

Running head: PRAGMATICS OF NEGATION

Negation is only hard to process when it is pragmatically infelicitous

Ann E. Nordmeyer and Michael C. Frank

Department of Psychology, Stanford University

Corresponding author: Ann E. Nordmeyer

Department of Psychology

Stanford University

Building 420 (Jordan Hall)

450 Serra Mall

Stanford, CA 94305

Phone: 650-721-9270

Email: anordmey@stanford.edu

Abstract

Negation is a fundamental element of language and logical systems, but processing negative sentences can be challenging. Early investigations suggested that this difficulty was due to the representational challenge of adding an additional logical element to a proposition, but in more recent work, supportive contexts mitigate the processing costs of negation, suggesting a pragmatic explanation. We make a strong test of this pragmatic hypothesis by directly comparing speakers and listeners. Speakers produce negative sentences more often when they are both relevant and informative. Listeners in turn are fastest to respond to sentences that they expect speakers to produce. Since negative sentences are only difficult in contexts when they are unlikely to be produced, representing negation is likely less difficult than previously supposed.

Keywords: Language, Psycholinguistics, Language Comprehension, Language Production, Pragmatics

Introduction

Language is a powerful tool that allows us to describe not only the state of the world as we see it, but also the world as it is not. Nevertheless, for human language users, processing negation is often slow and effortful. Deciding the truth value of a sentence like “star isn’t above plus” takes a lot longer than making the same decision about a positive sentence (Clark & Chase, 1972; Carpenter & Just, 1975; Just & Carpenter, 1971, 1976). And in language comprehension tasks, participants often show evidence consistent with having processed the positive components of a sentence prior to negating them, suggesting again that negation is challenging (Kaup & Zwaan, 2003; Kaup, Ludtke, & Zwaan, 2006; Hasson & Glucksberg, 2006; Fischler, Bloom, Childers, Roucos, & Perry, 1983; Lüdtke, Friedrich, De Filippis, & Kaup, 2008). Why do adults struggle to process negation despite spontaneously producing negative sentences with ease?

One explanation is that not all negations are equally felicitous. For example, it would be strange to say “my car isn’t purple”—unless we are in a parking lot where every car except mine *is* purple. And it would be even stranger (though still true) to follow up by saying “my car isn’t a parakeet.” On Gricean and neo-Gricean accounts of the pragmatics of language use in context, listeners expect speakers to produce informative and relevant utterances (Grice, 1975; Horn, 1984; Levinson, 2000; Sperber & Wilson, 1986). The fact that your car is not purple is uninformative if it doesn’t help me identify your car, and the subsequent remark about your car not being a parakeet is uninformative *and* irrelevant. Is this kind of pragmatic infelicity generally responsible for the processing cost of negation?

Consistent with this suggestion, presenting negative information in a supportive context can mitigate some of its processing costs (Wason, 1965; Glenberg, Robertson,

Jansen, & Johnson-Glenberg, 1999). When a negated feature is explicitly mentioned in preceding sentences (Lüdtke & Kaup, 2006), or when negation is presented within a dialogue (Dale & Duran, 2011), negative sentences tend to be processed faster. And in an ERP experiment, contextually-supported negations (e.g., “with proper equipment, scuba-diving isn’t very dangerous”) elicited smaller N400 responses—a marker of semantic processing costs—than unlicensed negations (e.g., “bulletproof vests aren’t very dangerous”; Nieuwland & Kuperberg, 2008).

Although this previous work supports the idea that some kind of contextual expectations are the source of negation’s processing cost, they do not specify the precise nature of these expectations. Our current experiment directly tests two hypotheses. First, speakers tend to produce negative sentences only when they are both relevant and informative given the context. Second, expectations about what speakers would likely say—and their match or mismatch with what the speaker in fact *does* say—are responsible for the processing costs of negation. To formalize this second hypothesis, we make use of recent probabilistic models of language comprehension, defining a listener’s pragmatic expectations as the probability that a speaker would utter a statement in order to convey a particular meaning (Frank & Goodman, 2012), and using *surprisal*, an information-theoretic measure of expectation-based processing costs (Levy, 2008), to predict processing times.

In our experiment, we show participants sets of characters who are identical except for the presence or absence of a feature (e.g., boys with or without apples). In the *speaker* condition, we ask participants to produce written descriptions that pick a particular target character out of the set, while in the *listener* condition, we ask participants to evaluate the truth value of negative or positive statements about the same pictures (see Figure 1). We predict that participants will take longer to comprehend negative sentences that are unlikely to be produced by speakers.

Method

Participants

We recruited a planned sample of 500 participants to participate in an online experiment through the Amazon’s Mechanical Turk (mTurk) website; 11 participants were rejected for indicating that they were under 18 after completing the experiment. Of the remaining 489 participants, 262 were male and 224 were female, three declined to report gender, and ages ranged from 18 – 65+. We restricted participation to individuals in the US and paid 50 cents for this 10 minute study. Participants were randomly assigned to either the speaker condition ($n = 296$) or the listener condition ($n = 193$); small differences in the amount of time participants took to complete each study resulted in more participants assigned to the speaker condition.

Stimuli

Thirty-two trial items were created in which characters were shown holding either two of the same common, recognizable objects (“target items”; e.g., two apples), or holding nothing. Within each trial, all characters were identical except for the presence or absence of objects; characters varied in appearance (e.g. skin tone, hair color, clothing, gender) across trials. Our previous work indicates that the results reported here are robust to a number of changes to the stimuli, including whether the context characters varied in appearance or not (Nordmeyer & Frank, 2014).

Each participant saw trials in which different proportions of characters were holding target items (context condition). These contexts showed $\frac{0}{4}$, $\frac{1}{4}$, $\frac{2}{4}$, $\frac{3}{4}$, or $\frac{4}{4}$ of the characters holding objects. The order of characters was shuffled on each trial, with the referent of the sentence appearing in a random position.

Participants in the speaker condition saw each image paired with an incomplete

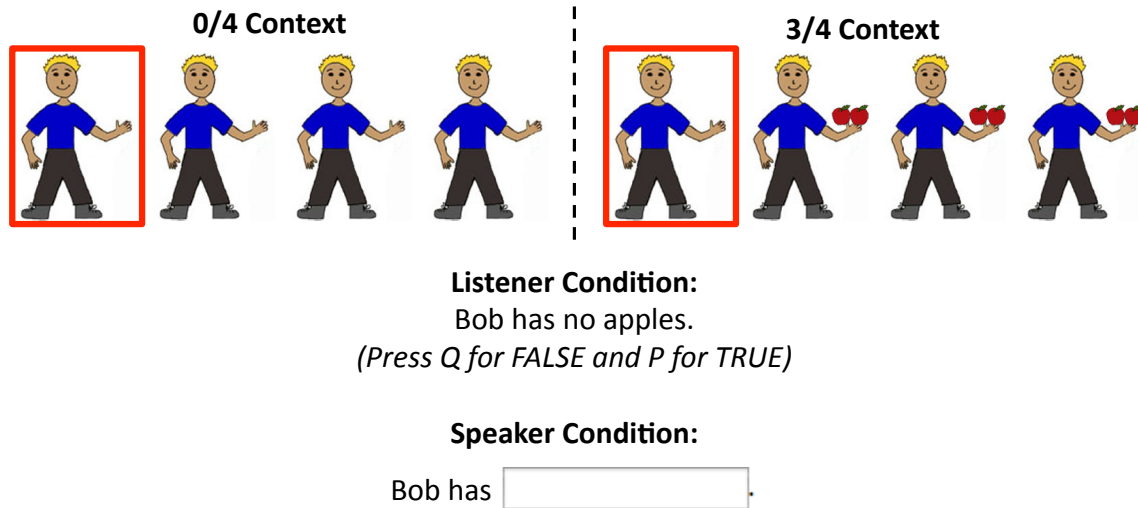


Figure 1. An example of a true negative trial with a 0/4 context (left) and a 3/4 context (right). The sentence “Bob has no apples” in the 0/4 context is both uninformative (because the sentence is true of all of the characters) and irrelevant (because apples are not present in the context), whereas the same sentence in the 3/4 context is informative and relevant.

sentence (e.g. “[NAME] has _____.”). In half of the trials, the highlighted character was holding target items (“item trials”), and in half of the trials, the highlighted character was holding nothing (“nothing trials”). The experiment was fully crossed such that target characters appeared with or without target items an equal number of times in each context type.

Participants in the listener condition saw the same set of images. On each trial a sentence of the form “[NAME] [has/has no] [ITEM]” appeared. Half of the sentences were positive and half were negative (sentence type), and they were paired with pictures such that half were true and half were false (truth value), resulting in four possible trial types (true positive, true negative, false positive, and false negative). Because true positive and false negative sentences cannot occur in a $\frac{0}{4}$ context (i.e. the referent must have the target item in these trials), and true negative and false positive sentences cannot occur in a $\frac{4}{4}$

context, each trial type occurred in four possible contexts. The experiment was fully crossed, with participants receiving eight true positive, eight false positive, eight true negative and eight false negative sentences distributed equally across context types in a randomized order over the course of the study.

Procedure

Participants were first presented with a brief overview screen which explained that they would play a language game. Once participants accepted the task, they were randomly assigned to either the speaker condition or the listener condition, and saw a more detailed instructions screen which explained the task and informed them that they could stop at any time. The speaker and listener conditions of the experiment can be viewed at

https://langcog.stanford.edu/expts/AEN/negatron_production2/negatron.html and <https://langcog.stanford.edu/expts/AEN/negatronv20/negatron.html>, respectively.

In the speaker condition, participants saw an array of four pictures on each test trial: The target pictures and three context pictures presented in a random horizontal arrangement. Participants were told to look at these pictures for four seconds, at which point a red box appeared around one of the pictures. One second later, an incomplete sentence appeared. Participants were told to finish the sentence (by typing into a small text box) using only a few words, in a way that would help someone else identify the character in the red box if they saw the pictures in a different order.

In the listener condition, participants first saw eight positive sentence practice trials with feedback about incorrect responses before beginning the test trials. In each test trial, participants saw an array of four pictures presented in a random horizontal arrangement. Participants were told to look at these pictures for four seconds, at which point a red box

appeared around one of the pictures. One second later, a sentence about that picture appeared. Participants were told to read the sentence and respond as quickly and accurately as possible with a judgment of whether it was true or false when applied to the highlighted picture. We recorded reaction times for each trial, measured as the time from when the sentence was presented to the moment when the response was made.

Data Processing

We excluded 18 participants who did not list English as their native language and two participants from the listener condition for having an overall accuracy below 80%, leaving a total of 469 participants for analysis (186 in listener condition, 283 in speaker condition).

In the speaker condition, we coded participant's productions. Affirmative responses labeling the target feature were coded as "positive" (e.g., "apples," "two apples," "red apples," etc.). Responses negating the target feature (e.g., "no apples") were coded as "negative." All other responses (e.g. descriptions of the characters' clothing or hair color) were coded as "other." Codes were hand-checked to ensure that label synonyms or spelling errors were coded correctly.

In the listener condition, we excluded trials with RTs greater than 3 standard deviations from the log-transformed mean, a criterion established in our previous experiments (Nordmeyer & Frank, 2014).

To predict responding in the listener condition, we used productions from the speaker condition. We calculated the proportion of positive sentences describing characters who possessed target items, and the proportion of negative sentences describing characters with nothing, creating probability distributions for true positive and true negative utterances in each context. We then used this distribution to calculate the surprisal of hearing a true positive or true negative sentence for each context. Surprisal is

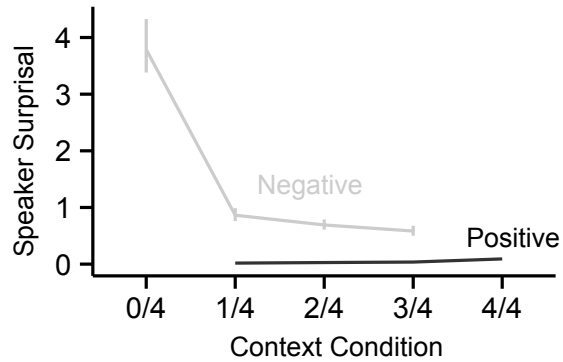


Figure 2. Surprisal for true positive and true negative sentences across different contexts. Negative sentences are shown in grey, and positive sentences in black. The context is notated by a fraction representing the number of characters in the context who held target items. Error bars show 95% confidence intervals computed by non-parametric bootstrapping.

an information-theoretic measure of the amount of information carried by an event (in this case, the amount of information conveyed by a sentence); in prior work on sentence comprehension it has been used successfully as a linking hypothesis between production probabilities and reaction times (Levy, 2008). Surprisal (or “self-information” I) for a sentence s is defined as

$$I(s) = -\log(P(s)). \quad (1)$$

Results

Speaker Condition

Participants were much more likely to produce negation when the target character was not holding anything (nothing trials), and produced negative sentences very rarely when the target character was holding target items (item trials). Consistent with our hypothesis that speakers will produce negative sentences that are relevant (i.e. negating a

salient feature of the context), negative sentences were produced very rarely on nothing trials in the $\frac{0}{4}$ context (e.g. “no apples” was a rare production on trials where there weren’t any apples in the context; see Figure 1), but became much more common on nothing trials in the $\frac{1}{4}$ context. In line with our hypothesis that speakers will produce informative utterances (i.e. produce sentences that are maximally effective at identifying the referent), participants in the speaker condition were increasingly likely to produce negation as the number of characters with target items increased, with the $\frac{3}{4}$ context eliciting the highest production of negation.

To evaluate the reliability of these patterns, we fit a binomial mixed-effects model to test the effect of trial type and context on the probability of producing a negative sentence. To test the quantitative effects of context, we coded this variable as numeric, e.g. the proportion of characters in the context with target items. In addition, because the difference between the $\frac{0}{4}$ and the $\frac{1}{4}$ contexts was so striking, we created a dummy code to separately test for the effects of the $\frac{0}{4}$ context compared to all of the other contexts. All mixed-effects models used the maximal convergent random effects structure and were fit using the lme4 package version 1.1-7 in R version 3.1.2. The model specification was as follows: `negation ~ context × trial type + dummy context + (1 | subject) + (1 | item)`. Raw data and analysis code can be found at <https://github.com/anordmey/negatron>.

All model coefficients are shown in Table 1. Main effects of trial type and dummy context confirm that participants were much more likely to produce negative sentences on nothing trials (i.e. trials where the target character held nothing) and much less likely to produce negation in the $\frac{0}{4}$ context. A marginally significant interaction between context and trial type suggests that there is a linear effect of context on nothing trials even after accounting for the effect of the $\frac{0}{4}$ context. Following up on this finding in an exploratory

Table 1

Coefficient estimates from a binomial mixed-effects model predicting speaker's productions in different contexts.

	Coefficient	Std. err.	z	$p(z)$
Intercept	-8.30	.82	-10.12	<.001
Context	.01	1.15	.01	0.99
Trial Type (nothing)	7.12	.81	8.76	<.001
Dummy context ($\frac{0}{4}$)	-4.74	.29	-16.30	<.001
Context \times Trial Type	2.18	1.17	1.86	0.063

analysis, we fit a model to only the data on nothing trials in the $\frac{1}{4}$, $\frac{2}{4}$, and $\frac{3}{4}$ contexts, and found a significant linear effect of context on the production of negation ($\beta = 2.28$, $p < .001$), suggesting that negative sentences were more likely to produced in contexts where they were more informative.

Figure 2 shows the surprisal of true positive (i.e. descriptions of the target noun on item trials) and true negative (i.e. negations of the target noun on nothing trials) sentences. Consistent with our predictions, surprisal is highest for true negative sentences in the $\frac{0}{4}$ context, and surprisal for true negative sentences decreases as the number of context characters with target items increases.

Listener Condition

Participants were fastest to respond to true positive sentences, and slowest to respond to true negative sentences, replicating previous findings (Clark & Chase, 1972). Listeners' responses to true negative sentences mirrored the surprisal of true negative sentences in the speaker condition, with participants responding slowest to true negatives

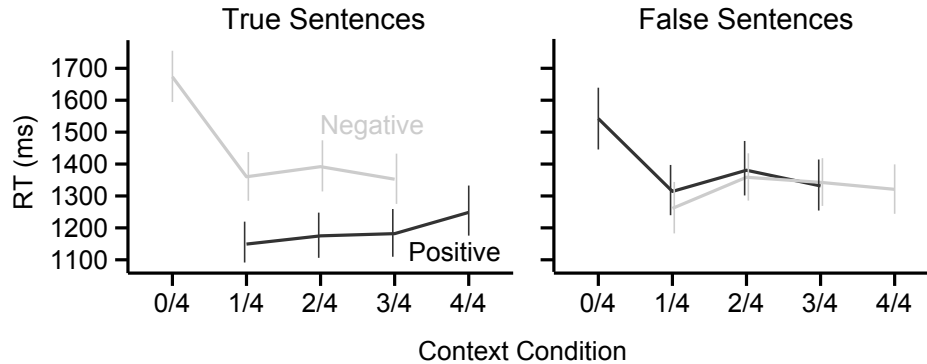


Figure 3. Reaction times for each trial type across different conditions. Responses to true sentences are shown on the left, and false sentences are shown on the right. Negative sentences are shown in grey, and positive sentences in black. The context is notated by a fraction representing the number of characters in the context who held target items. Error bars show 95% confidence intervals computed by non-parametric bootstrap.

in the $\frac{0}{4}$ context. The same pattern was seen in response to false positive sentences, suggesting that listeners expect speakers to describe relevant features of the context even when the sentence is false (see Figure 3).

We fit a linear mixed-effects model to examine the interaction between sentence type (positive or negative), truth value (true or false), and context as predictors of reaction time. The model specification was as follows: $RT \sim \text{sentence} \times \text{truth} \times \text{context} + (\text{sentence} \mid \text{subject}) + (\text{sentence} \mid \text{item})$. Significance was calculated using the standard normal approximation to the t distribution (Barr, Levy, Scheepers, & Tily, 2013).

All model coefficients are shown in Table 2. In addition to main effects of sentence type and truth value, the model showed an interaction such that true positive sentences elicited the fastest responses and true negative sentences elicited the slowest responses. The model showed a significant negative linear effect of context, with reaction times decreasing as the proportion of characters with target items increased. A significant

Table 2

Coefficient estimates from a mixed-effects model predicting listeners' reaction times in response to sentences in different contexts.

	Coefficient	Std. err.	<i>t</i>
Intercept	1482	39	37.57
Sentence (Negative)	-204	36	-5.61
Truth (True)	-369	36	-10.29
Context	-238	44	-5.44
Sentence × Truth	686	51	13.42
Sentence × Context	311	62	5.05
Truth × Context	366	61	5.97
Sentence × Truth × Context	-835	88	-9.52

three-way interaction between sentence type, truth value, and context indicates that this pattern was driven primarily by responses to true negative sentences, however, with true positive and false negative sentences showing a smaller positive linear effect of context.

Predicting Listeners with Speakers

To test the hypothesis that processing times are a function of listeners' expectations about what a speaker will say, we regressed the mean reaction time in response to true positive and negative utterances in each condition against the surprisal for the same utterances (Figure 4). There was a significant positive relationship between surprisal and reaction time for true negative sentences, $r^2 = .90$, $p < .001$, supporting our prediction that the effects of context on reaction time reflect differences in how speakers would describe the same stimuli.

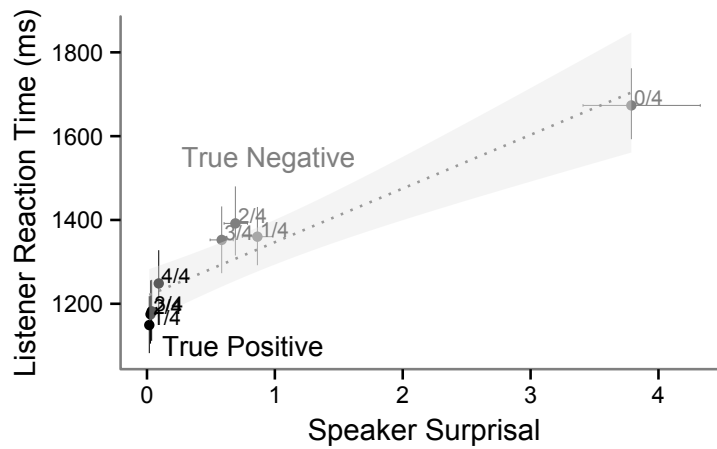


Figure 4. Reaction times in the listener condition plotted by surprisal in the speaker condition; each point represents a measurement for sentence type and context. Error bars on the horizontal and vertical axes represent 95% confidence intervals on their respective measures.

General Discussion

What makes negation hard to process? While previous work has proposed that processing negative elements is especially difficult because of features intrinsic to negation, our work here suggests instead that general pragmatic mechanisms are likely responsible. Negative sentences presented without context are uninformative and irrelevant; thus, they are unlikely to be produced by speakers. In turn, listeners respond to these unlikely utterances with increased processing times. In contexts where negation is more informative and relevant, processing costs are lower. Overall, this evidence supports a Gricean interpretation of the processing of negation.

While previous work has shown that contextual factors facilitate the processing of negation (Wason, 1965; Nieuwland & Kuperberg, 2008; Dale & Duran, 2011), our findings here go further. First, and most importantly, by using actual language productions as the

predictor of processing difficulty, our work strongly implicates specifically pragmatic (rather than representational) factors. Second, rather than treating pragmatics as a black box, we show that two different components—informativeness and relevance—each contribute to the relative (un-)likelihood of hearing a negation.

Our work here also uses surprisal, an information-theoretic measure of processing difficulty, as the linking hypothesis between speaker probabilities and reaction times (Levy, 2008). Although this link has substantial support in the realm of syntactic processing (Demberg & Keller, 2008; Boston, Hale, Kliegl, Patil, & Vasishth, 2008), to our knowledge, our findings are the first example of using surprisal over sentence-level pragmatic expectations, rather than word-level syntactic expectations. This success suggests that, in concert with the appropriate predictive models, surprisal theory could be productively applied to the prediction of processing difficulty beyond the level of syntax.

Although our focus here was on negation, our findings have implications for sentence processing more generally. Debates about the effects of pragmatics on linguistic processing exist in other domains, such as the processing of scalar implicatures (the pragmatic inference that e.g., “some” implicates “some but not all”; Huang & Snedeker, 2009, 2011; Grodner, Klein, Carbary, & Tanenhaus, 2010). Tomlinson, Bailey, and Bott (2013) provide an informative comparison between scalar implicature and negation, presenting mouse-tracking trajectories for each. Their negation data show the same pattern of processing difficulties we observe, and critically, their data on the processing of underinformative “some” utterances look almost identical. We hypothesize that, in both cases, participants’ processing difficulty is a function of the violation of their pragmatic expectations.

In sum, our findings here suggest that a large part of the processing difficulties of negative sentences arise from the relative pragmatic felicity of negation in context. They do not rule out the possibility that there is some cost to processing an additional logical

element, but this processing cost would have to be quite small with respect to the magnitude of the effects we observed (and previous measurements). This finding leads us to the following conclusion: When logical words are used in a communicative context, we have no difficulty understanding them.

Author Contributions

Both authors developed the study concept and contributed to the study design. Data collection was conducted by A. E. Nordmeyer. A. E. Nordmeyer performed the data analysis and interpretation under the supervision of M.C. Frank. Both authors contributed to the development of the manuscript and approved the final version of the manuscript for submission.

Acknowledgments

This work supported by the NSF GRFP and ONR N00014-13-1-0287.

References

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*, 255–278.
- Boston, M., Hale, J., Kliegl, R., Patil, U., & Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the potsdam sentence corpus. *The Mind Research Repository (beta)*.
- Carpenter, P., & Just, M. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*, 45–73.
- Clark, H., & Chase, W. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*, 472–517.

- Dale, R., & Duran, N. (2011). The cognitive dynamics of negated sentence verification. *Cognitive Science*, *35*, 983–996.
- Demberg, V., & Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, *109*, 193–210.
- Fischler, I., Bloom, P., Childers, D., Roucos, S., & Perry, N. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, *20*, 400–409.
- Frank, M., & Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*, 998.
- Glenberg, A., Robertson, D., Jansen, J., & Johnson-Glenberg, M. (1999). Not propositions. *Journal of Cognitive Systems Research*, *1*, 19–33.
- Grice, H. (1975). Logic and conversation. *1975*, 41–58.
- Grodner, D. J., Klein, N. M., Carbary, K. M., & Tanenhaus, M. K. (2010). Some, and possibly all, scalar inferences are not delayed: Evidence for immediate pragmatic enrichment. *Cognition*, *116*, 42–55.
- Hasson, U., & Glucksberg, S. (2006). Does understanding negation entail affirmation? An examination of negated metaphors. *Journal of Pragmatics*, *38*, 1015–1032.
- Horn, L. R. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42). Washington, D.C.: Georgetown University Press.
- Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive psychology*, *58*, 376–415.
- Huang, Y. T., & Snedeker, J. (2011). Logic and conversation revisited: Evidence for a division between semantic and pragmatic content in real-time language comprehension. *Language and Cognitive Processes*, *26*, 1161–1172.
- Just, M., & Carpenter, P. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*, 244–253.

- Just, M., & Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441–480.
- Kaup, B., Lüdtke, J., & Zwaan, R. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, 38, 1033–1050.
- Kaup, B., & Zwaan, R. (2003). Effects of negation and situational presence on the accessibility of text information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 439–446.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106, 1126–1177.
- Lüdtke, J., Friedrich, C., De Filippis, M., & Kaup, B. (2008). Event-related potential correlates of negation in a sentence-picture verification paradigm. *The Journal of Cognitive Neuroscience*, 20, 1355–1370.
- Lüdtke, J., & Kaup, B. (2006). Context effects when reading negative and affirmative sentences. In *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1735–1740).
- Nieuwland, M., & Kuperberg, G. (2008). When the truth is not too hard to handle. *Psychological Science*, 19, 1213.
- Nordmeyer, A. E., & Frank, M. C. (2014). A pragmatic account of the processing of negative sentences. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*.
- Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Oxford, UK: Blackwell Publishing.
- Tomlinson, J. M., Bailey, T. M., & Bott, L. (2013). Possibly all of that and then some: Scalar implicatures are understood in two steps. *Journal of Memory and Language*, 69,

18–35.

Wason, P. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, 4, 7–11.