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).

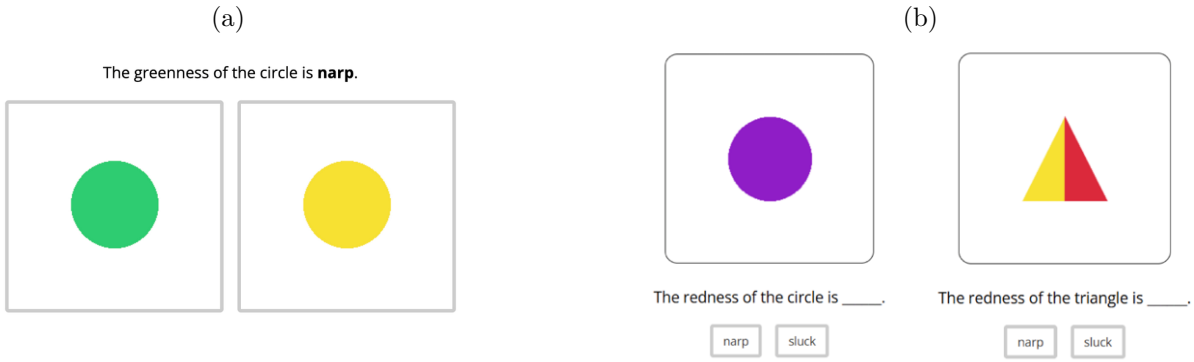

Figure 1: **Procedure (Adjective)**. (a) Sentence-picture matching task. If the word 'narp' was compatible with *all* situations, participants should select the image on the left. Otherwise, they should go for the image on the right. (b) Sentence-completion task. Left-picture is a **seen trial** involving 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.

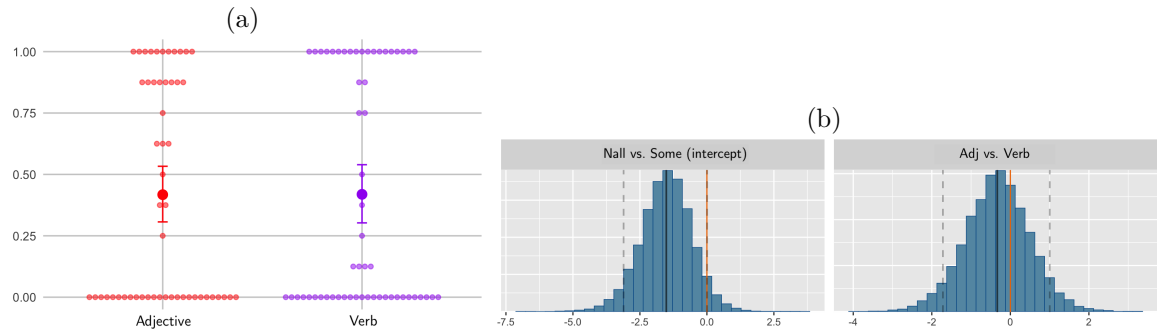

Figure 2: **Results.** (a) Proportion of *some*-consistent responses during testing. The small colored dots illustrate participants' individual means. Error bars are 95% bootstrapped confidence intervals. (b) Posterior distribution densities for the overall intercept and the difference between grammatical 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.