CHILD DEVELOPMENT

Child Development, xxxx 2017, Volume 00, Number 0, Pages 1–18

Getting to *No*: Pragmatic and Semantic Factors in Two- and Three-Year-Olds' Understanding of Negation

Tracy Reuter
Princeton University

Roman Feiman
University of California, San Diego

Jesse Snedeker
Harvard University

Although infants say "no" early, older children have difficulty understanding its truth-functional meaning. Two experiments investigate whether this difficulty stems from the infelicity of negative sentences out of the blue. In Experiment 1, given supportive discourse, 3-year-olds (N = 16) understood both affirmative and negative sentences. However, with sentence types randomized, 2-year-olds (N = 28) still failed. In Experiment 2, affirmative and negative sentences were blocked. Two-year-olds (N = 28) now succeeded, but only when affirmatives were presented first. Thus, although discourse felicity seems the primary bottleneck for 3-year-olds' understanding of negation, 2-year-olds struggle with its semantic processing. Contrary to accounts where negatives are understood via affirmatives, both sentence types were processed equally quickly, suggesting previously reported asymmetries are due to pragmatic accommodation, not semantic processing.

Natural language allows us to take a finite store of words and build an unbounded number of sentences. To do this, languages rely on both content and function words. Whereas content words refer to properties, events, and entities that we can think about in isolation, the meanings of function words lie in how they combine with content words and with each other. Take the sentence, "I have not been to the moon." A child might learn about the moon from seeing it, reading *Goodnight Moon*, learning facts about it ("it's *not* made of cheese"), and so on. But there is nothing like that to rely on when learning the meaning of "not." Instead, the child must compare the meanings of parallel sentences with and without "not" to try to close in on the contribution of this one little word to the whole. Because the meanings of function words like "not" are defined by their combinatorial properties, to ask how children learn function words is to ask how they learn some of the combinatorial rules that give language its full expressive power.

For this reason, it is curious that "no" is one of the first words that children learn. According to parental reports, "no" is in the top 10 most commonly produced words among 12-month-olds and is produced by over 95% of toddlers by 24 months (Dale & Fenson, 1996). For adults, "no" can serve as a truth-functional operator in a variety of syntactic and semantic contexts. It can be anaphoric (denying a preceding statement) and a single-word answer to a question. It can act as a quantifier, a sentential operator, or a metalinguistic objection, depending on the context in which it occurs (see Horn, 2001). When children first say "no," they do not appear to assign it its full truth-functional meaning, instead they use it primarily to reject offers and orders. Recent evidence from offline comprehension tasks shows children begin to understand *No* as a truth-functional operator by the beginning of their 3rd year (Austin, Theakston, Lieven, & Tomasello, 2014; Feiman, Mody, Sanborn, & Carey, in press). These experiments used a search task to assess how children interpret positive and negative statements. Children saw a toy and two possible hiding places (e.g., a bucket and a house). The experimenter hid the toy in one of the locations, behind a screen. Then, children heard either a positive or a negative statement about the toy's location (e.g., "It's in the bucket" or "It's not in the bucket"). They found that by 27 months, when

Correspondence concerning this article should be addressed to Roman Feiman, University of California, San Diego Department of Psychology McGill Hall, Room 3316 9500 Gilman Dr., La Jolla, CA, 92093-0109. Electronic mail may be sent to roman.feiman@gmail.com.

0009-3920/2017/xxxx-xxxx
DOI: 10.1111/cdev.12858

children heard, “It’s not in the bucket,” they correctly inferred that the toy was in the house. They reliably searched in the correct location given both affirmative and negative cues, with the same age of comprehension for “no” and “not.”

Although 2-year-olds succeed in using negation for search tasks, children as old as 4 or 5 struggle to understand negation in other contexts, resulting in poor performance in both online preferential looking tasks (Nordmeyer & Frank, 2014) and truth-value judgment tasks (Kim, 1985; see also Donaldson, 1970; Lloyd & Donaldson, 1976; Pea, 1980). For example, in Nordmeyer and Frank’s study, children were asked to “look at the boy who has no apples” when two boys—one with and one without apples—were present on the screen. In the truth-value judgment tasks, children were asked to judge whether a puppet who says that a banana is “not an apple” is right or wrong. Both 2- and 3-year-olds fail in these tasks, and while 4-year-olds have some success, they continue to show significant delays in online processing and higher error rates than adults. What explains children’s difficulties in processing negation in some paradigms, despite their success on other offline measures?

To understand why negation is hard for children, we might first ask why it is hard for adults. Perhaps the factors that pose a processing difficulty for adults are exacerbated to the point of total comprehension failure for children. Adults incur a cost in processing negation in a wide variety of tasks (Carpenter & Just, 1975; Clark & Chase, 1972, 1974; Just & Carpenter, 1971; Kaup, Lüdtke, & Zwaan, 2006; Kaup, Yaxley, Madden, Zwaan, & Lüdtke, 2007). At least some of these difficulties are due to the pragmatic infelicity of negated statements in contexts where they carry relatively little information (Nordmeyer & Frank, under review; Wason, 1965). All else being equal, negative statements are made more felicitous by the prior introduction—explicit or inferred—of the to-be-negated proposition as a question under discussion (QUD), with the possibility of it being true or false. In contrast, an affirmative statement can be a natural way of introducing a new QUD into a discourse (Tian, Breheny, & Ferguson, 2010).

If the prior presence of the relevant QUD is critical, then the prior mention of a proposition within a discourse would be just one way to make negation easier to process. Indeed, world knowledge can also serve to make a negative more plausible than an affirmative. Nieuwland and Kuperberg (2008) found that statements like: “With the proper equipment, bulletproof vests aren’t very dangerous” elicit a larger N400—an ERP signature of processing difficulty—than: “With the proper equipment, scuba diving isn’t very dangerous.” These findings follow from an account on which listeners expect speakers to convey informative messages, and the informativeness of negated statements can vary with world knowledge. Scuba diving is known to be a potentially dangerous activity, and stating that its risks can be mitigated is informative in that context. In contrast, bulletproof vests are designed for safety in dangerous circumstances. Stating that bulletproof vests are not themselves dangerous is at best a non sequitur that does not provide new or useful information, making that statement less felicitous.

Although pragmatic factors explain some of the difficulty of processing negation, the process of semantic composition involving the negation operator may impose a separate and additional processing cost. Even in cases where the relevant QUD is explicitly introduced (Tian et al., 2010) or when the discourse is manipulated to ensure that affirmative and negative statements are equally informative (Nordmeyer & Frank, under review), adults still process negative statements more slowly than their affirmative counterparts. This is just what one would expect on any account of semantic processing in which negating a proposition involves applying a truth-functional logical operator to the corresponding affirmative proposition to change its truth value, whereas affirming a proposition only involves representing it with no corresponding affirmation operator involved (e.g., Clark & Chase, 1972, 1974).

Given that both pragmatic and semantic factors seem to contribute to adults’ difficulties in processing negation, we might wonder to what extent each of these factors are responsible for children’s comprehension failures. DeVilliers and Flusberg (1975) adapted Wason’s (1965) paradigm, which had been used to show the importance of a pragmatically supportive context for adults. They showed children an array of toys and named them aloud. All toys except one were of the same object kind (e.g., seven bottles and one baby doll). The experimenter pointed to a toy, said either, “This is a. . .” or, “This is not a. . .,” and prompted the child to complete the sentence. Children had particular difficulty with “implausible” negatives—when the negation prompt was used while pointing to the more frequent toy (e.g., pointing to one of the bottles, so the correct answer would be, “This is not a baby doll”). On the other hand, when the negation prompt was used with the one exceptional toy, 2-year-olds made no more errors than 3- or 4-year-olds, and few errors overall. Whereas the difficulty of completing implausible negative sentences was reflected in a higher error rate for children, it was reflected in slower reaction times in adults (Wason, 1965). For children, as for adults, understanding negation is easier in a pragmatically supportive context with the rule in this case serving as a QUD (e.g., Is every toy here a bottle?). Negation is then used to point out the exception, giving a negative answer to the QUD.

None of the previous studies where 2- and 3-year-olds fail to understand negation have provided much in the way of pragmatic support; they have simply presented negative sentences out of the blue, together with correct or incorrect referents and with no preceding discourse (Donaldson, 1970; Kim, 1985; Lloyd & Donaldson, 1976; Nordmeyer & Frank, 2014; Pea, 1980). In the present experiment, we ask whether a lack of pragmatic support can fully explain toddlers' failure to process negation. Does easing the pragmatic load enable toddlers to process negated propositions successfully, or does a component of semantic difficulty persist and result in comprehension failure?

To examine the contributions of pragmatic and semantic factors, Experiment 1 looks at 2- and 3-year-olds' online processing of negation when given a pragmatically supportive context that is known to facilitate the processing of negation in adults (Snedeker, Lee, Reuter, & Jiang, 2012). This pragmatic support is provided by a narrative story that sets up a relevant QUD, a different one for each trial. If pragmatic accommodation is the sole reason why children fail to understand negation, then they should consistently succeed when there is a QUD established that clearly motivates the negated statement.

To look at the role of semantic processing, we manipulate the blocking and order of affirmative and negative trials. Affirmative and negative trials in Experiment 1 are intermixed. Although switching responses introduces conflict demands that are difficult for toddlers at this age in general (Diamond, 2013), there is also specific evidence that switching back and forth between affirmative and negative trials may impose a semantic processing cost independent of the pragmatic support given in each trial (see Austin et al., 2014; Feiman et al., in press). In Experiment 2, we block trials by polarity (affirmative vs. negative) and manipulate block order. As negation is an operator that switches the truth value of propositions, constructing syntactically and semantically similar affirmative propositions over a few trials in a row may facilitate the subsequent negation of other, similar propositions. In this case, we may find that seeing an affirmative block of trials first helps children process the negative block that follows.

Finally, to get more purchase on factors that might affect the comprehension of negation, we also looked at children's performance as a factor of their vocabulary. Recent work finds considerable variation in how efficiently young children process language moment to moment (Fernald, Perfors, & Marchman, 2006; Marchman & Fernald, 2008), with vocabulary size and processing efficiency being highly correlated (Borovsky, Elman, & Fernald, 2012).

Experiment 1

Method

Participants

Because previous studies (Austin et al., 2014; Feiman et al., in press) suggested that children begin to comprehend negation around 27 months, the 2-year-olds we tested were just older than this age. Twenty-eight children between 29 and 33 months old (M = 31 months, SD = 1.5 months) and 16 children between 36 and 48 months old (M = 42 months, SD = 3.5 months) were recruited from the greater Boston area and participated between June 2013 and October 2013. All had normal hearing and vision, and 23 were male. Four additional children were excluded from analysis because they were uncooperative during testing.

Materials and Stimuli

The experimenter controlled the study from a Dell Latitude E6410 host laptop with E-Prime 2.0 software (Psychology Software Tools, Pittsburgh, PA, USA). Participants viewed stimuli on a Tobii T-60 remote eye tracker with a 17-in. display (1,280 × 1,024 pixels).

Each session consisted of four unambiguous practice trials followed by eight critical trials. Figure 1 shows the structure of the critical trials. There were two target pictures for each trial: the affirmative target (e.g., the broken plate) and the negative target (e.g., the unbroken plate). Critical sentences used transitive verbs: *break*, *brush*, *close*, *color*, *dress*, *eat*, *fix*, *open*; and the following nouns: *plate*, *dog*, *can*, *star*, *teddy bear*, *apple*, *train*, and *box*. All nouns and verbs were chosen to be familiar; all are produced by over 80% of 30-month-olds according to MacArthur Communicative Development Inventory

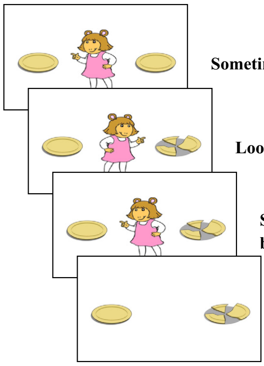

Figure 1. An example of a critical trial in Experiment 1, with the displays children saw on the left and the sentences they heard on the right. The first, second, and third display panels are accompanied by context sentences. The last display is accompanied by the interjection, then the critical sentence, and then the response prompt.

norms (Dale & Fenson, 1996). Critical trials consisted of context sentences spoken in a male voice (e.g., *Sometimes, DW accidentally breaks the dishes. Look! She broke the plate. She was going to break the other plate too, but her mom helped her instead*), followed by a different, female voice producing an interjection (*Oh, I know what happened!*), the critical sentence (*DW broke/didn't break one of the plates*), and finally, the response prompt (*Which one was it? Can you show me?*). Practice trials had a similar structure to critical trials. However, both the verb and final noun phrase in practice trial sentences identified the target (e.g., *DW climbed the big tree*). In contrast, the target in critical trial sentences was determined solely by verb phrase polarity (e.g., *DW broke/didn't break one of the plates*).

There were four experimental lists to counterbalance the polarity of each critical sentence (affirmative or negative) and correct target side (right or left). Within a list, the correct target side for each type of polarity was consistent throughout. This meant that if children showed side bias or perseverated from trial to trial, they would be at chance across the two trial types (i.e., success on one polarity and failure on the other). Whether the story featured Arthur or DW was counterbalanced relative to polarity. Each list had eight critical trials, divided into two blocks. Within each block, two trials were affirmative and two were negative, and trial order was randomized. Two researchers recorded the auditory stimuli. Both researchers were native speakers of English, and used child-directed speech. A male speaker recorded context sentences. The context sentences provided a pragmatically supportive discourse for both affirmative and negative critical sentences. A female speaker recorded the interjection, critical sentences, and response prompt. The response prompt directed the child to touch the target picture at the end of each trial (see Figure 1).

Procedure

Children sat in a chair or on their parent's lap. Before testing, the experimenter calibrated the eye tracker for each child using Tobii Studio software Tobii AB, Stockholm, Sweden, and participants listened to recorded task instructions. During testing, participants saw the four practice and eight critical trials, and were prompted to touch the screen to indicate their choice at the end of each trial (see Figure 1). Participants always received positive feedback (e.g., they saw a cartoon image and heard *Good job!*), regardless of response accuracy. After testing, the experimenter collected the vocabulary measure.

Results

Vocabulary

Children were given the Peabody Picture Vocabulary Test, 4th ed. (PPVT–IV; 3-year-olds: $M = 113$, $SD = 12.6$; 2-year-olds: $M = 101$, $SD = 22.6$). Seven children (four 3-year-olds, three 2-year-olds) did not complete the PPVT, due to time constraints (2) or refusal to participate (5). These children were nevertheless included in all analyses that did not involve these variables.

Online Processing

For coding looks during the critical sentence, the screen was split in half, with all on-screen looks coded as being either to the affirmative or negative target. We then took the proportion of looks to the affirmative target (e.g., the broken plate) as the dependent measure. Figure 2 shows 2- and 3-year-olds' looking in 100 ms segments. However, we did not analyze the data over these segments. As it takes at least 200 ms to plan a saccade (Matin, Shao, & Boff, 1993)—and additional time between saccades to process the visual stimuli and relate them to the critical sentences—where children are looking in one 100 ms period has to be tightly linked to where they were just looking in the preceding 100 ms.

As the correlation between neighboring chunks of time must decrease over longer periods, we aggregated participants' fixations over four longer

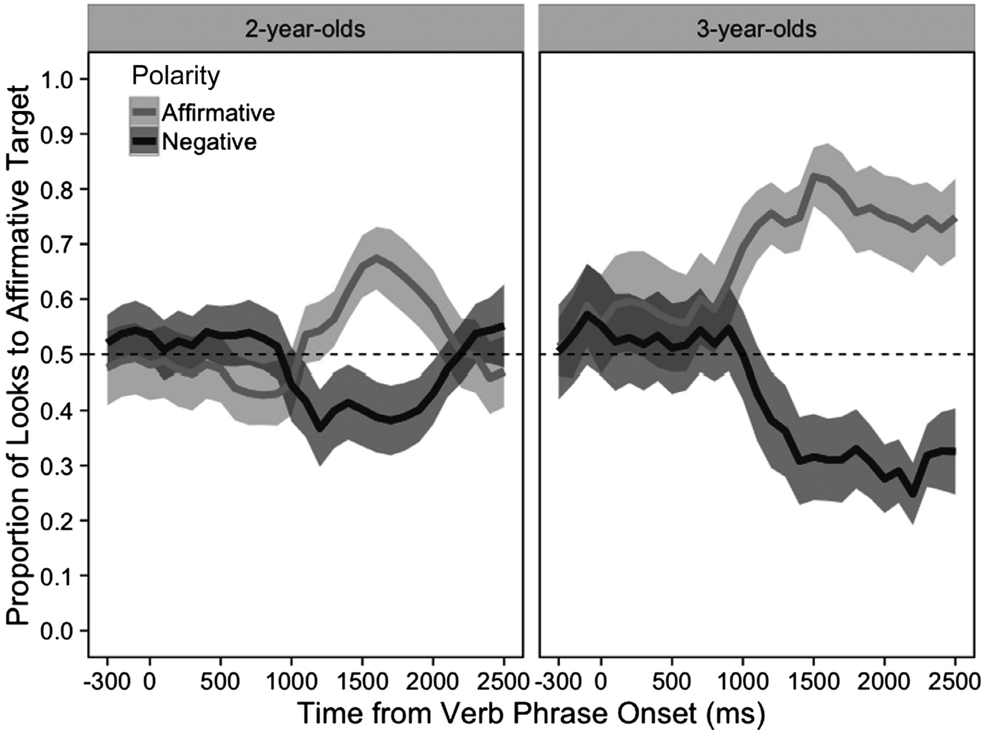

Figure 2. Polarity effect for 2-year-olds (n = 28) and 3-year-olds (n = 16) defined by the proportion of fixations to the affirmative target (e.g., the broken plate) during the critical sentence. Verb phrase onset is at 0 ms. Error bars show 1 *SE* from the mean within each 100 ms time bin, averaged by subjects.

time windows for our analyses (baseline: −500 ms to 0 ms; early: 0–1,000 ms; middle: 1,000–2,000 ms; late: 2,000–2,500 ms; see Table 1), with time 0 indicating the onset of the critical verb phrase (e.g., broke/did not break). The late window was bounded both by children hearing the response prompt after the critical sentence, and by the fact that there is less reason to expect their looking to be guided by the polarity of the sentence further away from that point in the critical sentence. A priori (before looking at the data in any way), time windows after the critical VP (Verb Phrase) were chosen to be 1,000 ms wherever possible, with shorter 500 ms windows where necessary. We expected that these time windows were nevertheless short enough that children would be looking to just one of the target or the distractor pictures within each window. Indeed, aggregating across all samples (60 per second) within a window, over half of children's looking within each time window was either 100% to the target or to the distractor. This justifies—indeed, necessitates—treating looks within each window as a binary variable. We therefore binarized looks within each window: the proportion of affirmative looks each window was coded as 1 if it was above .5, and 0 if below, averaging across all samples within a window. We excluded time bins where the proportion of affirmative looks was equal to exactly .5, because these samples cannot be categorized in the same binary fashion (i.e., they are neutral, and cannot be coded as either affirmative or negative). This excluded 11 of 1,222 time window samples, across all children and trials.

Using this binarized proportion of looks to the affirmative target within each time window as the dependent measure, the online fixation data were analyzed in the R programming language, v3.1.3 using the lme4 package, v4.1.1.7 (Bates, Maechler, Bolker, & Walker, 2015) to build a logit mixed effects model with a maximal random effects structure (see Barr, Levy, Scheepers, & Tily, 2013; Jaeger, 2008). The maximal model was then reduced according to the recommendations in Bates, Kliegl, Vasishth, and Baayen (2015). In all cases below, we describe only this final model and report its results after ensuring that the amount of variance accounted for is not significantly different either from the maximal model, if it converged, or else the near-maximal model containing no correlation coefficients between the random effects (which converged in all cases).

If participants can accurately process polarity information, we expect them to look more to the

Table 1
Results for 3-Year-Olds (n = 16) *and 2-Year-Olds* (n = 28) *During Time Regions Within the Critical Sentence*

	Early window	Middle window	Late window
	ms from verb phrase onset		
	0–1,000	1,001–2,000	2,001–2,500
3-year-olds			
Polarity	$\chi^2(1) = 2.45$	$\chi^2(1) = 7.31$	$\chi^2(1) = 7.29$
	$p = .12$	$p = .007$**	$p = .007$**
PPVT	$\chi^2(1) = 0.03$	$\chi^2(1) = 0.49$	$\chi^2(1) = 0.23$
	$p = .87$	$p = .49$	$p = .63$
Polarity × PPVT	$\chi^2(1) = 2.00$	$\chi^2(1) = 0.01$	$\chi^2(1) = 0.11$
	$p = .16$	$p = .94$	$p = .74$
2-year-olds			
Polarity	$\chi^2(1) = 0.007$	$\chi^2(1) = 17.21$	$\chi^2(1) = 3.35$
	$p = .93$	$p < .0001$***	$p = .07$†
PPVT	$\chi^2(1) = 0.62$	$\chi^2(1) = 0.75$	$\chi^2(1) = 5.91$
	$p = .24$	$p = .39$	$p = .02$*
Polarity × PPVT	$\chi^2(1) = 0.43$	$\chi^2(1) = 6.65$	$\chi^2(1) = 6.41$
	$p = .62$	$p = .01$*	$p = .01$*

Note. Results are from separate models within each age group and time window. PPVT = Peabody Picture Vocabulary Test. Statistical significance is marked as follows. †.1 > p > .05. *p < .05. **p < .01. ***p < .001.

affirmative target (e.g., the broken plate) when hearing affirmative sentences and to the negative target when hearing negative sentences. We refer to the difference in looking during an affirmative sentence relative to a negative sentence as the *polarity effect*. The first model included three binary predictors: *polarity* (affirmative, negative), *age* (older, younger), and *time window* (baseline, early, middle, and late). Polarity refers to whether the critical sentence was affirmative or negative. Age refers to whether participants were in the older (3-year-olds) or younger (2-year-olds) age groups. Time window refers to the periods over which looks were aggregated, with the onset of the critical VP at 0 ms. The p values for each effect were obtained by likelihood ratio tests of the model that included that term relative to a model without it. We explore our data using a series of models, starting from a model including all of the fixed effects of interest and looking for the highest order significant interaction. Where an interaction is significant, we explore what drives the effect by unpacking it into its component variables, first to lower order interactions, and then iterating the process down to pairwise contrasts of levels within the relevant variables. Unlike other variables, the vocabulary factor (high vs. low PPVT) depended on a median split within each level of another variable—age. Therefore, we analyze the effect of this variable only once significant higher order interactions justify looking at effects within each age group.

Following the reduction procedure from Bates et al. (2015), the final model specification in lme4 syntax was:

$$\text{Affirmative Looks} \sim \textit{Polarity} \times \text{Age} \times \text{Time Window} + (1 + \text{Polarity}||\text{Subject}) + (1|\text{Item})$$

Within this model, we were primarily interested in the effect of polarity and its interactions with age and time window. Indeed, the model revealed a significant three-way interaction of polarity, age, and time window, $\chi^2(3) = 9.64$, $p = .02$. To explore this interaction, we built separate models looking at the fixed effects of polarity and age within each time window after VP onset (early, middle, and late). While it would be appropriate to apply the model reduction procedure (Bates et al., 2015) to each model separately, this procedure could lead to different random effect structures across different time windows. It would then be possible that any differences in results across time windows could be due to differences between model structures. We therefore modeled each time window using the same model specification:

$$\text{Affirmative Looks} \sim \text{Polarity} \times \text{Age} + (1 + \text{Polarity}||\text{Subject}) + (1 + \text{Polarity}||\text{Item})$$

We find no significant effects in the early window, but a significant effect of polarity, $\chi^2(1) = 19.00$, $p < .0001$, and a marginal effect of age, $\chi^2(1) = 3.83$, $p = .05$, in the middle window. In the late window, we find highly significant effects of age, $\chi^2(1) = 10.34$, $p = .001$, polarity, $\chi^2(1) = 26.47$, $p < .0001$, and a significant Age × Polarity interaction, $\chi^2(1) = 18.87$, $p < .0001$. These polarity effects reflect that, within 1,000–2,500 ms after VP onset, both 2- and 3-year-olds look more at the affirmative target on affirmative sentences than on negative sentences (see Figure 2). The interaction of polarity with age reflects 2-year-olds' less robust differentiation of affirmative and negative sentences.

To explore this interaction and the main effect of age further, we look within each age group separately. Following Borovsky et al. (2012), we perform a median split by PPVT within each group, and then analyze the interaction of polarity and PPVT within each time window and age group separately. We again used an identical model specification for each time window, with no significant differences between any of these models and more maximal converging versions:

$$\text{Affirmative Looks} \sim \textit{Polarity} \times \text{PPVT} + (1|\text{Subject}) + (1|\text{Item})$$

Table 1 shows the results from each model.

Figure 3 shows high- and low-PPVT 2-year-olds' fixations in 100 ms segments. We find a significant interaction between polarity and PPVT only in the younger age group, in both the middle and late windows. To explore this interaction, we examined the performance of high-PPVT (M = 118, SD = 10.7) and low-PPVT (M = 85, SD = 18.6) 2-year-olds separately in each of these windows. We found that 2-year-olds with higher PPVT scores showed no polarity effect in the early 0–1,000 ms window, $\chi^2(1) = 0.01$, $p = .91$, whereas the polarity effect was robust in the middle 1,001–2,000 ms window, $\chi^2(1) = 16.11$, $p < .0001$, and marginal in the late 2,001–2,500 ms window, $\chi^2(1) = 3.32$, $p = .07$. In contrast, there was no polarity effect for 2-year-olds with lower PPVT scores in either the early, $\chi^2(1) = 1.59$, $p = .21$, or middle, $\chi^2(1) = 1.02$, $p = .31$, time windows, and a marginal effect in the opposite direction (i.e., more looks to the affirmative target on negative trials) in the late window, $\chi^2(1) = 3.05$, $p = .08$.

Offline Responses

Figure 4 shows children's offline action responses broken down by age and sentence polarity. We analyzed offline picture-choice response accuracy using nonparametric tests (Wilcoxon signed rank and Kruskal–Wallis tests), coding responses as binary—touching the side of the screen with either the correct or the incorrect picture. In affirmative trials, the affirmative referent (e.g., the broken plate) was coded as the correct picture for the behavioral response, whereas in negative trials, the negative referent (e.g., the unbroken plate) was coded as the correct picture.

For 3-year-olds, response accuracy was above chance in practice trials ($p < .001$), in affirmative trials ($p = .009$), and in negative trials ($p = .004$). For 2-year-olds, response accuracy was above chance in practice trials ($p < .001$) but at chance in both the affirmative trials ($p = .085$) and negative trials ($p = .60$). Overall, 3-year-olds made fewer errors than 2-year-olds on both affirmative trials ($p = .006$)

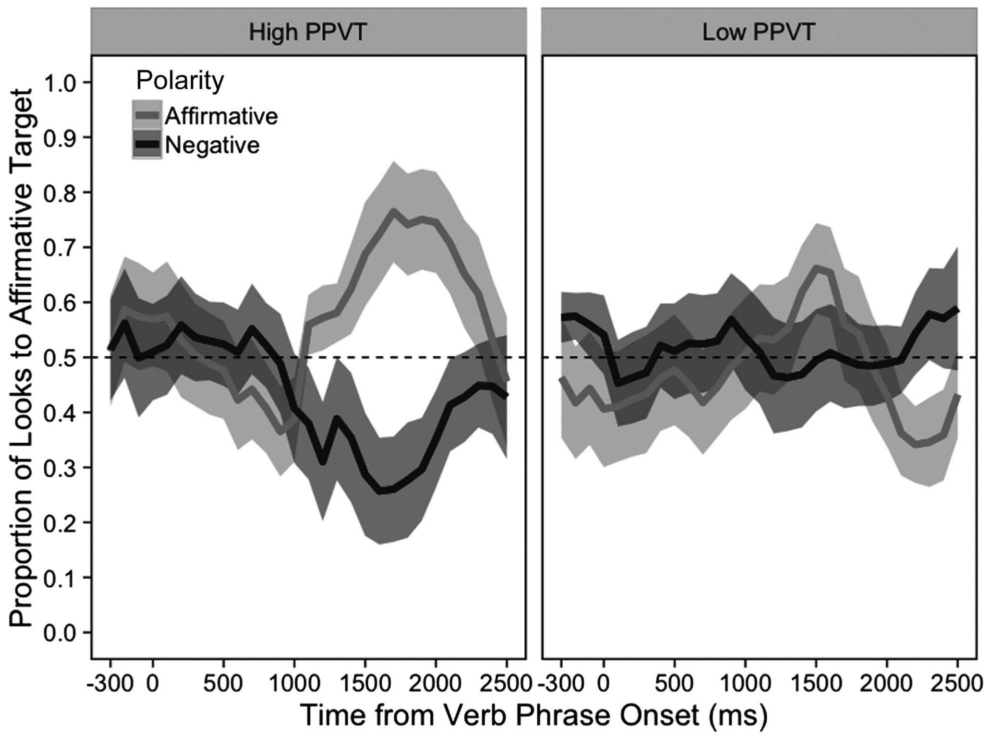

Figure 3. Condition effect for 2-year-olds with higher (n = 12) and lower (n = 12) Peabody Picture Vocabulary Test (PPVT) scores after a median split. Verb phrase onset is at 0 ms. Error bars show 1 *SE* from the mean, averaged by subjects.

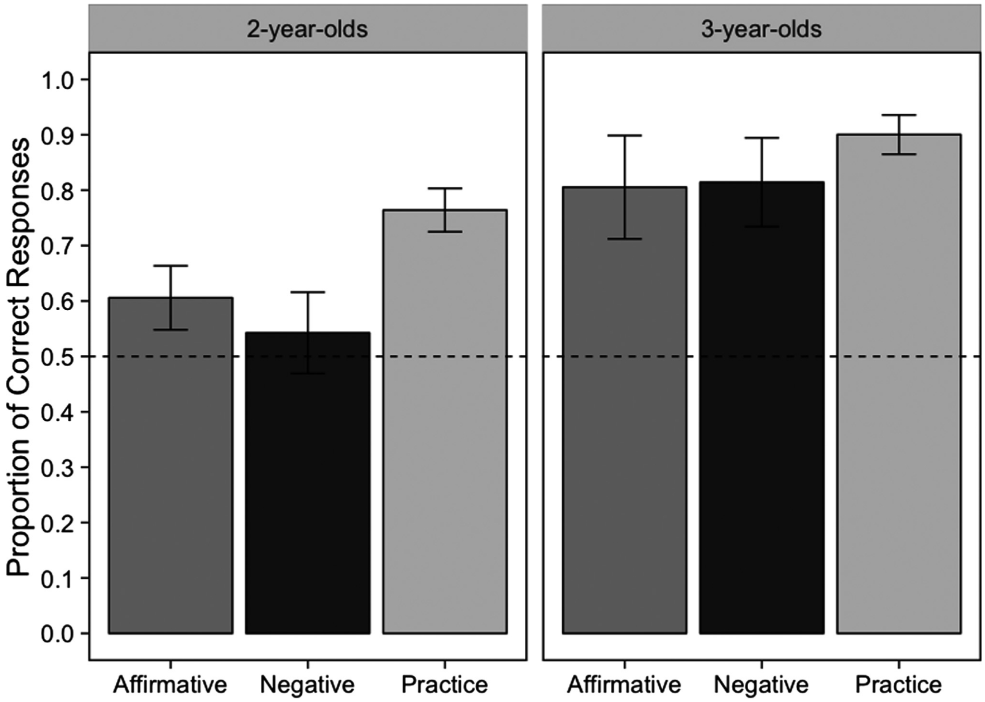

Figure 4. Offline responses for 2-year-olds (n = 28) and 3-year-olds (n = 16). *Y*-axis shows the percent correct choice of the affirmative target picture (e.g., the broken plate) in affirmative trials and the negative target picture (e.g., the unbroken plate) in negative trials. Error bars show 1 *SE* from the mean, averaged by subjects.

and negative trials (p = .02), and for both age groups, accuracy on affirmative trials did not differ from negative trials (2-year-olds p = .64; 3-year-olds p = 1).

As high-PPVT and low-PPVT 2-year-olds showed different patterns in their online processing of affirmative and negative sentence, but 3-year-olds did not, we also looked at the offline performance of high- and low-PPVT 2-year-olds separately. For 2-year-olds with higher PPVT scores, response accuracy was above chance in practice trials (p = .0009) but at chance in affirmative trials (p = .723) and negative trials (p = .245). For 2-year-olds with lower PPVT scores, response accuracy was above chance in practice trials (p = .006), above chance in affirmative trials (p = .042), and at chance in negative trials (p = .277). Within both the high- and low-PPVT group, children's accuracy on affirmative trials did not differ from their accuracy on negative trials, nor did high- and low-PPVT children's performance on negative trials differ from each other.

Discussion

Three-year-old children in Experiment 1 accurately used polarity information when interpreting both affirmative and negative sentences. As Figure 3 shows, they interpreted polarity information shortly after the disambiguating verb phrase, looking to the appropriate target. Picture-choice responses likewise indicated that 3-year-olds arrived at a correct interpretation for both affirmative and negative sentences, selecting the correct target in both cases. This adds to the growing body of recent evidence on 3-year-olds' success in processing negation in pragmatically supportive contexts (Austin et al., 2014; Feiman et al., in press; Snedeker et al., 2012).

Two-year-olds' ability to use polarity information was less robust. Those with high vocabularies differed significantly from those with low vocabularies. Only those with larger vocabularies made use of polarity information to guide their gaze while hearing the sentence, with the effect dissipating at the end of the sentence. However, both 2-year-olds with higher and lower vocabularies were at chance when giving an offline response to negative critical sentences (though those with lower vocabularies were better than chance on affirmatives). There are several reasons online and offline measures might differ. One is that the online measure has greater statistical power and precision of measurement, aggregating many looks within each time window, whereas the offline measure provides only one choice response per trial. Another possibility is that the difference in measures reflects a real difference in the demands made by the two types of responses. Offline responding requires a decision-making process, which translates a dynamically unfolding language comprehension process into a discrete, binary decision—which side of the screen to touch. In addition to language processing, this requires motor planning, a decision procedure, and the interaction and coordination of these processes. Any of these components could pose additional difficulty for 2-year-olds.

Importantly, there was no difference in accuracy between affirmative and negative trials. This pattern of results cannot be due solely to difficulty in processing negation—pragmatic or otherwise. If only negation was difficult, 2-year-olds would have performed at or below chance on negative sentences, but above chance on affirmative sentences. Instead, they were equally unsuccessful with both. This pattern of performance—equivalent success with both sentence types in 3-year-olds, equivalent failure in 2-year-olds—suggests that we succeeded in making negative sentences as pragmatically felicitous as affirmatives.

If the pragmatic infelicity of negation does not explain 2-year-olds' performance on this task, what does? One possibility is that 2-year-olds are able to comprehend one or both sentence types but have difficulty switching between them. Trials in Experiment 1 shifted in polarity every one or two trials and this may have made sentences of both types more difficult to process. Constructing semantic representations of negative sentences involves combining propositional content with a negation operator. Repeatedly applying this operator across multiple sentences may facilitate its processing on consecutive trials, whereas switching back and forth between constructions with and without the operator may require inhibiting its application on affirmative trials, then retrieving and applying it on negative trails despite that earlier inhibition.

Another possibility is that 2-year-olds might only be able to process negative sentences when they are preceded by a set of affirmatives rather than interspersed with them. If there is a cost both to constructing propositional representations and to applying the negation operator, then younger children with a lower semantic processing capacity may only be able to succeed at one task when the other is made much easier. If a child has already constructed a series of affirmative propositions, it may then be easier to construct the similar base proposition of the negated sentence, and then have resources left over for applying the negation operator.

One facet of our data supports the processing hypothesis. If 2-year-olds have difficulty inhibiting one interpretation, they may perseverate on the form they most recently encountered. Within a given list, the correct response for each type of sentence was always on the same side of the screen, so that perseverating by choosing either just the affirmative or just the negative target across both types of sentences would result in response-side perseveration (and chance performance relative to sentence polarity). Consistent with this possibility, we find 13 (46%) of the 2-year-olds selecting the same side for seven or eight of eight total critical trials.

In Experiment 2, we attempted to ease the processing demands as much as possible to see if this would improve 2-year-olds' comprehension of negation on both the online and offline measures. To do this we made four changes to the design. First, we introduced distinct visual referents (e.g., a bowl and a plate, instead of two plates) and repeated the same four verbs across affirmative and negative trials to allow better encoding of events during the discourse. Next, we simplified the dialog by combining the critical sentence and response prompt (e.g., *Show me the one DW ate/didn't eat*). Third, we included additional practice trials with response-contingent feedback to ensure children understood the task. Finally, and most importantly, Experiment 2 used a blocked design: Participants either heard four affirmative stories followed by four negative stories or vice versa.

In Experiment 1, we had found that pragmatic felicity alone was not sufficient for 2-year-olds to succeed in interpreting polarity. Our goal in this experiment was to ascertain whether easing the processing burden (in this same pragmatically felicitous context) would allow them to interpret these utterances correctly. We thought that the repetition of either affirmative or negative sentences within a given block might facilitate semantic processing for each type of sentence by reducing the need to inhibit the operator across trials. Alternatively (or additionally), if negative sentences are particularly difficult, and if processing similar affirmatives facilitates the processing of subsequent negatives, then children should do better when they hear the block of affirmative sentences first. If, on the other hand, affirmatives do not facilitate processing negatives, children may also perseverate in their affirmative responses through to the negative block, so that their performance on negatives would in fact be worse when the negative block comes second.

Experiment 2

Method

Participants

We recruited 48 children (M = 35 months, SD = 6 months) from the greater Boston area, and children participated between February 2014 and July 2014. All had normal hearing and vision, and 26 were male. We excluded three children from analysis because they were uncooperative during testing. As in Experiment 1, we tested two age groups. The younger group (N = 28) was 29–33 months old (M = 31 months, SD = 1.4 months) and the older group (N = 16) was 36–48 months old (M = 42 months, SD = 3.9 months).

Stimuli and Procedure

Figure 5 shows the structure of the critical trials, which was simpler in Experiment 2 than it had been in Experiment 1. A single female voice was used for all auditory stimuli. There was no interjection ("Oh, I know what happened") and no separate response prompt. Instead, the critical sentence cued the response (e.g., *Show me the one DW ate/didn't eat*). As in Experiment 1, there were two visual items in each trial. For critical trials, these were the affirmative target (e.g., the eaten apple) and the negative target (e.g., the uneaten banana).

Each session consisted of four touch practice trials, four story practice trials, and eight critical trials. The purpose of touch practice trials was to familiarize children with touching pictures on the screen to "show" the narrator the correct picture. Touch practice trials included two images, no preceding discourse, and simple commands (e.g., given an image of a chair and a box, the child was told to, *Show me the chair*). Story practice trials had set up a discourse parallel to that in the critical trials. However, instead of having a pronoun that can only be resolved by understanding the polarity of the sentence, the final noun in the story practice trials identified the target (e.g., *Show me the cat*).

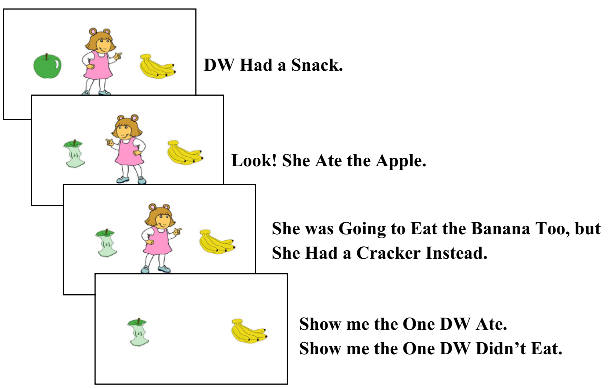

Figure 5. Example visual and auditory stimuli for Experiment 2 critical trials.

After the practice trials, children got eight critical trials. In contrast to practice trials, the target in critical trial sentences was determined solely by verb phrase polarity (e.g., *Show me the one DW ate/didn't eat*). Critical sentences used transitive verbs: *break*, *drink*, *eat*, and *wash.* Each verb occurred once in the negative and once in the affirmative block. Critical sentences also used familiar nouns: *plate*, *bowl*, *bus*, *train*, *water*, *milk*, *juice*, *soda*, *apple*, *banana*, *cookie*, *ice cream*, *duck*, *fish*, *truck*, and *car*. If children selected the correct image, they received positive feedback—smiling Dora the Explorer—and heard *Good job!* If children selected the distractor image, they saw a different picture of Dora and heard *Try again!* Unlike in Experiment 1, children received response-contingent feedback during the practice trials (but not the critical trials). Two children repeated a touch practice trial, and two repeated a story practice trial. All four children selected the target on the first repetition.

During testing, participants completed eight practice trials and then eight critical trials (four affirmatives followed by four negatives, or vice versa). On critical trials, children received positive feedback for each response, regardless of accuracy. There were four experimental lists, which counterbalanced the polarity of the first block (affirmative or negative), and the affirmative target noun (e.g., apple or banana) across children. Verb order, whether the story featured Arthur or DW, and the side of the correct response (right or left) were randomized within each block. After testing, the experimenter collected the PPVT vocabulary measure.

Results

Vocabulary

Children completed the PPVT–IV (3-year-olds: M = 123, SD = 14.6; and 2-year-olds: M = 112, SD = 12). Three children (two 3-year-olds, one 2-year-old) refused to complete the PPVT.

Online Processing

The dependent variable was operationalized just as in Experiment 1, taking the same approach to data analysis. Figure 6 shows the polarity effect within each age group and across time. The critical sentence in Experiment 2 was shorter than in Experiment 1, and there was no separate response prompt. Consequently, the time between the disambiguating VP and the child's action was much shorter than in the first study and variable across children. Thus, we chose to analyze fixations within a single time window that began at the VP onset ("ate" or "didn't"). To determine the endpoint of this window, we found the first 100 ms time window in which track loss was equal to or > 50% (suggesting that children were looking away from the screen or that tracking had been disrupted by their movement). The resulting window was 1,500 ms long. As in Experiment 1, we coded the fixations on each trial as being primarily to the affirmative (1) or to the negative side (0), because we once again found that the proportion of looking time on each trial was best characterized by a binary distribution. We coded every trial as a look to the affirmative target or a look to the negative target, depending on whether participants looked at the affirmative target more or less than half of the period. We excluded time bins where the proportion of affirmative looks was equal to exactly .5, which excluded 11 of 1,183 samples.

As in Experiment 1, we are primarily interested in the *polarity effect* and how it might vary depending on children's age and which type of sentence they saw first and any interactions between polarity and the other variables. We first look for the effects of age, polarity, and the manipulation of block order, using three binary predictors: *polarity* (affirmative, negative), *age* (older, younger), and *block order* (affirmatives first, negatives first). Following the same model reduction procedure as in Experiment 1, the final model specification was:

$$\text{Affirmative Looks} \sim \text{Polarity} \times \text{Age} \times \text{Block Order} + (0 + \text{Polarity}|\text{Subject}) + (0 + \text{Block Order}|\text{Item})$$

We find a highly significant three-way interaction of polarity, age, and block order, $\chi^2(1) = 9.21$, $p = .002$, as well as a significant two-way interaction between age and block order, $\chi^2(1) = 3.89$, $p = .048$, and marginal interactions between polarity and block order, $\chi^2(1) = 3.77$, $p = .05$, and polarity

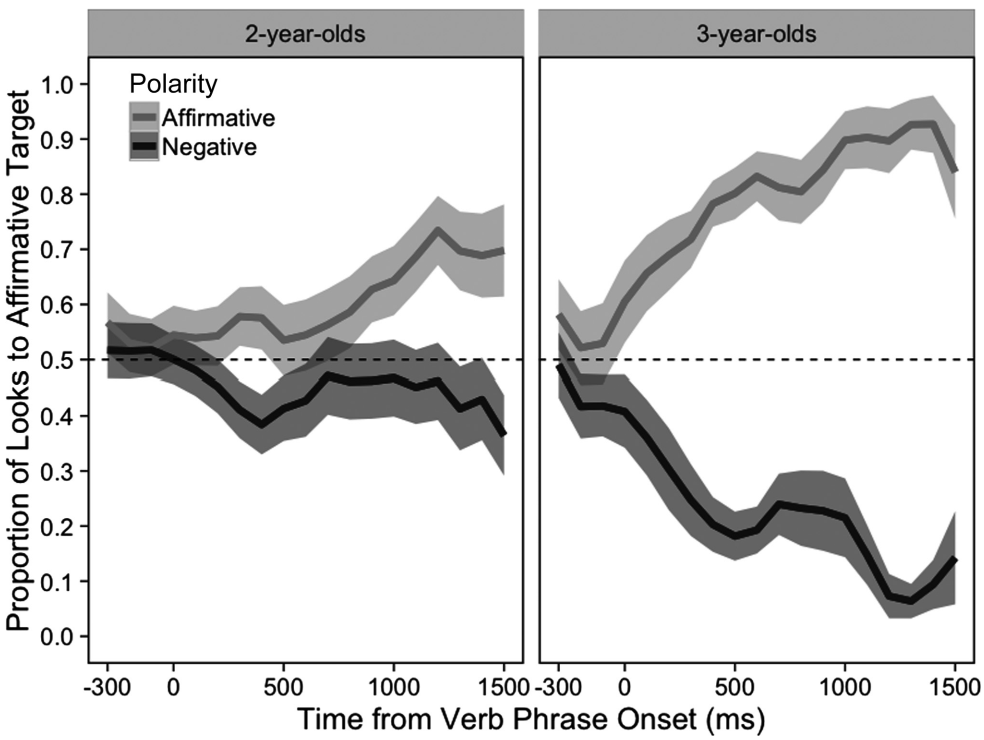

Figure 6. Condition effect for 2-year-old (n = 28) and 3-year-old (n = 16) children. Verb phrase onset is at 0 ms. Error bars show 1 *SE* from the mean, averaged by subjects.

and age, $\chi^2(1) = 3.31$, $p = .07$. To explore these interactions, we constructed separate models looking at the fixed effects of polarity and block order within each age group. We took the same approach as in Experiment 1, with the model within each age group specified as:

$$\text{Affirmative Looks} \sim \text{Polarity} \times \text{Block Order} + (1 + \text{Polarity}||\text{Subject}) + (1 + \text{Polarity}||\text{Item})$$

We found a significant effect of polarity, $\chi^2(1) = 11.74$, $p = .0006$, and a significant interaction between polarity and block order, $\chi^2(1) = 8.1$, $p = .004$, in the 2-year-olds. In the group of 3-year-olds, there was also a significant effect of polarity, $\chi^2(1) = 20.02$, $p < .0001$, and a marginal interaction with block order, $\chi^2(1) = 3.64$, $p = .06$. To explore these interactions further, we looked at the effect of polarity within each age group and within each block order separately. These models contained only random intercepts, following the usual reduction procedure. We found that, whereas 3-year-olds show a significant polarity effect both when they see affirmative trials first, $\chi^2(1) = 20.01$, $p < .0001$, and when they see negative trials first, $\chi^2(1) = 26.75$, $p < .0001$, 2-year-olds only show a polarity effect when seeing affirmative trials first, $\chi^2(1) = 12.48$, $p = .0004$, but not if negative trials come first, $\chi^2(1) = 0.17$, $p = .68$ (Figure 7).

Finally, as we did in Experiment 1, we look for the effects of vocabulary on processing of affirmatives and negatives. After a median split by PPVT within each group, we looked at an interaction of polarity and PPVT within each age group separately. Again, these models contained only random intercepts. Unlike in Experiment 1, we find no interaction of polarity and PPVT, and no main effect of PPVT in either age group (all ps > .05).

Offline Responses

Figure 8 shows offline response accuracy for each age group. As in Experiment 1, we analyzed offline response accuracy with Wilcoxon signed rank and Kruskal–Wallis tests. For 3-year-olds, response accuracy was above chance in practice trials, in affirmative trials, and in negative trials (ps < .001), with no difference in accuracy between affirmative and negative trials ($p = 1.00$). For 2-year-olds, response accuracy was above chance in practice trials ($p < .001$), in affirmative trials ($p = .014$), and in negative trials ($p = .001$). Again, there was no difference in accuracy between affirmative and negative trials ($p = .256$).

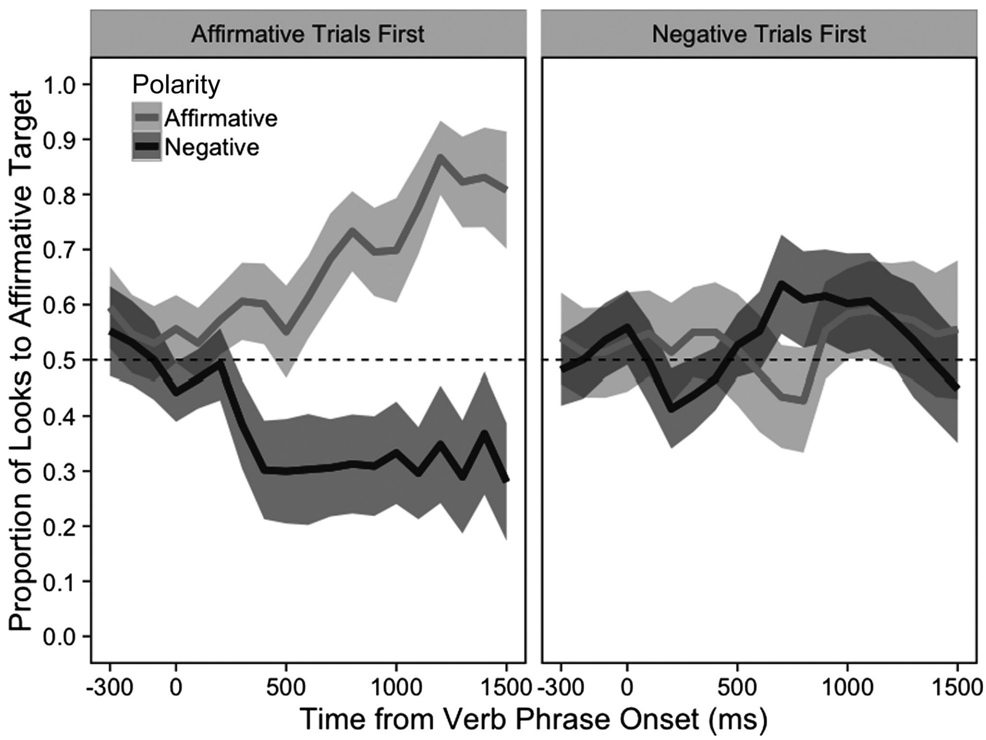

Figure 7. Condition effect for 2-year-olds that saw affirmative trials ($n = 14$) and negative trials ($n = 14$) in the first block. Verb phrase onset is at 0 ms. Error bars show 1 *SE* from the mean, averaged by subjects.

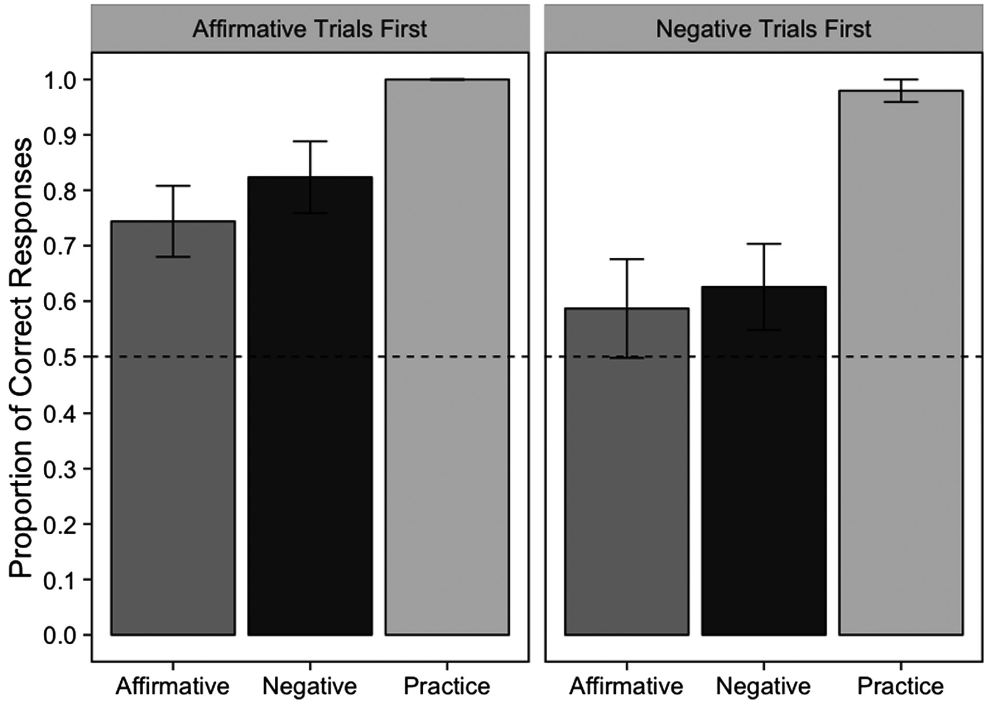

Figure 8. Responses for 2-year-olds that saw affirmative trials (n = 14) and negative trials (n = 14) in the first block. Error bars show 1 *SE* from the mean, averaged by subjects.

As only 2-year-olds who saw affirmative trials first succeeded on the online measure, we looked at their offline responses split by block order as well. For those who saw affirmative trials first, response accuracy was above chance on practice trials ($p < .001$), on affirmative trials ($p = .009$), and on negative trials ($p = .002$). However, for those who saw negative trials first, response accuracy was above chance on practice trials ($p < .001$) but at chance on both affirmative trials ($p = .43$) and negative trials ($p = .09$). In both cases, accuracy did not differ between the trial types. Offline and online success therefore appear to be driven by the same factor—2-year-olds succeed on both affirmative and negative trials only when affirmatives are presented first, on both online and offline measures.

Discussion

As in Experiment 1, we found that 3-year-olds could process negation within a supportive discourse context, this time when affirmative and negative trials were presented in blocks. Although 2-year-olds, as a group, struggled with both negative and affirmative sentences in Experiment 1, the blocking manipulation in Experiment 2 showed that their performance depends on the order of trials. Two-year-olds succeeded in understanding both forms when the affirmatives were presented first, but failed to comprehend either form when negatives were presented first. Affirmative trials facilitated—rather than interfering with—performance on subsequent negative trials.

Unlike in Experiment 1, vocabulary did not moderate 2-year-olds' success in Experiment 2. The key difference between Experiments 1 and 2 is the intermixing versus blocking of affirmative and negative trials. That we find an effect of vocabulary when trials are intermixed but not when they are blocked suggests that vocabulary may matter when there is a high processing load but not when negative sentences are easier to process. Note, however, that 2-year-olds in Experiment 2 also had a mean PPVT 1 *SD* higher than the ones in Experiment 1. This raises the possibility that vocabulary has the same effect on both tasks but that the effect of vocabulary on negation processing is nonlinear: Children in the low-average and high-average ranges may differ from one another, whereas mildly and moderately precocious children may be quite similar.

Another difference from Experiment 1 is that 3-year-olds in Experiment 2 appear to be showing a polarity effect by the time of the onset of the disambiguating VP. This is likely a consequence of blocking—over the course of a block, participants can learn to predict the correct response (the *verbed* or the *unverbed* object) before hearing the disambiguating phrase. Critically, because they do not receive any feedback on critical trials, they would only develop this expectation if they were able to process the polarity of the sentence in the first place. Note also that, because sentence polarity and the correct side of the screen were perfectly coupled in Experiment 1, children might have circumvented further processing of negation by simply associating the word "didn't" with the correct side of the screen (but again, only if they initially understood its meaning). However, 3-year-olds in both Experiment 2 and in Snedeker et al. (2012) succeed even when correct side and polarity are decoupled.

Experiment 2 was designed to ease the semantic processing load involved in comprehending negation. Modifying the design of Experiment 1, we shortened each trial by combining the critical sentence and the response prompt, included additional, response-contingent practice trials, and introduced visually distinct items for the target and distractor in each trial. We found that with all these changes, 2-year-olds' performance depended on the order in which they received the affirmative and negative trials. When negative trials came first, 2-year-olds still failed in processing both affirmative and negative sentences. When affirmative trials came first, they now succeeded on both.

General Discussion

Experiment 1 explored whether young children can interpret negation when the discourse context makes both negative and affirmative sentences pragmatically felicitous. In such contexts, 3-year-olds interpreted negative sentences as quickly and accurately as affirmatives, as evidenced in both their looking and their actions. Two-year-olds, however, struggled with both the affirmative and negative sentences. Two-year-olds with higher vocabularies did somewhat better on this task, distinguishing affirmative and negative sentences online, but still failed to respond correctly on the offline behavioral measure that followed. Experiment 2 revealed a factor that determined 2-year-olds' success: seeing the set of affirmative sentences before encountering the negatives. Under these circumstances, the 2-year-olds interpreted the negative utterances as quickly as the affirmatives (looking to the correct side of the screen in both cases) and were accurate more often than not in their action responses. However, when the negative block of sentences came first, 2-year-olds failed not only on this block, but also on the subsequent block of affirmatives.

The Acquisition of "No," "Not," and "Didn't"

Prior work investigating children's acquisition of the negation words "no" and "not" has produced a surprising finding: Despite differences between both the grammatical properties and input frequencies of these two words, children come to understand the logical, truth-functional meanings of "no" and "not" at the same age, shortly after they turn 2 (Austin et al., 2014; Feiman et al., in press). We tested comprehension of "didn't" here, and our findings allow us to tentatively add it to this list. The addition is only tentative because we did not find an age of total failure in younger 2-year-olds to contrast with an older age of success, so that we cannot be sure "didn't" is learned at just the same time as "no" and "not." The 2-year-olds we tested were 31 months old on average, whereas Austin and colleagues find earliest success with both "no" and "not" at 24 months, and Feiman and colleagues at 27 months. To allow for a more direct comparison across the three words, we would need to test the comprehension of "didn't" at these younger ages. If, however, we were to find that the logical negation meaning of all three words is acquired at the same age, it would provide further support for the thesis that there is a single limiting factor preventing the acquisition of these words earlier (see Feiman et al., in press). This limit may be purely linguistic. There may be a common difficulty in mapping all three words to the concept of negation. In that case, the linguistic difficulty would need to be general enough to explain why learning three quite different words is difficult in just the same way. Alternatively, the limit may be conceptual, so that children younger than two cannot yet think negated thoughts, and therefore cannot map any word to this meaning. The challenge in that case is to explain how they come to be able to think such thoughts.

Semantic Processing Limits 2-Year-Olds

The decisive effect of block order for 2-year-olds —only understanding negatives if they come after affirmatives—demonstrates that nonpragmatic factors play a role in their comprehension of negation. The pragmatic felicity of each critical sentence was set up by the preceding narrative discourse within that trial, which introduced the referents, the agent, and the relevant action. All these features were the same regardless of block order. There were no discourse connections that linked the events in different trials—no bridge between DW not eating an apple on one trial and breaking a plate on a later one—and thus no pragmatic reason why having some affirmative trials first would help process subsequent negative trials (but not vice versa). What the critical sentences in different trials did share is most of their lexical content, their syntactic and semantic structure. This suggests that success on the initial affirmative trials might affect subsequent performance on negative trials by facilitating other aspects of language processing.

Negation, in this context, involves applying a semantic operator to the proposition in the relative clause. To construct this proposition, the child must correctly determine the syntactic structure of the utterance and its logical form (its semantic structure). Prior work with both adults and children has shown that the structure of one sentence affects how we process subsequent sentences. Specifically, constructing a particular syntactic structure or logical form appears to facilitate subsequent construction of the same form, even when the referents and lexical content changes (Bock, 1986; Feiman & Snedeker, 2016; Huttenlocher, Vasilyeva, & Shimpi, 2004; Raffray & Pickering, 2010; Thothathiri & Snedeker, 2008; see Pickering & Ferreira, 2008 for review). In the present study, the affirmative trials may have served as structural primes for the closely matched negatives. Correctly interpreting "Show me the one DW ate" requires building much of the structure needed to understand "Show me the one Arthur didn't wash." Perhaps 2-year-olds typically fail to understand negated sentences because they lack the processing capacity to construct new propositions from the ground up and then negate them. When building the proposition is made easier (by practice or priming), this frees up the resources needed to successfully apply the negation operator. By the time children are 3, this bottleneck has disappeared, due perhaps to greater facility with negation, improved language processing abilities, or both. Note, however, that although 3-year-olds no longer need the semantic scaffolding of affirmatives to succeed, that does not mean they do not face any semantic processing difficulty at all. Indeed, adults incur some semantic processing cost to negation, even in supportive discourse contexts (Nordmeyer & Frank, under review; Tian et al., 2010). Where this difficulty is surmountable for both adults and 3-year-olds, it is sufficient to cause failure in 2-year-olds, unless there is additional semantic scaffolding, as from affirmatives.

Perhaps the most unexpected aspect of these results is that when children saw negative sentences first, it disrupted their processing not only of these sentences but also of the subsequent affirmatives. These are constructions that 2-year-olds clearly understand when they are not preceded or interspersed with negative sentences, as they succeed with affirmatives in the first block. We see two possible explanations for this strange failure. First, processing difficult negatives may require more warm-up and practice with the task. Easier affirmative sentences can provide this practice if they come first. If children are at first confused for several trials by a negative they cannot understand, they may give up before the easier affirmative sentences appear. We found no evidence to support this in our data; the children continued to attend to the stimuli and select answers on each trial. However, we cannot rule out the possibility that, despite their compliance, they had mentally disengaged—sitting through the task while being confused and error prone.

Another more intriguing possibility is that affirmatives are processed differently after a series of negations than they would be in other contexts. An affirmative coming out of the blue may simply be represented as a proposition. But a negative may only be felicitous in response to a QUD, so that processing a series of negatives involves considering a series of QUDs and marking each one as false. When processing affirmatives right after negatives, children may continue to represent the QUDs in each trial, with the affirmatives now causing them to mark each proposition as *true*. The difficulty with affirmatives following negatives may stem from a difficulty in providing any response whatsoever relative to a QUD—with marking a proposition *either* as true or false.

Three-Year-Olds Succeed When Felicity Conditions Are Met

In contrast to 2-year-olds, the semantic processing involved in negation does not seem particularly difficult for 3-year-olds, as long as the sentence is felicitous in the discourse. Our older children did not show an order effect in Experiment 2, but succeed no matter whether affirmative or negative trials come first. This is curious, given that 3-year-olds have failed in previous online studies of negation processing, as well as in some offline measures (Donaldson, 1970; Kim, 1985; Lloyd & Donaldson, 1976; Nordmeyer & Frank, 2014; Pea, 1980). If a pragmatically felicitous discourse is critical for their success in processing negative sentences, there should be pragmatic explanations for the previously reported failures.

In Kim's (1985) truth-value judgment task, children produced errors when judging a true negative statement (saying that a cup is "not a bear"), frequently responding that such statements were "wrong" instead of "right." The mention of bears in this context is surprising: A bear is just one of an infinite number of things that a cup is not, and Kim provided no discourse context that set up bears as particularly relevant to the conversation. In Nordmeyer and Frank's (2014) task, children were presented with a boy holding apples and another boy holding either a different object or nothing at all. They were then asked to "look at the boy who has no apples." Again, apples had never been mentioned previously. Older children and adults first looked to the boy who did have apples, before shifting to the correct referent. Younger children (2- and 3-year-olds) got stuck looking at the apple boy and never made it to the correct referent at all. Although we did not manipulate how much the discourse supported our negative sentences, the success of 3-year-olds in understanding negation in our task relative to the failure of this age group in tasks with less supportive contexts suggests that pragmatic felicity plays an important role in 3-year-olds' comprehension of negative sentences.

Processing the Affirmative to Process the Negative

It is worth dwelling on the fact that both adults and children in Nordmeyer and Frank's (2014) task look to the affirmative target first, with the younger children getting stuck there. Much of the recent work on negation has suggested that processing the affirmative is a critical first step in adults' semantic processing of the negative—first, you fully process and represent the affirmative—incorporating it into the discourse and evaluating it relative to the world—and only then negate it (Clark & Chase, 1972, 1974; Kaup et al., 2006, 2007; Orenes, Beltrán, & Santamaría, 2014). It is, however, also possible that an initial look to the affirmative reflects part of the pragmatic accommodation of negation rather than a step in semantic analysis. If the negative sentence comes out of context, we may construct its affirmative counterpart in an attempt to establish a relevant discourse—*aha! That boy has apples, so that's why we're calling the other boy the one who doesn't (as opposed to, say, the one in the blue shirt)*. If initial looks to the affirmative target are the consequence of an attempt to accommodate the negative statement into a sketchy discourse, we should see them disappear when the QUD is already clearly defined, making the negative independently plausible. However, if these looks are a necessary step in semantic processing, then we should observe them any time a negative sentence is encountered, independently of its pragmatic felicity.

In both of our experiments, when either 2- or 3-year-olds process the sentence successfully, we find that they interpret the affirmatives and negatives with equal speed and facility. Both 3-year-olds in Experiment 1 and 2-year-olds in Experiment 2 show no sign of looking to the affirmative target (e.g., *the broken plate*) before the negative target (*the unbroken plate*). In other words, when the QUD concerns which plates were being broken, processing a sentence about the plate that was *not* broken is as fast as processing a sentence about the plate that was, and no intervening looks to the broken plate are necessary. It therefore appears that the looks to the affirmative referent in prior studies are a consequence of pragmatic accommodation of the negative and not a necessary step in the semantic processing of negation.

Tian et al. (2010) report evidence from adults that is consistent with this claim. They present negative and affirmative sentences out of the blue, but they compare two constructions: simple declaratives, which require pragmatic accommodation, and cleft constructions, which directly mark the QUD. Their participants were presented with a negative sentence and either a matching (negative) or a mismatching (affirmative) picture and were asked to say whether the picture matches the sentence. When they are given ordinary negative sentences (e.g., "Jane didn't cook the spaghetti"), they are faster to respond to a mismatching picture (a picture of the affirmative—e.g., the cooked spaghetti), consistent with both the semantic processing hypothesis—one must build the affirmative on the way to processing the negative—and the pragmatic accommodation hypothesis. However, when the negative is presented in a cleft construction (e.g., "It was Jane who didn't cook the spaghetti"), participants are faster to respond to a *matching* picture (the negative referent—e.g., *uncooked* spaghetti). Tian and colleagues argue that the cleft construction sets up a QUD, which licenses the negation. These results, and ours, suggest that when negation is pragmatically felicitous, participants do not need to process the affirmative on the way to the negative.

Note that there is no contradiction between this claim and our preceding claim that semantic scaffolding of affirmatives aids in processing negatives. It may not be necessary to process the affirmative proposition *as a full-fledged, independent proposition*, complete with incorporation into the discourse and verification relative to the state of the world, in order to attach a negation operator to it. But it may nevertheless be the case that—in fact, it seems logically necessary that—processing an affirmative proposition involves almost all of the same semantic compositional steps as processing its negated counterpart (all, except for the inclusion of a costly negation operator). Thus, processing affirmatives may still facilitate processing similar subsequent negatives, even if the online processing of those negatives does not involve first forming the corresponding affirmatives. If children are to learn what "no," "not," or "didn't" mean, they must do so in the context of sentences in which these words are used. They must be able to understand the rest of the sentence well enough to figure out that the inclusion of these words serves to negate its meaning.

Conclusion

Both 2- and 3-year-olds are capable of processing negated sentences and responding to them appropriately, with evidence converging on two being the youngest age at which children understand three different expressions of negation: "no," "not," and "didn't." However, for children as for adults, there appear to be two separate and separable challenges in the processing of negation—one semantic and the other pragmatic. Three-year-olds readily process negated statements when they are an answer to a QUD already present in the discourse. This suggests that their poor performance in other tasks (e.g., Kim, 1985; Nordmeyer & Frank, 2014) reflects difficulty with the pragmatic accommodation rather than a problem in the construction of negated content. Two-year-olds, on the other hand, encounter substantial difficulty with the semantic processing of negation. They fail, even in a felicitous context, unless they are provided with a prior scaffold of parallel affirmative sentences. In contrast, for 3-year-olds, the order in which they hear affirmative and negative sentences makes no difference. As long as the felicity conditions are met, the interpretation of negatives is rapid and efficient. Taken together, our result suggests that, at least in the specific case of negation, semantic processing limitations are overcome earlier in language development than difficulties with pragmatic accommodation.

Additionally, our findings suggest that the two-step process that had been supposed to be a semantic signature of negation may merely reflect the pragmatic accommodation of infelicitous negation. If the semantic interpretation of negation was a two-step process, then we would expect that both 2- and 3-year-olds would consistently look to the affirmative target on first encountering negation and perform worse on negative trials than affirmative. Under pragmatically felicitous contexts, we instead find both the same time course of processing and the same success rate in behavioral responding for negatives as for affirmatives. We propose that children only entertain the affirmative counterpart of negative utterances when they must do so in order to establish a discourse where the negative statement is felicitous.

Acknowledgments

The authors thank all participating families, and Carmen Lucy Saura and Maria Massiani for their assistance in data collection. We thank Manizeh Kahn, Duygu Ozge, Tanya Levari, and Annemarie Kocab for helpful discussion. Michael Frank, Ann Nordmeyer, and an anonymous reviewer also provided insightful commentary.

References

Austin, K., Theakston, A., Lieven, E., & Tomasello, M. (2014). Young children's understanding of denial. *Developmental Psychology*, *50*, 2061. https://doi.org/10.1037/a0037179

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. doi: https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). *Parsimonious mixed models*. arXiv preprint arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01.

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355–387. https://doi.org/10.1016/0010-0285(86)90004-6

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology*, *112*, 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*, 45. https://doi.org/10.1037/h0076248

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology*, *3*, 472–517. https://doi.org/10.1016/0010-0285(72)90019-9

Clark, H. H., & Chase, W. G. (1974). Perceptual coding strategies in the formation and verification of descriptions. *Memory & Cognition*, *2*, 101–111. https://doi.org/10.3758/bf03197499

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*, 125–127. https://doi.org/10.3758/BF03203646

DeVilliers, J. G., & Flusberg, H. B. T. (1975). Some facts one simply cannot deny. *Journal of Child Language*, *2*, 279–286. https://doi.org/10.1017/S0305000900001100

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135. https://doi.org/10.1146/annurev-psych-113011-143750.

Donaldson, M. (1970). Developmental aspects of performance with negatives. In G. B. Flores D'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics* (pp. 397–410). Amsterdam, The Netherlands and London, UK: North-Holland.

Feiman, R., Mody, S., Sanborn, S., & Carey, S. (in press). What do you mean, no? Toddlers' comprehension of logical "no" and "not" . *Language Learning and Development*. Advance online publication.

Feiman, R., & Snedeker, J. (2016). The logic in language: How all quantifiers are alike, but each quantifier is different. *Cognitive Psychology*, *87*, 29–52. https://doi.org/10.1016/j.cogpsych.2016.04.002

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*, 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Horn, L. R. (2001). *A natural history of negation*. Chicago, IL: University of Chicago Press.

Huttenlocher, J., Vasilyeva, M., & Shimpi, P. (2004). Syntactic priming in young children. *Journal of Memory and Language*, *50*, 182–195.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446. https://doi.org/ 10.1016/j.jml.2007.11.007

Just, M. A., & Carpenter, P. A. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*, 244–253. https://doi.org/10.1016/s0022-5371(71)80051-8

Kaup, B., Lüdtke, J., & Zwaan, R. A. (2006). Processing negated sentences with contradictory predicates: Is a door that is not open mentally closed? *Journal of Pragmatics*, *38*, 1033–1050. https://doi.org/10.1016/j.pragma.2005.09.012

Kaup, B., Yaxley, R. H., Madden, C. J., Zwaan, R. A., & Lüdtke, J. (2007). Experiential simulations of negated text information. *The Quarterly Journal of Experimental Psychology*, *60*, 976–990. https://doi.org/10.1080/17470210600823512

Kim, K. J. (1985). Development of the concept of truth-functional negation. *Developmental Psychology*, *21*, 462–472. https://doi.org/10.1037/0012-1649.21.3.462

Lloyd, P., & Donaldson, M. (1976). On a method of eliciting true/false judgments from young children. *Journal of Child Language*, *3*, 411–416. https://doi.org/10.1017/S0305000900007273

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood. *Developmental Science*, *11*, 9–16. https://doi.org/10.1111/j.1467-7687.2008.00671.x

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information processing time with and without saccades. *Perception & Psychophysics*, *53*, 372–380.

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth is not too hard to handle an event-related potential study on the pragmatics of negation. *Psychological Science*, *19*, 1213–1218. https://doi.org/ 10.1111/j.1467-9280.2008.02226.x

Nordmeyer, A. E., & Frank, M. C. (2014). The role of context in young children's comprehension of negation. *Journal of Memory and Language*, *77*, 25–39. https://doi.org/10.1016/j.jml.2014.08.002

Nordmeyer, A. E., & Frank, M. C. (2015). *Negation is only hard to process when it is pragmatically infelicitous. In Proceedings of the 37th Annual Meeting of the Cognitive Science Society*.

Orenes, I., Beltrán, D., & Santamaría, C. (2014). How negation is understood: Evidence from the visual world paradigm. *Journal of Memory and Language*, *74*, 36–45.

Pea, R. D. (1980). The development of negation in early child language. In D. R. Olson (Ed.), *The social foundations of language and thought: Essays in honor of Jerome S. Bruner* (pp. 156–186). New York, NY: W. W. Norton.

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*(3), 427.

Raffray, C. N., & Pickering, M. J. (2010). How do people construct logical form during language comprehension? *Psychological Science*, *21*, 1090–1097. https://doi.org/10.1177/0956797610375446

Snedeker, J., Lee, M., Reuter, T., & Jiang, M. (2012). *Negation in children's online language comprehension: Evidence for rapid semantic analysis*. Talk at the Boston University Conference on Language Development. Boston, MA.

Thothathiri, M., & Snedeker, J. (2008). Syntactic priming during language comprehension in three-and four-year-old children. *Journal of Memory and Language*, *58*, 188–213. https://doi.org/10.1016/j.jml.2007.06.012

Tian, Y., Breheny, R., & Ferguson, H. J. (2010). Why we simulate negated information: A dynamic pragmatic account. *Quarterly Journal of Experimental Psychology*, *63*, 2305–2312. https://doi.org/10.1080/17470218.2010.525712

Wason, P. C. (1965). The contexts of plausible denial. *Journal of Verbal Learning and Verbal Behavior*, *4*, 7–11. https://doi.org/10.1016/S0022-5371(65)80060-3