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, $\Theta, \textit{cause1}, \textit{cause2} \Rightarrow \phi$.

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

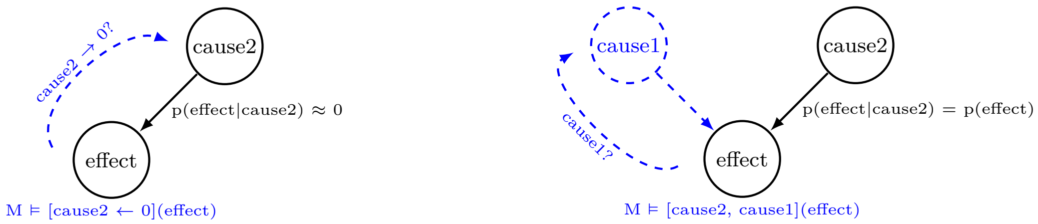

Figure 1: Anomalous (left) and novel (right) explanatory relation and their hypothesized repair strategies

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(effect|cause2) ≈ 0). If *cause2* were to be false, however, *effect would* be explainable (M ⊨ [cause2 ← 0](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(effect|cause2) = p(effect)). In this case, a more plausible strategy would seem to assume that there must be some additional eventuality influencing the *effect* (M ⊨ [cause2, cause1](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 ($\hat{\beta} = 1.92, z = 8.27, p < .01$).

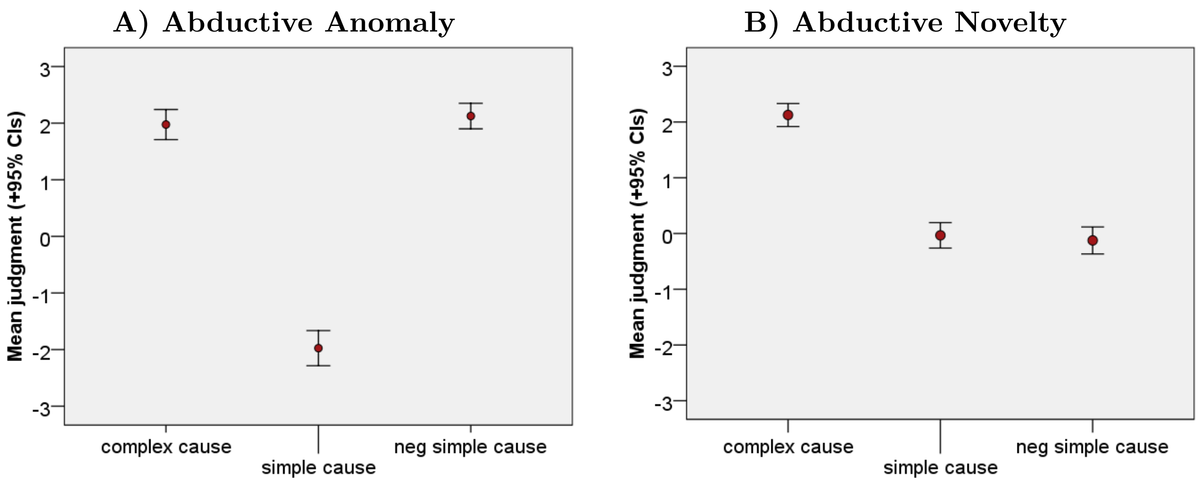

Figure 2: Mean contingency ratings and 95% confidence intervals in Exp. 1 for *anomaly* (A) and *novelty* (B). Positive values indicate "normal causal relations" and negative values indicate causal relations that are "opposite to normal". A value of 0 indicates no causal connection. 95% CIs estimated 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

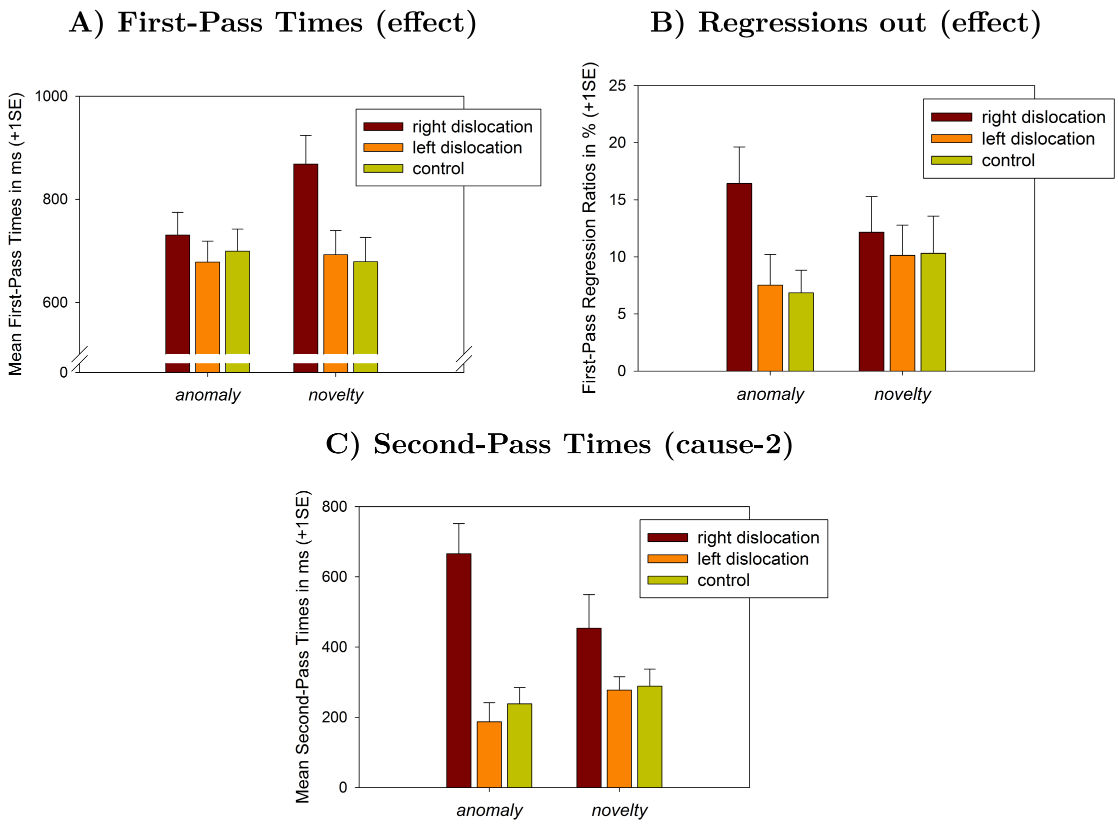

Figure 3: Mean first-pass times (top left) and first-pass regression ratios in the EFFECT region (top right) as well as second-pass times of the CAUSE-2 region (bottom row) in the eyetracking during reading experiment (Exp. 2). Error bars correspond to one standard error of the mean.

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.