

Running head: CONTEXTS OF NEGATION

Uninformative negation is infelicitous to both adults and children

Ann E. Nordmeyer and Michael C. Frank

Department of Psychology, Stanford University

Abstract

Why do some negative sentences sound strange, even when they are both true and grammatical? Past work suggests that both adults (e.g., Clark & Chase, 1972) and children (e.g., Kim, 1985) find negation difficult, but context can facilitate the processing of negative sentences (e.g., Wason, 1965). We argue that the effect of context on negation is driven by general pragmatic principles rather than the conceptual difficulty of negation. In Experiments 1 and 2, we found that more informative negative sentences elicited higher felicity ratings in adults. In Experiment 3, we showed that children are also sensitive to the contexts of true negative sentences, with three- and four-year-olds rating negative sentences higher in more informative contexts. We discuss the pragmatics of negation in light of these results, arguing that both adults' and children's difficulty processing negation is due to changes in the informativeness, and thus the felicity, of these sentences in different contexts.

Introduction

A sentence that is both grammatical and true can nevertheless sound odd in some contexts. If you open a mysterious box and find that it is empty, it's weird to say "This box doesn't have any chocolates in it," despite the truth of the proposition. But the same sentence becomes perfectly reasonable if you are in a chocolate store surrounded by boxes filled with chocolate. What is it about context that makes some sentences sound strange (*infelicitous*) in one situation, but normal in another?

Theories of pragmatics attempt to provide an account of how language is used for communication. For example, according to Grice's (1975) Cooperative Principle, speakers should produce utterances that are truthful, relevant (i.e., mention a salient feature of the context), and informative (i.e., effectively identify the speaker's intended referent). By assuming that speakers do so, listeners can make inferences about intended meaning that go beyond the sentence's literal meaning. Modern neo-Gricean theories tend to derive such inferences from the tension between being informative with respect to a communicative goal and minimizing the effort expended (Horn, 1984; Levinson, 2000; Frank & Goodman, 2012). These theories make predictions about sentence felicity: A sentence that is not truthful, relevant, or informative in the given context can sound strange even when it is both grammatical and true.

Here, we argue that both adults' and children's difficulty processing negative sentences can be explained by the relationship between contextual pragmatics and felicity. First we discuss past work on the role of context in adults' processing of negation, which motivates our desire to explore adults' explicit felicity judgments of negative sentences. Then we describe children's acquisition of negation, and what we can learn about children's pragmatic development by investigating whether children are sensitive to the

contexts of negation as well. Finally, we describe our current experiments, which test the hypothesis that the effect of context on negative sentences is driven by pragmatic principles that influence communication more generally.

The role of context in the processing of negation

Negative sentences can be difficult to process, even for adults. When presented without any context, people take longer to evaluate the truth of negative sentences such as “star isn’t above plus” compared to positive sentences (Clark & Chase, 1972; Carpenter & Just, 1975; Just & Carpenter, 1971, 1976). In EEG studies, participants who read true negative sentences like “A bird is not a truck” show an N400 response (a negative waveform that responds to semantically unexpected stimuli) similar to corresponding false positive sentences like “A bird is a truck” (Fischler, Bloom, Childers, Roucos, & Perry, 1983; Lüdtke, Friedrich, De Filippis, & Kaup, 2008). Based on these findings, it has been suggested that there is something specific to the representation of negation that makes processing negative sentences so challenging.

As our original example demonstrates, however, negative sentences are particularly sensitive to the effects of context. In a classic study, Wason (1965) found that when participants were asked to describe an array of seven red dots and one blue dot, participants who described the stimuli in terms of an exception to a rule (e.g., “Circle number four is blue and the rest are red”) were faster to process true negative sentences such as “Circle four is not red” compared to participants who described the stimuli in other ways (e.g., “Seven circles are red and one is blue”). Similarly, supportive written contexts, such as a narrative framing where the negated feature is made salient prior to reading the negative sentence, can facilitate the processing of negation (Glenberg, Robertson, Jansen, & Johnson-Glenberg, 1999; Lüdtke & Kaup, 2006; Nieuwland & Kuperberg, 2008; Dale & Duran, 2011), and reduce the N400 response to true negative

sentences (Nieuwland & Kuperberg, 2008). Although these findings suggest that context can have a powerful effect on the processing of negative sentences, it is not clear what mechanism leads to this facilitative effect.

One possible mechanism is that unsupportive contexts make negative sentences difficult because they are pragmatically uninformative. In recent work (Nordmeyer & Frank, under review), we demonstrated a graded effect of context on the processing of negation, with negative sentences processed faster as the base rate of the negated feature increased. Specifically, negative sentences that were more relevant and informative given the context were processed faster. We found that this effect of context could be reliably predicted by the probability of a speaker using the same sentence to describe the same picture. Speakers are unlikely to produce sentences that are uninformative and irrelevant, and correspondingly listeners take longer to respond to unlikely sentences.

Extending this previous work, in the current paper we hypothesize that general pragmatic principles can explain both the processing cost of negation *and* the supportive effect of context. If pragmatic infelicity is responsible for the context effects seen in negation processing, these context effects should be reflected in adults' explicit felicity judgments as well, with more informative sentences eliciting higher felicity ratings than less informative sentences. This finding would support our hypothesis that the effects of context on negative sentences are due to pragmatic principles that govern communication more generally, rather than some unique effect of negation. Furthermore, we propose that children's apparent difficulty with negation is in fact due to their sensitivity to these same pragmatic principles rather than an undeveloped semantic understanding of negation.

Children's understanding of negation

Although children begin producing negation spontaneously and accurately before age two (Bloom, 1970; Pea, 1980, 1982), preschoolers struggle to respond correctly to true

negative sentences. For example, the majority of three-year-olds and nearly half of four- and five-year-olds classified true negative sentences as “wrong” when an experimenter pointed to an apple and said “this is not a banana” (Kim, 1985), even though they had no difficulty classifying true and false positive sentences. This finding, along with others (Gilkerson, Hyams, & Curtiss, 2002; Loder, 2006; Nordmeyer & Frank, 2014b), suggests that children struggle to process negative sentences, despite the fact that they produce them so early.

There are several reasons why children may struggle to comprehend negative sentences. One possibility is that young children do not have a fully developed semantic understanding of words like “no” and “not.” Some functions of negation, such as refusal (e.g., “no go outside” to express a desire to stay indoors), are produced earlier than denial/truth-functional negation, which requires evaluating a proposition to determine its truth value. Children who fail comprehension tests of denial negation may not fully understand the semantics of negation, or may not have the cognitive abilities to evaluate the truth of a proposition. Conceptual differences in types of negation also covary with syntactic differences, at least in English (see Cameron-Faulkner, Lieven, & Theakston, 2007). For example, clear expression of denial negation often requires the use of “not” in a complex syntactic frame (e.g., “Nathaniels not a king”, Drozd, 1995). In contrast, rejection can easily be expressed in a single-word utterance (no!). This covariation further complicates inferences about the sources of developmental differences. Thus, children’s difficulty comprehending negative sentences could reflect a genuine failure to understand the semantics and/or the syntax of negation.

Another possible solution to this puzzle, however, is that children’s behavior is influenced by the pragmatics of the communicative situations that they are in. In most experiments of children’s comprehension of negation, such as the one described above, children hear negative sentences without any supportive pragmatic context. In these

contexts, although the sentences are not *false*, they are perhaps *wrong* according to the Cooperative Principle. In contrast, studies that have elicited spontaneous negations from children tend to use familiar contexts, such as reading picture books in an interactive, game-like setting (Pea, 1982; Hummer, Wimmer, & Antes, 1993). When children hear a true negative sentence and indicate that it is “wrong” (e.g., Kim, 1985), they may be reacting to the infelicity of the sentence rather than its truth value.

If the effect of context on negation is primarily pragmatic in nature, we can use negation as a case study to explore children’s sensitivity to contextual pragmatics. Evidence is mixed regarding whether children are sensitive to Gricean pragmatic principles. Some data suggest that children struggle to make scalar implicatures; for example, accepting a sentence like “Some of the horses jumped over the fence” in a scenario where *all* of the horses had jumped over the fence (Papafragou & Musolino, 2003, see also Noveck, 2001). Other evidence, however, suggests that children are sensitive to pragmatics and can make a number of pragmatic inferences about contextually-based implicatures (Papafragou & Tantalou, 2004; Barner, Brooks, & Bale, 2011; Yoon, Wu, & Frank, 2015; Stiller, Goodman, & Frank, 2015), adjective use (Horowitz & Frank, 2012, in press), and epistemic states (Hochstein, Bale, Fox, & Barner, 2014).

By examining children’s judgments of negative sentences in different contexts, we can ask whether children’s difficulty with negation is actually due to their sensitivity to pragmatics. When children say that a true negative sentence is “wrong,” they may be commenting on the felicity of the sentence rather than the truth value, demonstrating awareness that the same sentence can be a “good” thing to say in one context but a “silly” thing to say in another. We hypothesize that children’s late comprehension of negation is driven by the same pragmatic mechanisms that make negation challenging for adults.

The current studies

In our work we explore the effect of context on the explicit sentence felicity judgments of adults and children. We operationalize context as information that influences expectations about an upcoming utterance. In natural contexts, the information that constitutes “context” could include visual information, words, or actions in the setting of a conversation or another shared task Clark (1996). In our experiments, we use visual contexts in which we manipulate the base rate of certain features in the stimuli. For example, in a context where three people have apples and one person does not, it feels very natural to describe the last person by saying “He doesn’t have apples.” In contrast, in a context where no one is holding any objects (i.e., there are no apples present anywhere), the same sentence is uninformative, because it could refer to any of the people in the scene and there is no reason to discuss apples. In Experiments 1 and 2, we find that adults rate sentences in the former context as more felicitous than sentences in the latter context, supporting the findings of past research in the domain of sentence processing (e.g., Nordmeyer & Frank, 2014a).

We hypothesize that the effects of context on the felicity of negative sentences is driven primarily by Neo-Gricean pragmatic principles. To formalize this prediction, we compare the results of our adult data in Experiments 1 and 2 to the predictions of a Rational Speech Act model (Frank & Goodman, 2012), which predicts that listeners expect speakers to produce maximally informative utterances. In the above examples, the RSA model predicts that the sentence “he doesn’t have apples” is more informative (and therefore more likely to be produced) in the context where everyone else has apples, because in this context the utterance uniquely identifies the referent of the sentence. We find that sentences that are more probable according to this model are rated as more felicitous by our participants.

In Experiment 3, we examine whether preschoolers are sensitive to the contexts of

negative sentences. If children’s difficulty comprehending negative sentences is due to conceptual complexity or an undeveloped semantic understanding of negation, then these children should rate true negative sentences as being bad in any context. If, however, children are actually rating true negative sentences as “wrong” because of the infelicity of pragmatically unsupported negation, then we should see children’s performance improve when they are asked to rate true negative sentences in a supportive, informative context. We find evidence for the latter hypothesis, suggesting that children’s difficulty understanding negative sentences on classic tests of comprehension is due to pragmatic sensitivity rather than a failure to process negation.

In sum, the results of past research on negation suggest a paradox: Negative sentences are produced early by children and frequently throughout adulthood, yet both adults and children appear to find negation difficult on tasks that test their processing and comprehension. Our data suggest that the difficulty of negative sentences is due primarily to the infelicity of negative sentences in unsupported contexts, rather than any representational or syntactic difficulty specific to negation. Negative sentences are more likely to violate the Cooperative Principle, because a negative sentence can refer truthfully to a broader set of referents (e.g., at any given moment, there are more people in the world who are *not* holding apples than people who *are* holding apples, more boxes that do *not* contain chocolate than boxes that *do* contain chocolate, etc.), making negative sentences less informative by default. In the right contexts, however, negative sentences can be just as informative as positive sentences, suggesting that the “difficulty” of negation is actually due to a general difficulty understanding unlikely and therefore infelicitous utterances.

Experiment 1

In Experiment 1, we focused on two possible referents of negation. The same negative sentence can refer to the absence of a feature (nonexistence) or to an alternative

feature (alternative). For example, the sentence “This box doesn’t contain chocolate” might refer to an empty box (i.e., nonexistence) or a box containing an alternative object (e.g., broccoli instead of chocolate). Previous work suggests that adults are faster to identify the referent of nonexistence compared with alternative negation (Nordmeyer & Frank, 2013, 2014b). This finding could be due to processing demands: Identifying a missing feature is easier than identifying a changed feature. It could also reflect an expectation about the pragmatics of negation, however; adults may find alternative negation infelicitous because they expect the sentence to describe present, rather than absent, features.

Experiment 1 explored how different contexts affected participants’ felicity ratings for negative sentences. Half of the participants in Experiment 1 saw sentences presented in a context where none of the surrounding characters had any objects on their table (the none context), and the other half saw sentences presented in a context where everyone except for the target character possessed the negated object, e.g., apples (the target context). In half of the true negative trials, the referent of the negative sentence was a character who had nothing (nonexistence negation). In the other half of true negative trials, the referent of the sentence was a character who had some other object, e.g., a cat (alternative negation). Finally, we examined how different syntactic framings might influence sentence judgments (e.g., “has no apples” and “doesn’t have apples”). As previously mentioned, certain types of negation are often produced in particular syntactic frames; thus, it might be more natural to express denial negation using the “doesn’t have” frame. If the facilitative effect of context on negation is driven by the informativeness of negative sentences in different contexts, then these same effects should appear regardless of the syntactic framing of the sentences.

Method

Participants. We recruited a planned sample of 100 adults to participate in an online experiment through Amazon Mechanical Turk; six participants were excluded for indicating that they were under 18 after completing the experiment, and two participants who did not list English as their native language were excluded. Of the remaining 92 participants, 49 were male and 40 were female, three declined to report gender, and ages ranged from 18 – 65+. We restricted participation to individuals in the United States and paid 35 cents for the experiment, which took approximately five minutes to complete. Participants were randomly assigned to either the none context (n=47) or the target context (n=45).

Stimuli. We created 16 trial items. On each trial, four Sesame Street characters were shown standing behind tables. One character was randomly selected as the “target” character, designated by a red box (see Figure 1). The remaining three characters were designated as “context” characters.

In the none context, the context characters all stood behind empty tables. In the target context, each context character had an identical object on their table. The objects belonged to one of four categories: animals (cat, dog, horse cow), vehicles (car, bus, boat, truck), food (apple, banana, cookie, orange), and household objects (fork, spoon, bowl, plate). Stimuli were created using images from the Bank of Standardized Stimuli (Brodeur, Dionne-Dostie, Montreuil, & Lepage, 2010).

Below the characters was a sentence about the target character. Across the entire experiment, six of these sentences were positive sentences such as “[character] has a/an [object].” Five were negative sentences of the form “[character] has no [object]” and five were negative sentences of the form “[character] doesn’t have a/an [object].”

The target character had a target object, an alternative object (“alternative” trials),

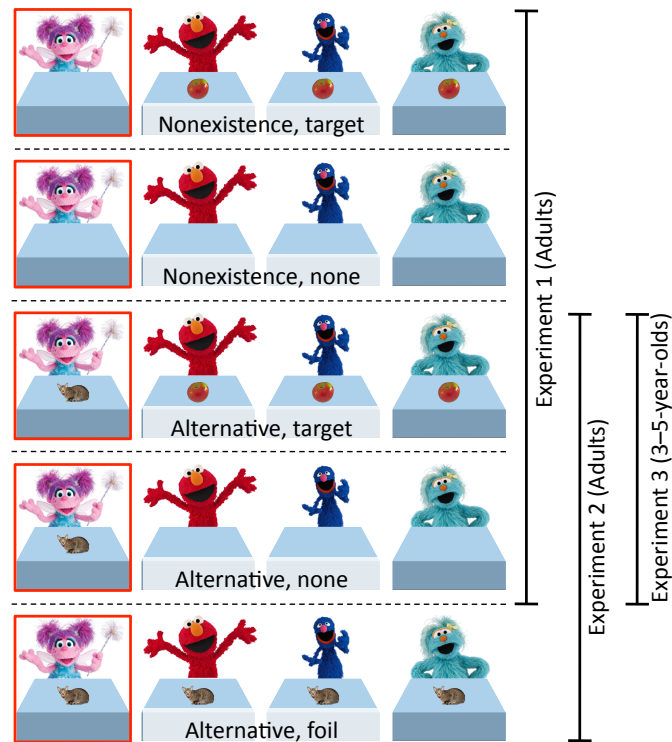


Figure 1. The five types of true negative trials across Experiments 1, 2, and 3. In this example, the target object is an apple and the alternative object is a cat. Below each set of pictures a sentence appeared about the character in red, e.g., “Abby doesn’t have an apple.”

or nothing (“nonexistence” trials), allowing us to examine two different negative concepts (see Figure 1 for depictions of these trials and the different context conditions). Trial conditions were crossed such that each participant saw six true positive trials, two false positive trials (one alternative and one nonexistence), two false negative trials (one “has no” sentence type and one “doesn’t have” sentence type), and eight true negative trials (two “has no”/nonexistence, two “has no”/alternative, two “doesn’t have/nonexistence,” and two “doesn’t have/alternative”). Each of these trial types was randomly assigned to a target object, and trials were presented in a random order.

A slider bar was positioned beneath the sentence, with a seven-point scale ranging from “Very Bad” to “Very Good.” A progress bar at the top of the screen informed participants how much of the experiment they had completed.

Procedure. Participants first saw an instructions screen that briefly described the task and informed them that they could stop at any time. Once participants agreed to participate, they saw an instructions screen that explained the task in more detail. Participants were asked to rate how “good” each sentence was based on how a hypothetical speaker might behave, e.g., “if no one would ever say a particular sentence in this context, or if it is just wrong, rank that as ‘Very Bad,’ but if something is right and sounds perfectly normal, mark it as ‘Very Good.’” Participants were encouraged to use the entire scale. On each trial the pictures, sentence, and slider bar appeared simultaneously, and participants had to make a selection on the scale in order to progress to the next trial.¹

Results and Discussion

Participants treated the scale as expected, consistently rating true sentences much higher than false sentences. Furthermore, participants consistently rated positive sentences higher than negative sentences. There was no effect of context, syntactic frame, or referent type on true positive sentences, false positive sentences, or false negative sentences, likely due to a ceiling effect for true positive sentences and a floor effect for false sentences (see Table 1).

We focused our primary analysis on true negative sentences, because true negative sentences are the most challenging sentence type in classic sentence verification tasks (e.g., Clark & Chase, 1972) and show the strongest context effects in sentence processing tasks (e.g., Nieuwland & Kuperberg, 2008; Nordmeyer & Frank, 2014a). As we predicted, these

¹The experiment can be viewed at <http://anordmey.github.io/neg-tablet/experiments/exp1/negadult.html>

sentences were rated significantly higher in a target context compared to a none context. For example, the sentence “Abby has no apples” was rated higher when all of the other characters *had* apples, compared to contexts where all of the other characters had nothing. This finding supports our hypotheses and replicates patterns previously seen in studies of processing time using explicit felicity judgments. Negative sentences expressing nonexistence were rated as more felicitous than negative sentences referring to an alternative object (see Figure 2).

To test the reliability of these findings, we fit a linear mixed-effects model examining the interaction between context, referent type (e.g., nonexistence or alternative), and negation framing (e.g., “has no X” vs. “doesn’t have X”) on felicity ratings.² Because we were primarily interested in the effects of context on true negative sentences, we focused on these trials in our analyses, though the effects reported here are significant in full models of all sentence types as well.

Examining the model, we found a main effect of context, with true negative sentences presented in a target context eliciting significantly higher ratings than true negative sentences presented in a none context ($\beta = .82, p < .05$). We found main effects of referent type, with nonexistence negation receiving higher ratings than alternative negation ($\beta = .65, p < .01$), as well as a significant effect of negation framing, with sentences of the form “doesn’t have X” receiving higher ratings than sentences of the form “has no X” ($\beta = .71, p < .01$). There were no interactions between negation frame, referent type, and context.

²The model specification was as follows: `rating ~ context × referent type × negation frame + (referent type × negation frame | subject) + (referent type × negation frame | item)`. Significance was calculated using the standard normal approximation to the *t* distribution (Barr, Levy, Scheepers, & Tily, 2013). Data and analysis code can be found at <http://github.com/anordmey/neg-tablet>.

Participants rated negative sentences with the framing “doesn’t have X” higher than sentences with the frame “has no X.” This finding might reflect an overall expectation for denial negation to be expressed using the “doesn’t have X” frame. This preference did not interact with context or referent type, however; participants preferred the “doesn’t have” framing for both alternative negation and nonexistence negation, and rated both sentence frames higher when they were presented in a target context. This finding suggests that effects of context on the felicity of negative sentences are not due to features of a specific syntactic frame.

In previous work we found that adults were somewhat faster to look at the referent of nonexistence negation compared with alternative negation (Nordmeyer & Frank, 2014b). In that experiment, the difference between nonexistence and alternative negation could have arisen because of superficial, stimulus-level differences (i.e., it might be easier to identify a character with nothing than one with an alternative object). Our replication of this result using explicit felicity judgments suggests that this difference is more likely to be due to pragmatic factors. One explanation for participants’ preference for nonexistence negation is that a sentence such as “Abby doesn’t have an apple” is more informative when Abby has nothing compared to when she has an alternative object. In a strong context (e.g., one where everyone else has apples), using negation to point out Abby’s lack of apples is informative because it uniquely identifies her character in the array. When Abby has some alternative object (e.g., a cat), however, there is a *more* informative utterance that a speaker could use (e.g., “Abby has a cat”). The existence of a more informative utterance makes these negative sentences less felicitous, even when they appear in context.

Overall, we found that a pragmatic context increases the felicity judgments of negative sentences. Across all types of negation, participants assigned higher ratings to sentences that were presented with a target context compared to sentences that were presented with a none context. This corroborates previous work in which more

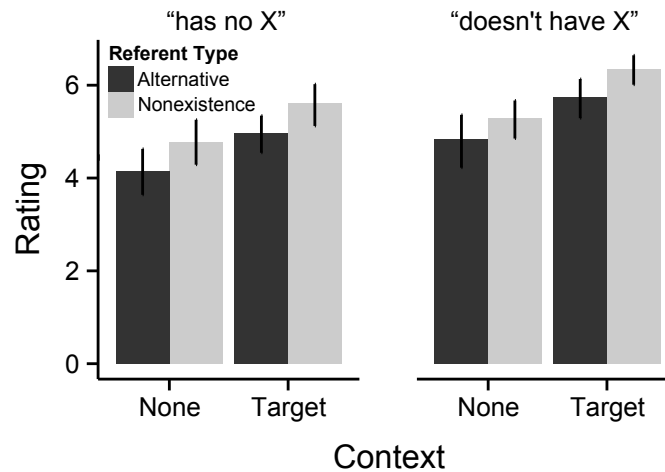


Figure 2. Ratings for different types of true negative sentences in Experiment 1. Sentences of the form “...has no X” are shown on the left, and sentences of the form “...doesn’t have X” are shown on the right. Negative sentences referring to an alternative object are shown in black, and negative sentences expressing nonexistence are shown in gray. Error bars show 95% confidence intervals.

informative negative sentences were processed faster than less informative negative sentences (Nordmeyer & Frank, 2014a). In the next section, we replicate this finding in a within-subjects experiment with an additional type of context to demonstrate that informativeness, rather than the exceptionality of the target picture (e.g., Wason, 1965), has the greatest influence on the felicity of negative sentences.

Experiment 2

In Experiment 2, we examined the effect of three contexts on alternative negation: A target context, in which all context characters had the negated target object (identical to Experiment 1), a none context, in which none of the context characters had any objects (identical to Experiment 1), and a foil context, in which all context characters had an alternative object (e.g., a different object than the one negated in the negative sentences;

see Figure 1). Participants saw all three context conditions throughout the experiment, allowing us to examine within-subject effects of context.

The addition of the “foil” context allowed us to test two competing hypotheses about the pragmatics of negation. In true negative trials with a foil context, all characters had the same objects on their table (e.g., cats), and the negative sentence referred to a different object (e.g., “Abby doesn’t have apples”). Some previous work suggests that a critical element of the effect of context on negative sentences is the fact that the referent of the negative sentence is the “odd one out” (e.g., the exceptionality hypothesis, Wason, 1965). If this is the case, the foil context might be even worse than the none context, because the target of the negative sentence does not stand out from the context. If, however, the effect of context is driven by the informativeness of negation, then there should be no difference between the foil and none contexts because negative sentences are, if anything, slightly more informative in the foil context.

Method

Participants. We recruited a planned sample of 200 adults to participate in an online experiment through Amazon Mechanical Turk. We excluded from our analysis six participants who indicated that they were under 18 after completing the experiment, four participants who did not list English as their native language, and six participants who had participated in a previous version of the experiment. Of the remaining 184 participants, 109 were male and 72 were female, three declined to report gender, and ages ranged from 18 – 65+. We restricted participation to individuals in the United States and paid 40 cents for the experiment, which took approximately seven minutes to complete.

Stimuli. Trials in Experiment 2 had the same structure as trials in Experiment 1, with a small set of exceptions. First, there were 24 trials. Second, all negative sentences were of the form “[character] doesn’t have a/an [object].” Third, on each trial, the target

character either had a target object on their table, or had an alternative object (eliminating the nonexistence trials). Each participant evaluated nine true positive sentences, three false positive, three false negative, and nine true negative trials.

In Experiment 2, context was a within-subjects factor with three levels. In the none context, context characters had nothing on their table, identical to the none context condition in Experiment 1. In the target context, context characters each had a target object on their table, same as the target context condition in Experiment 1. In the foil context, context characters had an alternative object on their table (e.g., all characters have a cat, but the sentence is about the presence/absence of apples; see Figure 1). Each context condition appeared an equal number of times within each trial type.

Procedure. The procedure was identical to Experiment 1.³

Results and Discussion

As in Experiment 1, participants rated true sentences much higher than false sentences, and consistently rated positive sentences somewhat higher than negative sentences. As before, context did not affect ratings for true positive sentences or false sentences (see Table 3). True negative sentences were rated significantly higher when they were presented in a target context compared to either the none context or the foil context (Figure 3), supporting our hypothesis that the effect of context on negative sentence ratings is based on the informativeness of the sentence in context. There was no difference between sentences presented in a foil context and sentences presented in a none context.

We fit a linear mixed-effects model to true negative sentence ratings to test the

³The experiment can be viewed at <http://anordmey.github.io/neg-tablet/experiments/exp2/negadult.html>.

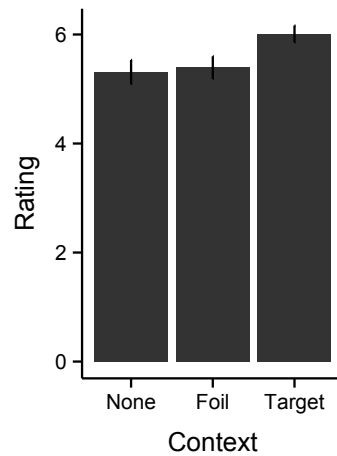


Figure 3. Ratings for true negative sentences in three context conditions in Experiment 2. Error bars show 95% confidence intervals.

effects of context.⁴ True negative sentences presented in a target context received significantly higher ratings than true negative sentences presented in a none context ($\beta = .70, p < .001$). There was no significant difference between sentences presented in a foil context and sentences presented in a none context ($\beta = .09, p = .22$).

In Experiment 2, negative sentences were rated as most felicitous in a context where all of the context characters possessed the negated object. Negative sentences presented in a foil context did not differ significantly from negative sentences in a none context. In the foil context, all characters (including the target character) had the same alternative object. If the pragmatics of negation require the referent of a negative sentence to be the “odd one out” (e.g., lacking a feature that everyone else has), we would expect sentences presented in a foil context to be rated lower than sentences presented in the none context. The fact that there was no difference between these two contexts suggests that this kind of exceptionality is not a necessary feature of a supportive pragmatic context for negation. Instead, negation appears to be pragmatically licensed in contexts where the negative

⁴The model specification was as follows: `rating ~ context + (1 | subject) + (1 | item)`

sentence is highly informative.

Model

Both of the preceding studies found a significant effect of context on participants' ratings of true negative sentences. Why does context have this effect on negative sentences? One hypothesis is that felicity ratings are influenced by the *informativeness* of negative sentences. According to theories of pragmatics, speakers should produce sentences that are appropriately informative based on the context. In a context where most characters have apples and one does not, it is informative to mention the latter character's lack of apples, because this feature is unique to the character being described.

To test this hypothesis we used a recent probabilistic model of pragmatics to make predictions about participants' felicity ratings. Details of this model—the “rational speech act” model of pragmatics—can be found in several previous publications (Frank & Goodman, 2012; Goodman & Stuhlmüller, 2013; Nordmeyer & Frank, 2014a). Here we give a brief sketch of the intuitions behind the model. The probability of a speaker making an utterance in context is defined as being proportional to the informativeness of the word in context minus its cost. Informativeness in context is calculated as the number of bits of information conveyed by the utterance. We assume that the utterance has a uniform probability distribution over its extension in context (e.g., “doesn't have apples” applies to any character without apples, leading to a uniform probability of picking out each individual character without apples). We defined cost as the number of words in the utterance multiplied by a cost-per-word parameter; in our simulations, we did not differentiate between different negative sentence frames, and treated negative sentences as having one word more than positive sentences. Probabilities were normalized over a sparse vocabulary of possible positive or negative utterances that could describe the characters.

Our model predicted the same qualitative pattern as the data from participants'

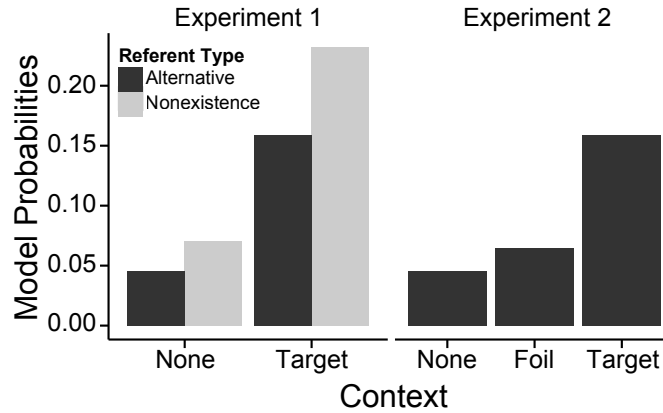


Figure 4. Predictions of a model of the informativeness of an utterance in context. The model predicts how probable a sentence is given a certain context. Best-fitting parameters were used for this simulation (cost = .8), but the qualitative pattern persists over a wide range of parameter values. Negative sentences expressing nonexistence are shown in black, and negative sentences referring to an alternative object are shown in gray.

felicity ratings (Figure 4). Sentences presented in a target context were preferred over sentences presented in a none context, because true negative utterances are more informative when everyone in the context possesses the negated object (e.g., the sentence “Abby doesn’t have an apple” uniquely identifies Abby when everyone else *does* have apples). Nonexistence negation was assigned higher probability than alternative negation, because alternative negation could be described by an equally informative and less costly positive utterance (e.g., “Abby has a cat”).

The model predicted that true negatives presented in a foil context were more probable than true negatives presented in a none context. This due to the relative informativeness of the positive utterance (e.g., “Abby has a cat”) compared to the negative utterance in each context. In the foil context, the positive utterance is true of every character, making it as uninformative as the true negative utterance. In the none

context, this positive utterance uniquely identifies the target character (because none of the other characters have cats), making it a much more informative utterance. Thus, in the none context, the fact that the positive utterance is so much more informative drives down the normalized informativeness of the true negative statement. This leads to a very small effect of the foil context according to the model; this difference is not significant in the experimental data.

Overall, participants' felicity ratings appear to reflect the informativeness of sentences in context. In the next experiment, we turn to children, and ask whether three- to five-year-old children also rate true negative sentences higher in more informative contexts.

Experiment 3

In Experiment 3, we examined whether children, like adults, find negative sentences more felicitous in informative contexts. Past work suggests that children produce negation very early (Bloom, 1970; Pea, 1980), but struggle on classic comprehension tasks for several more years (e.g., Kim, 1985). Children's success in these tasks is sensitive to the context that a negative sentence occurs in (e.g., de Villiers & Tager-Flusberg, 1975; Nordmeyer & Frank, 2014b), which could explain this apparent gap between production and comprehension. If children expect speakers to produce informative utterances, then on classic comprehension tasks they might be reacting to the infelicity of a negative sentence without context rather than its truth value. Under this account, children should perform differently when they are asked to evaluate negative sentences in a supportive context.

To simplify the experiment for children, we focused on alternative negation and varied the context (none vs. target) between subjects. Children were introduced to a puppet who was "just learning how to talk, and sometimes makes mistakes," and were asked to help the puppet learn by indicating on a 5-point smiley face scale whether the

puppet's sentence was "good" or "silly or a mistake." Children indicated their choice by selecting the face that corresponded to their choice on an iPad (for a discussion of using tablets to collect data from children, see Frank, Sugarman, Horowitz, Lewis, & Yurovsky, in press). We deliberately conflated felicity (i.e., "silliness") and truth value in order to encourage children to use the negative side of the scale for both incorrect sentences as well as silly sentences. In past studies using similar paradigms (i.e., Kim, 1985), children treated true negative sentences as "wrong;" here we evaluate if children are in fact reacting to the (in)felicity of these sentences by comparing their performance in infelicitous contexts (the none context) to their performance in felicitous contexts (the target context).

Method

Participants. Families visiting Children's Discovery Museum (CDM) in San Jose, CA were invited to participate in this study. Because our mission at CDM is to recruit a diverse sample of families, we used preset criteria for inclusion in the final sample but recruited inclusively regardless of language background. We also recruited from a broad range of ages, setting a target of 32 included children per age group (three- and four-year-olds) and ending data collection once this minimum had been reached in all groups after the rejection criteria were applied. In exchange for participation, children were given a sticker and a certificate.

Fourteen children were excluded for being outside the target age range (between three and five years old). An additional sixteen children were excluded whose parents indicated that they were exposed to English less than 75% of the time. One child was excluded for parental interference during the task. Twelve children declined to continue the experiment after going through the instructions and practice trials, and five additional children were excluded for completing fewer than half of the trials.

Children who failed to understand the scale, based on their performance on positive sentences, were rejected from the analysis. Children's responses to true positive sentences were coded as correct if they chose a happy face on the scale (4 or 5), and children's responses to false positive sentences were coded as correct if they chose a sad face on the scale (1 or 2). Children who chose an incorrect response on more than 2 trials were rejected from analysis. Eight children were rejected based on this criteria, four three-year-olds and four four-year-olds. These exclusions resulted in a final sample of 69 children whose data was analyzed: 35 three-year-olds (mean age = 3;7, range = 3;0 – 3;11, 20 female and fifteen male) and 34 four-year-olds (mean age = 4;5, range = 4;0 – 4;11, 17 female and 17 male).

Stimuli. Trials in Experiment 3 had the same structure as trials in the alternative/target and alternative/none contexts of Experiments 1 and 2 (see 1), but with only 16 trials. As in Experiment 2, all negative sentences were of the form “[character] doesn't have a/an [object].” Each child evaluated four true positive, two false positive, two false negative, and eight true negative trials. Context was a between-subjects factor, such that each child saw only target contexts or only none contexts.

Procedure. Parents and children were recruited on the floor of the CDM and were led to a small nearby research room. Children were invited to sit at a small table in front of an iPad, next to the experimenter. Parents sat slightly behind their children. Children were first encouraged to play a short pre-training “dots game” in which children needed to tap a set of five colorful dots in random locations; each dot transformed into an “X” when tapped. This pre-training sequence served both as a filler to engage children while answering any parent questions, and helped children select pictures effectively during the rest of the experiment.

Once this pre-training game was complete, children were introduced to the

experimenter's "puppet friend," Furble. The experimenter told children that Furble was just learning how to talk, and sometimes made mistakes. Children were then asked if they wanted to play a game that would help the puppet learn. Once children agreed, the experimenter advanced the iPad to an "instructions phase." During this phase, the experimenter explained that they would look at pictures with Furble, and Furble would say something about one of the pictures. If the puppet said something "really good, or right," children were told to select the happiest face. If the puppet said something "silly, or made a mistake," children were told to use the saddest face. Children were then told to select the slightly happy face if the puppet said something that was "a little bit good," to select the slightly sad face if the puppet said something "a little bit silly or a little bit of a mistake," and to select the neutral face if what the puppet said was "right in-between." After talking about the smiley-face scale, children were given two example trials (a true positive and a false positive), in which the experimenter told the child whether the sentence was good or silly and asked the child to select the corresponding smiley-face.

Once the children completed the instructions phase, they were given a chance to practice. The practice consisted of a true positive trial, two false positive trials, and a final true positive trial. On each trial, the four characters appeared and a red box appeared around the target character. The smiley-face scale appeared below the characters, and the target sentence was printed at the bottom of the page in small print for the experimenter to read (in the voice of the puppet). During the practice trials, the experimenter first identified the target character and then encouraged children to look at all of the other characters as well, e.g., "See Abby? Here's Abby with all of her friends. Let's hear what Furble has to say about Abby." The puppet would then produce a sentence, and children were asked whether Furble had said something good or silly. During the practice trials, if children didn't make a response, or made an incorrect response, the experimenter explained the scale again until children made their own response spontaneously. After

these four practice trials, children were asked if they wanted to continue to help the puppet learn, and children who agreed proceeded to the main experiment.

During the main experiment, trials proceeded in the same way as the practice trials, but the experimenter no longer drew children's attention to every context character. The experimenter would point to the target character and say, "Here's Abby! Let's hear what Furble has to say about Abby." Children were then encouraged to make a response. If children produced a word (i.e., "good" or "silly"), but did not make a response, they were reminded again about which face to choose for good or silly sentences. If children paused for more than five seconds, they were asked if they wanted to hear the sentence again. After each trial, a star appeared at the top of the screen, serving as a progress bar. At the end of the experiment, a large smiley face appeared and children were thanked for helping Furble.⁵

Results and Discussion

As with adults, children rated true sentences significantly higher than false sentences, indicating that the children in this sample were using the scale correctly. Children's ratings for positive sentences were higher than their ratings for negative sentences, though this is likely due partially to the fact that they were trained on positive sentences and partially to the context effects described below (see Table 4).

Focusing on children's performance on negative sentences, children in the target context rated true negative sentences significantly higher than children in the none context, supporting our hypothesis that children are sensitive to the felicity of negation in different contexts. Children rated false negative sentences significantly lower than true negative sentences in both context conditions, suggesting that children in both conditions

⁵The experiment can be viewed at <http://anordmey.github.io/neg-tablet/experiments/exp3/negkids.html>.

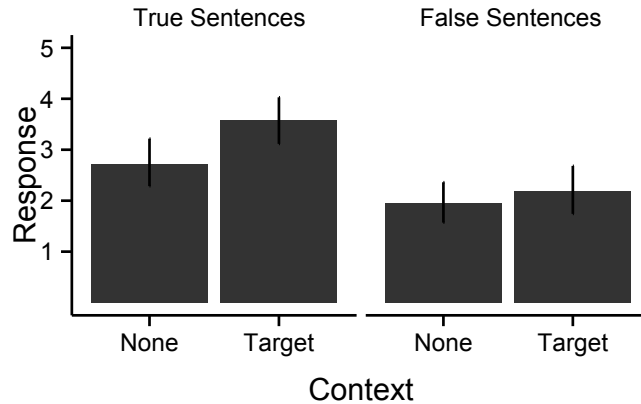


Figure 5. Children’s mean ratings for true and false negative sentences in the none and target contexts in Experiment 3. Responses to true negatives are shown on the left, and responses to false negatives are shown on the right. Error bars show 95% confidence intervals.

heard and processed the negation correctly, even when they rated true negative sentences as “very silly/a mistake” (see Figure 5).

We fit a linear mixed-effects model to test the effects of context and truth value on children’s negative sentence ratings.⁶ Because children were trained, with feedback, on positive sentences to familiarize them with the scale, and exclusion criteria was based on children’s performance on positive sentences, we focus on children’s responses to negative sentences in our analyses. Results of this model can be seen in Table 5. Children rated true negative sentences significantly higher than false negative sentences ($\beta = .74$, $p < .001$), indicating that they correctly perceived and processed the negation and recognized that false negative sentences were incorrect. Although there was no main effect of context on children’s responses to negative sentences, a significant interaction between truth value and context indicated that children in the target context rated true negative

⁶The model specification was as follows: `rating ~ context × truth + (1 | subject) + (1 | item)`

sentences significantly higher than did children in the none context ($\beta = .60, p < .05$), supporting our hypothesis that children's past difficulties responding to true negative sentences is due to the infelicity of these sentences, rather than a difficulty in processing negation.

A linear mixed-effects model that included children's age group did not yield a significant effect of age (three-year-olds vs. four-year-olds), and did not significantly improve model fit ($\chi^2(4) = 2.35, p = .7$). Figure 6 shows histograms of three-year-olds' and four-year-olds' responses to true negative trials, as well as the mean responses to true negative trials in each context condition. Both three- and four-year-olds were significantly more likely to rate a true negative sentence as "very good" in the target context, and were significantly more likely to rate a true negative sentence as "very silly / a mistake" in the none context. This effect of context, however, appears to be more pronounced amongst four-year-olds compared to three-year-olds. To examine the effect of context on true negative sentence ratings in each age group, we separately compared each age groups' responses to true negative sentences in the target context and the none context. A two-sample t-test showed that the difference between true negative ratings in the target context and the none context was not significant for three-year-olds ($t(32) = -1.56, p = 0.13$), but was significant for four-year-olds ($t(32) = -2.09, p < .05$), indicating that four-year-olds rated true negative sentences significantly lower in the none context compared to the target context.

Together these results suggest a pronounced effect of context on children's responses to true negative sentences. Three- and four-year-olds combined rated true negative sentences higher in the target context. In addition, children had no difficulty with false negative sentences, suggesting that their performance in response to true negative sentences is not due to confusion about how negation works or difficulty processing

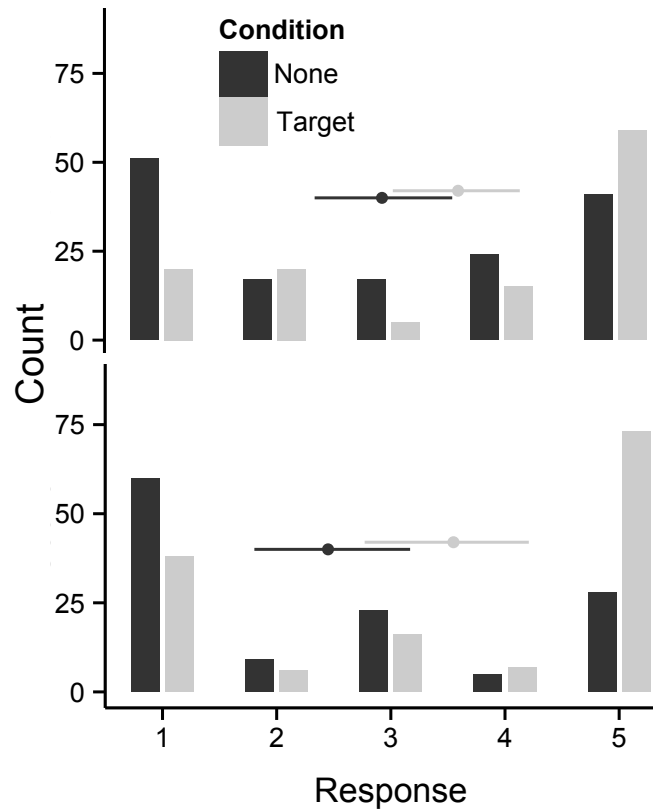


Figure 6. Histograms of children's responses to true negative trials in the target context and the none context. 3-year-olds are shown above, and 4-year-olds are shown below. Black bars represent responses in the none context, and gray bars represent responses in the target context. Points represent the mean response in each group, and horizontal error bars show 95% confidence intervals.

negation, but rather a sensitivity to the pragmatics of negation. Similar to adults, children found true negative sentences to be “silly” or a mistake when they were produced in the absence of an informative context, but were more likely to say that the same sentences were “good” in an informative context, suggesting that children’s responses to these sentences is due to their sensitivity to pragmatics, rather than a difficulty specific to negation.

General Discussion

The same negative sentence can sound perfectly fine in one context, but strange in another. In our experiments, we found that contextual differences led to significantly different pragmatic judgments for otherwise identical true grammatical negative sentences. What is it about the context of negative sentences that elicits these effects?

The negative sentences that received the lowest felicity ratings for both adults and children were alternative negations in a none context. On these trials, participants saw e.g., three characters with nothing, and a character with a cat (see Figure 1). The sentence “Abby doesn’t have apples” referred to the character with a cat. Although this sentence is true, it sounds very odd: Why is the speaker talking about Abby’s lack of apples, which is true of everyone in the context, instead of mentioning the cat? Compare this example to nonexistence negation in a target context: Three characters have apples, and Abby has nothing. Here, the same sentence sounds perfectly natural, because Abby’s lack of apples is unique, and there is little else to say about her. In this latter context, producing a negative sentence is reasonable and perhaps even expected.

Children’s responses to negative sentences in the none context replicate past findings (Kim, 1985): The majority of children said that the true negative sentences were “silly” or “a mistake,” despite the fact that the sentences were true. In contexts where all of the context characters possessed the negated item, however, children’s behavior shifted

substantially, with the majority of children rating the true sentences as “good.” These findings suggest that children’s past “failures” on similar comprehension tasks are likely due to the infelicity of pragmatically unsupported negation. Instead of interpreting children’s responses as “incorrect,” we can interpret their responses as a felicity judgment: Children in the none context are saying that the sentence is wrong or a mistake not because it is *false*, but because it is *infelicitous*.

In addition, children’s performance on this tasks suggests a sensitivity to the contexts of communication more generally. According to Neo-Gricean theories of communication, speakers should produce sentences that are maximally informative and relevant given the context. In our experiment, the none context was both uninformative and irrelevant. Children’s differential responses to true negative sentences in these two contexts suggests that children are aware of these communicative principles, and expect speakers (and puppets) to abide by them.

We did not find evidence for developmental change in children’s responses between three- and four-year-olds. Although the difference between contexts was not statistically significant for three-year-olds, children of both age groups rated true negative sentences higher in the target context compared to the none context. Based on the prior evidence from the negation acquisition literature, this finding is surprising: Past studies of children’s comprehension of denial and nonexistence negation suggest that two- and three-year olds struggle to respond correctly to true negative sentences, but improve by age four or five (Kim, 1985; Nordmeyer & Frank, 2014b). If, however, children’s performance on tests of negation comprehension is pragmatic in nature, it’s possible that even the youngest children in our sample already have the pragmatic abilities to be sensitive to this difference in context. Due to the metacognitive demands of the felicity judgment task, however, a different paradigm would likely need to be used to test younger children’s sensitivity to the contexts of negative sentences.

One consistent finding across all three experiments was the effect of context on true negative compared to false negative sentences. Our primary focus was on true negative sentences, because these have been shown to be influenced by context in the past (Wason, 1965; Glenberg et al., 1999; Nordmeyer & Frank, 2014b) and are the most difficult sentence type in classic sentence verification tasks (Clark & Chase, 1972). It is striking to note, however, that in every experiment we found significant effects of context on the evaluation of true negative sentences, but no effect of context on the evaluation of false negative sentences. One possibility is that this is a floor effect, and that participants rate false sentences at the lowest end of a felicity scale by default. Another possibility, however, is that hearing a false sentence is such a gross violation of the Cooperative Principle that speakers don't consider additional violations, such as an uninformative and irrelevant false sentence.

In the real world, the types of contexts that negative sentences occur in are rich and varied. Many kinds of information might influence the felicity of a negative sentence: Your real-world knowledge, the previous utterances in a conversation, the joint activity that you are engaged in with another speaker, etc. For the sake of these experiments, we chose to test the effects of a simplified set of contexts. The results that we have found here, however, are consistent with past data on the effects of a wider range of contexts on the processing of negative sentences (Wason, 1965; Glenberg et al., 1999; Nieuwland & Kuperberg, 2008; Dale & Duran, 2011; Nordmeyer & Frank, 2014a). This suggests to us that our key finding – that true negative sentences are more felicitous in more informative contexts – is robust to a wide range of contexts.

Most Neo-Gricean theories of communication are based on ideas about how real-time, face-to-face communication works. In our experiments, we used stripped-down contexts with little to no interaction between the “speaker” and the participant. Even in Experiment 3, although a puppet “spoke” the words, children were told to rate the

puppet's utterance on the iPad rather than speaking to the puppet directly, making this task very different from natural conversation. Despite this unnatural communicative setting, we found consistent evidence that both adults and children were sensitive to the context that a negative sentence occurred in. Our findings suggest that the pragmatics of language use are embedded in all contexts in which we use language, and are not merely an optional feature of face-to-face communication.

According to Grice's Cooperative Principle, speakers should produce utterances that are informative with respect to their intended meaning. If listeners expect speakers to abide by this principle, they should find uninformative sentences infelicitous. Our results support this view of communication: Negative sentences that were more informative (and therefore more likely to be produced by a speaker) were given higher felicity ratings. These data suggest that general pragmatic factors, rather than some specific quirk of negation, can explain the felicity of negative sentences in context.

Acknowledgments

This work supported by the NSF GRFP. Portions of this work were presented in (Nordmeyer & Frank, 2015a) and (Nordmeyer & Frank, 2015b).

References

- Barner, D., Brooks, N., & Bale, A. (2011). Accessing the unsaid: The role of scalar alternatives in childrens pragmatic inference. *Cognition*, *118*, 84–93.
- 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.
- Bloom, L. (1970). *Language development: Form and function in emerging grammars*. New York, NY: MIT Research Monograph, No 59.

- Brodeur, M. B., Dionne-Dostie, E., Montreuil, T., & Lepage, M. (2010). The bank of standardized stimuli (boss), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PloS one*, *5*, e10773.
- Cameron-Faulkner, T., Lieven, E., & Theakston, A. (2007). What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child Language*, *34*, 251–282.
- Carpenter, P., & Just, M. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*, 45–73.
- Clark, H. (1996). *Using Language*. Cambridge: Cambridge University Press.
- 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.
- de Villiers, J., & Tager-Flusberg, H. (1975). Some facts one simply cannot deny. *Journal of Child Language*, *2*, 279–286.
- Drozd, K. (1995). Child english pre-sentential negation as metalinguistic exclamatory sentence negation. *Journal of Child Language*, *22*, 583–610.
- 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. C., & Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*, 998.
- Frank, M. C., Sugarman, E., Horowitz, A. C., Lewis, M. L., & Yurovsky, D. (in press). Using tablets to collect data from young children. *Journal of Cognition and Development*.
- Gilkerson, J., Hyams, N., & Curtiss, S. (2002). ‘no i understand negation’: a preferential looking paradigm study of early knowledge of sentential and anaphoric negation.

- Unpublished master's thesis, University of Maryland, College Park.
- Glenberg, A., Robertson, D., Jansen, J., & Johnson-Glenberg, M. (1999). Not propositions. *Journal of Cognitive Systems Research*, *1*, 19–33.
- Goodman, N. D., & Stuhlmüller, A. (2013). Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science*, *5*, 173–184.
- Grice, H. (1975). Logic and conversation. *1975*, 41–58.
- Hochstein, L., Bale, A., Fox, D., & Barner, D. (2014). Ignorance and inference: Do problems with gricean epistemic reasoning explain childrens difficulty with scalar implicature? *Journal of Semantics*, ffu015.
- 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.
- Horowitz, A., & Frank, M. C. (2012). Learning from speaker word choice by assuming adjectives are informative. In *Proceedings of the 34th annual conference of the cognitive science society*.
- Horowitz, A., & Frank, M. C. (in press). Children's pragmatic inferences as a route for learning about the world. *Child Development*.
- Hummer, P., Wimmer, H., & Antes, G. (1993). On the origins of denial negation. *Journal of Child Language*, *20*, 607–607.
- 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.
- Kim, K. (1985). Development of the concept of truth-functional negation. *Developmental Psychology*, *21*, 462.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational*

- implicature*. Cambridge, MA: MIT Press.
- Loder, L. (2006). *Early understanding of negation: The word "not"*. Unpublished doctoral dissertation.
- 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. (2013). Measuring the comprehension of negation in 2-to 4-year-old children. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Nordmeyer, A. E., & Frank, M. C. (2014a). A pragmatic account of the processing of negative sentences. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*.
- Nordmeyer, A. E., & Frank, M. C. (2014b). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25–39.
- Nordmeyer, A. E., & Frank, M. C. (2015a). Children are sensitive to the contexts of negative sentences. In *Society for Research in Child Development 2015 Biennial Meeting*.
- Nordmeyer, A. E., & Frank, M. C. (2015b). Negation is only hard to process when it is pragmatically infelicitous. In *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*.

- Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*, 165–188.
- Papafragou, A., & Musolino, J. (2003). Scalar implicatures: Experiments at the semantics–pragmatics interface. *Cognition*, *86*, 253–282.
- Papafragou, A., & Tantalou, N. (2004). Children’s computation of implicatures. *Language Acquisition*, *12*, 71–82.
- Pea, R. (1980). The development of negation in early child language. In D. Olson (Ed.), *The social foundations of language and thought* (pp. 156–186). New York, NY: Norton New York.
- Pea, R. (1982). Origins of verbal logic: Spontaneous denials by two-and three-year olds. *Journal of Child Language*, *9*, 597–626.
- Stiller, A. J., Goodman, N. D., & Frank, M. C. (2015). Ad-hoc implicature in preschool children. *Language Learning and Development*, *11*, 176–190.
- Wason, P. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*, 7–11.
- Yoon, E. J., Wu, Y. C., & Frank, M. C. (2015). Childrens online processing of ad-hoc implicatures. In *Proceedings of the 36th Annual Meeting of the Cognitive Science Society*.

Table 1

Mean felicity rating for different trial types in Experiment 1, based on ratings given on a 7-point scale.

Context	Sentence type	Truth Value	Referent Type	Mean	Std. Dev.
None	Positive	True	Target item	6.51	0.89
None	Positive	False	Nothing	2.51	1.94
None	Positive	False	Alternative item	2.04	1.82
None	Negative (“doesn’t have”)	True	Nothing	5.29	1.44
None	Negative (“doesn’t have”)	True	Alternative item	4.83	1.92
None	Negative (“doesn’t have”)	False	Target item	2.00	1.64
None	Negative (“has no”)	True	Nothing	4.78	1.68
None	Negative (“has no”)	True	Alternative item	4.15	1.78
None	Negative (“has no”)	False	Target item	1.40	0.77
Target	Positive	True	Target item	6.42	0.76
Target	Positive	False	Nothing	1.88	1.47
Target	Positive	False	Alternative item	1.83	1.73
Target	Negative (“doesn’t have”)	True	Nothing	6.34	1.11
Target	Negative (“doesn’t have”)	True	Alternative item	5.74	1.41
Target	Negative (“doesn’t have”)	False	Target item	1.89	1.60
Target	Negative (“has no”)	True	Nothing	5.61	1.51
Target	Negative (“has no”)	True	Alternative item	4.96	1.41
Target	Negative (“has no”)	False	Target item	1.58	0.99

Table 2

Coefficient estimates from a mixed-effects model predicting true negative sentence ratings in Experiment 1.

	Coefficient	Std. err.	t value
(Intercept)	4.13	0.24	17.14
Context (Target)	0.82	0.33	2.46
Referent type (Nonexistence)	0.65	0.21	3.16
Frame (“doesn’t have”)	0.71	0.25	2.81
Context × Referent type	-0.002	0.27	-0.01
Context × Frame	0.05	0.33	0.14
Referent type × Frame	-0.22	0.24	-0.93
Context × Referent type × Frame	0.22	0.32	0.67

Table 3

Mean felicity rating for different trial types in Experiment 2, based on ratings given on a 7-point scale.

Context	Sentence type	Truth	Value	Mean Std. Dev.
None	Positive	True	6.68	0.54
None	Positive	False	1.62	1.32
None	Negative	True	5.31	1.52
None	Negative	False	1.74	1.55
Foil	Positive	True	6.63	0.59
Foil	Positive	False	1.73	1.52
Foil	Negative	True	5.40	1.42
Foil	Negative	False	1.70	1.49
Target	Positive	True	6.61	0.64
Target	Positive	False	1.72	1.53
Target	Negative	True	6.01	1.09
Target	Negative	False	1.71	1.44

Table 4

Children's mean felicity rating for different trial types in Experiment 3, based on ratings given on a 5-point scale.

Context	Sentence type	Truth Value	Mean	Std. Dev.
None	Positive	True	4.59	0.70
None	Positive	False	1.77	1.04
None	Negative	True	2.72	1.44
None	Negative	False	1.96	1.23
Target	Positive	True	4.61	0.57
Target	Positive	False	1.48	0.83
Target	Negative	True	3.58	1.37
Target	Negative	False	2.18	1.38

Table 5

Coefficient estimates from a mixed-effects model predicting negative sentence ratings in Experiment 3.

	Coefficient	Std. err.	t value
(Intercept)	1.98	0.23	8.55
Context (Target)	0.25	0.33	0.74
Truth Value (True)	0.74	0.17	4.45
Context×Truth Value	0.60	0.24	2.50