

Parents' and Children's Production of English Negation

Masoud Jasbi¹, Annika McDermott-Hinman, Kathryn Davidson, and Susan Carey²

1. Introduction

Previous research has proposed several stages for children's production of negative morphemes (Klima and Bellugi 1966; Wode 1977; Cameron-Faulkner, Lieven, and Theakston 2007; Thornton and Tesan 2013). For example, Cameron-Faulkner, Lieven, and Theakston (2007) proposed that *no* appears before *not*, and *not* before the contracted form *n't*. Klima and Bellugi (1966) proposed that negation appears as a sentence modifier (outside the sentence) before it moves between the subject and the verb. They also proposed that *can't* and *don't* are learned as unanalyzed wholes before their positive auxiliary variants. However, comprehension studies have not produced any evidence for such stages yet (Austin et al. 2014; Feiman et al. 2017; Reuter, Feiman, and Snedeker 2018). This discrepancy can be explained in two ways. First, the lack of evidence may be due to limitations in comprehension studies. Second, it is possible that the proposed stages are limited to production and cannot be generalized as stages in comprehension as well. This paper presents two corpus studies that support the second possibility. The results suggest that some previous stage hypotheses do not hold generally across children and may have been limited to a few children. Furthermore, for stages that do hold across children, there is evidence to suggest they are limited to production and may not generalize to comprehension. In the following section, we explain the previous stage hypotheses proposed for the development of negation. Section presents our first study that . Section presents our second study .

2. Background

Klima and Bellugi (1966) provided the first and most influential account of negation development. They used fortnightly recordings of mother-child conversations for three children in the Brown (1973) corpus: Eve (18-26 months), Adam and Sarah (26-50 months). They divided child utterances into three stages and suggested that in Stage 1, the syntactic category of negation (NEG) includes *no* and *not*, produced before or after a sentence "nucleus", i.e. noun and verb phrase without tense or inflection (NEG+S or S+NEG). Examples include: "No singing song", "No the sun shining", "No

¹University of California Davis, jasbi@ucdavis.edu

²Harvard University

money”, “No play that”, “Wear mitten no”, “No fall!”, and “Not a teddy bear”. It was hypothesized that auxiliary negatives like *don’t* and *can’t* are not produced or understood at this stage. In Stage 2, children add *can’t* and *don’t* as unanalyzed wholes to their list of negators, and move negation inside the sentence, between the subject and the verb phrase (NP+NEG+VP). The main evidence for *can’t* and *don’t* being unanalyzed wholes in this stage was the absence of positive auxiliary variants like *can* and *do*. Typical examples at this stage include “I can’t/don’t see you”, “I don’t want it”, “There no squirrels”, “He no bite you”, and “I no want envelope”. In Stage 3, auxiliary verbs like *can’t* and *don’t* are re-analyzed as AUX+NEG, additional negative auxiliaries like *won’t* and *isn’t* are produced, and positive auxiliaries like *can* and *do* are produced for the first time (NP+AUX+NEG+VP).

However, further investigations proved the first stage to be controversial. Bloom (1970) studied another three children (Kathryn, Eric, and Gia) between 19-27 months and did not find evidence for a sentence-external stage of negation (NEG+S / NEG+S). Children started with isolated *no* and once they produced multi-word utterances, they mostly combined *no* and *not* with noun and verb phrases (*no/not*+NP/VP). Nevertheless, Bloom (1970) reported that Kathryn produced some instances of sentence-internal negation with *no* such as “Kathryn no like celery”. Lord (1974) studied her own child Jennifer (19-26 months) and found no instances of sentence-external negation or sentence-internal *no*. She reported that her child started with single “no” utterances before 24 months and between 24-26 months started combining *no/not* with nominals, and *can’t/don’t* with verb phrases (*no/not*+NP and *can’t/don’t*+VP). Her conclusion was that the development of negation varies in children and not all of them go through the stages described by Klima and Bellugi (1966).

However, Wode (1977) used crosslinguistic data to support and expand Klima and Bellugi (1966)’s account. He compared productions of two German children (19-26 months), a Swedish child (20-42 months) from Lange and Larsson (1973), and English-speaking children from Bloom (1970) and Klima and Bellugi (1966). He proposed four stages: 1. one-word stage with only *nein*, *nä/nej*, or *no*; 2. multiword anaphoric stage where the single words from stage 1 are used as a response to a previous utterance followed by other words (e.g. “no, outside!” or “nein, Milch”); 3. multiword non-anaphoric stage where a single-word negative like *no* is used sentence-externally instead of sentence-internally (e.g. “nein sauber” for “I don’t want to be cleaned” or “no close” for “I can’t close the box”); 4. multiword intra-sentential negation where negation has moved inside the sentence (e.g. “Kathryn no like celery”, “I can’t open it”, or “ich habe nicht geschlafen”). In response to this proposal, Park (1979) argued that Wode (1977)’s account relied on insufficient evidence given that it used only 13 examples and no proper

distributional analysis. Furthermore, Park (1979) presented data from three German speaking children around 21-25 months that did not match Wode (1977)'s developmental stages.

The debate continued with de Villiers and de Villiers (1979) suggesting that previous studies provided little empirical evidence to support a general sentence-external stage. They investigated productions of Adam (27-31 months), Eve (18-22 months), and their own child Nicholas (23-29 months) and found very few sentence-external negatives with overt subjects that allowed for assessment of sentence boundary. They pointed out that even among these instances, many could plausibly be anaphoric. Despite these arguments, Déprez and Pierce (1993) used examples from children's productions in English, French, and German to provide a novel syntactic analysis for presentential negation in child language within the Principles and Parameters framework (Chomsky 1993). They argued that instead of negation moving from outside the sentence inside as Klima and Bellugi (1966) suggested, it is the subject NP that fails to move outside, from inside the VP. They suggested that child data is in line with the VP-internal subject hypothesis in adult grammar (Koopman and Sportiche 1991). However unlike previous studies, they had counted utterances with omitted subjects as instances of presentential negation (or rather VP-internal subjects) as well.

In response to Déprez and Pierce (1993), Stromswold and Zimmermann (2000) studied negation in five German-speaking children (Julia, Inga, Andreas, Kathrin, and Nicole) between 17 and 29 months. They found that out of 689 examples of negation, only one could plausibly support the hypothesis that at an early stage the negator can surface to the left of the subject and pre-sententially. Drozd (1995) provided a similar but large-scale analysis for English. Using data available from 123 children in CHILDES between the ages of 11 and 40 months, the study looked at utterances beginning with *no*, *not*, and *never* and used the available linguistic context to classify them as anaphoric or non-anaphoric. The study found a total of 456 instances of pre-sentential negation, out of which only 31 (6.7%) could be classified as instances of non-anaphoric pre-sentential negation. He argued that the best explanation for such rare distribution of presentential negation is that they are meta-linguistic uses of negation. In other words, the child's "no the sun shining" is similar to the adult version "don't say the sun is shining (that's wrong)", and not some stage in the development of negation.

Two relatively recent studies have focused on Klima and Bellugi (1966)'s second stage, where children are reported to produce non-adult-like infinitival negatives with *no*, *not*, and *don't* (e.g. "He no/not/don't bite you"). At this stage, children are hypothesized to not differentiate these forms and consider all as variants of negation. They are also hypothesized to not analyse *don't* as auxiliary plus negation and rather consider it as an unanalyzed whole.

The evidence was considered to be the absence of positive auxiliary forms like *do*. Schütze (2010) provided a quantitative analysis of negation in the speech of five children (Abe, Adam, Sarah, Nina, Ross) between 2 and 5 years of age. He showed that the non-adult-like infinitival negatives are quite rare, never exceeding 5% of children’s total productions. Instead he found that the only common error reaching about 10% of productions is non-agreeing *don’t* in sentences with third-person singular subjects (e.g. “He don’t bite you”). He proposed a grammatical account that could predict such errors.

Thornton and Tesan (2013) disagreed with Schütze (2010) and following Klima and Bellugi (1966) contended that children at the second stage have not yet identified *n’t* as a separate form of negation. To explain this stage, they proposed that children start with the hypothesis that negative words like *not*, *not*, and *don’t* are adverbs, thus producing sentences like “He no/not/don’t bite you”. Later in the third stage they realize that negation can also be a separate syntactic head showing agreement with subjects as in “He doesn’t bite you”. This analysis was inspired by Zeijlstra (2004)’s proposal for Negative Concord which divides languages into those with adverbial semantic negation and those with syntactic negation. Thornton and Tesan (2013) provided elicitation data on negative sentences with third-person singular subjects from four two-year-olds who had been recorded for about a year (Tesan 2005). They found that in line with their proposal, children produced *not* (e.g. “This not fits in here”) before producing *n’t* (e.g. “this does not fit in here”). A similar finding regarding the order in which negative morphemes emerge in English was reported by Cameron-Faulkner, Lieven, and Theakston (2007). They investigated the development of multiword negation in the speech of Brian (2;3-3;4, MLU 2.05-3.1) and reported that negative morphemes followed a *no*<*not*<*n’t* trajectory, mirroring their order of frequency in parents’ speech. Earliest multiword negation strategies were described as a combination of *no/not* with different types of phrases (*no/not*+XP), with *don’t* and *can’t* being the first contracted forms to emerge.

To summarize, previous research has suggested that English-speaking children learn to produce negative morphemes in a *no*<*not*<*n’t* order. Among negative auxiliary forms, *can’t* and *don’t* are learned first and before any positive auxiliary form including *can* and *do*. Klima and Bellugi (1966) argued that children’s production of English negation becomes adult-like after two non-adult-like stages: one in which they produce negation before a sentence and never inside it, and another in which they produce it within the sentence but do not use the right negative morpheme (e.g. “He not little” instead of “He isn’t little”). There has been an active debate over each of these stages. With the exception of Drozd (1995), previous studies

have mainly relied on data from a few available children often including the original data from Klima and Bellugi (1966). Given that over the years much more corpus data and computational tools have become available, it is important to revisit previous proposals and assess their current status.

3. Study 1

The aim of this study was to assess the overall production of negative morphemes *no*, *not*, and *n't* in parents' and children's speech. The study addresses the following questions: 1. Does the overall production of negation in children follow a *no*<*not*<*n't* cline (Cameron-Faulkner, Lieven, and Theakston 2007)? 2. Do children produce negative auxiliary forms such as *can't* and *don't* before their positive variants, suggesting that the negative forms are learned as unanalyzed wholes (Klima and Bellugi 1966)?

3.1 Methods

For samples of parents' and children's speech, we used the online database *chldes-db* and its associated R programming package *chldesr* (Sanchez et al. 2019). *Chldes-db* is an online interface to the child language components of TalkBank, namely CHILDES (MacWhinney 2000) and PhonBank. Two collections of corpora were selected: English-North America and English-UK. The dataset contained 14,195,967 tokens from 571 children, after necessary exclusions. We ran a token-based analysis of the corpora as well as an utterance-based analysis that could take utterance length and context into account. All data and analysis are available in the study's online repository³.

In our token-based analysis, all word tokens were tagged for the following: 1. the speaker (parent vs. child), 2. the age of the child when the word was produced in months, 3. whether the word was positive or negative, and 4. the type of negative morpheme produced. Here we report on the following classes of English negative morphemes in English: the forms *no* and *not*, all instances of negative auxiliary forms with *n't* as well as their positive forms without *n't* as controls. Unintelligible tokens were excluded (N = 402,117), as well as tokens that had missing information on children's age (N = 1,057,287). Third, tokens outside the age range of 1 to 6 years were excluded (N = 542,304) since children did not produce much outside this age range. Given these measures, data from 100 children were excluded from the final token-based analysis. Similarly, in our utterance-based analysis, each utterance was tagged for the following: 1. the number of tokens in the utterance 2. the speaker (parent vs. child), 2. the age of the child, 3. whether the utterance contained *no*, *not*, or *n't*. Unintelligible utterances (N = 177,804), utterances with missing information on children's age (N = 551,196) as well as those outside the age range of 1 to 6 years were excluded

³You can access the repository at https://github.com/jasbi/negation_production

($N = 99,069$). The final collection contained 3,729,241 utterances from 584 children.

3.2 Results

We first look at the proportions of different categories of negation in parents' and children's speech (Figure 1). The most frequent form in parents' speech was the contracted auxiliary negation *n't*, followed by *no*, and finally *not*. In children's productions and between the ages of 12-18 months, almost all negative forms were instances of *no*, with some contracted auxiliary negatives like *don't* and *can't*. As children grow older, the proportions of *not* and its contracted form *n't* increased while the proportion of *no* decreased. Similar to Cameron-Faulkner, Lieven, and Theakston (2007) we find that children start producing *no* earlier than other forms. However, we do not find evidence that the full form *not* is produced before its contracted form *n't*. The results in Figure 1 suggest that children start producing *not* and *n't* around the same time, if not slightly earlier for *n't*.

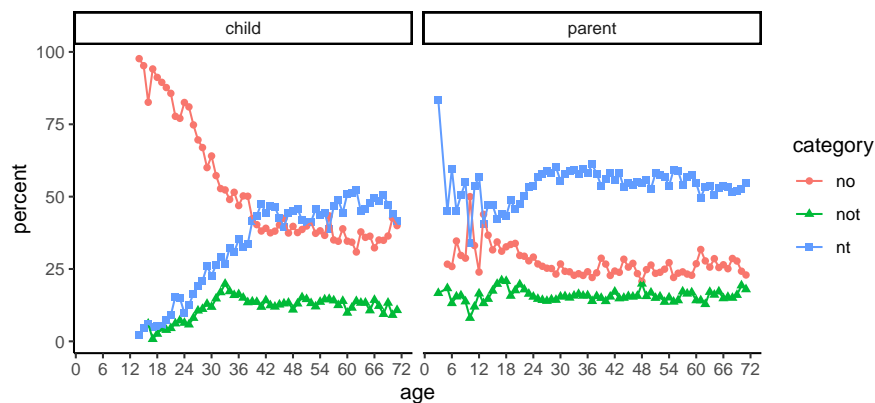


Figure 1: Proportion of different categories of negation in parents' and children's speech between 1 to 6 years of age.

Figure 2 shows the relative frequency of the morphemes *no*, *not* and *n't* per thousand words in the speech of parents and children. Children start producing *no* between 12-18 months and they immediately surpass their parents' rate of production for this morpheme. Between 18-42 months children produce two to three times more instances of *no* than their parents. This rapid increase and high frequency of *no* may be partly because parents ask many yes/no questions from children in this age range. After 42 months the frequency of *no* reduces substantially and gets closer to parents' level of 10 per thousand. For the negative morpheme *not*, children start their productions between 12-24 months and by 30 months of age, they are

producing *not* at the same rate as their parents (5 per thousand words). After 36 months children’s rate of *not* productions stay similar to their parents. Finally for the contracted form *n’t*, children’s productions start between 12-18 months and by 24 months they reach a rate of 5 instances per thousand words. They keep increasing this rate until they reach their parents’ rate of 15 instances per thousand words at 36 months. It is important to note that for all these negative forms, children reached an adult-level of production (in terms of relative frequency) by 30 months of age.

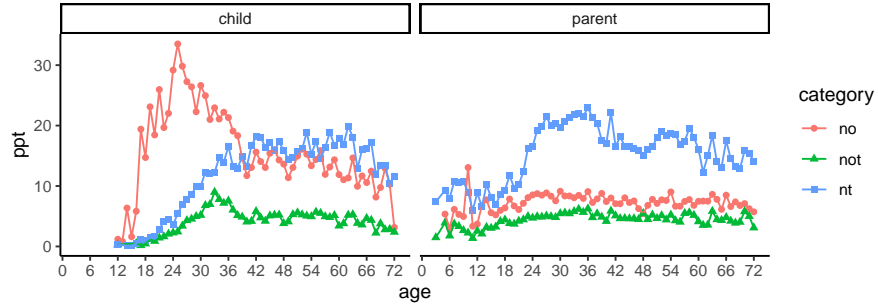


Figure 2: Relative frequency (parts per thousand) of the response particle *no*, verb phrase negation *not*, and its contracted form *n’t*

Stromswold and Zimmermann (2000) found that in German-speaking children, the word *nein* was produced before *nicht* and discussed three potential causes for this order of production: input frequency, phonetic complexity, and syntactic complexity. They explained that input frequency cannot be the cause because in German-speaking children’s input *nicht* was more frequent than *nein*. Similarly, English-speaking children hear more instances of *n’t* than *no* so input frequency cannot be the cause in English either. With respect to phonetic complexity, German *nicht* has a voiceless palatal fricative that can potentially be hard for children and delay its production. However, English *no* and *not* are quite similar and do not contain phones that are known to be particularly hard for children. This leaves us with syntactic complexity which is an obvious difference between isolated one-word negators like *no/nein* and multiword negators like *not/nicht*. Given that children start with shorter utterances (typically one word) and produce longer ones as they grow up, they may produce *no* earlier than *not* and *n’t* simply because *no* can appear as a single word utterance. In other words, even a hypothetical child that comprehends all negative morphemes may produce *no* earlier due to production limitations. We call this the “production bottleneck” hypothesis.

Given our dataset, we can test the production bottleneck hypothesis in two ways. First, we focus on children’s multiword utterances. Is the main

contributor to the high frequency of *no* in children’s speech the single-word “no” utterances? To answer this question we removed single-token utterances like “yes”, “no”, and “oh”, as well as utterances that combined such elements in a repetitive way like “no no” or “oh no” from children and parents’ speech. If early appearance and high frequency of *no* is mainly due to short and repetitive utterances produced by children early in their development, it should disappear once we focus on multiword utterances. As Figure 3 shows, this is largely what we found. While the frequencies of *not* and *n’t* in multi-word productions were similar to their overall frequencies seen before in Figure 2, the word *no* lost its large advantage in frequency and early occurrence, showing a very similar production trajectory as the other two negative morphemes.

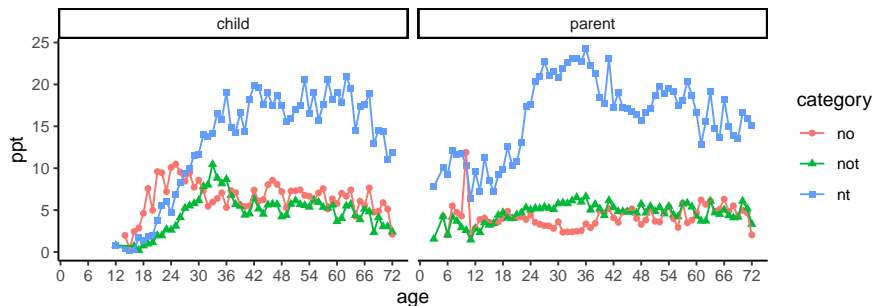


Figure 3: Relative frequency (parts per thousand) of the response particle *no*, verb phrase negation *not*, and its contracted form *n’t* in multiword utterances

The second way to test the bottleneck hypothesis is to artificially impose a production limitation on parents’ speech. To achieve this in our dataset, we grouped utterances into monthly age bins and sampled parents’ utterances in each age bin based on the utterance lengths produced by children in the same age bin. This way in each monthly age bin, we only included adult utterances that matched those of children’s in length. This approach limits parents’ speech to be shorter earlier and longer later, mimicking children’s productive development. The result of this artificial bottleneck on parents’ production of negative morphemes is shown in Figure 4 side by side with children’s negative productions. While previously parents produced *n’t* more frequently than *no* and *not* throughout children’s development (Figure 2), after introducing the bottleneck parents show a higher relative frequency for *no* than *not* and *n’t* in younger ages similar to what is seen in children. As children’s age increases, the relative frequency of *no* decreases and those of *not* and *n’t* increase in a way that mimicks the pattern seen in children’s production. Later and around 40 months, the order of production

reverses and adults produce *n't* more frequently than the other forms in this artificially induced bottleneck.

Taken together, the two tests indicate that the earlier emergence and high frequency of *no* in children's speech may be largely due to children's limited capacity in producing longer utterances and not necessarily earlier acquisition of this morpheme. Therefore, the question "which form is acquired earlier" may be better addressed by careful comprehension studies in the 12-24 month age range. It is important to note here that both Figure 3 and Figure 2 suggest the 12-24 months age range as a period where all three negative morphemes may receive their early form-meaning mappings. In order to better understand such early mappings of negation and their development we need more comprehension studies in this age range.

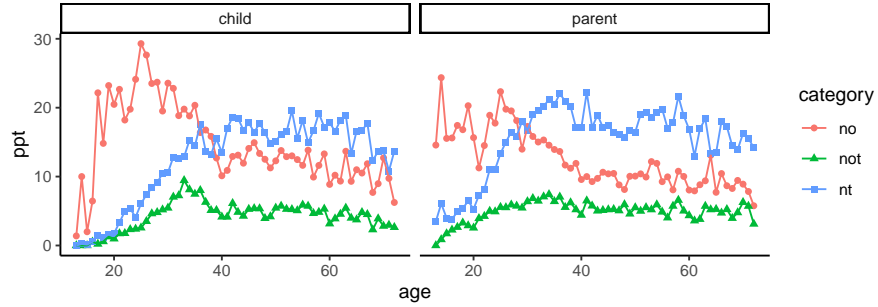


Figure 4: Relative frequency (parts per thousand) of the response particle *no*, verb phrase negation *not*, and its contracted form *n't* in children and parents' speech with an artificial bottleneck imposed on parents speech in each age bin

Moving to the second question: do negative auxiliaries appear before positive ones? Figure 5 shows the relative frequency of positive and negative auxiliary forms in the speech of children and their parents. Our results show that overall, children start producing the positive and negative auxiliary forms around the same time and produce the positive forms at a higher rate than negative ones. This is also true for individual auxiliary words such as *do/don't* and *can/can't* which are produced earlier than others. Therefore, the claim that negative auxiliary forms are produced before their positive counterparts is not supported by the available production data and consequently production data does not provide support for the hypothesis that auxiliary negative forms are learned as unanalyzed wholes.

3.3 Conclusion

4. Study 2

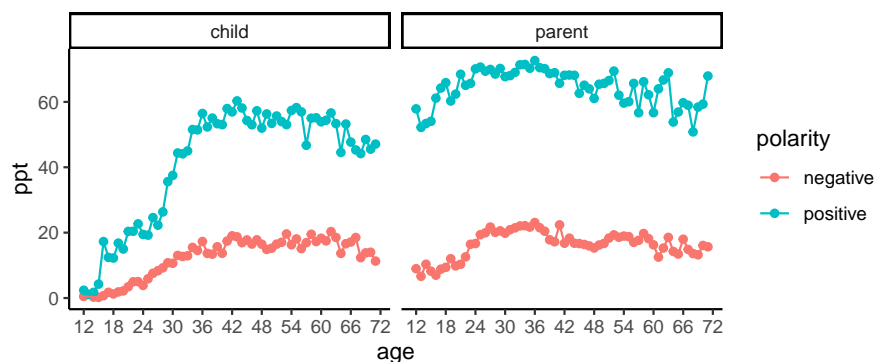


Figure 5: Relative frequency (parts per thousand) of positive auxiliary forms such as *do*, *are*, and *can* as well as their contracted negatives in the speech of parents and children.

4.1 Methods

4.2 Results

4.3 Conclusion

5. Discussion – Results supported early productions of *no* (stage 1) before *not* and *n’t*, but no evidence that *not* is produced before *n’t*. We did not find support for robust non-anaphoric pre-sentential negation (stage 3) or contracted forms like *don’t* and *can’t* as unanalyzed wholes (stage 4). The results suggest a great degree of variability in children’s productive development (de Villiers & de Villiers 1979). Our findings are compatible with simultaneous development of frequent negative forms with a production bottleneck that favors shorter utterances like *no* to appear earlier.

6. References

- Austin, Keith, Anna Theakston, Elena Lieven, and Michael Tomasello. 2014. “Young Children’s Understanding of Denial.” *Developmental Psychology* 50 (8): 2061.
- Bloom, Lois Masket. 1970. *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT press.
- Brown, Roger. 1973. *A First Language: The Early Stages*. Harvard U. Press.
- Cameron-Faulkner, Thea, Elena Lieven, and Anna Theakston. 2007. “What Part of No Do Children Not Understand? A Usage-Based Account of Multiword Negation.” *Journal of Child Language* 34 (2): 251.
- Chomsky, Noam. 1993. *Lectures on Government and Binding: The Pisa Lectures*. 9. Walter de Gruyter.
- de Villiers, Peter A, and Jill G de Villiers. 1979. “Form and Function in the Development of Sentence Negation.” *Papers and Reports on Child Language Development* 17: 57–64.

- Déprez, Viviane, and Amy Pierce. 1993. "Negation and Functional Projections in Early Grammar." *Linguistic Inquiry*, 25–67.
- Droz, Kenneth F. 1995. "Child English Pre-Sentential Negation as Metalinguistic Exclamatory Sentence Negation." *Journal of Child Language* 22 (3): 583–610.
- Feiman, Roman, Shilpa Mody, Sophia Sanborn, and Susan Carey. 2017. "What Do You Mean, No? Toddlers' Comprehension of Logical 'No' and 'Not!'" *Language Learning and Development* 13 (4): 430–50.
- Klima, Edward S, and Ursula Bellugi. 1966. "Syntactic Regularities in the Speech of Children." In *Psycholinguistics Papers*, 183–207. Edinburgh University Press.
- Koopman, Hilda, and Dominique Sportiche. 1991. "The Position of Subjects." *Lingua* 85 (2-3): 211–58.
- Lange, S, and K Larsson. 1973. "Syntactical Development of a Swedish Girl, Embla, Between 20 and 42 Months of Age, Part I: Age 20-25 Months." *Project Child Language Syntax*.
- Lord, Carol. 1974. "Variations in the Acquisition of Negation." *Papers and Reports on Child Language Development* 8: 78–86.
- MacWhinney, Brian. 2000. *The CHILDES Project: The Database*. Vol. 2. Mahwah, NJ: Erlbaum.
- Park, Tschang-Zin. 1979. "Some Facts on Negation: Wode's Four-Stage Developmental Theory of Negation Revisited." *Journal of Child Language* 6 (1): 147–51.
- Reuter, Tracy, Roman Feiman, and Jesse Snedeker. 2018. "Getting to No: Pragmatic and Semantic Factors in Two-and Three-Year-Olds' Understanding of Negation." *Child Development* 89 (4): e364–e381.
- Sanchez, Alessandro, Stephan C Meylan, Mika Braginsky, Kyle E MacDonald, Daniel Yurovsky, and Michael C Frank. 2019. "ChilDes-Db: A Flexible and Reproducible Interface to the Child Language Data Exchange System." *Behavior Research Methods* 51 (4): 1928–41.
- Schütze, Carson T. 2010. "The Status of Nonagreeing Don't and Theories of Root Infinitives." *Language Acquisition* 17 (4): 235–71.
- Stromswold, Karin, and Kai Zimmermann. 2000. "Acquisition of Nein and Nicht and the Vp-Internal Subject Stage in German." *Language Acquisition* 8 (2): 101–27.
- Tesan, Graciela Mariel. 2005. "What Do Children Have in Their Heads? Functional Heads and Parameter Setting in Child Language." PhD thesis.
- Thornton, Rosalind, and Graciela Tesan. 2013. "Sentential Negation in Early Child English." *Journal of Linguistics*, 367–411.
- Wode, Henning. 1977. "Four Early Stages in the Development of Li Negation." *Journal of Child Language* 4 (1): 87–102.
- Zeijlstra, Hedde. 2004. "Sentential Negation and Negative Concord." PhD thesis, Netherlands Graduate School of Linguistics.