Corpus evidence for the role of world knowledge in ambiguity reduction: Using high positive expectations to inform quantifier scope

Noa Attali, Lisa S. Pearl, and Gregory Scontras
Department of Language Science, UC Irvine

Investigations into interpretations of quantifier-negation utterances (e.g., *Every vote doesn't count*, which is ambiguous between *No vote counts* and *Not all votes count*) have found variation: child and adult interpretations of *every*-negation diverge (e.g., Musolino, 1999), adult interpretations of utterances with different quantifiers vary (e.g., *every*- vs. *some*- vs. *no*-negation; Attali et al., 2021), and even adult interpretations of different *all*-negation constructions alone (Carden, 1973) and in context (Heringer, 1970) show considerable disagreement. Can we concretely identify factors to explain some of this variation and predict tendencies in individual interpretations? Here, we show that a type of expectation about the world, which can surface in the linguistic contexts of *every*-negation utterances that we find in a speech corpus, predicts experimental preferences for the *not all* interpretation of different *every*-negation utterances. These findings suggest that world knowledge, as set up in a linguistic context, helps to effectively reduce the ambiguity of potentially-ambiguous utterances for listeners.

High positive expectations. In their computational cognitive model of this ambiguity, Scontras and Pearl (2021) demonstrate that a kind of world knowledge we term a "high positive expectation" (**hpe**) can explain some variation in behavior with *every*-negation utterances. For example, in *Every vote doesn't count*, an hpe is the prior belief that it's highly likely that every vote *does* count – that is, that the worlds consistent with the non-negated utterance (*Every vote does count*) are likely. This world knowledge quantitatively specifies a pragmatic factor in previous proposals meant to capture truth value judgment results (e.g., Musolino and Lidz, 2006; Gualmini, 2004).

In particular, an hpe could contribute to the felicity of using *every*-negation with a *not all* interpretation, thereby reducing the ambiguity of the utterance for listeners. For speakers, Scontras and Pearl's model predicts that they tend to endorse *every*-negation as a true description of a scenario consistent with the *not all* interpretation when *every*-negation conveys that an hpe is false (e.g., that some votes are, in fact, not counted). For listeners, Attali et al. (2021) find that *not all* interpretations are preferred on average for *every*-negation, and that when Scontras and Pearl's model is applied to predict these listener interpretations, it does so successfully if given an hpe: it accounts well for the qualitative and quantitative pattern of average cross-speaker interpretation preferences for the experimental, out-of-context sentences *Every/Some/No marble isn't red*. The modeled pragmatic listener prefers the *not all* over the *none* interpretation, when given an hpe, because the listener assumes a cooperative, efficient speaker. A cooperative speaker wants to say something true, and there are more ways for the *not all* interpretation to be true compared with the *none* interpretation; an efficient speaker wants to be informative, and it's highly informative to update a strongly biased, salient belief (e.g., that every vote does count; see Attali et al., 2021).

So here, we ask to what extent an hpe accounts for interpretations of different *every*-negation utterances in naturalistic contexts. As a case study, when a local linguistic context seems to express an hpe, is *not all* a more likely interpretation than *none*?

Corpus data and behavioral experiment. We identified 390 uses of *every*-negation in the radio and TV transcripts (1990-2012; $\approx$9 million clauses) in the Corpus of Contemporary American English (Davies, 2015). Following Degen (2015), we crowd-sourced interpretation preferences of these uses in their immediate contexts (three preceding sentences and one following sentence). For each item, participants (N = 208) completed a paraphrase-endorsement task (Scontras and Goodman, 2017), choosing on a sliding scale between *none* and *not all* paraphrases of the

potentially-ambiguous clause. In line with previous findings, we found a preference for *not all* interpretations and a high degree of variation (see Fig. 1).

Identifying high positive expectations in linguistic contexts. As a preliminary measure, the first author hand-coded categorically for the presence/absence of an overt hpe expression in each preceding context (finding that 59/390 (15%) had an hpe). For an automatic and principled measure of the expression of an hpe in the linguistic context, we calculated the degree of lexical overlap between the preceding linguistic context and a string representing the positive expectation (hpe). That is, for each item (e.g., *Every vote doesn't count*), we first coded hpe as the potentially-ambiguous utterance without negation (e.g., *Every vote does count*). We then coded for the extent to which the hpe appeared in the preceding context as the longest common substring (**LCS**) similarity (Needleman and Wunsch, 1970), calculated using the R stringdist package (van der Loo, 2014). Each LCS was equal to the longest sequence formed by pairing words from the preceding context string and hpe, while keeping their order intact; the dissimilarity d_{lcs} was then the number of unpaired words left over in both strings, and LCS similarity was $-d_{lcs}$. Thus LCS similarity ranges from 0 to $-\infty$, with higher values indicating a greater probability that the context contained a high positive expectation. For example, if the preceding context was *Every vote does count* for an utterance with the hpe *Every vote does count*, LCS similarity would equal 0. On the other hand, if the preceding context was *What is going on?*, LCS similarity would equal -8 (since all eight words in the two strings would be unpaired). The disadvantage of LCS similarity is its noisy potential to underestimate the presence of an hpe (e.g., it would discount the context *All votes should matter*); but it provides an automatic continuous measure. (Other lexical overlap implementations yield similar results.)

Results. Using the preliminary categorical hand-coding, we found that 50/59 (85%) of the utterances with hpes were on average better paraphrased by *not all* than *none*. Using the continuous and automatic LCS measure to assess if an hpe predicts a *not all* preference per item, we ran a linear mixed effects model predicting logit-transformed mean item responses by LCS similarity, with random intercepts for participants (see Fig. 2). To determine whether an hpe captures individual judgment variation, above and beyond mean item-level variation, we predicted logit-transformed item responses by LCS similarity, with random intercepts for participants and items. Both models found that LCS similarity was a significant predictor of a *not all* preference ($p < .001$ in both). Interestingly, a version of both models which calculated LCS similarity using overlap with the following – rather than preceding – context, found LCS similarity of the following context not to be a significant predictor of either item-level or judgment-level interpretations.

Conclusion. Our corpus analysis supports the plausibility of an hpe, expressed in the preceding linguistic context, playing a role in *not all* interpretation preferences for *every*-negation utterances. These results align with the previous modeling results and pragmatically-oriented proposals from truth value judgment studies. We note that we might underestimate the role of an hpe, because our automated method of identifying it is only one of many, and may be noisy; moreover, such an aspect of world knowledge could affect interpretations without necessarily receiving expression in the immediate discourse. In general, our findings support the theory that negation use is more felicitous in affirmative contexts (e.g., Wason, 1961), such as contexts containing an hpe.

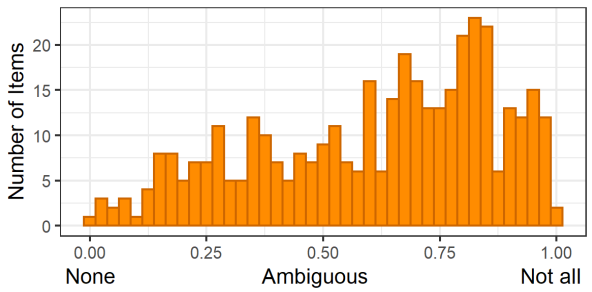

Figure 1: Histogram of average item interpretation from the corpus analysis.

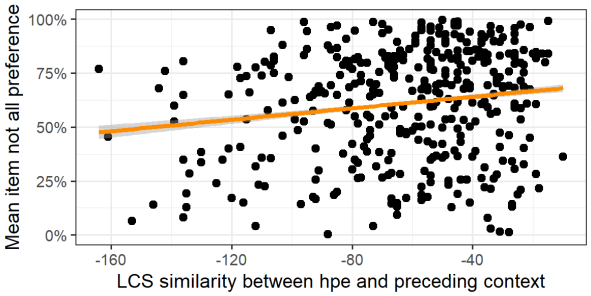

Figure 2: Preceding hpe and average *not all* item preference.

References

N. Attali, G. Scontras, and L. S. Pearl. Pragmatic factors can explain variation in interpretation preferences for quantifier-negation utterances: A computational approach. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 43, 2021.

G. Carden. Disambiguation, favored readings, and variable rules. *New ways of analyzing variation in English*, pages 171–82, 1973.

M. Davies. Corpus of Contemporary American English (COCA). 2015. URL `https://doi.org/10.7910/DVN/AMUDUW`.

J. Degen. Investigating the distribution of some (but not all) implicatures using corpora and web-based methods. *Semantics and Pragmatics*, 8:11–1, 2015.

A. Gualmini. Some knowledge children dont lack. *Linguistics*, 42(5):957–982, 2004.

J. T. Heringer. Research on quantifier-negative idiolects. In *Chicago Linguistic Society*, volume 6, page 95, 1970.

J. Musolino. Universal grammar and the acquisition of semantic knowledge: An experimental investigation into the acquisition of quantifier-negation interaction in english. 1999.

J. Musolino and J. Lidz. Why children aren't universally successful with quantification. *Linguistics*, 44(4):817–852, 2006.

S. B. Needleman and C. D. Wunsch. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of molecular biology*, 48(3):443–453, 1970.

G. Scontras and N. D. Goodman. Resolving uncertainty in plural predication. *Cognition*, 168:294–311, 2017.

G. Scontras and L. S. Pearl. When pragmatics matters more for truth-value judgments: An investigation of quantifier scope ambiguity. *Glossa: a journal of general linguistics*, 6(1), 2021.

M. P. van der Loo. The stringdist Package for Approximate String Matching. *The R Journal*, 6(1):111–122, 2014. doi: 10.32614/RJ-2014-011. URL `https://doi.org/10.32614/RJ-2014-011`.

P. C. Wason. Response to affirmative and negative binary statements. *British Journal of Psychology*, 52(2):133–142, 1961.