows: PPTs express *affective* states, which require acquaintance (Franzén 2018; Ninan 2024; Briesen n.d.). E.g., (1a) expresses *liking* pizza, and one cannot like something without trying it. In addition, Willer and Kennedy 2020 (hf. W&K) aim for a unified expressivist view of AI & PI. The thought for MPs is that they express *normative* states, which require "practical opinionatedness". E.g., (4a) expresses *accepting a norm* that sanctions tax fraud, which requires an opinion about whether or not to do it.

2.3.1 Willer and Kennedy 2020

W&K claim that different predicates come with specific GROUNDING REQUIREMENTS, that their use be based on certain aspects of the speaker's psychology, which they characterize

[4] In fairness, accounting for PI was never A&K's purpose.

as GROUNDING COMMITMENTS.[5] Grounding commitments are just specific sets of worlds representing propositions with special psychological status (similar to a set DOX representing plain belief):

- $\text{EXP}_{(a,w)}$: the propositions entailed by a's experiential attitudes at w
- $\text{NORM}_{(a,w)}$: the propositions entailed by a's norms at w

EXP/NORM *grounds* a proposition p iff $p \in$ EXP/NORM.

Next, predicates like *tasty* and *okay* carry specific grounding requirements:[6]

(10) a. $[\![a \text{ is tasty}]\!]^{\langle j,w\rangle}$ = defined only if $\text{EXP}_{(j,w)}$ grounds $[\![a \text{ is tasty}]\!]$. If so, $[\![a \text{ is tasty}]\!]^{\langle j,w\rangle} = 1$ iff a is tasty to j at w.

b. $[\![F\text{-ing is okay}]\!]^{\langle j,w\rangle}$ = defined only if $\text{NORM}_{(j,w)}$ grounds $[\![F\text{-ing is okay}]\!]$. If so, $[\![F\text{-ing is okay}]\!]^{\langle j,w\rangle} = 1$ iff F-ing is okay to j at w.

Then, W&K propose lexical entries for *not* and modals that respectively inherit and obviate the grounding requirement of their arguments

(11) a. $[\![\text{not-}p]\!]^{\langle j,w\rangle}$ = defined only if $[\![p]\!]^{\langle j,w\rangle}$ = defined. If so,...

b. $[\![\text{might } p]\!]^{\langle j,w\rangle}$ = defined only if $\mathcal{P}(W)$ grounds $[\![p]\!]^{\langle j,w\rangle}$. If so,....

Besides predicting only obviation under modals, a more specific concern with W&K's view is that it is somewhat *ad-hoc*. Recall that, for A&K, there is a principled reason why modals should overwrite the acquaintance presupposition, namely that modals require indirect evidence. W&K can co-opt that explanation for why modals overwrite *experiential* grounding requirements. But why should modals overwrite *normative* requirements too? Indirectness can't be the answer. This problem can be stated more dramatically: Why do modals, but not negation, overwrite grounding requirements? Why is it not the other way around? After all, we can define connectives/operators with switched grounding requirements:

(12) a. $[\![\text{not* } p]\!]^{\langle j,w\rangle}$ = defined only if $\mathcal{P}(W)$ grounds $[\![p]\!]^{\langle j,w\rangle}$ = defined. If so,...

b. $[\![\text{might* } p]\!]^{\langle j,w\rangle}$ = defined only if $[\![p]\!]^{\langle j,w\rangle}$ = defined. If so,....

The problem with W&K is that they provide no reason why, e.g., *might* behaves like (11b) and not like (12b).[7] So, W&K's descriptive breadth comes at some explanatory cost.

2.3.2 Ninan 2024

Ninan 2024 offers the most ambitious account of AI to date. His main empirical contribution is new data regarding PPTs under disjunction and quantifiers (see also Cariani 2021) which his theory predicts correctly. Behavior under modals is comparatively less prominent, and he inherits A&K's obviation mechanism. To get a grasp of his view, we need some definitions:

- A **generator** χ is a function from judges and worlds to categorical standards of taste

[5] W&K's view is couched in dynamic terms, but I present it in static terms to ease the presentation. I don't think anything crucial gets lost—and I hope the heresy is excused ;-)

[6] W&K claim that grounding requirements are not presuppositions, but I am formulating them as such to ease presentation.

[7] On the assumption that grounding requirements are not presuppositions, the same question arises for negation: why is negation (11a) rather than (12a)?

- A **categorical standard of taste** of judge j at world w, $\chi^{j,w}$, is a possibly partial function from D to {0,1} s.t.:
  - $\chi^{j,w}(o) = 1$ if j has tried and liked o
  - $\chi^{j,w}(o) = 0$ if j has tried and not liked o
  - not in the domain of $\chi^{j,w}$ if j has not tried o
- A **complete extension** σ of χ, $\sigma \succ \chi$, is a total function that agrees with χ on all the cases that χ decides, but decides all the other cases as well.

Sentences are evaluated at $\langle w, j, \sigma, g\rangle$-points. Here are the truth conditions for 'a is tasty' ($T(a)$):

(13) $[\![T(a)]\!]^{w,j,\sigma,g} = 1$ iff $\sigma^{w,j}(a^g) = 1$

The key to Ninan's account is the following assertability norm:

- **($\mathcal{A}$)ssertability norm**: $[\varphi]^c = \mathcal{A}$, iff $\forall\sigma \succ \chi$, $[\![\varphi]\!]^{w_c,j_c,\sigma,g_c} = 1$.

A sentence is assertable just in case it is true relative to every complete extension of the judge's categorical standard of taste. The following fall out of Ninan's account (*plus* Boolean negation):[8]

- $[T(a)]^c = \mathcal{A}$ iff $\chi^{w_c,j_c}(a) = 1$
- $[\neg T(a)]^c = \mathcal{A}$ iff $\chi^{w_c,j_c}(a) = 0$

To account for PPTs under modals, Ninan defines a *dispositional standard of taste* δ, defined as the things a judge is *disposed* to like. He then proposes that modals shift the taste parameter—roughly like A&K:

(14) $[\![\text{might } p]\!]^{\langle w,j,\sigma,g\rangle} = 1$ iff $\exists w' \in R(w) : [\![\text{might } p]\!]^{\langle w',j,\delta,g\rangle} = 1$ $(R(w) = \text{an epistemic state})$

2 problems remain: first, just like all other accounts, Ninan's view incorrectly predicts that AI disappears under modals, but not that it flips. Secondly, Ninan's view is not straightforwardly applicable to MP/PI. One can rehearse defining corresponding "moral" parameters (see below), but the result will be the same as in W&K: A good description of the facts, but at most a partial explanation for the interaction between modals & AI/PI.

In sum: there is no satisfactory account of the interaction between modals and expressive inferences. All accounts predict that modals *obviate* these inferences, but they don't predict the flip. Secondly, among available accounts, whereas A&K can explain why AI flips under modals (indirectness), there is no explanation for why PI would pattern similarly.

3 Expressivity & uncertainty

In what remains, I want to rehearse an explanation for the flipping behavior that can be extended beyond AI. Let us consider a different requirement of modals, namely *uncertainty*. The modals in (15) require that the speaker is uncertain about its prejacent p (Yalcin 2007, a.m.o.):[9]

[8] See Ninan 2024, fn. 31 & 34 for the relevant proofs.

[9] In fact, Goodhue 2017; Giannakidou and Mari 2016 argue that von Fintel and Gillies 2010's indirectness requirement boils down to uncertainty. We need not take a stance in that debate; it suffices to recognize that modals require uncertainty.

(15) It's raining.

a. # ...It might/must be raining.
b. # ...Probably, it is raining.
c. # ...If it is raining, then...
d. # ...Apparently, it is raining.

Uncertainty is a sufficiently general requirement to offer an explanation for the "flipping" behavior of AI and PI under modals. The idea is simple: uncertainty is incompatible with acquaintance / practical opinionatedness. Thus, PPTs/MPs under modals imply ¬AI / ¬PI.

Let us implement this idea enriching Ninan's account. We need to take 5 steps. First, define a categorical standard of morality and a complete extension thereof:

- A **categorical standard of morality** of judge j at world w, $\xi^{j,w}$, is a possibly partial function from D to {0,1} s.t.:
  - $\xi^{j,w}(F) = 1$ if j has a practical position and approves of F-ing
  - $\xi^{j,w}(F) = 0$ if j has a practical position and disapproves of F-ing
  - not in the domain of $\xi^{j,w}$ if j has not taken a stance on F-ing
- υ is a **complete extension** of ξ

Secondly, sentences are evaluated at $\langle w, j, \sigma, \upsilon, g\rangle$. Truth conditions for '$a$ is tasty' ($T(a)$) and 'F-ing is okay' ($O(F)$):

(16) a. $[\![T(a)]\!]^{w,j,\sigma,\upsilon,g} = 1$ iff $\sigma^{w,j}(a^g) = 1$ b. $[\![O(F)]\!]^{w,j,\sigma,\upsilon,g} = 1$ iff $\upsilon^{w,j}(F^g) = 1$

Thirdly, the assertability norm needs to be generalized too:

- **($\mathcal{A}$)ssertability norm (v2)**: $[\varphi]^c = \mathcal{A}$, iff $\forall\sigma \succ \chi, \forall\upsilon \succ \xi$, $[\![\varphi]\!]^{w_c,j_c,\sigma,\upsilon,g_c} = 1$.

Fourthly, by parallel reasoning as for $T(a)/\neg T(a)$, we predict that $O(F)/\neg O(F)$ are assertable iff the speaker has a practical position about F-ing:

- $[O(F)]^c = \mathcal{A}$ iff $\xi^{w_c,j_c}(F) = 1$
- $[\neg O(F)]^c = \mathcal{A}$ iff $\xi^{w_c,j_c}(F) = 0$

Lastly, turning to modals, we can write the uncertainty requirement as a presupposition:

(17) $[\![\text{might } p]\!]^{\langle w,j,\sigma,\upsilon,g\rangle}$ = defined only if j at w is uncertain whether p. If so,...

To generate the right predictions, we need the following psychological principles (for any x):[10]

- If x is uncertain whether $T(a)$, then x is not acquainted with a [ACQ ⇒ K]
- If x is uncertain whether $O(F)$, then x has not taken a stance on F-ing [PRAC ⇒ K]

It these principles are correct, then meeting the uncertainty requirement implies lacking acquaintance / practical stance. That is, the following come out true:

- [might $T(a)$]$^c = \mathcal{A}$ only if $a \notin dom(\chi^{w_c,j_c})$
- [might $O(F)$]$^c = \mathcal{A}$ only if $F \notin dom(\xi^{w_c,j_c})$

Proof for $T(a)$ (mutatis mutandis for $O(F)$): if [might $T(a)$]$^c = \mathcal{A}$, then $\forall\sigma \succ \chi, \forall\upsilon \succ \xi$, $[\![\text{might } T(a)]\!]^{w_c,j_c,\sigma,\upsilon,g_c} = 1$. If so, then by (17) j_c at w_c is uncertain whether $T(a)$. By [ACQ ⇒ K], j_c is not acquainted with a; that is, $a \notin dom(\chi^{w_c,j_c})$.

[10] These principles might need to be tweaked or made more precise—after all, it is possible to have tried a food and be unsure whether it's tasty, perhaps because it was neither clearly tasty nor not tasty, or because one does not remember its taste.

Acknowledgements. Thanks to Diego Feinmann and Mora Maldonado. This work is supported by grants BP-2022-00247 (funded by Generalitat de Catalunya & EU's Horizon 2020 research and innovation programme under MSCA grant agreement No 801370) and CEX2021-001169-M (funded by MICIU/AEI/10.13039/501100011033).

References

Anand, Pranav and Natasha Korotkova (2018). "Acquaintance content and obviation". In: *Proceedings of Sinn und Bedeutung*. Vol. 22. 1, pp. 55–72.

Briesen, Jochen (n.d.). "The Acquaintance Inference and Hybrid Expressivism". In: *Australasian Journal of Philosophy* ().

Bylinina, Lisa (2017). "Judge-dependence in degree constructions". In: *Journal of Semantics* 34.2, pp. 291–331. DOI: 10.1093/jos/ffw011.

Cariani, Fabrizio (2021). *The modal future: A theory of future-directed thought and talk*. Cambridge University Press.

Fintel, K von and A S Gillies (2010). "Must ... stay... strong!" English. In: *Natural Language Semantics* 18.4, pp. 351–383. DOI: 10.1007/s11050-010-9058-2.

Franzén, Nils (2018). "Aesthetic Evaluation and First-Hand Experience". In: *Australasian Journal of Philosophy* 96.4, pp. 1–14. ISSN: 0004-8402. DOI: 10.1080/00048402.2018.1425729. URL: https://www.tandfonline.com/doi/full/10.1080/00048402.2018.1425729.

Giannakidou, Anastasia and Alda Mari (2016). "Epistemic future and epistemic MUST: Nonveridicality, evidence, and partial knowledge". In: *Mood, aspect, modality revisited: New answers to old questions*, pp. 75–117.

Goodhue, Daniel (2017). "Must ϕ is felicitous only if ϕ is not known". In: *Semantics and Pragmatics* 10, 14–EA.

Lasersohn, Peter (2005). "Context dependence, disagreement, and predicates of personal taste". In: *Linguistics and Philosophy* 28.6, pp. 643–686. ISSN: 01650157. DOI: 10.1007/s10988-005-0596-x. URL: http://link.springer.com/10.1007/s10988-005-0596-x.

Ninan, Dilip (2014). "Taste predicates and the acquaintance inference". In: *Semantics and Linguistic Theory*. Vol. 24, pp. 290–309.

— (2024). "An Expressivist Theory of Taste Predicates". In: *Philosophers' Imprint* 24.1, pp. 1–23. ISSN: 1533-628X. DOI: 10.3998/phimp.2683. URL: https://journals.publishing.umich.edu/phimp/article/id/2683/.

Pearson, Hazel (2013). "A judge-free semantics for predicates of personal taste". In: *Journal of Semantics* 30.1, pp. 103–154.

Stephenson, Tamina C (2007). "Judge dependence, epistemic modals, and predicates of personal taste". In: *Linguistics and Philosophy* 30.4, pp. 487–525. ISSN: 01650157. DOI: 10.1007/s10988-008-9023-4. URL: http://link.springer.com/article/10.1007/s10988-008-9023-4.

Willer, Malte and Christopher Kennedy (2020). "Assertion, Expression, Experience". In: *Inquiry*.

Williamson, Timothy (1996). "Knowing and asserting". In: *The Philosophical Review* 105.4, pp. 489–523.

Yalcin, Seth (2007). "Epistemic modals". English. In: *Mind* 116.464, pp. 983–1026. DOI: Doi10.1093/Mind/Fzm983. URL: http://mind.oxfordjournals.org/content/116/464/983.full.pdf.

Loose nominals

Sreekar Raghotham
Rutgers University
sreekar.raghotham@rutgers.edu

Abstract

Proxy readings of nominals are a type of loose talk. They show the same empirical signature as other instances of loose talk. I model this behaviour by treating proxy readings as resulting from underspecification of individuals. This view helps better understand the behaviour of intensive reflexives and indexiphors.

1 Loose talk about nominals

Nominals can often be used to talk about entities that are representations of their referents — call such uses 'proxy readings'. Proxy readings don't seem to be available when the nominal is modified by an intensive reflexive: (1b) can only mean Ringo the drummer is on the second floor, while (1a) can be felicitously used to indicate the location of a statue of Ringo.

(1) a. Ringo is on the second floor
b. Ringo himself is on the second floor

This behaviour is reminiscent of loose talk (Sperber and Wilson 1985; Lasersohn 1999). There too, speakers felicitously deviate from the truth, with some modifiers ('regulators') curtailing allowable deviance: conversational goals allowing, (2a) is felicitous if John arrived at 9:05PM, but (2b) isn't.

(2) a. John arrived at 9PM
b. John arrived at 9PM exactly

Carter (2021) observes an asymmetry when there are two possibly loose expressions of the same kind (cf. Lewis 1979; Klecha 2018). The regulated instance of a loose utterance has to follow the unmodified instance, but not vice versa (3). Nominals behave similarly: proxy readings of nominals are allowed when they precede a regulated use of a nominal that refers to the same individual, but not vice versa (4).

(3) a. John left at 9 but he didn't leave at 9 exactly
b. #John left at 9 exactly but he didn't leave at 9.
c. John left at 9:03 exactly ... # However, John left at 9PM.

(4) a. Ringo is in the museum but Ringo himself isn't
b. #Ringo himself is in the museum, but Ringo isn't
c. Akhil said that he himself was in the green room.... # We found out that he$_{\mathbf{P}}$ was next to Chaplin

The presence of a regulating expression is crucial. A proxy reading of the anaphor is available even though it follows a standard interpretation of its antecedent (5). Indeed, (5a) is the standard example for proxy readings of anaphors (Jackendoff 1992, *a.o*). Moreover, the asymmetry is not limited to different instances of the same linguistic item: they survive so long as two nominals refer to the same individual (6).

(5) a. Ringo saw himself$_P$ in the wax museum
b. Each Beatle saw himself$_P$ in the museum

(6) a. Ringo$_P$ is in the museum but he himself isn't.
b. #Ringo himself is in the museum, but he$_P$ isn't.

It must be noted that regulators don't always curtail proxy readings, just as they don't always curtail loose readings of times (Lasersohn 1999, fn. 8). (7a) is felicitous even if the linguistics seminars don't start as the clock strikes 10. Similarly, in a context where there are only statues of Elvis and famous Elvis impersonators, (7b) is felicitous.[1]

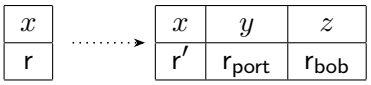
(7) a. While philosophy seminars start 10 mins late, linguistics seminars start at exactly 10 o'clock.
b. The impersonators are made of wax but Elvis himself is made of clay

These parallels lead me to conclude that proxy readings are a type of loose talk. The following sections outline a formal treatment of loose talk that preserves this parallel, and some empirical advantages of the view.

2 Analysis

Informally, loose talk is felicitous so long as the conversational goals allow it: If the differences between arriving at 9PM and 9:02PM are irrelevant, then (2a) is felicitous even if John arrived at 9:02 PM. Similarly, if differences between Ringo and his statue are irrelevant, then (1a) is a felicitous report of the location of Ringo's statue. Regulators make hitherto irrelevant distinctions relevant. For the case at hand, I suggest that the difference between an individual and its representation is made relevant by the intensifier.[2]

To model this informal view of loose talk, I will assume that discourse referents denote underspecified individuals (Shan 2015). When we utter sentences like (1a), being on the second floor is predicated of an underspecified Ringo Starr, whose possible precisifications include the drummer, his book, his portrait, a bobble-head in his likeness and so on. While world knowledge may whittle away some possibilities, discourse referents' denotations may remain underspecified felicitously. On this view, the contribution of an intensifier like *himself* in (1b) is to 'split' individuals. Intuitively, the denotation of the individual modified by the intensifier is sharpened, but may still remain underspecified. The individuals discarded are now values of other variables.



Formally, I adopt the framework in Shan 2015 and follow Carter in modeling the changes in relevance brought about by regulating expressions as modifications of accessibility relations (R). Accessibility relates two information states ($R \subseteq S \times S$, where S is a set of states). If state t is accessible from state s via the relation R, we write sRt. Truth at a state (s) for an assignment (g; at state s) is notated as $s, g_s \Vdash \phi$ (abstracting away from models). To model discourse

[1] Thanks to an AC reviewer for pointing this out. Another reviewer wonders whether the contrastive reduplication of Ghomeshi et al. (2004) is also a regulating expression, like the intensifier. At least to my ear, (7b) is infelicitous with 'RINGO ringo' in place of 'Ringo himself', which leads me to think that they are different phenomena with significant overlap in effects.

[2] Note that the distinctions can be made relevant by other means too. For instance, by means of an explicit reference to representations by means of 'a statue of x' or 'a picture of x'.

dynamics, we need a notion of a context. Contexts (σ), here, are pairs of a set of information states (S_σ), and an accessibility relation (R). Updating the context with a sentence with no regulating expressions, like (1a) whittles down the information states to those in which the sentence is true, leaving the rest of the context unchanged (8).

(8) $\sigma[\psi] := \langle \{t \mid s\, R\, t \wedge t, g_t \Vdash \psi\}, R\rangle$

Update with a sentence containing an intensifier ($\beta(\textsc{int}(\alpha))$) has an effect not just on the set of worlds, but also on the accessibility relation. The only states now accessible are the ones in which the discourse referent has split (modeling the intuition that regulation leads to sharpening), and the states that remain in the context are the R_α accessible ones at which $\beta(\alpha)$ is true (9a). In saying that splitting discourse referents discards individuals, I assume that variable names can be added, but not deleted or modified.[3] An accessible state in which an earlier discourse referent has split is captured by a requirement of multiple counterparts (9b; I and C are the familiar predicates from Lewis)

(9) a. $\sigma[\beta(\textsc{int}(\alpha))] := \langle \{t \mid s\, R_\alpha\, t \wedge t, g_t \Vdash \beta(\alpha)\}, R_\alpha\rangle$
b. $R_\alpha := \{s, t \mid I\, \alpha\, s \wedge I\, \alpha\, t \wedge \exists x\, [x \neq \alpha \wedge I\, x\, t \wedge C\, x\, \alpha]\}$

On this account, an update with an intensifier modifying 'Ringo' can in principle discard the drummer to be the value of a new variable, keeping only representations of Ringo in the earlier slot. An unfortunate consequence of this is that (1b) might be made true if Ringo's statue is on the second floor, contrary to intuition. To prevent this, I assume that there is a privileged precisification of an individual (an essence if you will) which can never be discarded.

The present treatment of loose talk about nominals is most compatible in spirit with an interval-based treatment of loose talk about times (Siu 2023). This is because unlike for Lasersohn (1999) and Carter (2021), a loose use of 'Ringo' does not make (1a) literally false on this account. Instead, there are various ways of making (1a) true, and intensification (and other means of sharpening conversational goals) reduces the number of such ways. The following section provides one justification for this move.

3 Empirical advantages

3.1 Intensifiers

This account improves on the standard account of adnominal intensive reflexive (Eckardt 2001). For Eckardt, the intensive reflexive is simply an identity function (10a), but triggers focus alternatives; the alternatives of the complex expression are not alternative individuals, like with emphatic focus, but individuals returned by alternative functions applied to the same individual (10b). While the resulting alternative set is the same, Eckardt suggests the latter derivation derives centrality effects not observed with plain emphatic focus: in a sense, the alternatives are 'grouped around' John, when modified by *himself*.

(10) a. $[\![\text{John himself}]\!]^o = \textsc{id}\,([\![\text{John}]\!]) = [\![\text{John}]\!] = \mathsf{j}$
b. $[\![\text{John himself}]\!]^f = \{g\,(\mathsf{j}) \mid g \in \textsc{alt}(\textsc{id})\}$ $\quad g :: \mathsf{e} \to \mathsf{e}$

Now consider that while (11a) is compatible with Ringo's statue being the sole occupant of the gallery, (11b) can only mean Ringo, the drummer is the sole occupant. This contrast cannot be explained by Eckardt's proposal, since the alternative set *Only* operates over is identical in both cases.

[3] This assumption will have to be suitably modified to account for discourse moves that 'merge' discourse referents, for instance in cases when conversational goals are loosened either through hedges, or some other linguistic means.

(11) a. Only Ringo is in the gallery
b. Only Ringo himself is in the gallery

On the present account however, since the use of an adnominal intensifier makes the distinctions between an individual and its representations contextually relevant, the representations too are members of the contextually determined alternative set. The centrality effects observed can still be derived à la Eckardt.

In the system sketched in the previous section, there is a natural notion of contextual relevance: being an active discourse referent. On an alternative view of loose talk like Lasersohn's (1999) or Carter's (2021), relevance is a matter of absence from a pragmatic halo, or due to being inaccessible from the current world. Such a view has trouble explaining the felicity of (11b) when a janitor is going about his job in the gallery, in addition to Ringo being there. By treating statues and janitors alike, we lose a formal footing for the notion of relevant individuals.[4]

3.2 Indexiphors aren't loose

We also gain purchase on the behaviour of indexiphors in Telugu — pronouns which are morphologically third person, but trigger first person agreement (Sundaresan 2012; Messick 2017; Messick 2023; Deal 2020). Indexiphors, like shifted indexicals require a *de se* interpretation. However, while shifted indexicals can receive proxy readings (12), indexiphors cannot (13a). Note that when the same pronoun triggers third person agreement, proxy readings are generally available (14). Indexiphors also show sensitivity to order: the continuation in (13b) is infelicitous even though the pronoun triggers third person agreement. Like we saw earlier, sensitivity to order is not linked to the use of the same expression more than once, but rather reference to the same individual. If *tanu* in (13b) refers to someone other than Akhil (or his representations), proxy readings are again available, like with 'Chaplin' in (13b).

(12) *Magahi*
anjaniyaa kahkai ki **ham** baṛhiyãã lagit hi
anjaniyaa said.3SG that 1SG good look be.1SG
'Anjani$_1$ said that she$_{1/P}$ looks good' Deepak Alok p.c.

(13) *Reporting on a trip Akhil and I took to a wax museum which has his statue* (Telugu)
a. Akhil **tanu** green room lo unnaa-**nu** ani annaaḍu
akhil 3SG green room LOC be-1SG COMP said
'Akhil$_1$ said that that he$_{1/*P/*2}$ was in the green room'
b. ... taravata, **tanu$_{*P}$** chaplin pakkana unnaa-ḍu ani čepææḍu
... later, 3SG chaplin next be-3MS COMP said.PST.3SG
'Later, he$_1$ said that he$_{1/*P/2}$ was next to Chaplin$_{3/P}$'

(14) *(same context as above)*
Akhil **tanu** green room lo unnaa-**ḍu** ani annaaḍu
akhil 3SG green room LOC be-3MS COMP said
'Akhil$_1$ said that that he$_{1/P/2}$ was in the green room'

Messick (2023) shows that it is the pronoun itself that triggers first person agreement, and argues that indexiphors have both third and first person features, with person features indexed to contexts (Schlenker 2003). Concretely, third person singular indexiphors are feature bundles of the following sort (D-BOUND indicates anaphoricity, –C is the embedded context and +C is the matrix context):

[4] A technique to record the history of halos or accessible worlds might answer this objection.

(15) *Feature structure of indexiphors* Messick 2023, p. 158
[⟨+AUTHOR, −C⟩, ⟨−AUTHOR, −ADDRESSEE, +C⟩, -PLURAL, D-BOUND]

To account for the lack of proxy readings here, I suggest that indexiphors are pronouns with built-in regulators. It seems reasonable to treat the third person features as intensive, since the pronoun can still be used in the language as an intensifier of nominals. The reason the pronoun surfaces as third person, following Messick 2023, is a morphological fact about the language.

4 Conclusion

Proxy readings of nominals are a type of loose talk. They show the same empirical signature as other instances of loose talk. I argued that this behaviour can be captured by appealing to underspecified individuals, that can be split as the discourse progresses. This view helps better understand the behaviour of intensive reflexives—they make representations of their associate relevant—and indexiphors—nominals with built-in regulators.

An obvious way of taking this approach forward is to tie in the present account of loose talk about nominals to other instances of loose talk (about times, fictional entities, heights) to arrive at a general theory of loose talk that does justice to its sensitivity to order. Another lacuna of the present account is a compositional implementation of regulation. I leave both for future work.

Acknowledgements. Thanks to Mark Baker, Simon Charlow, the four anonymous reviewers of AC2024, and audiences at Rutgers and Princeton universities for comments.

References

Carter, Sam (2021). "The dynamics of loose talk". In: *Noûs* 55.1, pp. 171–198. URL: https://doi.org/10.1111/nous.12306.

Deal, Amy Rose (2020). *A theory of indexical shift: Meaning, grammar, and crosslinguistic variation*. Cambridge, MA: MIT Press.

Eckardt, Regine (2001). "Reanalysing selbst". In: *Natural language semantics* 9.4, pp. 371–412.

Ghomeshi, Jila et al. (2004). "Contrastive focus reduplication in English (the salad-salad paper)". In: *Natural language & linguistic theory* 22, pp. 307–357.

Jackendoff, Ray (1992). "Mme. Tussaud meets the binding theory". In: *Natural Language & Linguistic Theory* 10.1, pp. 1–31. URL: https://doi.org/10.1007/BF00135357.

Klecha, Peter (2018). "On unidirectionality in precisification". In: *Linguistics and Philosophy* 41, pp. 87–124.

Lasersohn, Peter (1999). "Pragmatic halos". In: *Language*, pp. 522–551. URL: https://doi.org/10.2307/417059.

Lewis, David (1979). "Scorekeeping in a Language Game". In: *Journal of Philosophical Logic* 8, pp. 339–359. URL: https://doi.org/10.1007/BF00258436.

Messick, Troy (2017). "The morphosyntax of self-ascription: A cross-linguistic study". PhD thesis. Storrs, CT: University of Connecticut. URL: https://digitalcommons.lib.uconn.edu/dissertations/1621.

— (2023). "On apparent pronominal feature contradictions: Shifty agreement in Telugu". In: *Syntax* 26.2, pp. 137–182. URL: https://doi.org/10.1111/synt.12248.

Schlenker, Philippe (2003). "A plea for monsters". In: *Linguistics and philosophy* 26.1, pp. 29–120. URL: https://doi.org/10.1023/A:1022225203544.

Shan, Chung-chieh (2015). "Splitting hairs". In: *Proceedings of the 20th Amsterdam Colloquium*, pp. 363–367.

Siu, Charlie (2023). "Interval-based dynamics of loose talk". In: *Synthese* 202.1, p. 10.
Sperber, Dan and Deirdre Wilson (1985). "Loose talk". In: *Proceedings of the Aristotelian society*. Vol. 86, pp. 153–171. URL: https://www.jstor.org/stable/4545042.
Sundaresan, Sandhya (2012). "Context and (co)reference in the syntax and its interfaces". PhD thesis. Universitetet i Tromsø and Universität Stuggart. URL: https://hdl.handle.net/10037/4835.

Making a statement: eventuality denoting nominals

Peter R. Sutton
Universität Potsdam
Potsdam, Germany
peter.r.sutton@icloud.com

Hana Filip
Heinrich-Heine-Universität, Düsseldorf
Düsseldorf, Germany
hana.filip@gmail.com

1 Introduction

This paper has three principal aims. First, we address a question raised in Parsons 1990, §10.6, positively answered in Schwarzschild 2022: Do all common nouns (CNs) denote eventualities, namely only STATES? We argue that the answer to this question is *no*. To show this, we develop two novel diagnostic tests, and apply them to the data from Czech (Slavic) and English (Germanic), illustrating their cross-linguistic reach. One test is based on the light verb construction (LVC) and the other on genitive constructions. These tests determine if a CN has an eventuality-denoting sense, e.g., *belief*, *party* and *allegation*, and separate this class of CNs from those that do not (e.g., *boat*, *cat*).

Second, we apply well-known tests developed in event semantics for the classification of aspectual classes, fruitfully applied to multiple languages, to the class of eventuality-denoting CNs in order to identify if they have eventuality-denoting senses that are telic. To our knowledge, this is a novel extension of the use of these diagnostic tests to the nominal domain.

Third, for CNs that denote telic predicates, we motivate what counts as 'one CN-event' in their denotation, and so their telicity, by anchoring eventualities they denote to individuated entities (and times) via their thematic roles (see Grimm 2014 on anchoring for abstract CNs). By also controlling for polysemy, we thereby make successful predictions concerning the felicitous occurrence and interpretation of polysemous CNs in counting constructions. Take *allegation*, which is polysemous between an 'informational-entity' (propositional content) reading and an EVENT (telic) reading. In (1-a), we can directly count two informational entities denoted by *allegation*: *that Billie lied* and *that Dom swore*. In (1-b), the counting of two informational entities is ruled out, since there is only one thing alleged (*that Charlie lied*), however, the two Agents (Alex and Billie) allow *anchoring* to two discrete allegation-EVENTS, one performed by Alex, the other by Billie. In (1-c), neither counting of two informational entities nor counting of allegation-EVENTS is possible, and as a result (1-c) is uninterpretable.

(1) a. Alex's two allegations that Cal lied and Dom swore were true.
b. Alex and Billie's two allegations here at exactly 2:03pm that Cal lied upset Dom.
c. #Alex's two allegations here at exactly 2:03pm that Cal lied upset Dom.

2 Do all nouns denote eventualities?

Parsons (1990) asks whether regular CNs like *cat*, instead of denoting individuals, could denote STATES in which those individuals are participants (e.g., single cats as the Theme of cat states). He concludes that this view lacks evidential support, but Schwarzschild (2022) develops it explicitly. If such STATES are supposed to be part of the STATES, PROCESSES, EVENTS classification of eventuality types (see Mourelatos 1978), as also Parsons (1990) assumes, we think Schwarzschild (2022)'s move is wrong-footed, since it blurs two important distinctions: one between CNs that can denote EVENTS (e.g., *statement*, *party*), on the one hand, and (mental) STATES (e.g., *belief*, *fear*), on the other hand; and another between CNs that denote (mental) STATES (e.g., *belief*,

fear) and CNs like *boat* and *cat* whose STATE status is, at best, unintuitive. We devise two tests to demarcate eventuality-denoting CNs for Czech and English. The first test is based upon the light verb construction, and the second upon genitive constructions.

2.1 The Light Verb Construction (LVC) Test

If a CN can be felicitously used in a LVC, it has at least one sense in which it denotes (a set of) eventualities. In LVCs the verb is semantically bleached of its 'ordinary' meaning (see e.g., Pullum and Huddleston 2002, ch. 4, §7). This demarcates the CNs in (2) from those in (3) for English and likewise the CNs in (4) from those in (5) for Czech.

(2) a. Alex made that {allegation | claim | statement}. +LVC
b. Alex had that {belief | fear | hope | party}. +LVC

(3) a. Alex {gave someone | had} that {fact | information}. −LVC
b. Alex {made | had | took | gave someone} that {boat | cat}. −LVC

(4) a. Alex udělal toto {prohlášení | tvrzení}. +LVC
Alex do.PAST this {statement | claim}.
'Alex made this {statement | claim}.'
b. Alex měl {tento strach | tuto domněnku / pařbu / naději}. +LVC
Alex have.PAST {this fear | this belief / party / hope}
'Alex had this {fear | belief | party | hope}.'

(5) a. Alex {dal někomu | měl} {tento fakt | tuto informaci}. −LVC
Alex {gave.PAST someone | have.PAST} {this fact | this information}
'Alex {gave someone | had} this {fact | information}.'
b. Alex {dal někomu | měl} tuto {lod' | kočku}. −LVC
Alex {gave.PAST someone | have.PAST} this {boat | cat}
'Alex {gave someone | had} this {boat | cat}.'

2.2 Genitive Construction (GC) Tests

For English, we use the Saxon genitive construction, and for Czech, the genitive case construction to test whether a CN has an eventuality-denoting sense.

The GC Test. In a GC, A.GEN B, if B denotes an eventuality, then A can be a participant (e.g., Agent, Theme, Experiencer, Instrument) in that eventuality.

In English, the GC test demarcates the CNs in (6) from those in (7) (furthermore in (7b), the relation between A and B is highly underspecified):

(6) a. Alex's allegation/claim/party ≈ the alleging/claiming/party EVENT to which Alex stands in the Agent relation
b. Alex's belief/fear/hope ≈ the STATE of belief/fear/hope to which Alex stands in the Experiencer relation

(7) a. Alex's information ≉ the information STATE (or EVENT) in which Alex is the Experiencer/Agent/Theme. (If *information* denoted eventualities, and information were about Alex, then the Theme in that eventuality would be an INF(ormational) ENTITY, not Alex.)
b. Alex's boat/cat ≉ the boat/cat STATE to which Alex (as the owner participant) stands in the Experiencer relation

In Czech, similarly to English, the GC test demarcates the CNs in (8), for which the same approximate equivalences can be given as in (6), from those in (9), which, like the English

constructions, fail to have the relevant equivalences.[1]

(8) a. Alexův {argument | večírek}
Alex.M.GEN {argument | party}
Alex's {argument | party}

b. Alexova {víra | naděje}
Alex.F.GEN {belief | hope}
Alex's {belief | hope}

(9) a. Alexova informace
Alex.F.GEN information
Alex's information

b. Alexova {lod' | kočka}
Alex.F.GEN {boat | cat}
Alex's {boat | cat}

2.3 Summary

Our LVC and GC tests for both Czech and English clearly differentiate between CNs that intuitively have at least one sense that denotes eventualities (e.g., *allegation*, *belief*, *claim*, *fear*, *party*, *statement*), from those that do not (*boat*, *cat*, *fact*, *information*). Given the classification of aspectual classes into STATES-PROCESSES-EVENTS, this constitutes evidence against any theory that posits that all CNs, including e.g., *boat* and *cat*, denote STATES.

3 Telicity of eventuality-denoting CNs.

In this section, for CNs that have an eventuality-denoting sense, we motivate their aspectual class. (As is standard in event semantics, we assume a tripartite distinction into EVENTS, PROCESSES or STATES, originally set up by Mourelatos 1978.) This serves two purposes. First, it will strengthen our findings, summarized in section 2, that CNs like *boat* and *cat* are not eventuality-denoting, let alone STATE-denoting. Second, for CNs that are eventuality-denoting, it will allow us to investigate the effect of aspectual class on their countability property.

Most importantly, the above methodological assumptions lead to our novel proposal: namely, we propose the use of CNs (having an eventuality-denoting sense) in LVCs as a test for their aspectual class. To this goal, we show how we can capitalize on a battery of tests that are standardly used in event semantics to determine the aspectual class of expressions of various 'sizes' (verbs taken as lexical items, VPs or sentences) (see, e.g., Dowty 1979, p. 55ff, and Filip 2019 for an overview and a discussion). For reasons of space, here we only focus on English data, but these diagnostic tests have been applied extensively cross-linguistically. We also observe, as has also been cautioned by some (see Filip 2019, p. 282, and references therein), that these tests even in English do not converge on coherent and neatly separated categories, such as Vendler's (1957), but identify overlapping clusters which merely distinguish subsets of such categories or supersets. Moreover, not all these tests that were originally developed for English data are (directly) transferable to other languages, due to their language-specific properties. It may not be always entirely clear what the semantic properties of linguistic forms are that the intended analogous diagnostic tests (e.g., compatibility with temporal adverbials) used by various researchers identify in different languages, and some clarification is needed to determine whether they in fact access the same aspectually relevant properties in different languages.

In (10) and (11), '(?)' indicates that the construction is not felicitous unless coerced into an iterative interpretation. The predicates of eventualities in the LVCs in (10) pattern as EVENTS and those in (11) as STATES (see Table 1 for a results summary):

(10) a. Alex made that {statement | claim | allegation} {in under 2 minutes | (?)for 2 minutes/hours}.
b. Alex made that {statement | claim | allegation} {three times | (?)a lot}.

(11) a. Alex had that {belief | hope} {?in 5 years | for 5 years (while in grad school)}.

[1]We give fewer Czech examples, due to number and gender agreement between the head and genitive modifier.

Table 1: Results of applying our tests: whether nouns denote eventualities, and if so, which kind of eventualities (STATES (SC) or EVENTS (EV)).

Noun	boat	cat	fact	information	allegation	claim	party	statement	belief	hope	fear
LVC test	0	0	0	0	1	1	1	1	1	1	1
GCC test	0	0	0	0	1	1	1	1	1	1	1
Eventuality (if any)	–	–	–	–	EV	EV	EV	EV	ST	ST	ST

b. Alex had that {belief | hope} {?three times | ?a lot}.

Interestingly, we have not found cases of LVCs that suggest that the relevant noun denotes a PROCESS (e.g., the PROCESS denoting *sleep* patterns differently to both EVENT and STATE denoting expressions: *slept for/?in 2 hours, slept a lot/(?)three times*). Indeed, we observe that LVC constructions with nouns related to prototypical PROCESS VPs are not PROCESS-, but EVENT-denoting: *Alex jogged* (PROCESS) vs. *Alex took/had a jog* (EVENT). In the following, we restrict ourselves to discussing EVENTS and STATES, setting PROCESSES aside.

This evidence, combined with that from section 2, allows us to strengthen our conclusions. We now have a set of CNs that can denote eventualities, and, of these, some that can denote STATES (i.e., *belief* and *fear*). Our diagnostics clearly differentiate these CNs from CNs like *boat* and *cat*, suggesting that if there is a sense in which the latter denote 'states', these do not seem to be the same as the STATES that are distinguished in semantics as a subtype of eventuality. Our tests do not rule out views in which all CNs denote *situations*, provided that different sorts of entities (e.g., physical, informational, and eventualities) can witness these situations. See e.g., Sutton's (2022) account of polysemy which uses Cooper's (2023) Type Theory with Records.

4 Counting eventualities in the denotations of CNs

Our account expands the empirical coverage of the informal analyses in Sutton and Filip 2019 and Sutton and Filip 2020, as well as that of Grimm (2014). We claim: (i) EVENT-denoting senses of CNs are typically countable (e.g., we can count *allegation* qua EVENTS); (ii) What counts as 'one' such EVENT depends on anchoring relations to e.g., Agents, spatio-temporal locations, or Themes (see Grimm 2014); (iii) STATE-denoting senses of CNs are not countable, but anchoring to Experiencers or Stimuli can coerce their inherent mass senses into count (e.g., *?three fears*, see 14-b); (iv) INF ENTITIES in the denotations of polysemous CNs such as *belief* are countable without anchoring, even if they also have a (mass) STATE-denoting sense.

Thematic role assumption. EVENTS may be defined for the full range of thematic roles, and are homomorphically mapped to their temporal traces. Following Parsons 1990, we assume STATES have, at most, *Experiencer*, *Instrument*, and *Theme*.

Quantization assumption. Grammatical counting turns on enumerating quantized sets of entities relative to a context (we suppress details regarding contexts below). Quantized sets have no two members in a proper part relation: $QUA(P) \leftrightarrow \forall x, y[(P(x) \wedge P(y)) \rightarrow \neg x \sqsubset y]$ (see Krifka 1989).

EVENT-Anchoring The cardinality of a set of EVENTS in the denotation of a CN supervenes on the cardinality of a set of anchors: *two allegations* denotes two EVENTS only if there is a quantized set of two Agents, two temporal traces, or two locations. Counting with (predicates of) EVENTS via 'anchoring' to their participants, temporal traces or locations has its origins in Davidson 1969's view that we can identify eventualities in terms of the objects to which they are related. This is implemented in mereological terms by Krifka 1989, and has its first application to abstract CNs with Grimm's (2014) (coining of the term) *participant anchoring*. We give a set of anchors relative to a CN predicate, $\mathcal{P}$, and an anchoring relation, f in (12) and the cardinality

of EVENTS relative to $\mathcal{P}$ and f in (13).

(12) $ANCH(e_v, \mathcal{P}_{\langle v,t\rangle}, f_{\langle v,e\rangle}) \stackrel{\text{def}}{=} \lambda x.\exists e'[e' \sqsubseteq e \wedge \mathcal{P}(e') \wedge f(e') = x]$, where $f \in \{\text{AG}, \text{TH}, \tau, \text{LOC}\}$
The set of anchors of a sum eventuality e relative to an anchoring relation f and an eventuality-denoting predicate $\mathcal{P}$ is the set of f-participants of the $\mathcal{P}$-parts of e.

(13) $\mu_{ev}(e_v, \mathcal{P}_{\langle v,t\rangle}, f_{\langle v,e\rangle}) \stackrel{\text{def}}{=} |ANCH(e, \mathcal{P}, f)|$ if $QUA(ANCH(e, \mathcal{P}, f)), \perp$ otherwise.
A sum eventuality e counts as n $\mathcal{P}$s relative to anchoring relation f iff the cardinality of the set of f anchors of e for $\mathcal{P}$ is n, presupposing that this set is quantized.

Grimm (2014) claims that CNs like *despair* in *his despairs* (ex. 25a), are polysemous between a mass, STATE-denoting reading and "additional event-based readings, which are countable" (p.197). We disagree, there is only one mass STATE-denoting reading. If despair truly had a event-denoting countable sense, a numeral construction such as *??He had three despairs last year* should be straightforwardly felicitous (a signature property of count nouns), but it is not. Our analysis of Grimm's *his despairs* example is that it is STATE-denoting, but has a coerced count interpretation available via anchoring to discrete times (i.e. discrete periods of despair). This is analogous to a coerced *interruption reading* common for states denoted by individual-level predicates such as *Francis is occasionally blond* (Fernald 2000, p. 70).

Anchor blocking. We make the novel observation that polysemy blocks the use of an anchor: We cannot anchor a given sense of a CN via the lexical material of another sense of that CN. If N is polysemous between senses $S_{\langle v,t\rangle}$ and $S'_{\langle \sigma,t\rangle}$, then $S'_{\langle \sigma,t\rangle}$ cannot anchor $S_{\langle v,t\rangle}$. Take *allegation* which has two senses: (i) an INF ENTITY, that which is alleged (e.g., *A's two allegations (that B lied and that C swore)*; (ii) an EVENT of making an allegation (e.g., *A's allegation took a few seconds*). Understood as INF ENTITIES, *allegations* are counted directly. For instance, A's single utterance of *B and C both lied* (at one space-time location) can be described as *one allegation* or *two allegations* qua INF ENTITIES, depending on how *B and C both lied* is individuated in the context. However, A's single utterance of *B and C both lied* cannot be counted as *two allegations* qua EVENTS: *#A's two allegations each took a few seconds* is not felicitous in this context. Anchor blocking captures this. Understood as EVENTS, *allegation* must be counted relative to an anchor (see EVENT-anchoring, above). If INF ENTITIES could anchor *allegations*-qua-EVENTS, there should be a felicitous reading of *A's (two) allegations each took a few seconds* in the above context, contrary to fact. Furthermore, anchor blocking does not apply in the absence of polysemy. The EVENT-denoting sense of a CN such as *party* can be anchored via *Occasions* for a party (i.e., A' and B's PhD defences), or *Beneficiaries* (A and B qua their defences), because such CNs do not have a separate sense that denotes those Occasions, Beneficiaries etc.

5 Predictions

EVENT-denoting CNs. The cardinality of event-denoting senses of CNs supervene on cardinalities of quantized anchor sets. (1-b) and (1-c) are repeated below. In (1-b), the set of AGents $\{alex, billie\}$ is the only anchor set with cardinality 2. The events are otherwise concurrent and co-located. *Allegations*-qua-INF ENTITIES cannot count as *two* due to the single INF ENTITY denoted by *Charlie lied*. This correctly predicts that if we mention only one Agent (Alex), as in (1-c), the construction is not interpretable.

(1) b. Alex and Billie's two allegations here at exactly 2:03pm that Cal lied upset Dom.
c. #Alex's two allegations here at exactly 2:03pm that Cal lied upset Dom.

STATE-denoting CNs. CNs that only denote STATES are (typically) mass (see Mourelatos 1978). E.g., *fear* denotes STATES (of being in fear) and is mass. Mass nouns have cumulative

reference Quine 1960, and singular mass nouns e.g., *fear* can denote sums of states with both distinct Stimuli and Experiencers (14-a). Attempting to anchor STATES to Experiencers or Stimuli results in coercion: (14-b) and (15-b) are marked, but interpretable.

(14) a. Alex and Billie's fear of spiders and long flights are why they won't travel to Australia.
b. ?Alex's two fears of spiders and long flights are why she won't travel to Australia.

Interactions with polysemy. Mourelatos (1978) did not account for polysemy: *belief* is INF ENTITY/STATE polysemous, and the INF ENTITY sense (the Theme of the STATE) is countable without anchoring. Our *anchor blocking* blocks counting *belief* STATES via their Themes. One cannot select the mass (STATE) denoting sense of *belief* in counting constructions, because *belief* has a count INF ENTITY-denoting sense (*two beliefs* ≠ two *belief*-STATES). This explains the judgements in (15). (15-a) directly counts beliefs-qua-INF ENTITIES (*that Cal's birthday is tomorrow* and *that Dom's birthday is on Friday*). In (15-b), however, counting of two informational entities is ruled out, since there is only one thing believed (*that Cal's birthday is tomorrow*). The only recourse is try to anchor to Experiencers (Alex and Billie). This in turn just might make available a coerced count interpretation of what is inherently a mass STATE sense of *belief*. In (15-c), for *two beliefs* neither counting of two informational entities nor counting of (coerced) individuated STATES is available, and as a result (15-c) is uninterpretable.

(15) a. Alex's two beliefs that Cal's birthday is tomorrow and Dom's is on Friday are why they went shopping.
b. ?Alex and Billie's two beliefs that Cal's birthday is tomorrow are why they went shopping.
c. #Alex's two beliefs that Cal's birthday is tomorrow are why they went shopping.

EVENT-denoting CNs that are not relevantly polysemous (e.g., *party*) are not affected by anchor blocking. In (16-a), *party*-EVENTS are anchored to *Occasions* (the two defences of Alex and Bill) This is blocked in (16-b) resulting in uninterpretability.

(16) a. The two simultaneous parties for Alex and Billie's defences here at 2pm were attended by the same people.
b. #The two simultaneous parties for Alex's defence here at 2pm were attended by the same people.

6 Conclusion

Our novel and crosslinguistically applicable LVC and GC tests classify CNs that have an eventuality-denoting sense, for which we show that tests adapted from event semantics categorise their aspectual class. We argued, building on Grimm 2014, that counting with the eventuality-denoting senses of CNs requires *anchoring* to thematic roles, but also that aspectual class constrains what anchors are available. Given a contextually available anchor, we can motivate why the EVENT-denoting senses of *allegation* and *party* are easily countable, but also, in contrast to Grimm 2014, why this prompts coercion for STATE-denoting senses of *belief* and *fear*. We also identified that the polysemy of abstract CNs, e.g. *allegation*-qua-EVENT vs. *allegation*-qua-INF-ENTITY, is crucial to understanding their countability: (i) in contrast to Mourelatos 1978, we motivated why not all STATE-denoting CNs are uncountable; e.g., *belief* is countable when its INF-ENTITY-denoting sense is selected; (ii) we identified via *anchor blocking*, why an INF-ENTITY sense of a CN cannot anchor its EVENT-denoting sense; e.g., *allegation*-EVENTS cannot be anchored to *allegation*-INF-ENTITIES, because *allegation* is EVENT/INF-ENTITY polysemous.

Acknowledgements. This work has been supported by a grant awarded to Hana Filip by the German Research Foundation (Deutsche Forschungsgemeinschaft) for the project entitled "Individuation of Eventualities and Abstract Things" (2020-24). Peter Sutton additionally received funding from the University of Potsdam via a postdoctoral bridge stipend.

References

Cooper, Robin (2023). *From Perception to Communication: a Theory of Types for Action and Meaning*. Oxford University Press.

Davidson, Donald (1969). "The individuation of events". In: *Essays in honor of Carl G. Hempel*. Springer, pp. 216–234.

Dowty, David (1979). *Word Meaning and Montague Grammar*. Dordrecht: Reidel.

Fernald, Theodore B. (2000). *Predicates And Temporal Arguments*. Oxford University Press.

Filip, Hana (2019). "9. Aspectual class and Aktionsart". In: *Semantics - Noun Phrases and Verb Phrases*. Ed. by Paul Portner, Klaus Heusinger, and Claudia Maienborn. Berlin, Boston: De Gruyter Mouton. Originally published in Maienborn, von Heusinger, and Portner (eds.) "Semantics: An International Handbook of Natural Language Meaning", 2011, pp. 1186–1217., pp. 274–312. ISBN: 9783110589443. DOI: doi:10.1515/9783110589443-009.

Grimm, Scott (2014). "Individuating the abstract". In: *Proceedings of Sinn und Beduetung* 18. Ed. by Urtzi Etxeberria et al., pp. 182–200.

Krifka, Manfred (1989). "Nominal Reference, Temporal Constitution and Quantification in Event Semantics". In: *Semantics and Contextual Expression*. Ed. by R. Bartsch, J. F. A. K. van Benthem, and P. van Emde Boas. Foris, pp. 75–115.

Mourelatos, Alexander (1978). "Events, processes and states". In: *Linguistics and Philosophy* 2, pp. 415–434.

Parsons, Terence (1990). *Events in the Semantics of English*. Vol. 5. Cambridge, Ma: MIT Press.

Pullum, Geoffrey and Rodney Huddleston, eds. (2002). *The Cambridge grammar of the English language*. Cambridge University Press.

Quine, W (1960). *Word and object: An inquiry into the linguistic mechanisms of objective reference*. John Wiley.

Schwarzschild, Roger (2022). "Pure Event Semantics". Submitted. Manuscript made available on LingBuzz https://lingbuzz.net/lingbuzz/006888.

Sutton, Peter R. (2022). "Restrictions on copredication: a situation theoretic approach". In: *Semantics and Linguistic Theory (SALT)* 32, pp. 335–355.

Sutton, Peter R. and Hana Filip (2019). "Singular/Plural contrasts: The case of Informational Object Nouns". In: *Proceedings of the 22nd Amsterdam Colloquium*, pp. 367–376.

— (2020). "Informational Object Nouns and the mass/count distinction". In: *Proceedings of Sinn und Bedeutung 24* 2, pp. 319–335.

Vendler, Zeno (1957). "Verbs and Times". In: *The Philosophical Review* 66.2, pp. 143–160.

On Being Certain that Presuppositions don't Project Universally

Maik Thalmann
University of Göttingen
Göttingen, Germany
maik.thalmann@gmail.com

Andrea Matticchio
UMass Amherst
Amherst MA, USA
amatticchio@umass.edu

Abstract

In this paper, we present new experimental evidence that presupposition projection out of quantified environments can be modeled in a framework that uses a trivalent logic like Strong Kleene. Assuming a quantificational analysis, our experiment focuses on projection out of attitude reports, which are understudied from the empirical point of view. Theories that assume generalized universal projection, which are viewed as standard, cannot capture our data, which are instead compatible with the idea that presupposition failures can take part in the meaning computation. In the data analysis, we show how modeling standard deviations using a distributional model allows for new insights by enabling the detection of the squeamishness characteristic of presupposition failure in participants' response behavior.

1 Introduction

The projection problem for presuppositions is the problem of predicting what presupposition a complex sentence will impose on the context given the presupposition triggers in its constituents. A long line of research has addressed this problem as part of compositional semantics. Descriptively, logical operators have traditionally been associated with specific behaviors with respect to presupposition projection, stipulated as part of the information that comes with the lexical meaning of the operator (as in the classification of *holes*, *plugs*, and *filters* in Karttunen 1973).

Instead, we here pursue the hypothesis that the projection behavior of an operator does not have to be stipulated, but follows from the logical contribution of the operator and a logic dealing with truth-value gaps. We hypothesize that this logic is Strong Kleene for all domains of quantification. In this paper, we focus on presupposition projection from the scope of quantifiers because there is, as yet, no consensus about what empirical picture a theory of projection should derive for quantifiers over individuals. To see this, consider the variants of the sentence in (1), where the argument of the presupposition trigger *stop* contains a variable bound by a quantifier.

(1) All/Some/None of the students λx [x stopped smoking.]

At first glance, it seems reasonable to assume that the sentence in (1) with *all* presupposes that all of the students used to smoke, a universal presupposition. However, for existentially quantified presuppositions, as in (1) with *some*, Karttunen and Peters (1979) already suggest that a universal presupposition is too strong. A first proposal for these cases was fully spelled out by Heim (1983): for all the sentences above, the presupposition should carry universal quantificational force, but additionally, different strategies can apply and weaken the presupposition.

Chierchia (1995) suggests that the situation is more complex and that universal and existential presuppositions can follow from different logical forms of the same sentence. Later, Sudo et al. (2012) report a large degree of inter-speaker variation. Finally, work by Fox (2013) (building on George 2010) argues that a system based on the trivalent Strong Kleene logic allows for an adequate model of presupposition projection for (1) that is in line with the experimental results in Chemla (2009). Again, the predictions are qualified by the availability of mechanisms that can weaken (or sometimes strengthen) presuppositions—for example, *local accommodation*.

(2) Given a domain of quantification D, the truth value of a formula $\forall x\ \phi(x)$ is:
   a. `T` if $\phi(x) =$ `T` for all $x \in D$;
   b. `F` if there is an $x \in D$ such that $\phi(x) =$ `F`;
   c. `#` otherwise (if there is an $x \in D$ such that $\phi(x) =$ `#` and for no $x \in D$, $\phi(x) =$ `F`).

In (2), we report how truth values are assigned to a universally quantified formula according to the Strong Kleene logic, `#` being *neither true nor false*. Notably, importing Strong Kleene into natural language will not exactly predict presupposition projection, but rather trivalent truth conditions for a sentence.[1] The result is that for sentences with quantifiers, different requirements on the context might be predicted when the sentence is asserted or negated.

A proposal by Charlow (2009) further complicates the picture. According to his intuitions, *soft* and *hard* presupposition triggers (Abusch 2010; Abrusán 2016) lead to different projection patterns: the presuppositions of hard triggers (like *regret*, *too*, and *again*), unlike those of soft triggers (like *discover*, *win*, and *stop*), always project universally out of quantifiers. We will consider this potential difference between soft and hard triggers in our experimental design.

While most of the research on projection out of quantifiers has concentrated on quantifiers over individuals, projection out of modal domains has been at least partially neglected, to our knowledge. In remedy, here we focus on attitude reports. Seminal work by Heim (1992), grounded in dynamic semantics, derives that attitude predicates like *think* behave as plugs.

(3) Peter thinks that Jan stopped smoking.
   Presupposes: Peter thinks that Jan used to smoke. *(according to Heim 1992)*

Apart from approaches assuming that projection out of attitudes has a different profile altogether (as in Geurts 1999), the desideratum in (3) has not been questioned. More recent frameworks, like those by Schlenker (2009) or Mandelkern (2024), are still designed to meet this intuition. Instead, a very natural extension of the Strong Kleene approach to modality would predict that attitude predicates, or at least those with a quantificational semantics, behave just like quantifiers over individuals (translating (2) to a modal domain) for presupposition projection.

Before moving on to the predictions of this hypothesis, it is important to note that our prototypical attitude predicate *think* is a neg-raiser: this means that the falsity conditions of a *think* report might not be the Boolean complement of its truth conditions. For this reason, we will use the predicate *be certain*, which is not a neg-raiser and can give rise to indirect scalar implicatures like a universal quantifier when embedded under negation, as shown in (4).

(4) Peter is not certain that Jan smokes.
   $\not\rightsquigarrow$ Peter is certain that Jan doesn't smoke.
   $\rightsquigarrow$ Peter considers it possible that Jan smokes.

Assuming for the non-quantificational case that a sentence like *Jan canoed again* presupposes that *Jan canoed last time*, we give, in (5), the predictions of Strong Kleene for a *certain*-report embedding the presupposition trigger *again*. Being different from the predictions of a generalized Universal Projection hypothesis, in (6), this yields the starting point for our experiment.

(5) Peter is certain that Jan canoed again. *(Strong Kleene)*
   a. `T` if Peter is certain that Jan canoed last time and this time.
   b. `F` if Peter considers it possible that Jan canoed last time and didn't this time.
   c. `#` if Peter considers it possible that Jan didn't canoe last time, but he is certain that if Jan canoed last time, he canoed this time, too.

[1] Fox (2013) then adopts *Stalnaker's Bridge Principle* to conclude that the presupposition of a quantified formula is a disjunctive proposition; in our case, the negation of (2-c).

(6) Peter is certain that Jan canoed again. *(Universal Projection)*
   a. `T` if Peter is certain that Jan canoed last time and this time.
   b. `F` if Peter is certain that Jan canoed last time and considers it possible he didn't this time.
   c. `#` if Peter considers it possible that Jan didn't canoe last time.

2 Experiment: Quantifiers over worlds

2.1 Design and Materials

The experiment was set up as in (7) and carried out via PCIbex (Zehr and Schwarz 2022), with the `true`, `false`, and `undefined` scenarios serving as controls since for those contexts, both Strong Kleene and the Universal Projection hypothesis make the same predictions.

(7) 2×4 design (within-within; 48 items; 34 participants; $\overline{\text{age}} = 23.8$, SD $= 6.2$)
   a. NEGATION: without vs. with matrix negation
   b. SCENARIO: `true` vs. `false` vs. `undefined` vs. `critical`
   c. Presupposition TRIGGER as a pseudo-factor: *again* vs. *stop* (betw. items, 24 each)

Participants were asked to judge the truth of an attitude report in a context (the SCENARIO manipulation) using the scale in (8); this resulted in a continuous trivalent truth-value judgment task (see Križ and Chemla 2015 for the non-continuous version of this experimental task).

(8) (completely false) (neither completely true nor completely false) (completely true)



The items were structured as follows: a character expresses their doxastic state about a past and a present situation; in an (informationally redundant) image presented alongside the utterance, past and present situations are represented by pictures connected by an arrow. The partitioned cartoon bubble represents what the character considers possible. Before starting, participants were familiarized with the characters and the meaning of the arrows and the clouds. In (9), we list a sample item (translated from German) with *again*; the images are shown in Figure 1. For space reasons, we must leave out a detailed presentation of the *stop* items, though we show the images for one in Figure 1. The inclusion of the soft trigger *stop*, alongside the hard trigger *again*, allows us to see whether speakers differentiate between trigger types (cf. Charlow 2009).

(9) Peter is (not) certain that Jan canoed again.
   `true` "I'm certain Jan canoed last time, and I'm certain Jan canoed this time."
   `false` "I'm certain Jan canoed last time, but I have no idea if this time, Jan canoed or not."
   `undefined` "I have no idea if last time, Jan canoed or not, but I am certain that Jan canoed this time."
   `critical` "I have no idea if last time, Jan canoed or not, and I have no idea if this time, Jan canoed or not."

In the `critical` condition, the doxastic state of the character allows the possibility that the presupposition is not met, and also that the presupposition is met and the embedded proposition is false. If presuppositions project universally, as in Heim (1992), the possibility that the presupposition is not met should make the sentence undefined. If, on the other hand, truth values are as in (5), the possibility that the embedded proposition is false should make the sentence false, and its negation should be true. The predictions are summarized in Figure 2.

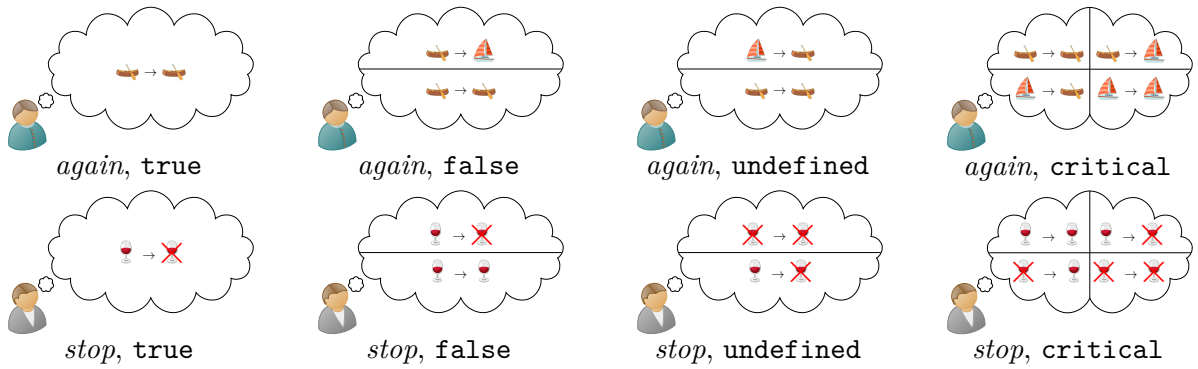


Figure 1: Images redundantly encoding the SCENARIO manipulation for the translated reports "Peter is (not) certain that Jan canoed again" (*top*) and "Markus is (not) certain that Sonja stopped drinking wine" (*bottom*).

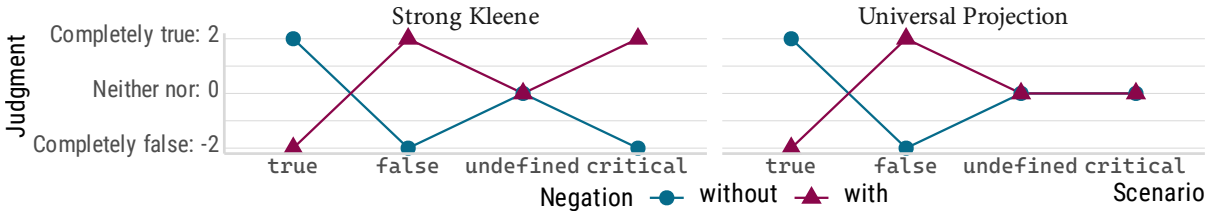


Figure 2: Predictions about presupposition projection from *be certain* for the two hypotheses for each SCENARIO.

This design crucially relies on the NEGATION factor to detect judgments of intermediate truth: if a sentence is neither true nor false, adding matrix negation should not affect the judgment.

We noted in (4) that *be certain* under negation gives rise to an indirect scalar implicature. To ensure that this inference does not affect our results, the character's doxastic state in every item allows for the possibility that the embedded proposition is defined and true. This way, no infelicity can arise because a scalar implicature is violated.

2.2 Results

We excluded 4.35% of the data due to the submission times. First, all observations where submission times exceeded 60 s, and then, all observations whose submission times exceeded the new mean ±3 standard deviations. The descriptive results after exclusion are shown in Figure 3.

We fit a Bayesian mixed distributed linear regression (using `brms` from Bürkner 2021 in R 4.4.1, R Core Team 2024) with 8 chains of 10,000 iterations each (5000 for warm-up). As is evident from the rating distribution in Figure 3, our data involve unequal variances between conditions: especially the `undefined` scenario shows large variances, while the `true` and `false`

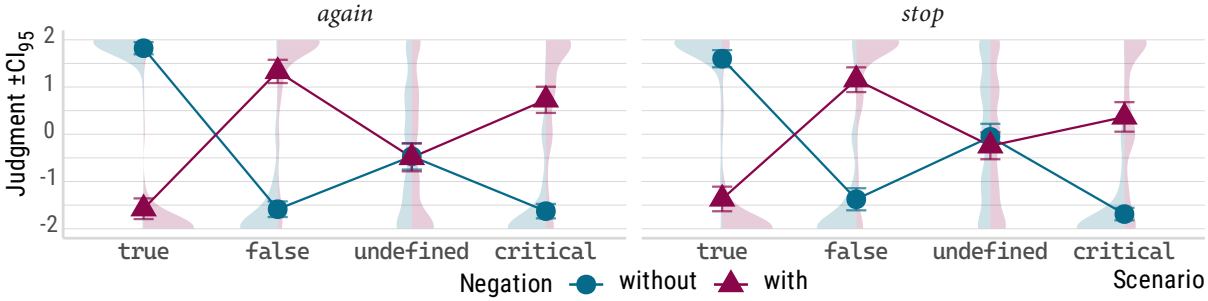


Figure 3: Results. The shaded areas indicate the raw ratings; the confidence intervals involve no by-participant clustering. The lower end of the scale ('completely false') was coded −2, the upper end ('completely true') as 2.

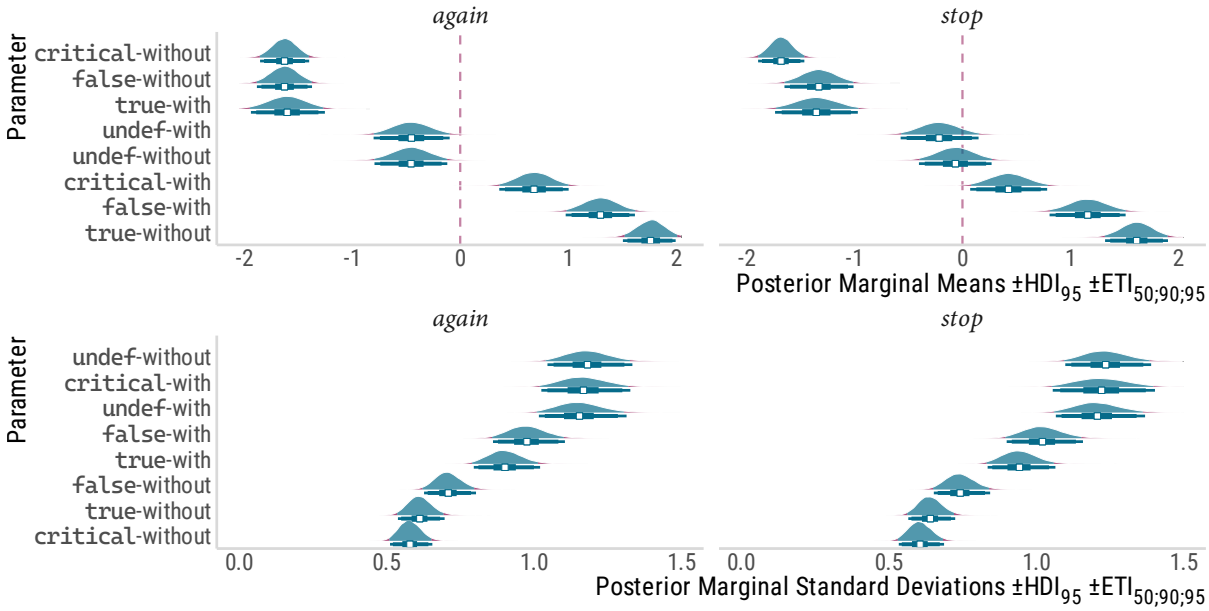


Figure 4: Posterior marginal effects for the location (*top*) and the scale parameters (*bottom*) in the model.

scenarios display much clearer rating clusters. For this reason, we estimated not just the location, as is standard, but also the scale parameter (the standard deviation) for each predictor, resulting in what is called a distributional model. This allows us to capture the unequal variances instead of violating the homogeneity of variance assumption (see, e.g., Kruschke 2015, Ch. 19). In (10), we show the formula for each sub-model, with the maximal random effect structure for the means (Barr et al. 2013). The standard deviations, being strictly positive, are modeled on the log scale, though we will present the back-transformed standard deviations on the response scale.

(10) a. $Y \sim \text{NEGATION} * \text{SCENARIO} * \text{TRIGGER} + (1 + \text{SCENARIO} * \text{TRIGGER} * \text{NEGATION} \mid \text{PARTICIPANT}) + (1 + \text{SCENARIO} * \text{NEGATION} \mid \text{ITEM})$

b. $\log \sigma \sim \text{NEGATION} * \text{SCENARIO} + \text{TRIGGER}$

The priors for model parameters (except those for correlations) were $N(0, x)$, with x ranging between .25 (for the log standard deviation) and 1 (for the location slopes). For the intercept and the slopes in the location sub-model, we also added lower and upper bounds to the priors to reflect the bounded dependent variable at -2 ('completely false') and 2 ('completely true').

The posterior marginal effects, which we use to show posterior predictions rather than relying on coefficients, were computed on the posterior draws using the `emmeans` package (Lenth 2024) and are shown in Figure 4 for both the means and the (back-transformed) standard deviations.

The controls all behave as expected, with the scenarios `true` and `false` showing an effect of negation in the expected direction, and `undefined` being insensitive to the polarity manipulation. As for the `critical` scenario, we see a clear effect of matrix negation and an overall similarity with the `false` scenario, suggesting that participants do not judge the items in these scenarios as involving presupposition failures. These results are supported by a Bayes Factor analysis using Regions of Practical Significance (ROPE) in Figure 5 (for an introduction, see Kruschke 2018). If the prior and posterior odds are the same (and there is no evidence for either difference or equivalence), then the (log) Bayes Factor is 0. If the posterior odds are higher than the prior odds (and the data thus support a difference hypothesis), the Bayes Factor will be larger than 0.

Interestingly, we also find stark contrasts with the standard deviations, Figure 4. Here, `true`

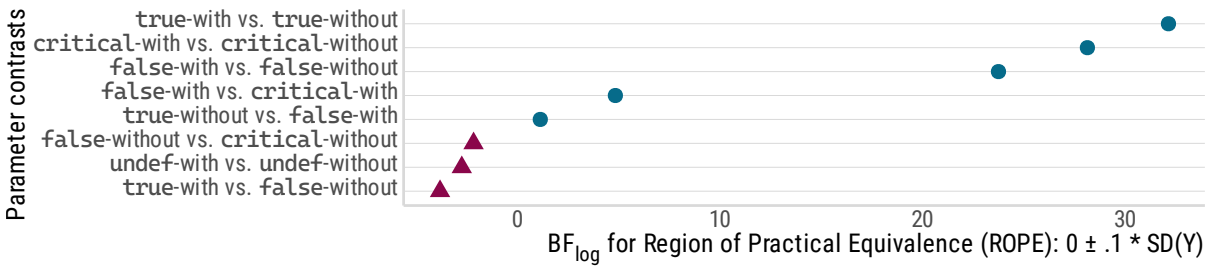


Figure 5: Bayes Factors. Comparisons where equivalence hypotheses are supported shown in red, others in blue.

and `false` scenarios show relatively small dispersion (with an increase due to negation), while the `undefined` scenario shows large estimates irrespective of negation. We take this to index presupposition failure-induced 'squeamishness' (Strawson 1964). By contrast, the `critical` scenario patterns with `true` and `false` without negation, though negation impacts dispersion.

3 Discussion and conclusion

In accordance with a presuppositional view of *stop* and *again*, the `undefined` scenario does not show an effect of negation. In the `critical` scenario, negation has an effect, suggesting no presupposition failure. This is in line with the generalized Strong Kleene hypothesis, but clashes with the Universal Projection hypothesis. Additionally, there is no evidence that the two types of triggers show different projection patterns (contra Charlow 2009). More generally, we take our results to support a view where presupposition failures can take part in the meaning composition, and presupposition projection is determined by the contribution of each logical operator—including attitude predicates—provided a suitable trivalent logic like Strong Kleene.

To our knowledge, we provided the first application of distributional models to detect intuitions of presupposition failure via standard deviations. In our data, presupposition failures increased standard deviations more than negation, thus modeling the well-known squeamish reaction (Strawson 1964; von Fintel 2004). This result opens up new analytical avenues for experiments on truth-value gaps, even when the dependent variable does not target truth values.

There is an unexpected difference between the `critical` and `false` scenarios with negation, both in terms of means and standard deviations. In speculation about this pattern, we offer two possible explanations. It is a fact that the `critical` scenario with negation is the most complex condition of the design: it involves the highest number of partitions in the character's doxastic state and requires extra computational cost because of negation. It is possible that judgments become less sharp and variance increases as a result of participants' uncertainty or cognitive load. Alternatively, it might be the case that the participants, who are fully informed about the character's doxastic state, judge the negated report in the `critical` condition as underinformative, even if true. Consider the sentence in (11) together with its scalar implicature:

(11) Peter is not certain that Jan canoed again.
   a. *Entails:* Peter considers it possible that Jan didn't canoe again.
   b. *Implicates:* Peter considers it possible that Jan canoed again.

The entailment and the implicature of the sentence address only two out of four of Peter's doxastic possibilities in the `critical` scenario. Instead, these two doxastic possibilities are all that features in the `false` scenario, where the sentence is judged as true with little uncertainty.[2]

[2]Clemens Mayr (p.c.) suggests that embedded polar interrogatives, which are felicitous with negated *be certain* (Mayr 2019), might compete with our items. A trivalent theory of embedded questions is needed to explore this.

Acknowledgements. This project was carried out within the RTG 2636 "Form-Meaning Mismatches," funded by the *Deutsche Forschungsgemeinschaft*, project number: 429844083.

References

Abrusán, Márta (2016). "Presupposition cancellation: Explaining the 'soft–hard' trigger distinction". In: *Natural Language Semantics* 24.2, pp. 165–202. DOI: 10.1007/s11050-016-9122-7.

Abusch, Dorit (2010). "Presupposition triggering from alternatives". In: *Journal of Semantics* 27.1, pp. 37–80. DOI: 10.1093/jos/ffp009.

Barr, Dale J. et al. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal". In: *Journal of Memory and Language* 68.3, pp. 255–278. DOI: 10.1016/j.jml.2012.11.001.

Bürkner, Paul-Christian (2021). "Bayesian Item Response Modeling in R with brms and Stan". In: *Journal of Statistical Software* 100.5, pp. 1–54. DOI: 10.18637/jss.v100.i05.

Charlow, Simon (2009). "'Strong' predicative presuppositional objects". In: *Proceedings of ESSLLI 2009*.

Chemla, Emmanuel (2009). "Presuppositions of quantified sentences: Experimental data". In: *Natural Language Semantics* 17.4, pp. 299–340. DOI: 10.1007/s11050-009-9043-9.

Chierchia, Gennaro (1995). *Dynamics of meaning: Anaphora, presupposition, and the theory of grammar*. Chicago: University of Chicago Press.

von Fintel, Kai (2004). "Would you believe it? The King of France is back! (Presuppositions and truth-value intuitions)". In: *Descriptions and Beyond*. Ed. by Marga Reimer and Anne Bezuidenhout. Oxford: Oxford University Press, pp. 315–341.

Fox, Danny (2013). "Presupposition projection from quantificational sentences: Trivalence, local accommodation, and presupposition strengthening". In: *From Grammar to Meaning*. Ed. by Ivano Caponigro and Carlo Cecchetto. Cambridge University Press, pp. 201–232. DOI: 10.1017/cbo9781139519328.011.

George, B. (2010). "A new case for an old logic: Reviving Strong Kleene approaches to presupposition projection". Unpublished ms. UCLA.

Geurts, Bart (1999). *Presuppositions and Pronouns*. Amsterdam: Elsevier.

Heim, Irene (1983). "On the projection problem for presuppositions". In: *WCCFL 2: Second Annual West Coast Conference on Formal Linguistics*. Ed. by Michael Barlow, Daniel P. Flickinger, and Michael T. Wescoat, pp. 114–125.

— (1992). "Presupposition projection and the semantics of attitude verbs". In: *Journal of Semantics* 9, pp. 183–221. DOI: 10.1093/jos/9.3.183.

Karttunen, Lauri (1973). "Presuppositions of compound sentences". In: *Linguistic Inquiry* 4.2, pp. 169–193.

Karttunen, Lauri and Stanley Peters (1979). "Conventional Implicature". In: *Presupposition*. Ed. by David A. Dineen and Choon-Kyu Oh. New York: Academic Press, pp. 1–56.

Križ, Manuel and Emmanuel Chemla (2015). "Two methods to find truth-value gaps and their application to the projection problem of homogeneity". In: *Natural Language Semantics* 23.3, pp. 205–248. DOI: 10.1007/s11050-015-9114-z.

Kruschke, John K. (2015). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Oxford: Academic Press.

— (2018). "Rejecting or Accepting Parameter Values in Bayesian Estimation". In: *Advances in Methods and Practices in Psychological Science* 1.2, pp. 270–280. DOI: 10.1177/2515245918771304.

Lenth, Russell V. (2024). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R package version 1.10.0. URL: https://CRAN.R-project.org/package=emmeans.

Mandelkern, Matthew (2024). *Bounded meaning: The dynamics of interpretation*. Oxford: Oxford University Press. DOI: 10.1093/oso/9780192870049.001.0001.

Mayr, Clemens (2019). "Triviality and interrogative embedding: context sensitivity, factivity, and neg-raising". In: *Natural Language Semantics* 27.3, pp. 227–278. DOI: 10.1007/s11050-019-09153-8.

R Core Team (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Schlenker, Philippe (2009). "Local Contexts". In: *Semantics & Pragmatics* 2. DOI: 10.3765/sp.2.3.

Strawson, P. F. (1964). "Identifying reference and truth-values". In: *Theoria* 30.2, pp. 96–118. DOI: 10.1111/j.1755-2567.1964.tb00404.x. URL: https://doi.org/10.1111%2Fj.1755-2567.1964.tb00404.x.

Sudo, Yasutada et al. (2012). "Presupposition projection out of quantified sentences: Strengthening, local accommodation and inter-speaker variation". In: *Logic, Language and Meaning. 18th Amsterdam Colloquium*. Ed. by Maria Aloni et al. Berlin: Springer, pp. 210–219. DOI: 10.1007/978-3-642-31482-7_22.

Zehr, Jérémy and Florian Schwarz (2022). *PennController for Internet Based Experiments (IBEX)*. DOI: 10.17605/OSF.IO/MD832.

Distributivity facilitates ACD resolution

Jad Wehbe
Massachusetts Institute of Technology
Cambridge, MA, US
jadwehbe@mit.edu

Kate Kinnaird
Massachusetts Institute of Technology
Cambridge, MA, US
kinnaird@mit.edu

Martin Hackl
Massachusetts Institute of Technology
Cambridge, MA, US
hackl@mit.edu

Abstract

Theories of plural predication disagree on the division of labor between the lexical semantics of predicates and various syntactic operators, e.g. the distributivity operator (Krifka 1992; Kratzer 2007). A central issue is whether the LFs giving rise to certain distributive and cumulative readings involve syntactic movement (Beck and Sauerland 2000; Winter 2000). In this project, we adopt the logic of Hackl, Koster-Hale, and Varvoutis (2012) to diagnose movement indirectly via a facilitation effect on ACD resolution. We show that, similar to *every*, definite plurals under a particular distributive construal decrease the cost of ACD resolution compared to a definite singular baseline. This provides evidence that the distributive reading requires LF-movement of the plural DP hosting ACD above the matrix antecedent of the ellided constituent. We argue that this follows naturally only from theories of distributivity/cumulativity on which there is a syntactically represented distributivity operator that the plural DP needs to outscope in order to get the desired reading.

1 Introduction

Plural predication gives rise to cumulative inferences, as illustrated in (1): if (1-a) is true and (1-b) is true, we can infer (1-c). One perspective attributes these inferences to the lexical semantics of predicates. For a 1-place predicate, this amounts to assuming that whenever the predicate is true of parts of a plurality it has to also be true of the resulting plurality. This property can be generalized to n-place predicates as in (2) and has been argued to be a universal property of all lexical predicates (Krifka 1992; Kratzer 2007).[1]

(1) a. Jack invited Student 1 (yesterday).
b. Jack invited Student 2 (today).
c. Jack invited the two students.

(2) Lexical Cumulativity Universal:
$\forall a_1, a_1', ..a_n, a_n' : \forall P : P(a_1)(..)(a_n) = 1 \wedge P(a_1')(..)(a_n') = 1 \Rightarrow P(a_1 + a_1')(..)(a_n + a_n') = 1$

To spell out such an account in more concrete terms, we can assume that a sentence like (1-a) has the LF in (3-a). We assume that perfective aspect denotes an existential quantifier over subparts of the time interval that tense picks out and that predicates have a time argument in addition to their individual arguments. The resulting truth-conditions for (1-a) are given in (3-e). The LF for (1-c), even when understood distributively, would be parallel to (3-a) in that the predicate can combine directly with the plural DP, (4-a), and still give rise to cumulative

[1] We illustrate cumulative inference over the object DP in line with our experimental items. However, parallel consideration arise for DPs in other positions including subject DPs.

inferences. I.e. as long as the temporal argument of *invite* falls under (2), (1-c) is true if Student 1 was invited by Jack at some time and Student 2 at some, possibly different time.

(3) **Baseline assumptions:**

a. LF for (1-a): [[Past t][PFV [Jack [invited Student 1]]]]
b. $[\![\text{PFV}]\!] = \lambda P_{\langle i,t \rangle}.\lambda t.\exists t' \subseteq t : P(t')$
c. $[\![\text{Past}]\!]^c = \lambda t : t < t_c.t$
d. $[\![\text{invited}]\!] = \lambda y.\lambda x.\lambda t.x$ invited y at t.
e. $[\![$ (3-a) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1)(j)(t')$, where t is a past time interval.

(4) **Lexical cumulativity account:**

a. LF for (1-c): [[Past t] [PFV [Jack [invited the two students]]]]
b. $[\![$ (1-c) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1 + s_2)(j)(t')$, where t is a past time interval.
c. $[\![\text{invited}]\!](s_1)(j)(t_1) \wedge [\![\text{invited}]\!](s_2)(j)(t_2) \Rightarrow [\![\text{invited}]\!](s_1 + s_2)(j)(t_1 + t_2)$

A prominent alternative to the lexical perspective is to assume that these inferences are not due to the predicates themselves, but can be attributed to independently needed pluralization operators that modify the predicate (e.g. Beck and Sauerland 2000). On this view, the expression that furnishes cumulative inferences is syntactically complex and we need to ask what the syntactic distribution of the relevant pluralization operators are. One version of the syntactic analysis assumes that pluralization operators apply to every predicate of the right type, be they lexical or syntactically derived. Cumulative inferences for (1-c) may then simply arise from applying a pluralization operator directly to the predicate. A more restrictive perspective, argued for in Wehbe (2023), building on Ionin and Matushansky (2002), is to assume that pluralization operators can only apply to predicates that are derived by movement. Wehbe (2024), moreover, argues that in order to ensure that a sentence like (1-c) is true when Student 1 and Student 2 were invited by Jack at different times, the plural DP and the accompanying pluralization operator have to take scope above Aspect, thus allowing us to distribute over the existential quantifier over times. The resulting movement-based analysis is given in (5).

(5) **Movement analysis:**

a. LF: [[Past t] [the two students [* [1 [PFV [Jack [invited t_1]]]]]]]
b. $[\![$ (5-a) $]\!]$=1 iff $\exists t' \subseteq t : [\![\text{invited}]\!](s_1)(j)(t') = 1 \wedge \exists t'' \subseteq t : [\![\text{invite}]\!](s_2)(j)(t'') = 1$
c. $[\![*]\!] = \lambda f_{\langle e,t \rangle}.\lambda x.\exists g_{\langle e,t \rangle}[[\forall y[g(y) = 1 \rightarrow f(y) = 1] \wedge x = \bigoplus g]$

We see that both approaches succeed in deriving cumulative inferences for cases like (1-c). However, they differ with respect to what role the syntax plays. On the lexical cumulativity view, the syntactic structure of the sentence that gives rise to the cumulative inference is not different from the one that is needed for singular object DPs or, for that matter, the one that gives rise to a collective reading on which there are no inferences to the parts of the object plurality. In all three cases, the object DP can be interpreted in situ. For the syntactic approach sketched above, however, deriving the distributive reading involves a logical form in which the object DP is not interpreted in its base position. Rather, on this account, it has to take scope above Aspect to support inferences to its parts.

Since both approaches are equally successful at the basic task of deriving distributive truth-conditions for sentences like (1-c) we need to look elsewhere for evidence that might help identify the correct logical form of such sentences under this interpretation. One fruitful avenue is to look for correlations between the presence/absence of DP movement and the presence/absence of cumulative inferences. Wehbe (2023), for instance, observes that in the Lebanese Arabic double-subject construction cumulative inferences are present for plural DPs that have moved but absent for those that are interpreted in situ. Another approach, pursued in the present

study, is to ask whether sentences that are interpreted distributively and thus exhibit cumulative inferences have processing signatures that are indicative of DP movement.

2 Probing for the correct LF via ACD resolution

The present study follows the approach of Hackl, Koster-Hale, and Varvoutis (2012) to look for a processing signature of covert DP-movement. The basic idea in Hackl et al. is to use the relative ease with which comprehenders are able to resolve antecedent contained deletion (ACD) as a proxy for how and when they can determine the correct LF-position of the DP hosting the ACD site. To appreciate their logic, consider the factorial template in (6) in which the presence or absence of ACD is crossed with two types of determiners on the hosting DP - the quantificational determiner *every* or the non-quantificational determiner *the*.

(6) Jane read {every$_{\textbf{+quant.}}$, the$_{\textbf{-quant.}}$} book that Mary { did$_{\textbf{+ACD}}$, wrote$_{\textbf{-ACD}}$}.

For a sentence to exhibit ACD it has to have a structural description (LF) where the elided VP is not included inside the matrix VP, which serves as its antecedent (Sag 1976). This can be achieved, on canonical assumptions, if the object DP hosting the ACD site is moved to a position higher than the matrix VP, thereby "undoing" antecedent containment (Sag 1976, Williams 1977; see Jacobson 1992 for an alternative). Importantly, while both quantificational and non-quantificational DPs need to be moved when they host an ACD site, Hackl et al. assume that only the former require it on their own, irrespective of whether or not they host an ACD site. From a procedural perspective, this means that the [+quant, +ACD] variant of (6) contains two triggers for the parser to identify the correct LF while the [-quant,+ACD] variant contains only one, namely the ACD site itself.

Adopting a linking hypothesis that connects the number of cues for a particular LF with the ease with which that LF can be identified leads Hackl et al. to expect a facilitation effect for *every* on ACD resolution. More specifically, since resolving ACD is a complex process, which involves postulating an LF without antecedent containment (i.e. by assumption covert movement of the object DP), identifying the antecedent VP and retrieving it from memory when computing the meaning of the relative clause, the [+ACD] variants of (6) are expected to be more costly than their [-ACD] counterparts. It is also expected, however, that when the host DP is projected from *every* this cost is reduced and ACD resolution is facilitated since a critical component of resolving ACD - identifying an LF without containment - is triggered independently from and earlier than the ACD site. By contrast, projecting the host DP from *the* does not provide a separate trigger for moving the object DP above the matrix VP. Thus, in this condition the cue for such an LF is provide by the ACD site alone, at a stage when a parse without movement is likely entrenched and so relatively costly to abandon. Combining these two types of considerations leads to the prediction of a particular type of interaction such that the [+quant,+ACD] condition is relatively less costly than its [-quant,+ACD] counterpart. This facilitation effect has been documented, among other things and most relevantly for our purpose here, in naturalness ratings in Hackl et al. and it has been replicated in a series of subsequent work including Gibson et al. (2015), who offer an alternative account of the facilitation effect.[2]

The present project extends this logic to distributive readings of plural definites. Concretely, if distributive readings are due to a syntactically represented pluralization operator that applies to a derived predicate, then a definite plural object DP, just like *every* in (6), triggers movement

[2]The facilitation effect has also been documented in self-paced reading data in Hackl et al. and both types of observations have been replicated and extended to a variety of other expressions that are hypothesized to trigger upstream movement in their own right (see Hackl, Koster-Moeller, and Gottstein 2009, Breakstone et al. 2012, Kotek 2019).

in order to get the desired reading. Thus, we expect a similar facilitation effect to Hackl et al.'s observation for *every* for distributively interpreted definite plurals.

3 Distributivity and ACD resolution: an experimental study

3.1 Task and Design

We adopt the logic of Hackl et al. to probe whether definite plural DPs in object position undergo movement when they are interpreted distributively. We focus on a particular instance of a distributive reading where the different atomic parts of the relevant plural DP participated in the event in question at different times. We will refer to this as a *one-at-a-time* reading. Our target sentences, which parallel the format in (6), are paired with contexts that strongly encourage a *one-at-a-time* interpretation. An example is given in (7), where we crossed +/-ACD and the type of DP (definite plural vs. singular).

(7) **Context:** 25 international students are invited for a campus visit at an Ivy league university, and they get the chance to speak with professors one-on-one.
Sentence: The English literature professor met with the humanities {students$_{\mathbf{+PL}}$, student$_{\mathbf{-PL}}$} that the philosophy professor {did$_{\mathbf{+ACD}}$, interviewed $_{\mathbf{-ACD}}$} due to the dean's insistence.

Items were presented in PCIbex (Zehr and Schwarz 2023) using a counterbalanced Latin Square design, with each participant being assigned to one of four lists. The experiment consisted of 32 fillers and 32 target items. Item order was randomized for every participant. Filler items were of similar complexity and length to target items. As exemplified in (7), Participants first saw a context sentence followed by a target sentence, and then were prompted to rate the target sentence's naturalness relative to the context on a 7 point Likert scale. Following their selection, the context sentence and target sentence were removed from the screen. Participants were then given a Yes/No comprehension question asking about information conveyed by the target sentence. To probe for whether participants are paying attention to the crucial distinction between plural and singular definite DPs, some of our comprehension questions are sensitive to this distinction (e.g. for (7): *Did the English literature professor meet with only one student?*). To establish a baseline for our design (seeking to replicate Hackl et al.), we also ran a variant of the study where the definite plural in (7) is replaced with *every* (e.g. *every humanities student*), keeping everything else constant.

3.2 Results

For each of our two experiments, we recruited 80 monolingual English speakers using crowdsourcing platform Prolific. Participants were excluded if they failed any of our three practice items or if they answered less than 8 out of our 12 attention check items correctly (36 for the baseline and 19 for the target experiment). Participants were paid approximately 5 USD.

Mean ratings for the baseline and target experiments, with 95 percent confidence interval-based error bars, are displayed in Figure 1. We report the maximally-specified linear mixed-effects model (conditioned on convergence), predicting naturalness ratings as a function of [+/-ACD], DP type (every vs. definite singular for the baseline experiment and definite plural vs. definite singular for the target experiment), and their interaction. We include random slopes of DP type and [+/-ACD] for participants and a random slope of [+/-ACD] for items. The model reveals a significant interaction between [+/-ACD] and DP type for both the baseline experiment and the target experiment ($p<0.0001$ and $p<0.05$, respectively). Analyses were performed in R Statistical Software (v 4.2.2; R Core Team 2022) using the lme4 package (Bates et al. 2015).

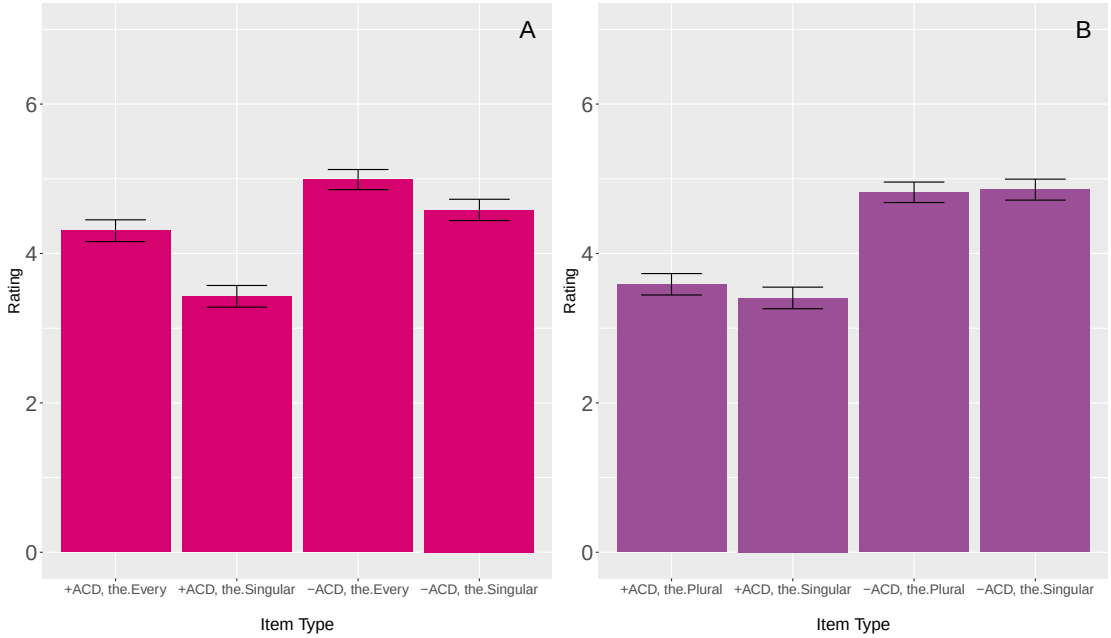


Figure 1: Mean naturalness ratings for Baseline Experiment in A (N=80) and Target Experiment in B (N=80)

3.3 Discussion

Our baseline experiment with *every* replicates the results in Hackl et al. In particular, we see that the presence of ACD results in a marked degradation of naturalness, but critically, we observe an interaction indicating that ACD resolution is facilitated with *every*, when compared to the definite singular. Following Hackl et al., we take this to support the conclusion that *every* in object position has to move for reasons independent of ACD.

Turning to the target experiment, we see a parallel facilitation effect of the definite plural on ACD resolution. Adopting, again, the reasoning in Hackl et al., we take this to show that the distributive reading over a plural object DP in our environment requires movement of that DP. Interestingly, as is evident in Figure 1, the size of the interaction effect with the definite plural is smaller than that with *every* (Model estimate= 0.2188 for the target experiment and 0.47188 for the baseline experiment). This difference in effect size could be due to several factors. One possibility is that that the distinction between *the students* and *the student* is easier to miss for our participants than the distinction between *every student* and *the student*. This would result in more noise in the target experiment, thereby lowering the size of the effect. To explore this possibility, we excluded participants who failed to correctly answer more than one-third of comprehension questions which were sensitive to the singular/plural distinction. This resulted in the effect size going up to 0.26906 and the p-value going down to 0.0194 from 0.0457.[3] This is as expected on the view that part of the difference in effect size is due to participants missing the singular/plural distinction in the target experiment.

Another possible contributing factor to the difference in effect size is that the information needed to recognize the trigger for movement in the two experiments is different. In the case of *every*, it is simply recognizing that *every* occurs in object position. On the other hand, the relevant information in the case of the definite plural crucially includes the distributive context (assuming that definite plurals have to move only under the distributive construal). Facilitation is therefore only predicted for those participants who keep the context in mind and attempt to

[3]Excluding participants at 25% and 50% thresholds resulted in similar changes in effect size and p value, although simpler models were required for convergence.

generate a matching distributive reading for the target sentence.

4 Conclusion

Just like Hackl et al., we observe a link between the nature of the DP hosting an ACD site (definite singular vs definite plural) and the ease of ACD resolution. Specifically, definite plurals under our *one-at-a-time* contexts exhibit a facilitation effect. We take this to show that the definite plural is perceived by comprehenders as a trigger for movement in our environment, allowing the parse where antecedent containment has been undone to be identified before the ACD site is reached. This follows from theories where pluralization operators are always syntactically represented and where movement of the object DP to a position above the relevant operator is required to generate the desired distributive reading (Beck and Sauerland 2000; Wehbe 2023; Wehbe 2024). On the other hand, it is difficult to explain this facilitation effect if our distributive reading is due to lexical cumulativity. It is nevertheless possible that the lexical cumulativity universal holds and that it is the particular type of distributive reading over times that we target here that requires movement. Such a mixed view, however, raises doubts about the explanatory utility of lexical cumulativity. In particular, giving up an account of our cases in terms of lexical cumulativity can only be achieved at the cost of a stipulation, e.g. that temporal specifications of predicates are not covered by lexical cumulativity.

Our results raise interesting questions in the context of previous processing studies of distributivity. For example, Dotlačil and Brasoveanu (2021) argue for a distinction between lexical and phrasal distributivity based on differences in how easy the corresponding readings are to access. In particular, they show that the distributive reading over an indefinite exemplified in (8-a) is dis-preferred relative to the collective reading, while there is no parallel difficulty for the distributive reading in (8-b). Dotlačil and Brasoveanu (2021) take this as evidence that these two kinds of distributive readings are due to distinct sources: the reading in (8-a) requires a distributivity operator to apply to a derived predicate that includes the indefinite, while the distributive reading in (8-b) follows from lexical cumulativity.

(8) a. The students read a book./ Dist. reading: The students each read a book.
b. The students lifted the piano./Dist. reading: The students each lifted the piano.

Given that our diagnostic indicates that the *one-at-a-time* reading is due to phrasal distributivity, if the characterization of Dotlačil and Brasoveanu (2021) above is correct, we would expect our sentences to pattern with (8-a) rather than (8-b). It is not clearthat the *one-at-a-time* distributive reading is in fact dis-preferred in the same way as (8-a). If it turns out that our sentences pattern with (8-b), this would require a different explanation for the processing difference in (8) which nevertheless maintains that both LFs involve a distributivity operator.

To conclude, our study lays the foundation for a new tool to compare frameworks where plural predication is mediated by syntactic operators (e.g. Beck and Sauerland 2000), in which movement plays a crucial role, to frameworks where we don't expect to see systematic syntactic correlates of plural predication (e.g. Winter 2000; Schmitt 2019). Looking ahead, we hope to use this new tool to investigate the role of movement in generating different readings in plural predication. For example, non-local cumulative readings have been argued to require movement of the relevant DPs to a derived scope position (Beck and Sauerland 2000), but certain accounts, including the more recent plural projection framework (Schmitt 2019) are able to generate these non-local readings without movement. Our tool can help to distinguish the predictions of these two theories. Another direction for future research is to investigate whether distributivity is in fact crucial for the facilitation effect with definite plurals. We therefore plan to test whether plural DPs under collective construals give rise to the same facilitation effect we observed here.

References

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using lme4". In: *Journal of Statistical Software* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Beck, Sigrid and Uli Sauerland (2000). "Cumulation is needed: A reply to Winter (2000)". In: *Natural language semantics* 8.4, pp. 349–371.

Breakstone, Micha Yochanan et al. (2012). "On the Analysis of Scope Ambiguities in Comparative Constructions: Converging Evidence from Real-Time Sentence Processing and Offline Data". In: *Proceedings of SALT*. Vol. 21, pp. 712–731.

Dotlačil, Jakub and Adrian Brasoveanu (2021). "The representation and processing of distributivity and collectivity: ambiguity vs. underspecification". In: *Glossa: A journal of general linguistics* 6.1.

Gibson, Edward et al. (2015). "A pragmatic account of complexity in definite antecedent-contained-deletion relative clauses". In: *Journal of Semantics* 32.4, pp. 579–618.

Hackl, Martin, Jorie Koster-Hale, and Jason Varvoutis (2012). "Quantification and ACD: Evidence from real-time sentence processing". In: *Journal of Semantics* 29.2, pp. 145–206.

Hackl, Martin, Jorie Koster-Moeller, and Andrea Gottstein (2009). "Processing Opacity". In: *Proceedings of Sinn und Bedeutung 13*. Vol. 13. 1, pp. 171–186. DOI: https://doi.org/10.18148/sub/2009.v13i1.524.

Ionin, Tania and Ora Matushansky (2002). "DPs with a twist: A unified analysis of Russian comitatives". In: *Proceedings of FASL*. Vol. 11.

Jacobson, Pauline (1992). "Antecedent Contained Deletion in a Variable-Free Semantics". In: *Proceedings of SALT 2*, pp. 193–213. DOI: https://doi.org/10.3765/salt.v2i0.3027.

Kotek, Hadas (Feb. 2019). *Composing Questions*. The MIT Press. ISBN: 9780262351065. DOI: 10.7551/mitpress/10774.001.0001. URL: https://doi.org/10.7551/mitpress/10774.001.0001.

Kratzer, Angelika (2007). "On the plurality of verbs". In: *Event structures in linguistic form and interpretation* 269.300, p. 2.

Krifka, Manfred (1992). "Thematic Relations as Links between Nominal Reference and Temporal Constitution/Manfred Krifka". In: *Lexical matters* 24, p. 29.

R Core Team (2022). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Sag, Ivan A (1976). "Deletion and logical form." PhD thesis. Massachusetts Institute of Technology.

Schmitt, Viola (2019). "Pluralities across categories and plural projection". In: *Semantics and Pragmatics* 12, pp. 17–1.

Wehbe, Jad (2023). "Against the lexical view of cumulative inferences". In: *Proceedings of Sinn und Bedeutung*. Vol. 27, pp. 693–711.

— (2024). "Covert Reciprocals: A scope-based account of reciprocal alternations". In.

Williams, Edwin S (1977). "Discourse and logical form". In: *Linguistic inquiry*, pp. 101–139.

Winter, Yoad (2000). "Distributivity and dependency". In: *Natural language semantics* 8.1, pp. 27–69.

Zehr, Jérémy and Florian Schwarz (Jan. 2023). *PennController for Internet Based Experiments (IBEX)*. DOI: 10.17605/OSF.IO/MD832. URL: osf.io/md832.

Believe is not a propositional attitude verb

Simon Wimmer
Department of Philosophy, Heinrich-Heine-University Duesseldorf
Duesseldorf, Germany
simon.wimmer@hhu.de

Abstract

I develop a challenge for the view that *believe* is a propositional attitude verb based on two observations: (i) *believe* can embed *in O*, and (ii) *in O* does not denote a proposition. To develop my challenge, I argue (section 2) that *believe* is not homonymous or polysemous between a propositional belief-that and non-propositional belief-in interpretation, and (section 3) that type-shifting *in O*'s denotation to a proposition falsely predicts that belief-in and belief-that reports are equivalent.

1 Introduction

What does *believe* contribute to the logical form of a belief-that report like *Ayesha believes that Rebecca swims*? A relation to a proposition, say propositional attituders (e.g. Schiffer 2003; King 2007; Forbes 2018).[1,2] For them, *believe* is a propositional attitude verb.

I argue that *believe* doesn't contribute a relation to a proposition to the logical form of belief-that reports. My argument is based on two observations. First, as belief-in reports like *Boris believes in his friends* and *Carol believes in sets* illustrate, *believe* can also embed the prepositional phrase *in O* instead of a *that*-clause.[3,4] But, second, *in O* intuitively doesn't denote a proposition, but a property of an eventuality, just like other prepositional phrases headed by *in* (e.g. *in the middle of the night* and *in the pool*).[5] Given these two observations, my argument goes as follows. Suppose that, in belief-in reports, *believe* contributes a relation to a proposition, just as it's meant to do in belief-that reports. Then *in O* leaves *believe*'s proposition slot unsaturated. But suppose that nothing else saturates that slot. Then we get a semantically ill-formed output. Yet belief-in reports are well-formed. Something has gone wrong.

The remainder of this paper defends the two claims we just supposed for the sake of the argument. Section 2 argues that, in belief-in reports, *believe* makes the same contribution as in belief-that reports: I suggest that relevant diagnostics support neither the homonymy nor the polysemy of *believe* across the two environments. Section 3 then argues against one way to get something other than *in O*'s denotation to saturate *believe*'s proposition slot: I suggest that type-shifting the property denoted by *in O* to a proposition falsely predicts that belief-in reports are equivalent to belief-that reports.[6]

[1] By 'proposition' I mean whatever is the object or content of belief-that, e.g. a set of worlds (e.g. Stalnaker 1984), a structured proposition (e.g. King 2007), or a property of individuals (e.g. Pearson 2013).

[2] Forbes' belief predicate in the metalanguage isn't a propositional attitude verb, as it's a monadic predicate of a state. However, *believe* in the object language is still treated as a propositional attitude verb.

[3] For discussion of belief-in see Price 1969; Szabó 2003; Textor 2013; Kriegel 2018; Wimmer forthcoming.

[4] Other reports where *believe* doesn't embed *that*-clauses include *Ayesha believes Rebecca* and *Sue believes the claim that sets exists*. For reasons of space, I focus on belief-in reports, but see e.g. Moltmann 2022; Djärv 2023; 2024 for discussion of the former, and Moulton 2009; Uegaki 2016 for discussion of the latter kind of report.

[5] I assume that *in O* is, and *believe in* (without *O*) isn't, a constituent. Results from constituency tests in Osborne 2015, pp. 252–5 and Müller 2023, pp. 6–11 either support this assumption (e.g. coordination, cf. examples 1 and 2) or can be explained away (e.g. question formation). For reasons of space, I don't apply these tests here.

[6] As an alternative to the type-shifter, propositional attituders might say that *S believes in O*, together with further context, merely makes salient a proposition about O that saturates *believe*'s proposition slot. In this

2 *Believe* Makes The Same Contribution

I now argue that *believe* makes the same contribution to the logical form of belief-in and belief-that reports. I focus on two ways to say that *believe* makes different contributions across these environments: the first is to say that *believe* is homonymous; the second is to say that it's polysemous.[7,8] Either way, *believe* would be ambiguous: there would be one reading or sense—the belief-that reading or sense—that denotes a relation to a proposition, and another—the belief-in reading or sense—that denotes something else. My argument works by highlighting the lack of evidence for saying that *believe* makes different contributions across belief-in and belief-that reports: I argue that relevant diagnostics support neither the homonymy nor the polysemy of *believe* across these environments. Given Grice 1978's modified Occam's razor, we thus have reason to deny that *believe* makes different contributions.

More specifically, I consider results from three diagnostics: the zeugma, contradiction, and pedantic correction tests.[9] Take the zeugma test first. Here, we form a pair of sentences containing the target word such that the two sentences get plausible interpretations only given different interpretations of the target. We then conjoin these sentences so as to elide one of the constituents containing the target word. If the resulting sentence is odd, that's evidence that the target is ambiguous, i.e. homonymous or polysemous. If not, that's evidence that the target is unambiguous or polysemous. For instance, *This bottle is made from beautiful glass and very tasty* sounds odd. That's evidence for the homonymy or polysemy of *bottle*.[10]

Failing the zeugma test doesn't distinguish unambiguous from polysemous expressions because polysemous expressions can fail the test. Consider a case of 'co-predication', where *bottle* occurs once, but with two senses: *David drank and smashed the bottle* (Felappi 2019, p.5). Here, *drank* triggers *bottle*'s content-sense; *smashed* its container-sense.

Crucially, *believe* fails the zeugma test. If we control for confounds, we get sentences that don't sound odd. Thus *believe* is either unambiguous or polysemous. Consider:

1. (a) Sue believes in the abundance of human error and (Salma) that everyone is wrong about sets.
   (b) Sue believes that everyone is wrong about sets and (Salma) in the abundance of human error.

2. (a) John believes in all the teachings of his church and (Jamal) that he is a prophet.
   (b) John believes that he is a prophet and (Jamal) in all the teachings of his church.[11]

way, they might argue, *S believes in O* functions analogously to *S believes*, which is acceptable in contexts that make salient a proposition that saturates *believe*'s proposition slot. One issue for this reply resembles that for the type-shifter section 3 considers. If a belief-in report's truth-conditions merely concern believing the salient proposition about O, propositional attituders falsely predict that belief-in reports are equivalent to belief-that reports. So they must say that *in O*'s denotation contributes something more. But if they say this, a theory that just relies on *in O*'s denotation, and not on a proposition made salient by *in O*, is potentially just as predictive, but simpler. Relatedly, relying on context in the way the reply does is risky: what if context makes salient a proposition about something other than O? Which proposition is now predicted to saturate *believe*'s proposition slot, the one about O, raised to salience by *in O*, or the one not about O?

[7]What polysemy is is much debated (cf. Vicente and Falkum 2017), with some even denying that polysemy differs from monosemy (e.g. Fodor and LePore 2002; Brody and Feiman 2023). I grant propositional attituders their preferred account of polysemy, as long as it's consistent with my use of the diagnostics below.

[8]By focusing on homonymy and polysemy, I leave open whether *believe* makes a distinct contribution in a belief-in report because it's context-dependent. For lack of space, I can't discuss this option here, though I'm not aware of diagnostics (see e.g. Viebahn and Vetter 2016) that favour it over the claim that *believe* makes the same contribution across belief-in and belief-that reports.

[9]My discussion of the zeugma and contradiction tests are deeply indebted to Mankowitz 2024's helpful summary of how they work and what limitations they have. My use of the pedantic correction test follows Felappi 2019.

[10]For recent discussions of the zeugma test see Liu 2023; Liebesman and Magidor 2024; Mankowitz 2024.

[11]Due to confounds, other elided conjunctions sound at least somewhat odd. Consider pseudogapping:

Unlike with the zeugma test, failing some other tests does distinguish unambiguous from polysemous expressions, allowing me to argue that *believe* isn't polysemous. Take the contradiction test.[12] We take a sentence with our target word and first affirm, then deny it. If we can get a context where this is acceptable, that's evidence our target is ambiguous. If we can't, that's evidence our target is unambiguous, i.e. neither homonymous nor polysemous. For instance, in a context where you drank the content of a bottle of juice, *What you drank was a bottle, but it wasn't a bottle* is acceptable. That's evidence that *bottle* is homonymous or polysemous.

Crucially, *believe* fails the contradiction test. Even in contexts where belief-that and belief-in are both salient, first affirming, then denying belief is unacceptable.

Context 1: Boris believes in his friends but doesn't believe that he'll need their help soon.
Context 2: Carol believes in sets but doesn't believe that everyone is wrong about them.

3. # {Boris/Carol} believes, but doesn't believe.

Another test to distinguish a polysemous expression like *bottle* from an unambiguous one is the pedantic correction test used by Felappi (2019, p. 5) (see also Textor 2011). In cases of co-predication, a pedantic speaker (imagine a strict English teacher) may correct one's use of the polysemous word, e.g. by replying to *David drank and smashed the bottle* that *David didn't drink the BOTTLE: he drank its CONTENT*. Yet even a pedantic speaker may not correct the uses of *believe* in 1 and 2 in the way they could correct our use of *bottle*: responding to 1a with *Sue doesn't BELIEVE that everyone is wrong about sets: she THINKS so* or to 2a with *John doesn't BELIEVE in all the teachings of his church: he has FAITH in them* would be odd.

- ?? John believes in all the teachings of his church and Jamal does that he is a prophet.
- ?? John believes that he is a prophet and Jamal does in all the teachings of his church.

But pseudogapping is generally infelicitous with belief reports:

- ?? John believes in all the teachings of his church and Jamal does in all the teachings of his.
- ?? John believes that he is a prophet and Jamal does that he is another prophet.

Topic shift also gets us odd-sounding sentences:

- – ? Boris believes in his friends and that it's raining.
  – ?? Boris believes that it's raining and in his friends.
- – ? Carol believes in sets and that it's raining.
  – ?? Carol believes that it's raining and in sets.

But topic shift also makes elided conjunctions of two belief-that reports odd-sounding:

- ? Boris believes that his friends are trustworthy and that it's raining.
- ? Carol believes that sets exist and that it's raining.

Changing the order of conjuncts can also get us odd-sounding sentences, even if we stay on topic:

- – Boris believes in his friends and that his mum is wrong about them.
  – ?? Boris believes that his mum is wrong about his friends and in his friends.
- – Carol believes in sets and that everyone else is wrong about them.
  – ?? Carol believes that everyone else is wrong about sets and in them.

But the contrast between the first and second members of each pair suggests that the second's oddity isn't due to ambiguity in *believe*, but due to something that affects only the second member. This suggestion is supported by corresponding non-elided conjunctions that sound similarly odd:

- ?? Boris believes that his mum is wrong about his friends and he believes in his friends.
- ?? Carol believes that everyone else is wrong about sets and she believes in them.

Whatever goes wrong with the non-elided conjunctions likely goes wrong with their elided counterparts too. Perhaps, asserting the first conjunct triggers some form of inference to the second, rendering assertion of the second redundant. Believing that everyone is wrong about sets might (given contextual factors we assume in our interpretation) presuppose believing in them; likewise for believing that one's mum is wrong about one's friends and believing in one's friends. If so, the second conjuncts assert what's presupposed by the first, as in

- ?? Sue quit smoking and used to smoke.

[12] Viebahn (2018) argues that the contradiction test doesn't distinguish unambiguous from polysemous expressions. But see Mankowitz 2024 for a reply.

In sum, *believe* fails the zeugma, contradiction, and pedantic correction tests.[13] That's evidence that *believe* doesn't make different contributions across belief-that and belief-in reports.

3 A Type-Shifter Does Not Help

Given that *believe* makes the same contribution across belief-that and belief-in reports, propositional attituders must say that *in O* somehow indirectly provides a propositional argument to saturate *believe*'s proposition slot. One way to do this is to appeal to a type-shifter, a function that takes the property denoted by *in O* as input and returns a proposition as output; which proposition is returned can be constant or context-dependent. One key consequence of using a type-shifter is that we predict belief-in reports to just attribute a belief-that. For if (in a context) the type-shifter converts the target property F into a proposition p, we predict the equivalence (in that context) of *S believes in O*, where *in O* (in that context) expresses F, and *S believes that P*, where P (in that context) expresses p.

Based on some of the observations I've made in Wimmer forthcoming, section 4, I argue against this equivalence: belief-in reports don't just attribute a belief-that. My argument works by comparing Price's (1965; 1969) type-shifter-friendly theory with part of my forthcoming type-shifter-unfriendly theory. I argue that my theory is preferable to Price's: it covers the same variety of cases, but doesn't multiply senses.[14,15]

Price distinguishes two senses of belief-in reports. Based on differences between examples like 4 and 5, he speaks of a 'factual' and an 'evaluative' sense (1965, pp. 12–3).

4. Carol believes in sets.
5. Daniele believes in public transport.

For Price, 4 is, at least by default, interpreted factually. That's because, at least by default, it seems equivalent to *Carol believes that sets exist.*[16] By contrast, 5 is, at least by default, interpreted evaluatively. The evaluative sense is meant to require that one positively evaluates what one believes in. For Price (1965, p. 18), that evaluation takes the form of an evaluative belief-that: 5 attributes to Daniele the belief that public transport is a highly efficient way of transporting persons and commodities, and that that's a good thing too. A belief-in report's evaluative sense thus attributes a belief that what one believes in has a certain property, and that it's a good thing that it does so. For Price, then, any belief-in report—whether factual or evaluative—is equivalent to a belief-that report, just as a type-shifter predicts.

I've previously argued that belief-in reports don't have distinct factual and evaluative senses, and that their different meanings are due to context-dependence instead (Wimmer forthcoming, sections 3 and 4). On my view, if one believes in O, one must (i) believe O to have a contextually salient property, e.g. to be a highly efficient way of transporting persons and commodities. And it must (ii) be the case that one would have contextually relevant positive feelings (e.g. satisfaction

[13] *Believe* also fails three further tests Cruse (1986, pp. 59–66) uses to distinguish homonymous from unambiguous and polysemous expressions. For reasons of space, I cannot go through these tests here.

[14] Further observations in Wimmer forthcoming, section 5 suggest an evidential requirement on belief-in reports, on which if one believes in O, one either has reason to form a belief-that before gathering easily available sensory evidence or cannot easily gather that sensory evidence. Unfortunately, presenting the argument against type-shifter-friendly theories based on these observations is beyond the scope of my present discussion.

[15] For lack of space, I can't test for distinct factual and evaluative belief-in readings or senses and so can't argue that there's no necessity to multiply senses here. But see Wimmer forthcoming, section 2.

[16] At best, this apparent equivalence only holds by default. In a context where it's salient whether sets can act as the foundation for mathematics, 4 can appear equivalent to *Carol believes that sets can act as the foundation for mathematics*. For Price, this would still be a factual interpretation. And in a context where it's also salient whether it's a good thing that sets can do this, 4 can be interpreted as also attributing the belief that it's a good thing that sets can do this. Here, Price would interpret 4 evaluatively. Parallel points also apply to 5.

or gratitude) should one learn that O has the property one believes it to have or would have contextually relevant negative feelings (e.g. disappointment or a sense of betrayal) should one learn that O doesn't have that property.[17]

(ii) requires that one is somehow affectively invested in what one believes in. But that investment neither entails nor is entailed by an evaluative belief-that. Thus, my 'affective requirement' is distinct from Price's 'evaluative belief requirement' and entails that belief-in reports aren't equivalent to belief-that reports.[18] For this reason, I use the remainder of this paper to argue for my requirement. I highlight that it's satisfied across the cases that force Price, given his evaluative belief requirement, to distinguish factual and evaluative senses. Given this, the affective requirement has the benefit of allowing us not to multiply senses.

Price's evaluative belief requirement for 5 explains a key intuition. Suppose Daniele becomes a 'localist' and now believes that it's a bad thing to transport persons and commodities—people and goods better stay where they are. Now, 5 isn't true, at least not on its default reading (Price 1965, p. 18). For Price, that's because Daniele now lacks the required evaluative belief that it's a good thing that public transport is a highly efficient way of transporting persons and commodities. But my affective requirement explains this intuition too. For if Daniele is a localist, he wouldn't be satisfied (disappointed) should he learn that public transport is (isn't) an effective way of transporting persons and commodities.

Crucially, my affective requirement also covers cases Price's evaluative belief requirement doesn't cover. Sometimes subjects have a belief-in, yet believe what they believe to be the case to be neutral or even bad, rather than good. These subjects don't satisfy Price's evaluative belief requirement. This forces Price to posit a distinct factual sense of belief-in reports that doesn't have the evaluative belief requirement. As an example, consider:

> Context: 6-year-old Gino believes it's bad to have monsters in one's wardrobe. Despite his parent's efforts to convince him otherwise, he believes there are monsters in his wardrobe.
>
> 6. Gino believes in monsters. True

Gino is like localist Daniele in believing what he believes to be the case to be bad, rather than good. Yet unlike localist Daniele, Gino has a belief-in. 6 is true.

Unlike Price's evaluative belief requirement, my affective requirement covers cases like Gino's. Thus, on my view, such cases don't force us to multiply senses. My requirement's crucial innovation is that it's possible that one would have some positive feeling should one learn that a proposition one believes is true or some negative feeling should one learn that that proposition isn't true, even if one doesn't believe its truth to be a good thing.[19] With this in mind, consider Gino. Though he wouldn't be grateful if he learned that there are monsters in his wardrobe, he'd feel a kind of satisfaction. For he'd be proven right, against the 'better judgment' of his parents. Conversely, Gino would feel a kind of disappointment if he learned that no monsters are in his wardrobe. For he'd be proven wrong, in line with the 'better judgment' of his parents. The satisfaction or disappointment Gino would feel in these cases would concern his role as a knower: they would be 'epistemic' or 'intellectual' feelings.[20]

[17] Conditions (i) and (ii) are individually necessary, but not jointly sufficient conditions for believing in O. To get jointly sufficient conditions we also need the evidential requirement mentioned in fn. 14.

[18] More precisely, the affective requirement only entails that belief-in reports are sometimes not equivalent to belief-that reports. See Wimmer forthcoming, fn.12 for discussion.

[19] It's also possible if one doesn't believe that it would be a good thing to learn that that proposition is true.

[20] One might reply that Gino believes that it would be good for him, in his role as a knower, to learn that there are monsters. This would allow Price to say that even Gino satisfies a version of the evaluative belief requirement. Thus, Price wouldn't need to multiply senses. However, this reply raises difficult questions. Must Gino have such a belief to believe in monsters? And if he had such a belief, wouldn't he also have a belief that it would be good for him, in his role as a knower, to learn that no monsters are in his wardrobe? After all, learning the truth is obviously good for one in one's role as a knower. But if so, could his intellectual satisfaction be due to his

Given what I've said about Gino, one might worry that the affective requirement predicts incorrect truth-conditions for 5. At least by default, 5 entails that Daniele would be satisfied in his role as a public transport advocate, not in his role as a knower, should he learn that public transport is an effective way of transporting persons and commodities. But, to cover Gino's case, didn't I weaken the affective requirement to merely entail satisfaction in some role? No. Context determines in what role one would have to be satisfied (disappointed) should one learn that the things one believes in do (don't) have the property they're believed to have. This yields the right predictions about Gino and Daniele. Gino's role as a knower is made salient by his parents' contextually given attempt to convince him that there are no monsters in his wardrobe. By contrast, Daniele's role as a public transport advocate is made salient by the default context we assume when hearing of someone believing in public transport.

In sum, my theory covers the same variety of cases as Price's, but doesn't multiply senses. And crucially, its affective requirement entails that belief-in reports aren't equivalent to belief-that reports. My theory is type-shifter-unfriendly.

4 Conclusion

I've argued that *believe* isn't a propositional attitude verb based on two observations: (i) *believe* can embed *in O*, and (ii) *in O* doesn't denote a proposition. To develop my argument, I suggested (section 2) that *believe* isn't homonymous or polysemous between a propositional belief-that and non-propositional belief-in interpretation, and (section 3) that type-shifting *in O*'s denotation to a proposition falsely predicts that belief-in and belief-that reports are equivalent.

As presented here, my argument isn't decisive. Fns. 6 and 20 noted two possible replies, along with some key questions for them. But other replies are possible too. For instance, propositional attituders might argue that positing homonymy or polysemy without independent evidence is still worth it, because it's necessary for getting the overall best compositional semantics of belief-in and belief-that reports. I haven't said anything to rule out this reply so far. But, I want to close by sketching, albeit very briefly, a semantics that puts pressure on it.

Let *believe* contribute the property of being a belief state to belief-that and belief-in reports (cf. e.g. Borer 2005; Pietroski 2005; Lohndal 2014; Elliott 2017). And let its embedded *that*-clause denote the property of having a certain propositional content (e.g. Kratzer 2006; Moltmann 2013; Moulton 2015; Elliott 2017; Bassi and Bondarenko 2021). Finally, drawing on Wimmer forthcoming, but omitting the evidential requirement mentioned in fn. 14, let *in O* denote the property of (i) having the propositional content that O has a contextually salient property and (ii) requiring a contextually salient affective investment in O having that property. *Believe* can now compose with a *that*-clause and with *in O* via predicate modification.[21] This gives propositional attituders a compositional semantics to contend against.[22]

Acknowledgements. I thank Guy Longworth, Kristina Liefke, Giulia Martina, the audience at PLM6, and anonymous reviewers for AC2024 and several journals for discussion. Some of the work on this paper was supported by a postdoc stipend from the Fritz Thyssen Foundation.

belief that it would be good for him, as a knower, to learn that there are monsters? If it was due to it, his belief that it would be good for him, as a knower, to learn that there are no monsters would be predicted to also give rise to intellectual satisfaction upon learning that there are no monsters. But would he have that intellectual satisfaction? Unfortunately, pursuing issues like these is beyond the scope of my present discussion.

[21] Even if an embedded *that*-clause was to denote a proposition, a dedicated syntactic head could convert that proposition to the property of having that proposition as content. Cf. Borer (2005) and Lohndal (2014): according to them, the internal argument of a verb is introduced by an additional syntactic head Asp.

[22] This semantics also explains restrictions on composition: e.g. the oddness of *S believes that P in O* is, roughly speaking, explained by two facts: (i) per eventuality, only one proposition can fill the thematic role of being its content; (ii) both *that P* and *in O* give us a proposition to fill that role.

References

Bassi, Itai and Tatiana Bondarenko (2021). "Composing CPs: Evidence from disjunction and conjunction". In: *Semantics and Linguistic Theory* 30. DOI: 10.3765/salt.v30i0.4837.

Borer, Hagit (2005). *The normal course of events*. Structuring Sense 2. Oxford: Oxford University Press.

Brody, Gabor and Roman Feiman (2023). "Polysemy does not exist, at least not in the relevant sense". In: *Mind & Language*. DOI: 10.1111/mila.12474.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

Djärv, Kajsa (2023). "Knowing and Believing Things: What DP-Complementation Can Tell us about the Meaning and Composition of (Factive) Attitudes". In: *Journal of Semantics*. DOI: 10.1093/jos/ffac015.

— (2024). "Source and Content DPs: Belief-Verbs as Optionally Ditransitive Hintikkan Attitudes". In: *Proceedings of the 39th West Coast Conference on Formal Linguistics*. Ed. by Robert Autry et al. Somerville, MA: Cascadilla Proceedings Project, pp. 124–134.

Elliott, Patrick D. (2017). "Explaining DPs vs. CPs without syntax". In: *Proceedings of the Chicago Linguistic Society*. Chicago, pp. 171–185. URL: https://patrl.keybase.pub/papers/cls52_final.pdf.

Felappi, Giulia (2019). "The Face-Value Theory, Know-that, Know-wh and Know-how". In: *Thought: A Journal of Philosophy* 8.1, pp. 63–72. DOI: 10.1002/tht3.406.

Fodor, Jerry A. and Ernest LePore (2002). *The compositionality papers*. Oxford: Clarendon Press.

Forbes, Graeme (2018). "Content and Theme in Attitude Ascriptions". In: *Non-propositional Intentionality*. Ed. by Michelle Montague and Alex Grzankowski. Oxford: Oxford University Press.

Grice, H. Paul (1978). "Further Notes on Logic and Conversation". In: *Pragmatics (Syntax and Semantics 9)*. Ed. by Peter Cole and Morgan. New York: Academic Press, pp. 113–127.

King, Jeffrey C. (2007). *The nature and structure of content*. Oxford: Oxford University Press.

Kratzer, Angelika (2006). *Decomposing attitude verbs*. Hebrew University, Jerusalem.

Kriegel, Uriah (2018). "Belief-That and Belief-In: Which Reductive Analysis?" In: *Non-Propositional Intentionality*. Ed. by Alex Grzankowski and Michelle Montague. Oxford University Press. DOI: 10.1093/oso/9780198732570.003.0008.

Liebesman, David and Ofra Magidor (2024). "Ambiguity Tests, Polysemy, and Copredication". In: *Australasian Journal of Philosophy*, pp. 1–10. DOI: 10.1080/00048402.2024.2310193.

Liu, Michelle (2023). "How to Think about Zeugmatic Oddness". In: *Review of Philosophy and Psychology*. DOI: 10.1007/s13164-023-00718-5.

Lohndal, Terje (2014). *Phrase structure and argument structure: a case study of the syntax-semantics interface*. Oxford studies in theoretical linguistics 49. Oxford, United Kingdom: Oxford University Press.

Mankowitz, Poppy (2024). "One "True" Meaning". In: *Ergo an Open Access Journal of Philosophy* 11. DOI: 10.3998/ergo.6783.

Moltmann, Friederike (2013). *Abstract Objects and the Semantics of Natural Language*. Oxford: Oxford University Press. DOI: 10.1093/acprof:oso/9780199608744.001.0001.

— (2022). "Empathetic attitude reports". In: *Proceedings of Sinn und Bedeutung* 26, pp. 609–621. DOI: 10.18148/sub/2022.v26i0.1020.

Moulton, Keir (2009). "Clausal Complementation and the *Wager*-class". In: *Proceedings of the 38th Annual North East Linguistics Society* 2, pp. 165–178.

— (2015). "CPs: Copies and Compositionality". In: *Linguistic Inquiry* 46.2, pp. 305–342. DOI: 10.1162/LING_a_00183.

Müller, Stefan (2023). *Grammatical theory*. 5th ed. Language Science Press. DOI: 10.5281/zenodo.7376662.

Osborne, Timothy (2015). "Diagnostics for Constituents: Dependency, Constituency, and the Status of Function Words". In: *Proceedings of the Third International Conference on Dependency Linguistics*, pp. 251–260.

Pearson, Hazel (2013). "The Sense of Self". Dissertation. Cambridge, Mass: Harvard University.

Pietroski, Paul M. (2005). *Events and Semantic Architecture*. Oxford: Oxford University Press.

Price, H. H. (1965). "Belief 'In' and Belief 'That'". In: *Religious Studies* 1.1, pp. 5–27. DOI: 10.1017/S0034412500002304.

— (1969). *Belief*. Gifford Lectures. London: George Allen & Unwin Ltd.

Schiffer, Stephen R. (2003). *The things we mean*. Oxford: Oxford University Press.

Stalnaker, Robert C. (1984). *Inquiry*. Cambridge, Massachusetts: The MIT Press.

Szabó, Zoltán Gendler (2003). "Believing in Things". In: *Philosophy and Phenomenological Research* 66.3, pp. 584–611. DOI: 10.1111/j.1933-1592.2003.tb00280.x.

Textor, Mark (2011). "Knowing the Facts". In: *Dialectica* 65.1, pp. 75–86.

— (2013). "Seeing something and believing IN it". In: *Perspectives on Perception*. Ed. by Mary Margaret McCabe and Mark Textor. Berlin: Walter de Gruyter, pp. 65–86.

Uegaki, Wataru (2016). "Content Nouns and the Semantics of Question-Embedding". In: *Journal of Semantics* 33.4, pp. 623–660. DOI: 10.1093/jos/ffv009.

Vicente, Agustín and Ingrid L. Falkum (2017). "Polysemy". In: *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. DOI: 10.1093/acrefore/9780199384655.013.325.

Viebahn, Emanuel (2018). "Ambiguity and Zeugma". In: *Pacific Philosophical Quarterly* 99.4, pp. 749–762. DOI: 10.1111/papq.12229.

Viebahn, Emanuel and Barbara Vetter (2016). "How many meanings for 'may'? The case for modal polysemy". In: *Philosopher's Imprint* 16.10. URL: http://hdl.handle.net/2027/spo.3521354.0016.010.

Wimmer, Simon (forthcoming). "Belief-in is belief-that with affectivity and evidentiality". In: *Proceedings of Sinn und Bedeutung 28*. URL: https://philpapers.org/go.pl?id=WIMBIB&u=https%3A%2F%2Fphilpapers.org%2Farchive%2FWIMBIB.pdf.

On the availability of inverse scope reading in Dutch doubly-quantified thetic sentences

Hongchen Wu
Georgia Institute of Technology
Atlanta (GA), United States of America
hwu480@gatech.edu

Maxime A. Tulling
Université de Montréal
Montréal (QC), Canada
maxime.tulling@umontreal.ca

Abstract

This study investigates the availability of inverse scope readings in Dutch doubly quantified sentences, focusing on thetic (general rule-based) versus categorical (individual-focused) contexts. Although previous research suggests that inverse scope readings are rarely available in Dutch simple transitive sentences, we show that such readings are available and that the availability of such readings correlates with the theticity and tense of the doubly quantified sentences. Building on Mandarin experiments by Wu (2019, 2023), the study adopts a 2x2 design for Dutch, manipulating utterances' tense (present vs. past) and context (thetic vs. categorical). The results of a scope judgment task (N=113) show that inverse scope is significantly more acceptable in thetic contexts than in categorical ones, and marginally less acceptable in sentences about the past compared to the present. These results challenge the isomorphic scope interpretation typically proposed for Dutch doubly-quantified simple transitions and contradict the claim that indefinite subjects cannot be interpreted non-specifically in Dutch. Furthermore, this study provides cross-linguistic evidence for the impact of the thetic-categorical distinction on the quantifier scope and calls for further investigation into the interaction between temporal marking, thetic-categorical distinction, and scope interpretations.

1 Introduction

Unlike the ambiguity of scope observed in English (1), canonical Dutch transitive sentences with an indefinite subject and a universally quantified object (2) strongly reject the inverse scope reading (Philip 2005; Hendriks, Koops van 't Jagt, and Hoeks 2012).

(1) A bear tickles every turtle.
a. (surface scope) a > every: 'There is a specific bear who tickles all the turtles.'
b. (inverse scope) every > a: 'For every turtle, there is a different bear who tickles it.'

(Hendriks et al. 2012, ex. 1)

(2) Een beer kietelt elke schildpad.
a bear tickles every turtle;
'A bear tickles every turtle. (a > every; *every > a) (Hendriks et al. 2012, ex. 2)

However, in certain contexts, such as in (3), Dutch speakers find the doubly quantified sentence fully ambiguous, similar to its Mandarin counterpart, despite claims that both Mandarin and Dutch typically exhibit an isomorphic scope interpretation for doubly quantified simple transitives.

(3) Zoals opgelegd door de wet, bewaken drie politieagenten iedere uitgang.
as required by the law, guard three policemen every exit
'As required by law, three policemen should be guarding every exit. (3 > every, every > 3)

(Wu 2019, ex. 239)

(4) (Anzhao falü guiding) san-ming jingcha kanshou meige chukou.
(As law demands) Three-CL policemen guard every exit
'As required by law, three policemen guard every exit.' (3 > every; every > 3)

The context in (3) and (4), which allows inverse scope readings in Dutch and Mandarin, is known as a thetic environment. According to Kuroda 1972 and his subsequent work, thetic sentences assert general regularities or the existence of events. In contrast, categorical sentences attribute properties to specific

individuals. In the Japanese data that Kuroda 1972 examined, a categorical sentence in Japanese is grammatically marked through 'topic marking' and is composed of a thetic predication combined with a direction of attention, interest or consciousness toward an entity. Interestingly, it has been observed that inverse scope readings are not available in categorical sentences in Mandarin (Larson and Wu 2018) and in English (Basilico 1998; Basilico 2017). However, what exactly the relationship is between the thetic-categorical distinction and quantifier scope remains an intriguing yet understudied question.

Although very few studies have experimentally tested scope judgments in thetic sentences, Wu (2019; 2023) has confirmed the high availability of inverse scope readings in Mandarin thetic sentences (4). However, the thetic environment does not always yield a high availability of inverse scope readings, as it may be constrained by other syntactic cues. For instance, Wu and Larson (2023) observed that in thetic sentences, the accessibility of inverse scope readings appears to vary depending on the presence of the aspect marker, being more accessible when the aspect marker is absent and less accessible when it is present. Ladusaw (1994) following Kratzer (1989) and Diesing (1992), argues that temporal orientations reflected in the INFL domain can cue the thetic-categorical distinction. It also aligns with our basic language intuition that sentences with past temporal orientation (e.g., aspect marker is present) often express completed events and completed events often involve presupposing specific individuals (i.e., more compatible with categorical environments). But, it remains unclear if and how theticity and tense affect scope judgments in languages and if these patterns generalize across languages.

Questions & Hypotheses

The present study aims to experimentally investigate if and how theticity and tense affect quantifier scope readings in Dutch. Unlike Mandarin Chinese which relies on grammatical aspects to indicate temporals, Dutch obligatorily marks tense on the inflected verb, making it a great test case for disentangling the effects of theticity and temporal orientation.

Based on existing research and our native intuitions, two main hypotheses and predictions emerge. First, we hypothesize that inverse scope readings are more available under thetic contexts than under categorical contexts. We predict that this pattern holds for Dutch doubly quantified sentences irrespective of the temporal orientation markings of the sentence. Second, we expect that temporal marking affects the degree of availability of inverse scope. We anticipate that doubly quantified simple transitives with past tense in Dutch allow less for inverse scope readings compared to the ones with present tense, since descriptions of past events are more likely to be associated with specific individuals.

2 Methodology

Materials

Building on Wu's (2019; 2023) experiments with Mandarin thetic sentences, we designed a 2x2 experiment to investigate scope judgments in Dutch. We created 4 sets of stories that manipulated the TENSE and CONTEXT of the target utterances. Each target utterance contained a transitive verb with an indefinite numerical plural subject and a universally quantified singular object (e.g., *drie studenten versieren elk klaslokaal* 'three students decorated every classroom'). We manipulated the TENSE of the target utterance to be PRESENT or PAST. In addition, we provided context for the utterances that were either THETIC (descriptions of general rules or traditions) or CATEGORICAL (about specific individuals and a specific event) and matched the temporal orientation of the target utterance.

An example item set can be found in Figure 1. The situation contexts were displayed in blue, and target utterances in red. After each target utterance, participants were asked to complete a binary forced-choice task. A question was displayed (matching the tense and theticity of the context), asking the participant how many people were involved with the described action, given that there was X number of objects: 'X' or 'more than X' (e.g., *Er zijn vijf klaslokalen. Hoeveel studenten zijn er aan het versieren? Drie studenten of meer dan drie studenten?* 'There are five classrooms. How many students are decorating? Three students or more than three students?'). The choice of 'X' corresponds to surface scope reading while the choice of 'more than X' corresponds to inverse scope reading.

In addition to the target utterances, we created 12 filler items that required participants to read the context carefully and use the information provided in the context to answer questions based on quantity. For example, indicating the maximum or minimal amount a person wants to spend on a gift or indicating

the minimal or maximal time that someone is required to work. These filler items were used as attention checks to make sure participants paid attention to the instructions when completing the tasks. We also randomized the presentation order of the two given choices for filler items, in order to counterbalance potential biases that may arise from a default tendency to select the first or second option.

	Thetic present	**Thetic past**	**Categorical present**	**Categorical past**
Situation:	Volgende maand moet de tweede klas de klaslokalen voor het eerste leerjaar inrichten voor de feestdagen. De mentor van de tweede klas, die alles graag goed geregeld heeft, vraagt de directrice naar de regels en richtlijnen voor het inrichten van de lokalen, zodat alles volgens de traditie verloopt. De directrice antwoordt:	Tijdens een schoolreünie bespreekt een docent met een oud-leerling hoe de schooltradities door de jaren heen zijn veranderd. Nu is het gebruikelijk dat heel de eerste klas de versiering regelt tijdens de feestdagen, maar vroeger was het anders. De docent vraagt hoe dit in voorgaande jaren geregeld werd. De oud-leerling antwoordt:	Het is 1 april en een groepje studenten heeft bij hun school binnengebroken om een grap uit te halen waarbij ze de klaslokalen van de onderbouw helemaal volhangen met slingers en ballonnen. Door hun acties is echter het automatische alarmsysteem afgegaan, en een medewerker van het beveiligingsbedrijf heeft de conciërge opgebeld om te vragen wat er aan de hand is. De conciërge antwoordt:	Op 1 april heeft een groepje studenten bij hun school binnengebroken om de klaslokalen helemaal vol te hangen met slingers en ballonnen. De volgende dag ziet de directeur alle rommel en vraagt aan de conciërge wat er gebeurt is. De conciërge antwoordt:
	Next month, the second-year students need to decorate the first-year classrooms for the holidays. The second-year mentor, who likes to keep everything well-organized, asks the principal about the rules and guidelines for decorating the classrooms to ensure everything goes according to tradition. The principal replied:	*During a school reunion, a teacher discusses with a former student how school traditions have changed over the years. Nowadays, it's customary for the entire first-year class to handle the decorations during the holidays, but it used to be different. The teacher asks how this was organized in previous years. The former student replied:*	*It's April 1st, and a group of students broke into their school to pull a prank by filling the classrooms of the lower grades with streamers and balloons. However, their actions triggered the automatic alarm system, and a security company employee called the janitor to ask what was going on. The janitor replied:*	*On April 1st, a group of students broke into their school to fill the classrooms with streamers and balloons. The next day, the principal sees the mess and asks the janitor what happened. The janitor replied:*
Target:	Drie studenten versieren elk klaslokaal. *Three students decorate every classroom.*	Drie studenten versierden elk klaslokaal. *Three students decorated every classroom.*	Drie studenten versieren elk klaslokaal. *Three students are decorating every classroom.*	Drie studenten versierden elk klaslokaal. *Three students decorated every classroom.*
Question:	Er zijn vijf klaslokalen. Hoeveel studenten zijn er normaal gesproken nodig om ze te versieren? *There are five classrooms. How many students are usually needed to decorate them?*	Er waren vroeger vijf klaslokalen. Hoeveel studenten waren er normaal gesproken nodig om ze te versieren? *There used to be five classrooms. How many students were usually needed to decorate them?*	Er zijn vijf klaslokalen. Hoeveel studenten zijn er aan het versieren? *There are five classrooms. How many students are decorating?*	Er zijn vijf klaslokalen. Hoeveel studenten waren er aan het versieren? *There are five classrooms. How many students were decorating?*
Response:	A. 3 B. more than 3	A. 3 B. more than 3	A. 3 B. more than 3	A. 3 B. more than 3
Scope:	3 > every every > 3	3 > every every > 3	3 > every every > 3	3 > every every > 3

Figure 1: Example stimulus set (n = 4) displaying each experimental condition. Stories manipulated the context (thetic or categorical) and tense (present or past) of the target utterances. All items are available online.

Data acquisition, Pre-processing & Analysis

We recruited 145 adult Dutch-speaking participants (18 - 62 years old, M = 32.6, SD = 9.36) via Prolific 2024 for a 5 to 10-minute study conducted on Qualtrics. After completing two practice items, participants were presented with 4 target items (one per condition, drawn from four different sets) and 12 filler items in a randomized order. Participants viewed one item at a time and were not allowed to revise previous responses. To ensure coverage of all 16 target utterances, the experiment had 4 versions, each containing a different subset of the target items. We excluded 32 participants who scored below 85% accuracy on the filler items, leaving 113 participants. These participants collectively provided 19 to 23 observations per target utterance.

We calculated the response percentages of the forced choice task (choosing 'X' or 'more than X') for each condition, aggregated across items to visualize the data. To determine whether participants' responses were influenced by our experimental manipulations, we analyzed the binary response data with a generalized linear mixed-effects model. Our maximal converging model included CONTEXT (thetic or categorical) and TENSE (present or past) and their interaction as fixed effects, with random intercepts for items and participants. To test the contribution of our fixed effects, we performed a likelihood ratio test comparing our model and a nested model leaving out the variable of interest. The analysis was performed using the glmer-function from the 'lme4' package in R (Bates et al. 2015; Team 2021). We used the 'emmeans' package (Russell 2022) to generate predicted probabilities and standard errors so that we could visualize the model's findings while accounting for variation across participants and items.

3 Results

The nested GLMM model comparisons revealed that CONTEXT was a significant predictor of participants' responses to the forced choice task ($\chi^2(1) = 6.76, p < .01$). The odds of selecting 'more than X' (inverse scope) increased for utterances in a thetic context compared to a categorical one (OR = 1.99, CI = 1.17–3.38). The utterance's TENSE was not a significant predictor ($\chi^2(1) = 3.29, p = .07$), although the odds of selecting 'more than X' marginally increased for present compared to past utterances (OR = 1.58, CI = 0.96–2.61). The predicted probabilities of choosing 'more than X' in the different experimental conditions are displayed in Figure 2. There was no significant interaction between the two predictors ($\chi^2(1) = 0.19, p = 0.66$).

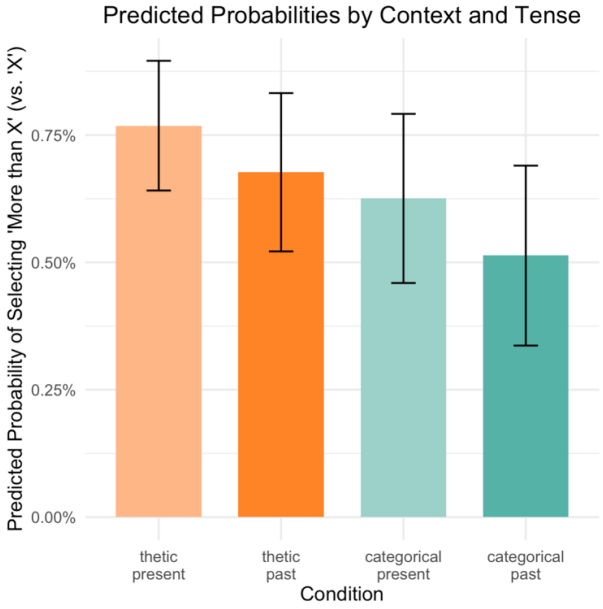


Figure 2: Predicted probability of selecting 'more than X' (inverse scope) over 'X' (surface scope) for experimental conditions, based on GLMM model output.

Predicted probabilities of choosing 'more than X' were higher than anticipated, ranging from 50% (categorical past) to 75% (thetic present). Two factors contributed to this. First, among our 113 participants, 23 participants always picked 'more than X' and 8 only picked 'X', which may indicate a response bias or misunderstanding of the task instructions. For the 82 remaining participants who did differentiate between conditions, we plotted their average responses in Figure 3, aggregated by experimental condition and item set (1–4).

As can be observed from the grouped responses to different experimental item sets, a second contributing factor to the high proportion of predicted 'more than X' responses is the fact that half of the experimental items elicited a high proportion of 'more than X' responses in general. Observationally, high-stakes items (e.g., police officers guarding exits) elicited more 'more than X' responses compared to casual situations (e.g., children decorating classrooms).

4 Discussion

All in all, our experimental design was effective in showing that thetic contexts increase the availability of inverse scope judgments in Dutch, just as previously observed for Mandarin Chinese (Wu 2019; Wu 2023). We also found that tense independently affects sentence interpretation, such that inverse scope is less available in past tense sentences, even though this effect was only marginally significant. Besides the effects of our experimental manipulations, we also observed high variability between participants and experimental items. The majority of participants (n = 82) distinguished between experimental conditions, while 23 consistently chose responses aligned with an inverse scope reading and 8 exclusively favored surface scope interpretations. Future research should explore whether these patterns reflect genuine population-level differences in scope assignment or can be attributed to task-specific effects. In

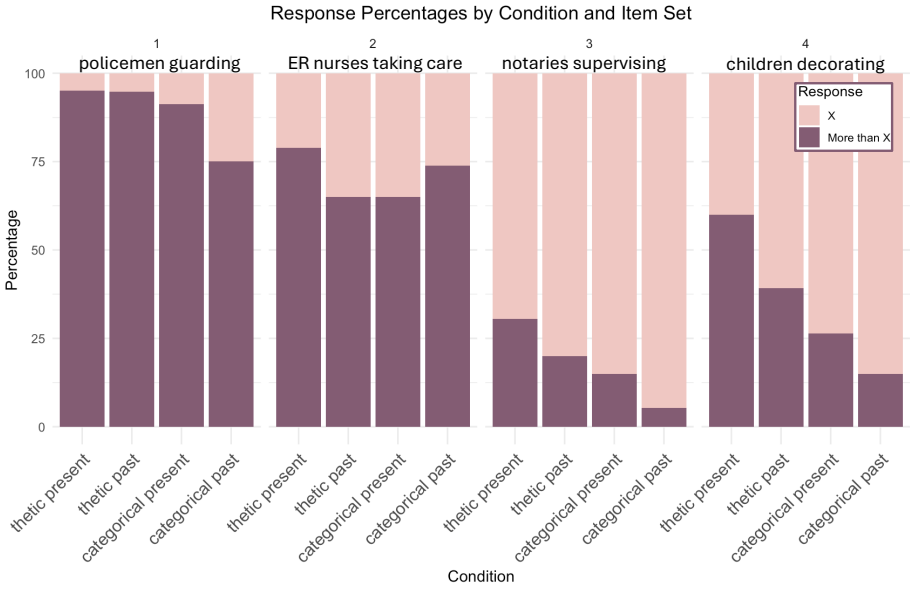


Figure 3: Percentage of average responses per condition and item set (1-4) across the 83 participants that differentiated between experimental conditions.

addition to this, high-stake scenario items (e.g., nurses manning intensive care units) elicited a much higher proportion of 'more than X' responses (compatible with an inverse scope reading) than items describing more casual situations (such as children decorating classrooms). We believe this might be imposed by world knowledge, e.g., the knowledge that it is more likely and more necessary to have many nurses in the intensive care unit. The exact influence of world knowledge on scope judgments should be explored further.

An important avenue for further research is the role of temporal markings in the accessibility of inverse scope. While the current study found only a marginal effect of tense, a pilot study (n = 12) indicated that the perfective aspect (e.g., *Drie kinderen hebben elk klaslokaal versierd* 'Three children have decorated every classroom') almost exclusively biases interpretations toward surface scope. This finding aligns with our native speaker judgments and suggests that the bounded timeframe associated with the present perfect may encourage a categorical perspective. In subsequent studies, we aim to manipulate tense and aspect in Dutch and Mandarin to compare the scope effects of temporal markings across these languages.

5 Theoretical Implications

The results yield several theoretical implications. First, the availability of the inverse scope reading in thetic contexts provides evidence against the isomorphic scope interpretation (i.e., surface scope only) typically argued for Dutch doubly-quantified simple transitives, and questions the accounts that have been proposed or employed to explain the absence of inverse scope reading in Dutch simple transitives (Beghelli and Stowell 1997, Philip 2005, Hoop and Krämer 2005, Hendriks, Koops van 't Jagt, and Hoeks 2012). The syntactic account (Beghelli and Stowell 1997) attributes the issue to covert movement, arguing that inverse scope reading is only possible when the subject is reconstructed to a scope position lower than Spec of DistP where the universal quantified object needs to land, but this covert movement is not accessible to sentences like (2). However, this account fails to explain why covert movement is accessible to sentences like (3) to yield inverse scope reading but not accessible in (2) in Dutch. In contrast, the pragmatic account (Philip 2005) argues that the interpretation of indefinite subjects in Dutch transitive sentences with a sentence-initial subject and a universally quantified direct object is guided by a pragmatic rule favoring the strongest meaning consistent with contexts. Take (2) as an example, the surface scope reading implies the inverse scope reading but not vice versa: if a specific bear tickles every turtle, then it is also true that each turtle is tickled by some bear, but not the other way around. Therefore, Dutch speakers would always choose the surface scope reading according to the Strongest Meaning Hypothesis. This account fails to explain why Dutch speakers would not obey the pragmatic rule and choose inverse

scope readings for sentences like (3). Likewise, Hoop and Krämer 2005's semantic constraint account also fails to account for the high availability of inverse scope readings in Dutch thetic sentences since this constraint argues that subjects should receive a specific interpretation and object receive a non-specific interpretation, making the inverse scope reading impossible. Moreover, none of these accounts fully explains the considerable variation among adult native speakers of Dutch as observed in the experiment, nor what conditions contribute to these differences.

Second, the availability of inverse scope readings contradicts the general claim that indefinite subjects cannot be interpreted non-specifically in Dutch (Hoop and Krämer 2005, Hendriks, Koops van 't Jagt, and Hoeks 2012). Instead, the results support the theoretical conjecture of Philip (2005: 274) that "it is always in principle possible for a high indefinite subject to be nonspecific as long as the verb is transitive.". An overlooked observation in the literature is that Dutch adults accepted the inverse scope reading for sentences like (2) about 20% of the time in a sentence-picture matching task (Hendriks, Koops van 't Jagt, and Hoeks 2012). However, the previous accounts did not answer the question of under what conditions can the indefinite subject be nonspecific reading, and a doubly-quantified sentence can have inverse scope reading (Wu and Larson 2023). Instead, the thetic-categorical framework by Kuroda 1972 offers a clear explanation of when an indefinite subject can have a nonspecific reading. In thetic contexts, the identity of the indefinite subject is irrelevant to the goal of describing the regularities, allowing for a nonspecific reading and, thus, the possibility of scope permutation with the universal quantifier.

Third, the significantly higher acceptability of inverse scope readings in thetic contexts compared to categorical contexts, especially in present inflected sentences in Dutch, provides important cross-linguistic evidence for the influence of the thetic-categorical distinction as well as tense marking on quantifier scope. Wu and Larson 2023 argues that the thetic-categorical distinction is encoded in the topic-prominence property of Mandarin. In Mandarin, subjects are by default categorical subjects and simple transitive sentences are typically categorical sentences. However, subjects in Mandarin can become thetic subjects in a de-topicalized structure to express a thetic sentence, thus allowing for scope permutations with quantified objects. Dutch, while not a prototypical topic-prominence language, shares similarities to Mandarin in that speakers strongly prefer interpreting grammatical subjects as definite but readily allow nonspecific readings for thetic readings. It is thus worthwhile to explore the interaction between temporal marking and the thetic-categorical distinction, as well as their combined influences on scope interpretations from a cross-linguistics perspective.

Acknowledgements: We thank all participants who took the time to complete the experiment. This research was funded by Faculty Professional Development funds from the School of Modern Languages and Ivan Allen College at the Georgia Institute of Technology, awarded to Hongchen Wu.

Data Availability: All stimuli, data, and analysis files are available at: `https://osf.io/ztxcu/`.

References

Basilico, David (1998). "Object Position and Predication Forms". In: *Natural Language & Linguistic Theory* 16.3, pp. 541–595.

— (2017). "Transitivity and Theticity". In: *Understanding Predication*. Ed. by Piotr Stalmaszczyk. Peter Lang GmbH, Internationaler Verlag der Wissenschaften, pp. 143–177.

Bates, Douglas et al. (2015). "Fitting Linear Mixed-Effects Models Using **lme4**". English. In: *Journal of Statistical Software* 67.1. ISSN: 1548-7660. DOI: 10.18637/jss.v067.i01. URL: http://www.jstatsoft.org/v67/i01/ (visited on 07/22/2018).

Beghelli, Filippo and Tim Stowell (1997). "Distributivity and Negation: The Syntax of Each and Every". In: *Ways of Scope Taking*. Ed. by Anna Szabolcsi. Dordrecht: Kluwer, pp. 71–107.

Diesing, Molly (1992). *Indefinites*. Cambridge, MA: MIT Press.

Hendriks, Petra, Roberta Koops van 't Jagt, and John Hoeks (2012). "Restricting Quantifier Scope in Dutch: Evidence from Child Language Comprehension and Production". In: *Empirical Approaches to Linguistic Theory: Studies in Meaning and Structure*. Ed. by Britta Stolterfoht and Sam Featherston. Berlin, Boston: De Gruyter Mouton, pp. 147–168. DOI: 10.1515/9781614510888.147.

Hoop, Helen de and Irene Krämer (2005). "Children's Optimal Interpretations of Indefinite Subjects and Objects". In: *Language Acquisition* 13.2, pp. 103–123.

Kratzer, Angelika (1989). "Stage-Level and Individual-Level Predicates". Manuscript, University of Massachusetts.

Kuroda, S.-Y. (1972). "The Categorical and the Thetic Judgment". In: *Foundations of Language* 9.2, pp. 153–185.

Ladusaw, William (1994). "Thetic and Categorical, Stage and Individual, Weak and Strong". In: *Proceedings from Semantics and Linguistic Theory IV*. Ed. by Mandy Harvey and Lynn Santelmann. Ithaca, NY: Cornell University, pp. 220–229.

Larson, Richard and Hongchen Wu (2018). "Quantifier Scope and Topicality in Mandarin". In: *Presentation at IACL-26, University of Wisconsin-Madison, May 6, 2018*.

Philip, William (2005). "Pragmatic Control of Specificity and Scope: Evidence from Dutch L1A". In: *Proceedings of Sinn und Bedeutung 9*. Ed. by Emar Maier, Corien Bary, and Janneke Huitink. Nijmegen: Radboud University, pp. 271–285.

Prolific (2024). *Prolific*. First release: 2014. London, United Kingdom. URL: `https://www.prolific.co`.

Russell, V. Lenth (2022). *emmeans: Estimated Marginal Means, aka Least-Squares Means*. R Package. URL: `https://CRAN.R-project.org/package=emmeans`.

Team, R Core (2021). *R: A Language and Environment for Statistical Computing*. Vienna, Austria. URL: `https://www.R-project.org/`.

Wu, Hongchen (2019). "Quantifier Scope in Mandarin". PhD thesis. Stony Brook University.

— (2023). "Three-Way Variance in Scope Judgments in Mandarin". Presentation at the Workshop on Variation and Syntactic Theory at NWAV 51, The City University of New York-Queens College, October 14, 2023.

Wu, Hongchen and Richard Larson (2023). "Quantifier Scope and Subjecthood in Mandarin". Manuscript, Georgia Institute of Technology and Stony Brook University.

Toward a Semantic Theory of Universal NCIs

Yusuke Yagi
University of Connecticut
Storrs, Connecticut, USA
yusuke.yagi@uconn.edu

Takanobu Nakamura
ILLC, University of Amsterdam
Amsterdam, The Netherlands
t.nakamura@uva.nl

Kenta Mizutani
Aichi Prefectural University
Aichi, Japan
kmizutani@for.aichi-pu.ac.jp

Abstract

Negative Concord Items (NCIs) are quantificational expressions that must be licensed by a local negation. Though the licensing condition has often been characterized by syntactic terminologies, Kuhn (2022) pursued semantic characterization of the condition. This paper extends Kuhn's analysis to NCIs with universal quantificational force, focusing on Japanese *dare-mo*. This is achieved by augmenting Hamblinian alternative semantics in Kuhn's dynamic setup. The success of the analysis provides further support for merging dynamic semantics and alternative semantics.

1 Introduction and Summary

Negative Concord Items (NCIs) are quantificational expressions that exhibit the three distinguishing properties summarized in (1). (Suppose the sentence in the `typewriter font` is a quasi-English.)

(1) a. "`Mary didn't see person`$_{\text{NCI}}$" is interpreted as "*Mary saw no one*".
b. NCIs must be licensed by a negation.
c. The licensing negation must be local to an NCI.

Although the accumulated literature pursues *syntactic* analyses of (1), Kuhn (2022) proposes a semantic analysis of these properties with the apparatus of dynamic semantics, which is reviewed in section 2. Despite an analytical success, Kuhn (2022) bases the proposal on the assumption that NCIs are *existential* quantifiers that take scope below negation, and this assumption is ill-suited to languages whose NCIs are *universal* quantifiers (Shimoyama 2011; see also Giannakidou 2000). This study aims to overcome this dilemma by augmenting Kuhn's dynamic analysis with Hamblinian alternative semantics (Hamblin 1973).

The focus of this paper is a Japanese NCI *dare-mo*, which Shimoyama (2011) convincingly argued to be a universal quantifier. The arguments by Shimoyama are laid out in section 3. Following Shimoyama (2011), we assume that *dare-mo* semantically takes scope over the licensing negation via syntactic movement. The movement leaves a higher-type trace, which lets *dare-mo* take split scope (Kuhn 2022) below and above negation. Below negation, *dare-mo* introduces Hamblinian alternatives to the composition. Above negation, it universally quantifies over the alternative. Furthermore, *dare-mo* enforces the empty value condition on a variable, notated as $\mathbf{0}_x$, which is the pivotal aspect of the semantic theory of NCI licensing put forth by Kuhn (2022).

The proposal, laid out in section 4, inherits Kuhn's correct predictions on the licensing condition of NCIs. Furthermore, it correctly predicts the scope interaction between *dare-mo* and quantificational adverbs observed in Shimoyama (2011). The success of the analysis offers

further support for the augmentation of dynamic semantics and plural alternative semantics, independently pursued by Li (2021).

2 Kuhn (2022)

The gist of Kuhn's (2022) proposal is summarized as follows.

(2) a. NCIs are existential quantifiers that undergo Quantifier Raising (QR) above the licensing negation.

b. They take split scope below and above the negation: their existential force takes scope below the negation, and they impose the empty value condition on a variable above it.

Kuhn (2022) formalizes the idea with a dynamic system. To assimilate the notation with our proposal, we describe it here with Plural Dynamic Logic (PDL; van den Berg 1996). PDL works with *plural information states* (PISs), i.e., sets (G, H, ...) of variable assignments functions $g, g', ..., h, h', ...$ Existential quantification $\exists_x \phi$ is interpreted in the standard way in dynamic literature, namely as random (re)assignment of the value x. We assume that a discourse starts with an empty state G such that $g(\alpha) = \star$ for every $g \in G$ and any variable α. $\star$ is the designated individual that makes every predicate undefined (van den Berg 1996). The formal details are laid out in section 4.

Consider now the formulae in (3) and (4). The intended interpretation is spelled out in words in the second line, and their dynamic effects are visualized in Figures 1 and 2. In (3), the existential force takes scope below the negation to derive the intended "*saw no one*" reading, satisfying (1a). Since negation insulates the dynamic effect of existential quantification, the first conjunct of (3), $\neg(\exists_x \wedge \texttt{saw}(\texttt{m}, x))$, is externally static. Hence, after interpreting the first conjunct, the value under x is still empty ($h(x) = \star$ for any $h \in H$). The contribution of the empty value condition $\mathbf{0}_x$ is the underlined part of the spelled-out interpretation. It tests if the value(s) stored under the variable x is empty. The condition is satisfied in (3) due to negation, and the sentence is grammatical. Since (4) does not contain negation, the dynamic effect of the existential quantification is passed onto the evaluation of the last conjunct, $\mathbf{0}_x$. The empty value condition is not satisfied because some values are stored under the variable x due to the dynamicity of existential quantification. Hence, the sentence is ungrammatical.

(3) $\neg(\exists_x \wedge \texttt{saw}(\texttt{m}, x)) \wedge \mathbf{0}_x$ (`Mary didn't see person`$_{\text{NCI}}$.)
$\rightsquigarrow$ It is not the case that there is a person x such that Mary saw x. <u>There is no such x.</u>

(4) $\exists_x \wedge \texttt{saw}(\texttt{m}, x) \wedge \mathbf{0}_x$ (`Mary saw person`$_{\text{NCI}}$.)
$\rightsquigarrow$ There is a person x such that Mary saw x. #<u>There is no such x.</u>

G	x	y	....
g_1	$\star$	$\star$	....
g_2	$\star$	$\star$	....
g_3	$\star$	$\star$	....

$\xRightarrow{\neg(\exists_x \wedge \texttt{saw}(\texttt{m},x))}$

H	x	y	....
h_1	$\star$	$\star$	....
h_2	$\star$	$\star$	....
h_3	$\star$	$\star$	....

$\Longrightarrow$ $\checkmark \mathbf{0}_x$

Figure 1: The dynamic effect of the grammatical use of NCI in (3)

G	x	y	....
g_1	$\star$	$\star$	....
g_2	$\star$	$\star$	....
g_3	$\star$	$\star$	....

$\xRightarrow{(\exists_x \wedge \texttt{saw}(\texttt{m},x))}$

H	x	y	....
h_1	a	$\star$	....
h_2	b	$\star$	....
h_3	c	$\star$	....

$\Longrightarrow$ $\# \mathbf{0}_x$

Figure 2: The dynamic effect of the ungrammatical use of NCI in (4)

The locality requirement of the licensing negation, although Kuhn (2022) himself does not commit to any theory, can be attributed to (2a). QR across the negation is a prerequisite for an NCI to take a split scope so that the existential force can take below the scope of negation and impose the empty value condition above it. The clause-boundedness of QR then accounts for the locality requirement (= (1c)).

3 Shimoyama (2011)

The Japanese *dare-mo* (WHO-ALSO/EVEN) is often analyzed as an existential quantifier scoping under negation like the English NPI *anyone* (e.g. Kawashima 1994 and Kishimoto 2008). Shimoyama (2011), however, claims that *dare-mo* is a universal quantifier scoping over negation. The argument comes from the way how *dare-mo* scopally interacts with a quantificational expression Q and negation, the combination of which is a Non-Anti-Additive (NAA) function $Q\neg_{NAA}$. A $Q\neg_{NAA}$ function does not validate anti-additivity, i.e. $f(A \vee B) \not\Rightarrow f(A) \wedge f(B)$, and separates otherwise indistinguishable $\forall > \neg$ and $\neg > \exists$ readings: $\forall > Q\neg_{NAA}$ is not equivalent to $Q\neg_{NAA} > \exists$. Shimoyama (2011) utilizes the quantificational adverb *taitei* (usually) to create a NAA function because it takes wide scope over clausemate negation as in (5).

(5) *Taro-wa taitei sankasi-nakat-ta.*
Taro-TOP usually participate-not-PAST
'It was usually the case that Taro did not participate.' (usually > not)
*'It was not the case that Taro usually participated.' (not > usually)

By employing the NAA function $Q_{usually}\neg_{NAA}$, Shimoyama (2011) demonstrates that *dare-mo* in Japanese indeed has the $\forall > \neg$ reading. Consider (6).

(6) *Nihonjin gakusei-no* ***dare-mo***$_{NCI}$ ***taitei*** *sankasi-****nakat****-ta.*
Japanese student-GEN NCI usually participate-not-PAST
'For every Japanese student, it was usually the case that s/he did not participate.'

We assume with Shimoyama (2011) that '*taitei*(ϕ)' is true for a proposition ϕ iff ϕ is true more than 50% out of all situations. Suppose then that there are three Japanese students and six meetings and that each student missed four out of the six meetings, as depicted in Table 1 (✓ marks attendance). Since every student misses more than three meetings, the $\forall >$ *taitei* $> \neg$ reading is true, and the sentence is indeed judged to be true under this scenario. On the other hand, the *taitei* $> \neg > \exists$ reading (i.e. it was usually the case that no students participated) is false under the same scenario because there were only two meetings in which no student participated. The truth of the $\forall >$ *taitei* $> \neg$ reading, therefore, indicates that *dare-mo* is a universal quantifier.

	MTG1	MTG2	MTG3	MTG4	MTG5	MTG6
STD1					✓	✓
STD2	✓					✓
STD3	✓	✓				

Table 1: Scenario for (6)

Coupled with the fact that *dare-mo* obeys the same licensing conditions as NCIs (see Watanabe 2004; Kataoka 2006), Shimoyama's (2011) observation suggests that *dare-mo* should be analyzed as a universal NCI.[1] Shimoyama (2011), however, does not commit to how the NCI

[1]We do *not* claim that every NCI in language is a universal quantifier. See, for example, Kamali and Zeijlstra (2024) for fine-grained discussions.

is licensed. At the same time, we cannot directly apply Kuhn's (2022) analysis to *dare-mo*, because NCIs are existential quantifiers under his analysis. Our solution augments Kuhn's (2022) analysis with *Hamblinian alternatives* (Hamblin 1973), which is independently motivated by the fact that *dare-mo* involves the indeterminate pronoun *who* (see Kratzer and Shimoyama 2017).

4 Proposal

Following Brasoveanu (2008), we adopt a logical object language with type t (truth values), type e (individuals) and type s (*states*), which models variable assignments. We take a *discourse referent* (dref) u_n as a function from states to individuals.[2] We abbreviate $\langle se\rangle$ (dref) as E, $\langle st, \langle st, t\rangle\rangle$, (*dynamic proposition*, i.e. a relation between PISs) as T and $\langle\langle st, \langle st, t\rangle\rangle, t\rangle$ (a set of dynamic propositions) as $\{T\}$. We adopt the PDL update rules in (7).

(7) a. $G[\![P(u_1)...(u_n)]\!]H$ iff $G = H$ & $\forall g \in G : \langle u_1(g), ..., u_n(g)\rangle \in I(P)$ (lexical relation)
b. $G[\![u = u']\!]H$ iff $G = H$ & $\forall g \in G : u(g) = u'(g)$ (dref identity)
c. $G[\![u = a]\!]H$ iff $G = H$ & $\forall g \in G : u(g) = a$ (value identity)
d. $G[\![\phi \wedge \psi]\!]H$ iff $\exists K : G[\![\phi]\!]K$ & $K[\![\psi]\!]H$ (conjunction)
e. $G[\![\neg\phi]\!]H$ iff $G = H$ & $\neg\exists K : H[\![\phi]\!]K$ (negation)
f. $G[\![\exists_u]\!]H$ iff $\forall g \in G : \exists h \in H : g[u]h$&$\forall h \in H : \exists g \in G : g[u]h$ (existential quantifier)
g. $G[\![\mathbf{0}_u]\!]H$ iff $\{d \mid d \neq \star$ & $\exists h \in H : d = u(h)\} = \emptyset$ (empty value condition)

We follow the semantic composition in Ciardelli, Roelofsen, and Theiler (2017). A sentence without a set-expanding item ends up denoting a dynamic proposition.

(8) a. $[\![\text{student}]\!] = \lambda v_E : \{\text{student}(v)\}$ / $[\![\text{participate}]\!] = \lambda v_E : \{\text{participate}(v)\}$
b. $[\![\text{a}]\!] = \lambda P_{\langle E\{T\}\rangle} : \lambda Q_{\langle E\{T\}\rangle} : \{\exists_u \wedge P(u) \wedge Q(u)\}$
c. $[\![\text{EC}]\!] = \lambda S_{\{T\}} : \cup S$ (existential closure)
d. $[\![\text{EC}]\!]([\![\text{A student participates}]\!]) = \cup\{\exists_u \wedge \text{student}(u) \wedge \text{participate}(u)\}$
$= \exists_u \wedge \text{student}(u) \wedge \text{participate}(u)$

Some expressions, e.g., *wh*-indefinites, expand alternatives, and the resultant sentential denotation is a non-singleton set of dynamic propositions.[3]

(9) a. $[\![\text{who}]\!] = \lambda Q_{\langle E\{T\}\rangle} : \{\exists_u \wedge \text{person}(u) \wedge Q(u) \wedge u = x \mid x \in D_e\}$
b. $[\![\text{Who participates}]\!] = \{\exists_u \wedge \text{person}(u) \wedge \text{participate}(u) \wedge u = x \mid x \in D_e\}$

We define universal quantification over alternatives as union-formation of PISs: $\Cup S$ produces H such that each dynamic proposition in S is true with respect to a subset of H.

(10) a. $G[\![\Cup S]\!]H$ iff $H = \cup\{K \mid \exists\phi \in S : G[\![\phi]\!]K\}$ (state aggregation)
b. $[\![\forall]\!] = \lambda S_{\{T\}} : \Cup S$ (universal closure)
c. $[\![\forall]\!]([\![\text{Who participates}]\!]) = \Cup\{\exists_u \wedge \text{person}(u) \wedge \text{participate}(u) \wedge u = x \mid x \in D_e\}$
where its output PIS is $\cup\{K \mid \exists x \in D_e : G[\![\exists_u \wedge \text{person}(u) \wedge \text{participate}(u) \wedge u = x]\!]K\}$

Lastly, we define negative morpheme and *taitei* (usually) in an alternative-sensitive way.[4] We take USUALLY as a placeholder for an analysis of dynamic adverbial quantification, which we put aside here. Any theory of it may work as long as it is defined for dynamic propositions.

[2] The idea is originally from Muskens (1996), but he takes a dref as primitives as well.

[3] One may use this non-singleton set of dynamic propositions to model questions (Li 2021), and one may indeed take our analysis as an extension of his theory. However, we remain neutral about how one should model questions.

[4] This choice is vital for us to define universal quantification over dynamic propositions that contain negation. The standard negation in Hamblin semantics takes the set complement of the prejacent, but this removes the structure of alternatives, which is the prerequisite for $\forall$.

(11) a. $[\![\text{not}]\!] = \lambda S_{\{T\}} : \{\neg p \mid p \in S\}$
b. $[\![\text{usually}]\!] = \lambda S_{\{T\}} : \{\textsc{usually}(p) \mid p \in S\}$

Coming back to Japanese, we define *dare-mo* with *wh*-indefinite, 0_u and $\forall$.

(12) $[\![\text{dare-mo}]\!] = \lambda Q_{\langle E\{T\}\rangle} : \Cup\{\exists_u \wedge \text{person}(u) \wedge Q(u) \wedge u = x \mid x \in D_e\} \wedge \mathbf{0}_u$ (*to be revised*)

We assume that *dare-mo* leaves the higher-order trace (Kuhn 2022), lifting its type. This achieves the split scope: $\Cup$ and $\mathbf{0}_u$ takes the highest scope while $\exists_u$ takes the narrowest scope.

(13) $[\![\text{dare-mo}]\!] =$
$\lambda R_{\langle\langle E\{T\},\{T\}\rangle,\{T\}\rangle} : \Cup R(\lambda Q_{\langle E\{T\}\rangle} : \{\exists_u \wedge \text{person}(u) \wedge Q(u) \wedge u = x \mid x \in D_e\}) \wedge \mathbf{0}_u$

Now, *dare-mo* leaves type $\langle E\{T\}, \{T\}\rangle$ trace. A simplified logical form is given below.

(14) $[_{\text{d}}$ `Dare-mo`^{u_1} $[_{\text{c}}$ λG_i $[_{\text{b}}$ `usually not` $[_{\text{a}}$ G_i `participate`]]]]

a. $G_i(\lambda v : \{\text{participate}(v)\})$
b. $[\![\text{usually}]\!]([\![\text{not}]\!](G_i(\lambda v : \{\text{participate}(v)\})))$
c. $\lambda G_{\langle E\{T\},\{T\}} : [\![\text{usually}]\!]([\![\text{not}]\!](G(\lambda v : \{\text{participate}(v)\})))$
d. $\Cup[\![\text{usually}]\!]([\![\text{not}]\!](\{\exists_{u_1} \wedge \text{person}(u_1) \wedge \text{participate}(u_1)) \wedge u_1 = x \mid x \in D_e\})) \wedge \mathbf{0}_{u_1}$
$= \Cup[\![\text{usually}]\!]\{\neg(\exists_{u_1} \wedge \text{person}(u_1) \wedge \text{participate}(u_1)) \wedge u_1 = x)) \mid x \in D_e\} \wedge \mathbf{0}_{u_1}$
$= \Cup\{\textsc{usually}(\neg(\exists_{u_1} \wedge \text{person}(u_1) \wedge \text{participate}(u_1)) \wedge u_1 = x)) \mid x \in D_e\} \wedge \mathbf{0}_{u_1}$

(14d) is informally paraphrased as (15).

(15) $\Cup\{\{K_1 \mid \text{Taro}^{u_1}$ usually does not participate in $K_1\}, \{K_2 \mid \text{Jiro}^{u_1}$ usually does not participate in $K_2\}, ...\} \wedge \mathbf{0}_{u_1}$

Crucially, the universal closure, i.e., PDL universal quantification over alternatives, takes scope over *taitei*, which creates $Q\neg_{NAA}$ with negation. Thus, the reading of (6) is correctly predicted. Since our analysis also assumes that an NCI takes split scope via QR, it correctly predicts the clause-boundedness of NCI licensing (1c). At the same time, negation is obligatory because of the empty value condition $\mathbf{0}_u$, which is evaluated at the end of dynamic conjunctions due to higher-order QR. Consider Figure 3 without negation and Figure 4 with negation.

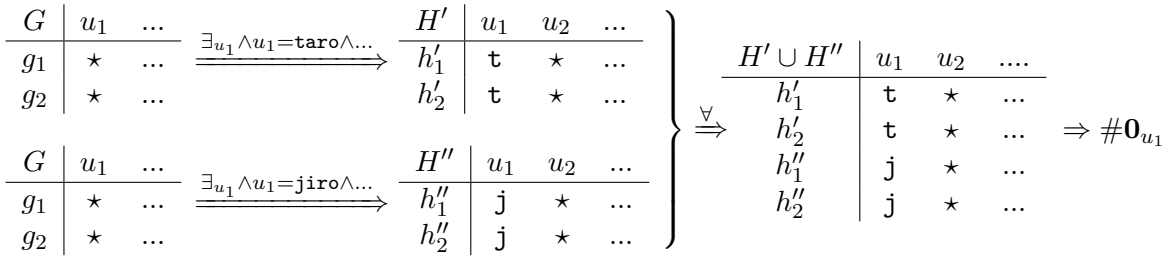


Figure 3: The dynamic effect of the ungrammatical use of universal NCI *dare-mo*

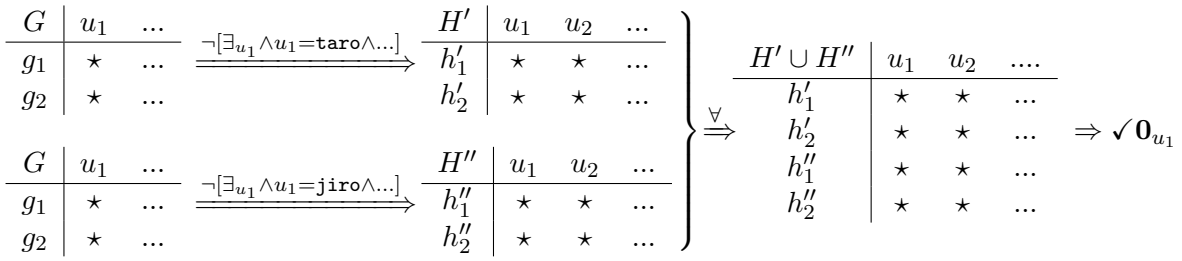


Figure 4: The dynamic effect of the grammatical use of universal NCI *dare-mo*

In Figure 3, the universal closure stores individuals who participated under the dref u_1, and

the empty value condition is not satisfied. In Figure 4, on the other hand, negation voids the dynamic effect of $\exists_{u_1}$ in each member of the set of dynamic propositions. Thus, no individual is stored under u_1 after the application of universal closure, and $\mathbf{0}_{u_1}$ is satisfied.

Notice that negation has to be applied to each member of the set of dynamic propositions. Otherwise, the cancellation of dynamic effect by negation collapses the structure of Hamblin alternatives, which is necessary to achieve (non-trivial) universal closure. Suppose that negation is defined for dynamic propositions and the alternatives expand after negation applies to the prejacent as shown in (16a). Then, one may not obtain (16b) from (16a). The value of u is assigned under negation, and thus, its value is empty outside the scope of negation. Thus, $u = x$ outside the scope of negation fails to expand the set of dynamic propositions, i.e., the value of u is $\star$ under any dynamic proposition, and (16b) is a singleton set of dynamic propositions.

(16) a. $\neg(\exists_u \wedge \text{person}(u) \wedge \text{participate}(v)))$
b. $\{\neg(\exists_u \wedge \text{person}(u) \wedge \text{participate}(v)) \wedge u = x \mid x \in D_e\}$

This poses a problem for an alternative in which the non-Hamblin system is default and ⟦dare-mo⟧ introduces Hamblin alternatives as shown in (17).

(17) $[\![\text{dare-mo}]\!] = \lambda R_{\langle\langle ET,T\rangle,T\rangle} : \Cup\{R(\lambda Q_{\langle ET\rangle} : \exists_u \wedge \text{person}(u) \wedge Q(u) \wedge u = x) \mid x \in D_e\} \wedge \mathbf{0}_u$

Because of its higher-order type, alternative-expansion takes split scope in (17), too, which leads to the problematic configuration in (18c) and (18d).

(18) $[_{\text{d}}$ `Dare-mo`^{u_1} $[_{\text{c}}$ λG_i $[_{\text{b}}$ `usually not` $[_{\text{a}}$ G_i `participate`]]]]
a. $G_i(\lambda v$:participate$(v))$
b. USUALLY$(\neg(G_i(\lambda v$:participate$(v))))$
c. $\lambda G_{\langle ET,T\rangle}$:USUALLY$(\neg(G(\lambda v$:participate$(v))))$
d. $\Cup\{\text{USUALLY}(\neg(\exists_{u_1} \wedge \text{person}(u_1) \wedge \text{participate}(u_1)) \wedge u_1 = x)) \mid x \in D_e\} \wedge \mathbf{0}_{u_1}$

Therefore, while the alternative (17) would be an attractive alternative, this runs into a problem in regard to the interaction between alternative expansion and dynamic negation.[5]

Acknowledgements. The first author is partially supported by Fulbright Grant. The second author would like to acknowledge support from the Dutch Research Council (NWO) as part of project 406.18.TW.009 *A Sentence Uttered Makes a World Appear - Natural Language Interpretation as Abductive Model Generation.* The third author is supported by the JSPS Grant-in-Aid for Early-Career Scientists, Grants Numbers JP22K13117.

[5]One may restore the value of a dref assigned under negation by enriching our analysis with bilateralism (Krahmer and Muskens 1995 among others), projective contents (Mandelkern 2022) or flat-update system (Hofmann 2022; Hofmann 2024). However, any successful theory still has to explain that anaphoric reference to a dref under negation is impossible from outside the scope of negation, which causes the problem observed in (16a) and (16b).

References

Brasoveanu, Adrian (2008). "Donkey pluralities: Plural information states versus non-atomic individuals". In: *Linguistics and Philosophy* 31.2, pp. 129–209. DOI: 10.1007/s10988-008-9035-0.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler (2017). "Composing alternatives". In: *Linguistics and Philosophy* 40.1, pp. 1–36. DOI: 10.1007/s10988-016-9195-2.

Giannakidou, Anastasia (2000). "Negative...Concord?" In: *Natural Language and Linguistic Theory* 18.3, pp. 457–523. DOI: 10.1023/a:1006477315705.

Hamblin, Charles (1973). "Questions in Montague English". In: *Foundations of Language* 10.1, pp. 41–53.

Hofmann, Lisa (2022). "Anaphora and negation". PhD thesis. University of California, Santa Cruz.

— (2024). "Anaphoric accessibility with flat update". Ms., University of Stuttgart.

Kamali, Beste and Hedde Zeijlstra (2024). "Negative Dependencies in Turkish". In: *Languages* 9.11, pp. 342–376. DOI: 10.3390/languages9110342.

Kataoka, Kiyoko (2006). *Nihongo hitei- bun-no koozoo: Kakimazebun-to hitei- kooo-hyoogen. (Syntactic structure of Japanese negative sentences: scram- bling construction and negation-sensitive elements)*. Tokyo: Korosio.

Kawashima, Ruriko (1994). "The structure of noun phrases and the interpretation of quantificational NPs in Japanese". PhD thesis. Cornell University.

Kishimoto, Hideki (2008). "On the variability of negative scope in Japanese". In: *Journal of Linguistics* 44.2, pp. 379–435. DOI: 10.1017/S0022226708005161.

Krahmer, Emiel and Reinhard Muskens (1995). "Negation and disjunction in discourse representation theory". In: *Journal of Semantics* 12.4, pp. 357–376. DOI: 10.1093/jos/12.4.357.

Kratzer, Angelika and Junko Shimoyama (2017). "Indeterminate pronouns: The view from Japanese". In: *Contrastiveness in Information Structure, Alternatives and Scalar Implicatures*. Ed. by Chungmin Lee, Ferenc Kiefer, and Manfred Krifka. Springer, pp. 123–143. DOI: 10.1007/978-3-319-10106-4_7.

Kuhn, Jeremy (2022). "The dynamics of negative concord". In: *Linguistics and Philosophy* 45.1, pp. 153–198. DOI: 10.1007/s10988-020-09318-3.

Li, Haoze (2021). "Mandarin *wh*-conditionals: A dynamic question approach". In: *Natural Language Semantics* 29.3, pp. 401–451. DOI: 10.1007/s11050-021-09180-4.

Mandelkern, Matthew (2022). "Trivalence and anaphora". Ms., New York University.

Muskens, Reinhard (1996). "Combining Montague semantics and discourse representation". In: *Linguistics and Philosophy* 19.2, pp. 143–186. DOI: 10.1007/BF00635836.

Shimoyama, Junko (2011). "Japanese indeterminate negative polarity items and their scope". In: *Journal of Semantics* 28.4, pp. 413–450. DOI: 10.1093/jos/ffr004.

van den Berg, Martin (1996). "Some aspects of the internal structure of discourse. The dynamics of nominal anaphora". PhD thesis. University of Amsterdam.

Watanabe, Akira (2004). "The genesis of negative concord: Syntax and morphology of negative doubling". In: *Linguistic Inquiry* 35.4, pp. 559–612. DOI: 10.1162/0024389042350497.

Testing theories of SDA via acquisition

Eleonora Zani
Goethe-University Frankfurt
Franfurt am Main, Germany
zani@lingua.uni-frankfurt.de

Ivano Ciardelli
University of Padua
Padova, Italy
ivano.ciardelli@unipd.it

Emanuela Sanfelici
University of Padua
Padova, Italy
emanuela.sanfelici@unipd.it

Abstract

The aim of this study is to examine how children and adults interpret indicative and counterfactual conditionals with disjunctive antecedents, such as "if A or B, C". Disjunctive antecedent conditionals (DACs) are associated with a debated inference known as the *simplification of disjunctive antecedents* (SDA), where a DAC is interpreted as equivalent to the conjunction of its *simplifications*: "if A, C" and "if B, C".

To explore this, we administered a picture-supported Truth Value Judgment Task to 140 children (aged 5;0–9;11) and 28 adults. Participants were asked to evaluate the truth or falsity of 16 DACs in both indicative and counterfactual mode, with position effects controlled (e.g., testing both "if A or B" and "if B or A"). Our results reveal that the SDA interpretation emerges as early as age 5. Moreover, a growing preference for the SDA interpretation is associated with a decreasing tendency to adopt an alternative interpretation, in which a DAC is analyzed as the disjunction of the two simplifications.

1 Introduction

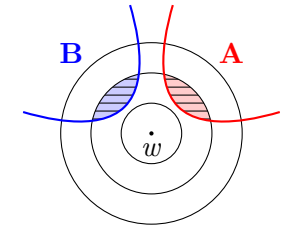
The present study examines *disjunctive antecedent conditionals* (henceforth, DACs), which are conditionals of the form "If A or B, C". These structures are commonly interpreted as equivalent to the conjunction of their *simplifications*, "If A, C" and "If B, C". This interpretation is known as *SDA interpretation*, which stands for *simplification of disjunctive antecedents*. Consider the following example:

(1) a. If it had rained or snowed, the hike would have been cancelled.
b. If it had rained, the hike would have been cancelled, **and** if it had snowed, the hike would have been cancelled.

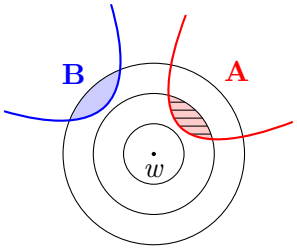
Sentence (1-a) intuitively implies both "if it had rained, the hike would have been cancelled" and "if it had snowed, the hike would have been cancelled", as expressed in (1-b). However, this intuition stands in contrast to the predictions of the classical framework for counterfactuals, known as *Minimal Change Semantics* (Lewis, 1973). This framework introduces a key parameter in the evaluation of counterfactual conditionals: *comparative similarity*. Given an evaluation world w, this parameter ranks possible worlds according to their similarity to w.

Assuming that an A-world is a world in which the proposition A is true and the set $\min_w(\mathsf{A})$ is a subset of the set of A-worlds which differ minimally from w, a counterfactual conditional statement "if A, C" is judged true at w only if all closest A-worlds are also worlds where the consequent C holds. For a DAC such as "if A or B, C", this framework yields different interpretations depending on the configuration of the two disjuncts:

- If the closest A-worlds and B-worlds are equally realistic relative to w, the SDA interpretation is predicted, as we consider some A-worlds and some B-worlds and require that C holds in both (as in Figure 1a). The red area represents the closest A-worlds ($\min_w(\mathsf{A})$) relative to w, and the blue area the closest B-worlds ($\min_w(\mathsf{B})$). The regions with horizontal lined patterns depict the worlds that are closest to w within the union $\mathsf{A} \cup \mathsf{B}$.
- Conversely, if, for instance, A-worlds are considered more realistic than B-worlds, the selection function driven by comparative similarity will favor A-worlds, disregarding B-worlds. Thus, the evaluation reduces to assessing "if A, C", as B-worlds are not considered (as illustrated in Figure 1b). We will refer to this interpretation as the *asymmetric reading* (AR).



(a) Case 1



(b) Case 2

Figure 1

Nevertheless, even when judging (1-a) with respect to an evaluation world set in summer and assuming that we enjoy hiking even when it rains but that snow is too dangerous, it appears that we are considering both disjunct worlds despite the fact that snow-worlds are typically perceived as less realistic than rain-worlds during this season.

To explain the robustness of SDA, a number of accounts have been proposed. Semantic accounts, pioneered by Alonso-Ovalle (2009) (also Ciardelli, 2016; Ciardelli et al., 2018; Santorio, 2018; Khoo, 2018; Cariani and Goldstein, 2020), exploit the idea that a disjunctive antecedent "if A or B" is associated with a set of two propositions, $\{|\mathsf{A}|, |\mathsf{B}|\}$ (a result obtained using *alternative*, *inquisitive*, or *truth-maker semantics*), and that the truth of the conditional requires the consequent to follow on the supposition of each proposition in the set. To this, some accounts (Santorio, 2018; Cariani and Goldstein, 2020) add an assumption of *homogeneity*: when the antecedent is associated with multiple propositions, the conditional is *true* if all of them imply the truth of the consequent, *false* if all of them imply the falsity of the consequent, and lacks a truth-value in mixed cases. A different, pragmatic account (Bar-Lev and Fox, 2020) views SDA as an implicature, arising through a grammatically encoded, non-Gricean exhaustification process. Further, most authors regard SDA as connected to the phenomenon of *free-choice inferences*, whereby from "you may A or B" one infers both "you may A" and "you may B". In this case, too, the inference is not predicted by standard modal semantics, and the same ideas discussed above have been invoked to explain it. As our discussion indicates, the question about the nature and source of SDA is tied to some important theoretical debates in semantics: SDA is taken by some to motivate a refinement of intensional semantics, by others to provide evidence for a particular kind of non-Gricean strengthening recipe. To shed light on this key question, we conducted an empirical study.

2 Experiment

2.1 Research questions

We now present the empirical questions addressed in our study and discuss their implications for the theoretical debate.

1. *At what age does SDA first arise?*
   Recent experimental works (Zhou et al., 2013; Tieu, Romoli, et al., 2015; Huang and Crain, 2019) have found that 5-year-old children can compute free-choice inferences. If SDA and free-choice are manifestations of the same phenomenon, as widely assumed, we would expect the SDA interpretation to emerge early as well.

2. *Do children shift from AR to SDA as they grow?*
   According to Bar-Lev and Fox 2020, SDA results from the pragmatic strengthening of a sentence whose literal meaning follows the Lewisian AR.

   If the relevant strengthening mechanism is not yet universally acquired at age 5, we expect to see some proportion of AR readings in younger children, decreasing with age in favor of an SDA interpretation. While the absence of such a trend would not necessarily be evidence against the account of Bar-Lev and Fox, it would imply that the pragmatic abilities involved in deriving SDA are universally present in children already at age 5.

3. *Do children shift from a Disjunctive Conditional Reading (DCR) to SDA?*
   According to Santorio (2018) and Cariani and Goldstein (2020), DACs involve a homogeneous quantification over the alternatives introduced by the disjunction, similar to the one commonly associated with plural definites (Löbner, 2000; Križ, 2019). Recent experimental work (Tieu, Križ, et al., 2019) has shown that a notable number of younger children display an existential reading of plural definites, which shifts to a universal reading with age. If DACs involve homogeneity, we may expect a similar developmental pattern: younger children might exhibit an existential reading of DACs, accepting them when *some* disjunct implies the consequent. This interpretation, called the *disjunctive conditional reading* (DCR), treats DACs as the *disjunction* of their simplifications. Over time, this reading may give way to a universal interpretation, equivalent to SDA, where *every* disjunct implies the consequent. This shift is not predicted by other accounts.

4. *Does SDA arise earlier or more strongly for indicatives than for counterfactuals?*
   Some authors (e.g., Stalnaker, 1968; Adams, 1975; Khoo, 2022) support a uniform view of conditionals, while others (e.g., Veltman, 1985; Gauker, 2005) argue that indicatives and counterfactuals follow different logics: indicatives are strict, validating SDA, whereas counterfactuals are variably strict, potentially invalidating SDA. If SDA appears earlier or more robustly in indicatives than in counterfactuals, this would support the latter view. Conversely, a parallel emergence of SDA across both types of conditionals would suggest a unified logical status for SDA.

2.2 Methodology

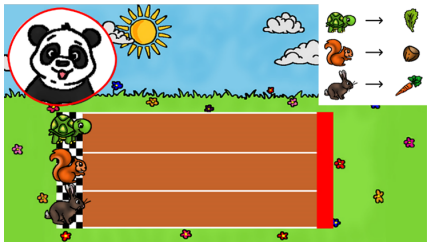
A total of 140 Italian monolingual children, aged between 5;0 and 9;11, and 28 adults participated in a picture-supported Truth Value Judgment Task. Participants were asked to evaluate statements made by Panda, a character who occasionally lies. The task included 4 pre-test items, 16 DACs, 16 *control* items featuring simple conditionals (e.g., "if A, C"; 8 requiring a true judgment and 8 requiring a false judgment), and 4 *bonus* items with conditionals containing negated antecedents (e.g., "if not A, B"). Eligibility criteria were as follows: participants had to correctly complete the pre-test, which assessed task understanding, interpretation of simple present indicative conditionals, and comprehension of disjunctions in main clauses; correctly judge at least 14 out of 16 control items; and complete the test without interruptions. The experiment involved four race scenarios, each featuring three animals of different speeds (e.g., tortoise, squirrel, hare) competing for unique prizes (e.g., lettuce leaf, hazelnut, carrot). To avoid the possibility that children might mistakenly interpret the disjunction as a conjunction (Singh et al., 2016; Tieu, Yatsushiro, et al., 2016), only one animal could win each race.

In addition to the pre-test items, the mode was manipulated across two sets of scenarios: 2 scenarios (18 items) presented in indicative mode (see item (1) for an example of a DAC in indicative mode) and 2 scenarios (18 items) presented in counterfactual mode (see item (2) for an example of a DAC in counterfactual mode).

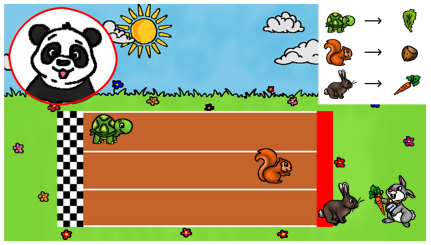
(1) Se lo scoiattolo o la tartaruga vincerà la gara, avrà in premio una nocciola.
If the squirrel or the tortoise win.FUT.3SG the race, get.FUT.3SG as prize a hazelnut.
If the squirrel or the tortoise wins the race, it will get a hazelnut.

(2) Se lo scoiattolo o la tartaruga avesse vinto la gara, avrebbe avuto in premio una nocciola.
If the squirrel or the tortoise have.SBJV.IPFV.3SG win.PTCP.PST the race, have.COND.3SG get.PTCP.PST as prize a hazelnut.
If the squirrel or the tortoise had won the race, it would have gotten a hazelnut.

Two of the four scenarios displayed all competitors at the starting line (Figure 2a), involving indicative conditionals. The other two scenarios showed the end of the race, with the fastest animal winning (Figure 2b), and involved counterfactuals. Participants were divided into two groups: List 1 judged the first and third scenarios in indicative mode and the second and fourth in counterfactual mode, while List 2 received the reverse order. Each scenario contained 4 DACs, 4 control items, and a bonus item, presented in that order. To avoid facilitating access to the simplifications of the DACs, control sentences were presented after the target items. We manipulated position in DACs at a design level: both "if A or B" and "if B or A" were tested to ensure a balanced evaluation of the disjuncts.



(a) Indicative scenarios



(b) Counterfactual scenario

Figure 2: Comparison of indicative and counterfactual scenarios in the experiment.

The interpretation of the DACs was determined by the pattern that emerged from the four target items presented in each scenario to each participant. The SDA pattern requires a false judgment ("F") of all four items, as the consequent must hold for each disjunct in the antecedent. The DCR interpretation, on the other hand, arises when all target items are judged as truth ("T"), since it only requires that at least one of the two disjuncts follow the consequent. An AR interpretation arises only if the most realistic disjunct (in this case, the squirrel) is linked to the prize in the consequent.

Table 1: Interpretation patterns of *target sentences*.

	SDA if A, C **and** if B, C	DCR if A, C **or** if B, C	AR if A, C*
(1) if **Squirrel** or Tortoise, **Hazelnut**	F	T	T
(2) if Tortoise or **Squirrel**, **Hazelnut**	F	T	T
(3) if Squirrel or **Tortoise**, **Lettuce**	F	T	F
(4) if **Tortoise** or Squirrel, **Lettuce**	F	T	F

* Assuming A is regarded as the most realistic disjunct.

Note: In the antecedent, the unique competitor among the two disjuncts correctly associated with the prize in the consequent is highlighted in bold, as is the prize itself.

Additionally, each scenario included a *bonus item* (e.g., 3), designed to assess participants' evaluations of the relative closeness of the disjuncts in a DAC. A *true* response indicated that the participant regarded one disjunct in the target DAC as closer than the other.

(3) Se non vincerà/avesse vinto la lepre, vincerà/avrebbe vinto lo scoiattolo.
If the hare doesn't win/hadn't won, the squirrel will win/would have won.

2.3 Results

Figure 3 presents a bar chart illustrating the distribution of interpretative patterns derived by participants across different scenarios. The data is segmented by age group, revealing a significant proportion of SDA patterns starting at age 5 (33.0%). Data were fitted to a Generalised Mixed Model (multinomial)[1]. The dependent variable was "Pattern", consisting of four values: SDA, AR, DCR and *Others*. The predictors were mode (indicative, counterfactuals), Age (in months) and their interaction. The model shows that SDA is derived more than AR, DCR and Others. The unique effect which is significant either including adults (Table 2) or excluding from the sample (Table 3) is Age between DCR and SDA (p = 0.002 including adults; p = 0.013 excluding them).

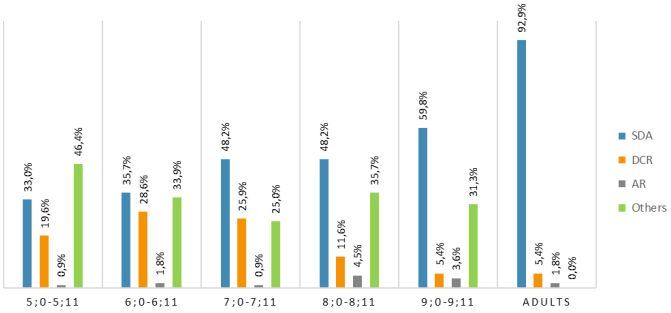


Figure 3: Interpretations of DACs among age groups

Response	Names	Effect	Estimate	SE	Exp(B)	95% Confidence Intervals		z	p
						Lower	Upper		
AR - SDA	(Intercept)	(Intercept)	-3.474	0.484	0.031	0.012	0.080	-7.179	< .001
AR - SDA	Mode	Ind. - Ctf.	-0.729	0.828	0.482	0.095	2.444	-0.880	0.379
AR - SDA	Age	Age	-0.013	0.011	0.987	0.967	1.008	-1.194	0.233
AR - SDA	Mode * Age	Ind. - Ctf.	-0.013	0.017	0.987	0.955	1.021	-0.759	0.448
DCR - SDA	(Intercept)	(Intercept)	-2.745	0.461	0.064	0.026	0.159	-5.958	< .001
DCR - SDA	Mode	Ind. - Ctf.	0.843	0.673	2.323	0.621	8.683	1.253	0.210
DCR - SDA	Age	Age	-0.038	0.012	0.963	0.940	0.986	-3.155	0.002
DCR - SDA	Mode * Age	Ind. - Ctf.	0.020	0.015	1.020	0.990	1.050	1.291	0.197
Others - SDA	(Intercept)	(Intercept)	-1.443	0.341	0.236	0.121	0.461	-4.234	< .001
Others - SDA	Mode	Ind. - Ctf.	0.260	0.537	1.297	0.453	3.713	0.484	0.628
Others - SDA	Age	Age	-0.030	0.010	0.971	0.952	0.989	-3.082	0.002
Others - SDA	Mode * Age	Ind. - Ctf.	0.012	0.013	1.012	0.987	1.038	0.946	0.344

Table 2: Parameter Estimates (Fixed Coefficients) - children ad adults. (Full Model Summary: AIC = 989.743; BIC = 1097.989; LRT X^2 = 921.437; Dev = 941.743)

Response	Names	Effect	Estimate	SE	Exp(B)	95% Confidence Intervals		z	p
						Lower	Upper		
AR - SDA	(Intercept)	(Intercept)	-3.109	0.394	0.045	0.021	0.097	-7.892	< .001
AR - SDA	Mode1	Ind. - Ctf.	-0.238	0.718	0.788	0.193	3.217	-0.331	0.740
AR - SDA	Age	Age	-0.002	0.027	0.998	0.947	1.052	-0.071	0.944
AR - SDA	Mode * Age	Ind. - Ctf.	-0.023	0.043	0.978	0.899	1.063	-0.529	0.597
DCR - SDA	(Intercept)	(Intercept)	-0.800	0.214	0.449	0.295	0.683	-3.743	< .001
DCR - SDA	Mode	Ind. - Ctf.	0.079	0.362	1.082	0.532	2.200	0.217	0.828
DCR - SDA	Age	Age	-0.045	0.018	0.956	0.922	0.991	-2.485	0.013
DCR - SDA	Mode * Age	Ind. - Ctf.	0.020	0.022	1.020	0.977	1.065	0.906	0.365
Others - SDA	(Intercept)	(Intercept)	-0.516	0.208	0.597	0.397	0.898	-2.478	0.013
Others - SDA	Mode	Ind. - Ctf.	-0.171	0.334	0.843	0.438	1.623	-0.511	0.609
Others - SDA	Age	Age	-0.028	0.018	0.973	0.940	1.007	-1.582	0.114
Others - SDA	Mode * Age	Ind. - Ctf.	0.016	0.017	1.016	0.983	1.051	0.955	0.340

Table 3: Parameter Estimates (Fixed Coefficients) - only children. (Full Model Summary: AIC = 1009.269; BIC = 1113.139; LRT X^2 = 591.381; Dev = 961.269)

[1]Model: mclogit::mblogit Pattern ~ 1 + Mode + Age + Mode:Age + (1 | Scenario_List) + (0 + Age | ID participant). "Scenario_List" is the scenario combined with the list assigned:i.e. *A.1* is the scenario A of List 1.

The absence of a mode effect is further supported by an individual analysis. The two tables below provide a detailed examination of each participant's responses (see Table 4 for children and Table 5 for adults). These tables illustrate whether participants exhibit consistency across both scenarios presented in the indicative mode and both scenarios in the counterfactual mode. In addition to the four response categories previously discussed—SDA, DCR, AR, and *Others*—a fifth category, "Mix", has been introduced. This category captures instances in which a participant derives two different interpretations within scenarios of the same mode. Upon excluding the "Mix" column and row, it is notable that all but one participants are positioned along the diagonal of the tables, thereby confirming their tendency to generate consistent interpretations across modes, irrespective of the specific interpretation chosen.

Counterfactual \ *Indicative*	**SDA**	**DCR**	**AR**	**Others**	**Mix**
SDA	**50**	0	0	0	3
DCR	0	**14**	0	0	3
AR	0	0	**2**	0	0
Others	0	0	0	**28**	7
Mix	8	0	0	9	8

Table 4: Children's profiles (total: 140)

Counterfactual \ *Indicative*	**SDA**	**DCR**	**AR**	**Others**	**Mix**
SDA	**26**	0	0	0	0
DCR	0	**1**	0	0	0
AR	0	**1**	**0**	0	0
Others	0	0	0	**0**	0
Mix	0	0	0	0	0

Table 5: Adults' profiles (total: 28)

We now turn to the analysis of the results related to the bonus item. The inclusion of this item was primarily motivated by the need to examine whether an equally realistic condition is a necessary prerequisite for SDA. The data in Table 6 have been filtered to include only those participants who answered "T" for the corresponding bonus item within the same scenario. It is important to note that answering "T" corresponds to the computation of a non-equally realistic condition, which, according to Minimal Change Semantics, leads to AR. However, despite this, SDA remains the preferred interpretation among these participants, indicating that SDA does not depend exclusively on equidistant conditions.

Pattern	**Children**	**Adults**
SDA	41.1%	87.8%
AR	3.0%	4.1%
DCR	20.8%	8.2%
Others	35.1%	0.0%
Total	100.00%	100.00%

Table 6: Interpretative patterns of participants who answered "T" to the bonus item

3 Conclusion

Our results indicate that the SDA interpretation of DACs emerges early, with a preference for this interpretation evident as early as age 5. This supports the view that SDA is linked to free-choice inferences, which also develop early (Zhou et al., 2013; Tieu, Romoli, et al., 2015; Huang and Crain, 2019). While SDA rates increase with age, this increase is not accompanied by a decrease in AR rates; in fact, AR is nearly unattested, even among participants whose answer to the bonus item indicated that they regarded one disjunct as strictly closer than the other. On the account of Bar-Lev and Fox, the absence of AR can be explained by assuming that the strengthening mechanism responsible for SDA is fully acquired at age 5; this, however, leaves the subsequent increase in SDA unexplained. This increase is accompanied by a decrease in DCR rates, in line with the predictions of homogeneity-based accounts (Santorio, 2018; Cariani and Goldstein, 2020), that can ascribe this trend to the independently attested shift from existential to universal interpretations of homogeneous predication. Finally, SDA rates are not significantly different between indicatives and counterfactuals, in line with the predictions of uniform accounts of conditionals. Future research could test the hypothesis that the SDA interpretation is linked to homogeneity by testing the same children on both plural definites and DACs and comparing their answers.

References

Adams, Ernest (1975). *The logic of conditionals: An application of probability to deductive logic*. Vol. 86. Springer Science & Business Media.

Alonso-Ovalle, Luis (2009). "Counterfactuals, Correlatives, and Disjunction". In: *Linguistics and Philosophy* 32 (2), pp. 207–244.

Bar-Lev, Moshe E and Danny Fox (2020). "Free choice, simplification, and innocent inclusion". In: *Natural Language Semantics* 28.3, pp. 175–223.

Cariani, Fabrizio and Simon Goldstein (2020). "Conditional heresies". In: *Philosophy and Phenomenological Research* 101.2, pp. 251–282.

Ciardelli, Ivano (2016). "Lifting conditionals to inquisitive semantics". In: *Semantics and Linguistic Theory (SALT) 26*. Ed. by Mary Moroney et al. Ithaca, NY: LSA and CLC Publications.

Ciardelli, Ivano et al. (2018). "Two switches in the theory of counterfactuals". In: *Linguistics and Philosophy* 41.6, pp. 577–621. ISSN: 1573-0549. DOI: 10.1007/s10988-018-9232-4. URL: 10.1007/s10988-018-9232-4.

Gauker, Christopher (2005). *Conditionals in Context*. MIT Press, Cambridge, MA.

Huang, Haiquan and Stephen Crain (Aug. 2019). "When OR is assigned a conjunctive inference in child language". In: *Language Acquisition* 27, pp. 1–24. DOI: 10.1080/10489223.2019.1659273.

Khoo, Justin (2018). "Disjunctive antecedent conditionals". In: *Synthese* 198.8, pp. 7401–7430. ISSN: 1573-0964. DOI: 10.1007/s11229-018-1877-6.

— (2022). *The meaning of* if. Oxford University Press.

Križ, Manuel (2019). "Homogeneity effects in natural language semantics". In: *Language and Linguistics Compass* 13.11. e12350 10.1111/lnc3.12350, e12350. DOI: https://doi.org/10.1111/lnc3.12350. eprint: https://compass.onlinelibrary.wiley.com/doi/pdf/10.1111/lnc3.12350. URL: https://compass.onlinelibrary.wiley.com/doi/abs/10.1111/lnc3.12350.

Lewis, David (1973). *Counterfactuals*. Blackwell.

Löbner, Sebastian (2000). "Polarity in natural language: Predication, quantification and negation in particular and characterizing sentences". In: *Linguistics and philosophy* 23, pp. 213–308.

Santorio, Paolo (2018). "Alternatives and truthmakers in conditional semantics". In: *The Journal of Philosophy* 115.10, pp. 513–549.

Singh, Raj et al. (2016). "Children interpret disjunction as conjunction: Consequences for theories of implicature and child development". In: *Natural Language Semantics* 24.4, pp. 305–352.

Stalnaker, Robert (1968). "A theory of conditionals". In: *Studies in Logical Theory*. Ed. by N. Rescher. Oxford: Blackwell.

Tieu, Lyn, Manuel Križ, et al. (2019). "Children's acquisition of homogeneity in plural definite descriptions". In: *Frontiers in Psychology* 10, p. 463805.

Tieu, Lyn, Jacopo Romoli, et al. (2015). "Children's Knowledge of Free Choice Inferences and Scalar Implicatures". In: *Journal of Semantics* 33.2, pp. 269–298.

Tieu, Lyn, K. Yatsushiro, et al. (2016). "On the Role of Alternatives in the Acquisition of Simple and Complex Disjunctions in French and Japanese". In: *Journal of Semantics* 34.1, pp. 127–152.

Veltman, Frank (1985). "Logics for Conditionals". PhD thesis. University of Amsterdam.

Zhou, Peng et al. (2013). "Children's knowledge of free choice inferences". In: *Semantics and Linguistic Theory* 23, pp. 632–651. URL: https://api.semanticscholar.org/CorpusID:14043324.

An experimental assessment of the nall lexical gap

Ruizhe Zhou
École Normale Supérieure-PSL
Paris, France
ruizhezhou1999@gmail.com

Benjamin Spector
École Normale Supérieure-PSL
Paris, France
spector.benjamin@gmail.com

Mora Maldonado
Laboratoire de Linguistique de Nantes, CNRS
Nantes, France
mora.maldonado@univ-nantes.fr

Abstract

Across the world's languages, the quantificational concept *not all* (*nall*) is never lexicalized. This also holds for temporal (*not always*), modal (*not must*) and locative (*not everywhere*) domains. Existing accounts for these lexical gaps fall into two broad categories. According to Cognitive Markedness accounts, *nall* is not lexicalized because it expresses a cognitively unnatural concept. According to Communicative Efficiency accounts, *nall* is not intrinsically marked, but is not very useful for efficient communication. Across two artificial language learning experiments, we found no cognitive bias against *nall*, suggesting that its absence from the lexicon cannot be attributed to cognitive markedness.

1 Introduction

Despite substantial variation, human languages exhibit striking regularities in their lexica, including gaps where logically possible meanings lack a dedicated lexical form (Horn 1972). For example, a quantifier meaning *not all* (henceforth, *nall*) is not attested across natural languages as a single lexical unit, whereas quantifiers meaning *all*, *some* (i.e., *not none*) and *none* are. This observation holds not only for quantification over individuals, but also for quantification over times (*not always*), locations (*not everywhere*), and possible worlds (*not must*). The source of this lexical gap is not immediately clear: a quantifier meaning *nall* would satisfy all of the formal constraints that universally hold for mono-morphemic quantifiers (e.g., conservativity and monotonicity; Barwise and Cooper 1981) and has the same computational complexity as *some* measured in terms of semantic automata (Steinert-Threlkeld and Szymanik 2019). As noticed by Horn (1972), both *nall* and *some* are, in most contexts, typically interpreted as meaning 'some but not all' (due to pragmatic strengthening), which renders one of them redundant. The question is then why it is systematically *some* rather than *nall* that is lexicalized.

Two types of explanations have been proposed to account for this typological gap. Under the first type of account, *nall* is not lexicalized because it expresses a cognitively unnatural concept (Horn 1972; Katzir and Singh 2013). We will refer to these as the Cognitive Markedness accounts. Arguments supporting these accounts have been based on evidence that quantificational expressions involving negative meanings are harder to process and reason about than those involving positive meanings (Geurts and Der Slik 2005). However, in all these studies, the negative expressions tend to also be morphosyntactically more complex than their positive alternatives, which makes it unclear whether the observed difficulty is due to their semantic or to their structural complexity.

In contrast, recent alternative accounts have been proposed to not rely on cognitive markedness but instead explain the lexical gap by appealing to communicative pressures (Enguehard and Spector 2021; Uegaki 2023; Bar-Lev and Katzir 2023). We will group these as Communicative

Efficiency-based accounts. Very generally, these accounts argue that *nall* is not lexicalized because it has low communicative utility in real-world situations. In the specific case of Enguehard and Spector (2021), the *nall* lexical gap is thought to come from the fact that statements involving *nall* are very uninformative, given the frequencies of different situations in the actual world. Imagine a room where no object is red versus another room where all objects are red. Intuitively, the former is more likely than the latter. If this is indeed the case, a *some* statement (which negates *no*) is comparatively more surprising and thus more informative than a *nall* statement (which negates *all*). The absence of a word for *nall* can therefore be explained by the fact that *nall* would convey, most of the time, less information than *some*, without the assumption that its meaning is cognitively marked.

To investigate the role of cognitive markedness in the *nall* lexical gap, we draw inspiration from studies that used artificial language learning and rule learning experiments to test similar markedness-based accounts for a number of lexical constraints in different domains (e.g., quantification, person systems, clausal embedding predicates; Chemla, Buccola, and Dautriche 2019; Maldonado and Culbertson 2022; Maldonado, Culbertson, and Uegaki 2022). The underlying idea behind these studies is that words or rules involving cognitively unnatural meanings should be harder to learn than those involving natural ones. In this study, we conduct two artificial language learning experiments to test whether learners are more likely to infer that a novel logical word means *some* as opposed to *nall*. An account based on cognitive markedness predicts that learners should prefer to lexicalize meanings akin to *some* rather than meanings akin to *nall*, in line with the typological generalization. Instead, accounts based on communicative efficiency would predict learners to be as likely to adopt a system that lexicalizes *some* as one that lexicalizes *nall*.

2 Methods

We used an artificial language learning paradigm, where learners are trained on input that is in principle ambiguous, and must generalize beyond this input in a way that is compatible with only one of the hypotheses of interest (aka extrapolation paradigm, Culbertson and Adger 2014). Learners were first taught two novel words that described colored shapes: one was compatible with *all* scenarios, and the other was compatible with *none* situations. Crucially, participants were not told which word to use in a 'mixed', *some-but-not-all* scenarios (henceforth, *sbna* scenarios). At test, learners had to decide which of the two novel words they would use to describe these *sbna* situations. If participants prefer to have a word that means *some*, they will use the same word for both *all* and *sbna* scenarios. Conversely, if they prefer to lexicalize the meaning *not all*, they will use the same word for both *no* and *sbna* scenarios. We manipulated whether novel words were adjectives or verbs. The motivation for using two different grammatical categories and sentence types is to have some confidence that whatever effect we find is not specifically due to the fact that the particular syntactic construction used might favor a certain interpretation of the target word (cf. Discussion).

2.1 Materials

In the Adjective condition, participants were exposed to sentences like *'The redness of the triangle is narp'* and the structure [The <COLOR>-ness of the <SHAPE> is <NOVEL ADJECTIVE>] was used. In the Verb condition, sentences like *'The circle pakes blue'* which follow the structure [The <SHAPE> <NOVEL VERB> <COLOR>]. Shapes could be circles, triangles, squares, and stars, and colors could be green, red, yellow, purple, red and orange.

Sentences always appeared paired with pictures that could make them either true or false. Pictures could instantiate three different situations: (i) *all* situations (where the shape was

entirely filled with the color specified by the sentence), (ii) *none* situations (where the shape was filled with an incorrect color), and (iii) 'mixed', *some-but-not-all* situations (where the shape was half-filled with the color mentioned in the sentence and half-filled with another color).

In both conditions, the two novel words were randomly selected for each participant from a list of four nonce words that followed English phonotactic rules for the given grammatical category (Novel adjectives included the words 'sluck', 'zeck', 'teeb' and 'narp'; novel verbs included the words 'fick', 'pake', 'feep' and 'ralk').

2.2 Procedure

There were two training phases and a test phase. The first training phase consisted of a sentence-picture matching task, as exemplified in Figure 1a. Participants encounter a sentence involving one of the two novel words and had to choose between two pictures (a *all* and a *none* picture) the one that was best described the sentence. Through feedback, they learned that one of the target words was compatible with an *all* situation (e.g., 'narp' in Fig. 1a) while the other word was compatible with a *none* situation (e.g., 'sluck' in Fig. 1a). There were 16 training trials (8 for each word).

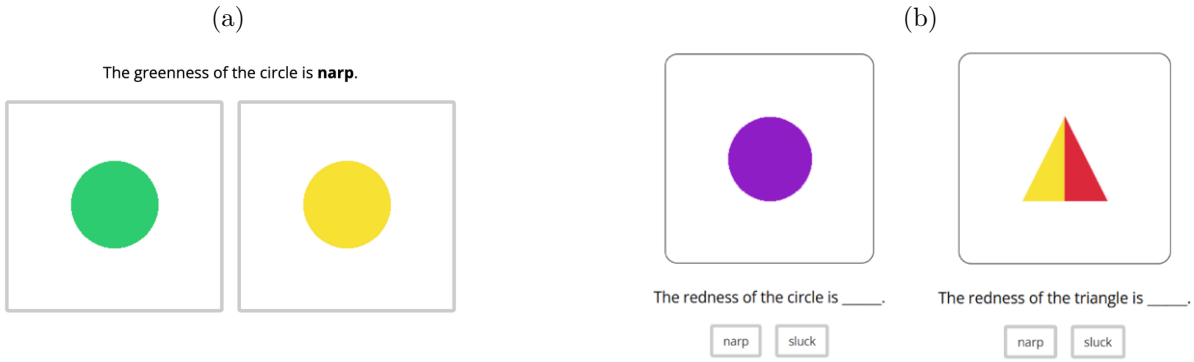


Figure 1: **Procedure (Adjective)**. (a) Sentence-picture matching task. If the word 'narp' was compatible with *all* situations, participants should select the image on the left. Otherwise, they should go for the image on the right. (b) Sentence-completion task. Left-picture is a **seen trial** involving a *none* situation, as the circle is not red. Right-picture is a **target trial** involving a *snba* situation, as some part of the triangle is red.

The second training phase involved a sentence-completion task, where participants had to use one of the novel words to complete a description of a picture, as in Figure 1b. Pictures always involved *all* and *none* situations. There were 16 training trials in this phase (8 per situation). Participants received feedback on their answers.

The testing phase also involved a sentence-completion task, but this time no feedback was provided. This phase included trials involving *all* and *none* situations, similar to those used during the second training phase (henceforth, *seen* trials). Crucially, we also included *target* trials involving a *some-but-not-all* (*sbna*) scenario, where the mentioned shapes were half-filled with the mentioned color and half-filled with another color (see Fig. 1b). There were 16 seen trials and 8 target trials.

We predict that if *nall* is cognitively marked, as proposed by Cognitive Markedness accounts, participants would prefer to have a word that is compatible with both *all* and *sbna* scenarios, similarly to the word 'some' in English. More specifically, we would expect that learners prefer a system that lexicalizes a *some* and a *none* meaning, rather than one that lexicalizes a *nall* and an *all* meaning. In our experiment, this preference would manifest as participants selecting the word that they learned for *all* meanings in the new *sbna* situations during testing. This preference

would be expected to hold regardless of the grammatical category of the novel words. However, if, contrary to Cognitive Markedness accounts, *nall* is not cognitively marked, participants should show no such preference in either experiment, essentially choosing one of the two words by chance.

2.3 Participants

121 English-speaking adults were recruited through Prolific (61 for the Adjective condition and 60 for the Verb condition). Participants were paid 1.4 pounds for taking part in the experiment, which lasted approximately 10 minutes. As per our preregistration (`https://osf.io/eyms8`), participants were excluded if they completed fewer than 11 out of 16 correct trials for either *all* or *none* situations during sentence-completion tasks in training and testing (16 trials per situation). This was meant to ensure that participants had correctly learned the situations compatible with each of the words. This resulted in the analysis of 53 participants in the Adjective condition and 57 participants in the Verb condition.

3 Results

Recall that in the experiment participants are initially taught two novel words: one compatible with *all* situations and one compatible with *none* situations. During the test phase, participants are asked to decide which of the two words they would use to describe new *mixed* scenarios. Figure 2 shows the proportion of responses where participants selected the word compatible with *all* situations to describe novel, *sbna* situations during testing. This response pattern is taken to indicate that participants are reinterpreting one of the novel words to mean *some*, thereby showing a preference for lexicalizing *some* rather than the alternative *nall*.

We use *R*'s *brms* package (Bürkner 2018) as an interface to *Stan* (Carpenter et al. 2017) to run a Bayesian binomial mixed-effects regression model evaluating predicting *some*-like responses by grammatical category. This model is meant to evaluate whether the proportion of responses compatible with a *some* meaning is above what one would expect by chance, as predicted by Cognitive Markedness accounts. Our dependent variable is whether participants' responses are compatible with a *some* interpretation, as explained above (coded as 1 if compatible, 0 if not). The model included the grammatical category of the novel words (Adjective vs. Verb), sum-coded, as fixed effect. We include random intercepts for participants and for the color words. We set the same student-t prior on our fixed effect as well as on the intercept ($DF = 6, \mu = 0, \sigma = 1.5$); for the random effects, we set a half-Cauchy prior with scale parameter 10.

Our model suggests that participants are much more likely to assign a *nall* than a *some* meaning to a novel word (intercept $= -1.515, 90\%\text{CI} = [-3.109, 0.006], \text{SE} = 0.96, \text{P}(\text{intercept} > 0) = 0.05$). This counts as strong evidence for a preference for *nall* meanings ($\text{P}(\text{intercept} < 0) = 0.95$). This preference was not influenced by the grammatical category of the novel word (i.e., Adjective vs. Verb; $\beta = -0.282, 90\%\text{CI} = [-1.587, 0.975], \text{SE} = 0.787, \text{P}(\beta > 0) = 0.338$). Posterior densities visualizing these results are shown in Figure 2b.

4 Discussion

Across both experiments, participants showed no preference for a *some* meaning, challenging accounts of the *nall* lexical gap that attribute it to the cognitive markedness of this meaning. Surprisingly, if anything, we found the opposite pattern: a preference for a *nall* meaning. Although our experimental design does not allow us to pinpoint the origins of this preference, we explore two possible explanations.

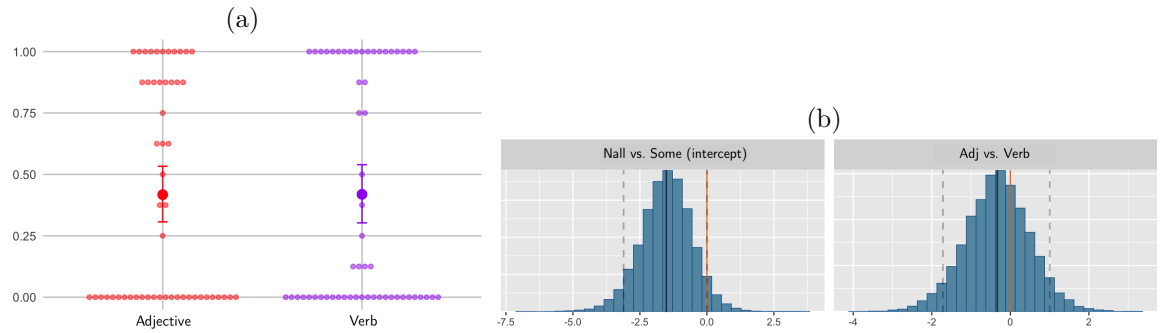


Figure 2: **Results.** (a) Proportion of *some*-consistent responses during testing. The small colored dots illustrate participants' individual means. Error bars are 95% bootstrapped confidence intervals. (b) Posterior distribution densities for the overall intercept and the difference between grammatical category (Adjective vs. Verb), their mean point estimates (solid orange line) and 90% credible intervals (dashed grey lines).

A first possible interpretation of the *nall* preference is that participants may have used their knowledge of English to perform the task. Indeed, one possible completion, in English, for *The redness of the circle is* [*novel adjective*] might be with 'complete' and 'incomplete'. Given that the 'incomplete' is actually compatible with *sbna* scenarios, this interpretation might lead to something that looks like a *nall* preference. We should note that only a few participants, when translating the novel words at the end of the experiment, reported taking them to mean 'complete' and 'incomplete'. Instead, most participants in the Adjective condition took the words to mean 'correct' and 'incorrect'. And, as far as we can see, there is no reason why 'incorrect' would be a better description of a *sbna* situation than 'correct'.

The preference is also detected in the Verb condition, which used *The circle* [*novel verb*] *red*. A natural way to do this task would be to analyze the target words as meaning 'is' or 'is not' (as suggested by most participants). Could any of these completions look like a *nall* meaning? The literature on homogeneity, both theoretical and experimental (Križ and Chemla 2015), suggests that *The circle is red* tends to mean that the circle is entirely red (*all* meaning), while *The circle is not red* tends to mean that the circle is not red at all (*no* meaning). This strategy, therefore, is not expected to produce any bias towards *nall*. However, preliminary experimental data that we obtained might suggest that English speakers do have a tendency, to some extent, to prefer using sentences like *The triangle is not red* to describe an *sbna*-situation.

Another possibility is that, in our experiment, participants prefer to lexicalize more frequent meanings. As observed in the introduction, situations where an object is fully red are intuitively less frequent than those where an object is not red at all (since objects can be many colors other than red). As a result, situations where an object is not fully red are more common than situations where an object is not fully non-red. In principle, a word describing not fully red situations conveys a *nall* meaning, whereas a word describing not fully non-red situations conveys a *some* meaning. This was the key argument behind communicative efficiency-based accounts such as Enguehard and Spector (2021): a quantifier meaning *nall* is less efficient than a quantifier meaning *some* because it would serve to describe situations that are more common, and thus ends up being less informative than the *some* alternative. One possibility is that the lexicalization preferences observed in our experiment do not reflect the pressure for informativity that Enguehard and Spector (2021) argue shapes this lexical gap. Instead, learners in our experiment might simply prefer to lexicalize meanings that are more frequent. The reason for this preference is unclear. It is possible that a communication task would better capture a pressure for informativity.

Overall, across two artificial language learning experiments, participants showed a pref-

erence for lexicalizing *nall* rather than *some*. These results speak against accounts of the non-lexicalization of *nall* that are based on cognitive markedness. Several avenues for future research emerge from our study. First, as discussed earlier, a task more oriented to communication could allow us to explicitly assess the predictions of accounts based on communicative efficiency. Moreover, to better understand the impact of the native language translations, we need to further test different grammatical categories and explore how they are interpreted in English.

Acknowledgements. This project has benefited from the support of ANR project ComCogMean (ANR-23-CE28-0016; PI: Benjamin Spector, co-PI: Mora Maldonado).

References

Bar-Lev, Moshe E and Roni Katzir (2023). "Communicative stability and the typology of logical operators". In: *Linguistic Inquiry*, pp. 1–23.

Barwise, Jon and Robin Cooper (1981). "Generalized Quantifiers and Natural Language". In: *Linguistics and Philosophy* 4.2, pp. 159–219. DOI: 10.1007/BF00350139.

Bürkner, Paul-Christian (2018). "Advanced Bayesian Multilevel Modeling with the R Package brms". In: *The R Journal* 10.1, pp. 395–411. DOI: 10.32614/RJ-2018-017.

Carpenter, Bob et al. (2017). "Stan: A probabilistic programming language". In: *Journal of statistical software* 76.1.

Chemla, Emmanuel, Brian Buccola, and Isabelle Dautriche (2019). "Connecting Content and Logical Words". In: *Journal of Semantics* 36.3, pp. 531–547. DOI: 10.1093/jos/ffz006.

Culbertson, Jennifer and David Adger (2014). "Language learners privilege structured meaning over surface frequency". In: *Proceedings of the National Academy of Sciences* 111.16, pp. 5842–5847.

Enguehard, Émile and Benjamin Spector (2021). "Explaining gaps in the logical lexicon of natural languages: A decision-theoretic perspective on the square of Aristotle". In: *Semantics and Pragmatics* 14, pp. 1–31.

Geurts, Bart and Frans van Der Slik (2005). "Monotonicity and processing load". In: *Journal of semantics* 22.1, pp. 97–117.

Horn, Laurence R. (1972). "On the Semantic Properties of Logical Operators in English". In: *Doctoral Dissertation*. University of California, Los Angeles.

Katzir, Roni and Raj Singh (2013). "Constraints on the Lexicalization of Logical Operators". In: *Linguistics and Philosophy* 36, pp. 1–29. DOI: 10.1007/s10988-013-9135-9.

Križ, Manuel and Emmanuel Chemla (2015). "Two methods to find truth-value gaps and their application to the projection problem of homogeneity". In: *Natural Language Semantics* 23.3, pp. 205–248. DOI: 10.1007/s11050-015-9114-z.

Maldonado, Mora and Jennifer Culbertson (2022). "Person of interest: Experimental investigations into the learnability of person systems". In: *Linguistic Inquiry* 53.2, pp. 295–336.

Maldonado, Mora, Jennifer Culbertson, and Wataru Uegaki (2022). "Learnability and constraints on the semantics of clause-embedding predicates". In: *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 44. 44.

Steinert-Threlkeld, Shane and Jakub Szymanik (2019). "Learnability and Semantic Universals". In: *Semantics and Pragmatics* 12, pp. 1–36. DOI: 10.3765/sp.12.8.

Uegaki, Wataru (2023). "The informativeness/complexity trade-off in the domain of Boolean connectives". In: *Linguistic Inquiry* 55.1, pp. 174–196.