

The production of negation in parents' and children's speech

Masoud Jasbi<sup>1</sup>, Annika Mcdermott-Hinman<sup>2</sup>, Kathryn Davidson<sup>2</sup>, & Susan Carey<sup>2</sup>

<sup>1</sup> UC Davis

<sup>2</sup> Harvard University

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

Correspondence concerning this article should be addressed to Masoud Jasbi, 469 Kerr Hall, University of California, One Shields Avenue, Davis, CA, 95616. E-mail: jasbi@ucdavis.edu

12

Abstract

13   this is the abstract

14       *Keywords:* negation, production, child language

15       Word count: X

DRAFT

The production of negation in parents' and children's speech

## Introduction

Under Construction!

## Previous Studies

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: "the first stage from the first month of study for each child; the last from the month in which the MLUs approach 4.0 for each of the three children; and the second stage between the two." They 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 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 is the absence of positive auxiliary variants like *can* and *do*. Typical examples at this stage include "I can't see you", "I don't want it", "He not little, he big", "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). Klima and Bellugi (1966)'s account was in line with the transformational theory of negation at the

time, which hypothesized that adult negation starts as a sentence modifier at deep structure, and a transformation later moves it inside the sentence (Klima, 1964). Therefore, children's first stage was interpreted to reflect a universal innate initial state, and the second and third stages the process of learning the required surface-level transformations.

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 & 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 Zeijlstra2004'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 “*no/not*+XP” utterances, 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.

## Current Study

This paper builds on the rich literature on children's production of negation by bringing the largest dataset of English parent-child interactions available in CHILDES (MacWhinney, 2000). Study 1 investigates the frequency of production for each negative morpheme in parents' and children's speech during the first 6 years of children's development. More specifically the study checks to see if there is an overall *no* < *not* < *n't* cline (Cameron-Faulkner et al., 2007), and if children produce negative auxiliary forms like *can't* and *don't* before their positive variants like *do* and *can*. Study 2 focuses on non-adult-like productions of negation in children and examines the prevalence of pre-sentential negation (e.g. "no the sun shining") as well as infinitival negation (e.g. "He no/not/don't bite you"). It also discusses the concept of stagehood and whether observed prevalences of such non-adult-like productions are best captured as a developmental stage. Study 3 looks for the most common negative syntactic structures present in parents' and children's speech. It uses a current state-of-the-art parsing algorithm to provide a large-scale estimate of syntactic structures negation participates in as well as its productivity in parent-child interactions. We end with a discussion of English negation's combinatorial capacity in children's early speech and its possible origins.

### Study 1: Development of Negative Morphemes

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



## 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.

In our token-based analysis, all word tokens were tagged for the following information: 1. the speaker role (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 word produced. For this study we considered the following classes of 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. These exclusions resulted in the exclusion of data 100 children from the final 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 role (parent vs. child), 2. the age of the child, 3. whether the utterance contained *no*, *not*, or *n't*. Utterances with missing information on children's age ( $N = 379,399$ ) as well as those outside the age range of 1 to 6 years were excluded ( $N = \text{NULL}$ ). The collection contained  $X$  utterances from  $N$  children.

## 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 are 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* increase while the proportion of *no* decreases. Similar to Cameron-Faulkner et al. (2007) we find that children start productions of *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*.

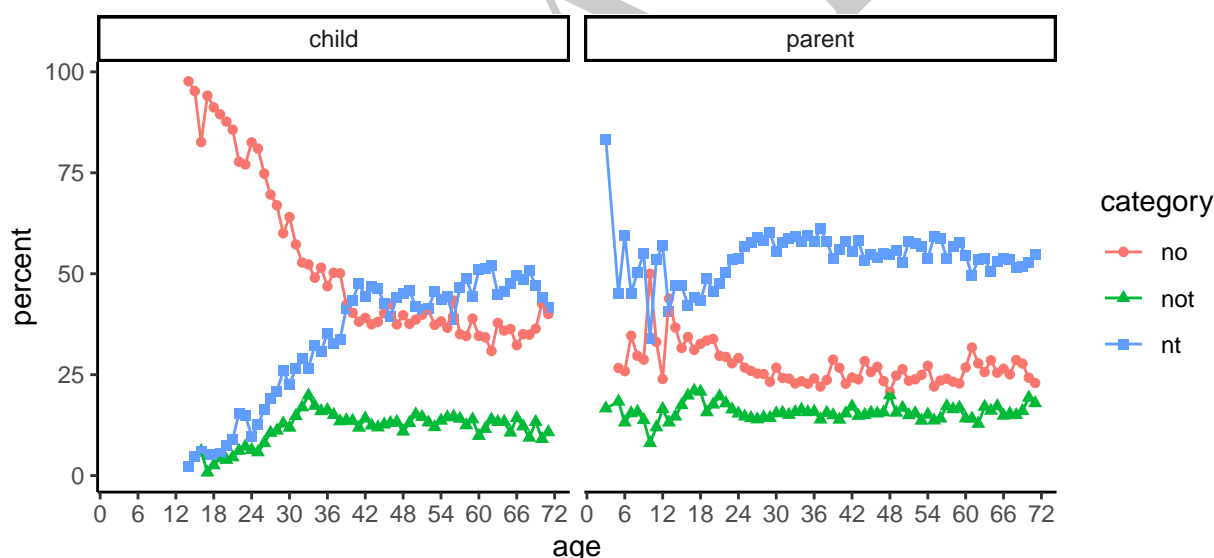


Figure 1. Proportion of different categories of negation in parents' and children's speech between 1 to 6 years of age.

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 at age 36 months. It is important to note that for all these negative forms, children have reached a substantial level of production by 30 months of age.

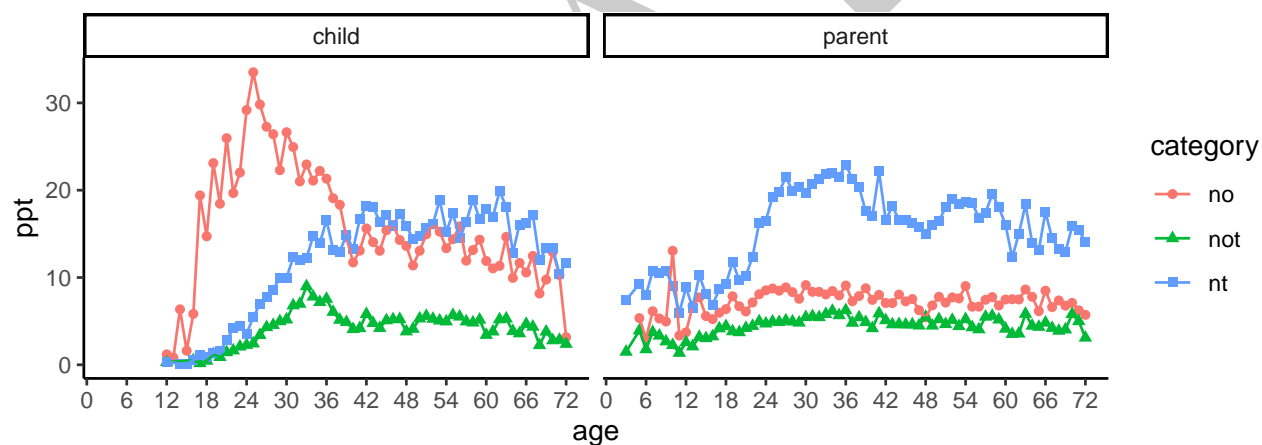


Figure 2. Relative frequency (parts per thousand) of the response particle *no*, verb phrase negation *not*, and its contracted form *n't*

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 forms *no*, *not*, and *nt* may produce *no* earlier than the other two because of production limitations.

One way to address this issue in our dataset is to focus on children's multiword utterances. Is the main contributor to the high frequency of *no* single word 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* largely lost its advantage in frequency and early occurrence, showing a very similar developmental trajectory as the other two negative morphemes. This provides an indication that the earlier emergence of *no* may be due to children's production limitations and does not necessarily reflect earlier acquisition of this morpheme. The question of 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 that the 12-18 months age range is a crucial period where all three negative morphemes receive their early form-meaning mappings.

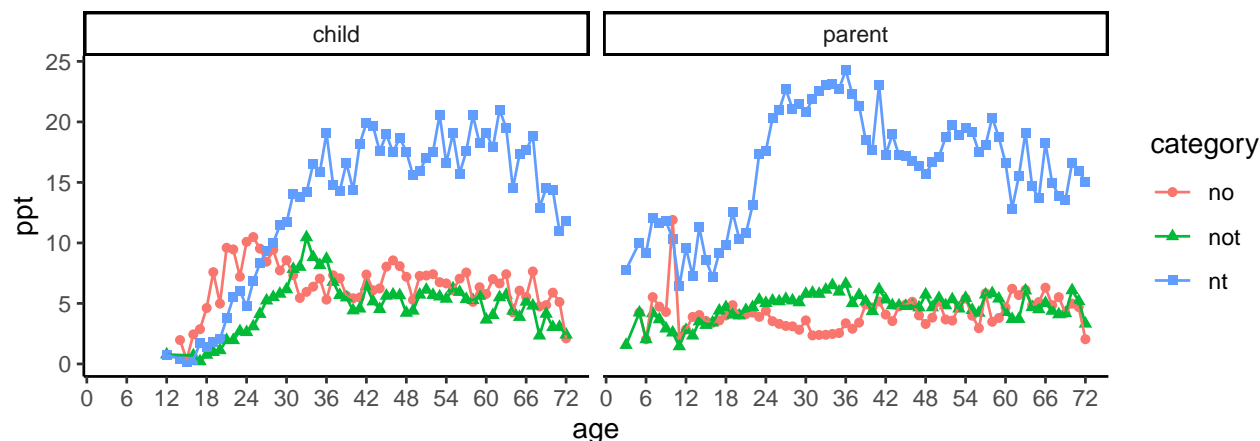


Figure 3. Relative frequency (parts per thousand) of the response particle *no*, verb phrase negation *not*, and its contracted form *n't* in multiword utterances

Moving to the second question: do negative auxiliaries appear before positive ones?

Figure 4 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 always 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 support the hypothesis that auxiliary negative forms are learned as unanalyzed wholes.

## Conclusions

The results of our corpus study suggest that children produce *no* earlier and more frequently than *not* and *n't*. We did not find strong evidence for *not* appearing before *n't* and the general trends suggested that they appear around the same time. While earlier emergence of *no* in production may suggest its earlier acquisition and comprehension, it is also possible that early emergence of *no* is due to its syntactic and usage properties. The

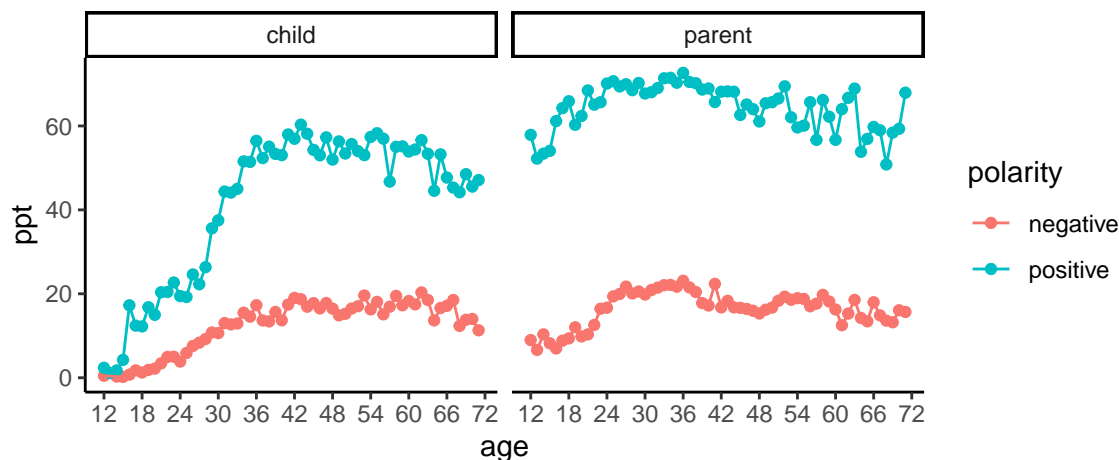


Figure 4. Relative frequency (parts per thousand) of positive auxiliary forms such as *do*, *are*, and *can* as well as their contracted negatives in the speech of parents and children.

morpheme *no* can be produced on its own as a speech act in response to speech participants, while negative forms *not* and *n't* require composition with other words. We found that when we consider only children's multiword utterances, the early emergence and advantage of *no* disappears. Therefore, the production data does not provide support for some strong order or stage of acquisition in children's acquisition of negative morphemes. We believe it is more appropriate for future comprehension research to adjudicate this matter.

We also investigated whether negative auxiliary forms such as *can't* and *don't* emerge before their positive counterparts such as *do* and *can*. Previous research had cited such a phenomenon as evidence for the proposal that negative auxiliary forms are first learned as unanalyzed wholes. Contrary to previous reports, our data showed that the positive auxiliary forms emerge around the same time as the negative ones but produced much more frequently. Therefore, production data does not provide evidence for negative auxiliaries being learned as unanalyzed forms.

## Study 2: Non-adult-like Productions

The aim of this study was to assess the prevalence of non-adult-like utterances in children's productions. Klima and Bellugi (1966)'s account of negation development relies on two types of non-adult-like utterances: 1. Pre-sentential instances where a negative morpheme is modifying a sentence (NEG + NP + VP); and 2. Intra-sentential instances where in some cases fails to show the adult-like agreement (NP + *no/not/don't* + VP). How prevalent are such errors and what are their most likely sources?

### Method

### Results

### Conclusion

## Study 3: Early Syntactic Structures for Negation

In this study we ask what syntactic analyses can be attributed to children's early negative utterances. We also ask what common syntactic structures are present in parents' speech. The goal is to provide an account of children's syntactic development in the spirit of Klima and Bellugi (1966)'s pioneering work with data and computational tools available today.

289 **Method**

290 **Results**

291 **General Discussion**

292       So far: Our results are compatible with emergence of negation around 18-24 months,  
293       converging with studies on comprehension of negation.



## Acknowledgments

## Declaration of Interest Statement

## References

## Appendices

The auxiliary category in our previous figures lump together a wide variety of auxiliary verbs that develop at different rates. Figure 5 shows the production of common negative auxiliary verbs in the speech of children and parents, sorted from top-left to bottom-right based on frequency. The most frequent negative auxiliary form in child-directed speech is *don't* and it is also the earliest and most frequent auxiliary form in children's speech. Children start producing it between 12-24 months and they quickly reach the parents' rate at 36 months. Perhaps the fastest development occurs with the auxiliary *can't*. Children start producing it between 18-24 months and very quickly surpass their parents' rate.

Figure 6 shows the development of negative and positive indefinite pronouns: *everything*, *nothing*, *something*. Children start producing these words quite early as well, with *nothing* reaching the parent level of production at 30 months.

### Adverbs of frequency

Bloom, L. M. (1970). *Language development: Form and function in emerging grammars*. Cambridge, MA: MIT press.

Brown, R. (1973). *A first language: The early stages*. Harvard U. Press.

Cameron-Faulkner, T., Lieven, E., & Theakston, A. (2007). What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child*

315 *Language*, 34(2), 251.

316 Chomsky, N. (1993). *Lectures on government and binding: The pisa lectures*. Walter  
317 de Gruyter.

318 de Villiers, P. A., & de Villiers, J. G. (1979). Form and function in the development of  
319 sentence negation. *Papers and Reports on Child Language Development*, 17, 57–64.

320 Déprez, V., & Pierce, A. (1993). Negation and functional projections in early grammar.  
321 *Linguistic Inquiry*, 25–67.

322 Drozd, K. F. (1995). Child english pre-sentential negation as metalinguistic  
323 exclamatory sentence negation. *Journal of Child Language*, 22(3), 583–610.

324 Klima, E. (1964). Negation in english. *The structure of language*, ed. By ja fodor and  
325 jj katz, 246–323. Englewood Cliffs, NJ: Prentice-Hall.

326 Klima, E. S., & Bellugi, U. (1966). Syntactic regularities in the speech of children. In  
327 *Psycholinguistics papers* (pp. 183–207). Edinburgh University Press.

328 Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua*, 85(2-3),  
329 211–258.

330 Lange, S., & Larsson, K. (1973). Syntactical development of a swedish girl, embla,  
331 between 20 and 42 months of age, part i: Age 20–25 months. *Project Child Language Syntax*.

332 Lord, C. (1974). Variations in the acquisition of negation. *Papers and Reports on*  
333 *Child Language Development*, 8, 78–86.

334 MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Mahwah, NJ:  
335 Erlbaum.

336 Park, T.-Z. (1979). Some facts on negation: Wode's four-stage developmental theory of

negation revisited. *Journal of Child Language*, 6(1), 147–151.

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). Chldes-db: A flexible and reproducible interface to the child language data exchange system. *Behavior Research Methods*, 51(4), 1928–1941.

Schütze, C. T. (2010). The status of nonagreeing don't and theories of root infinitives. *Language Acquisition*, 17(4), 235–271.

Stromswold, K., & Zimmermann, K. (2000). Acquisition of nein and nicht and the vp-internal subject stage in german. *Language Acquisition*, 8(2), 101–127.

Tesan, G. M. (2005). *What do children have in their heads? Functional heads and parameter setting in child language* (PhD thesis).

Thornton, R., & Tesan, G. (2013). Sentential negation in early child english. *Journal of Linguistics*, 367–411.

Wode, H. (1977). Four early stages in the development of li negation. *Journal of Child Language*, 4(1), 87–102.

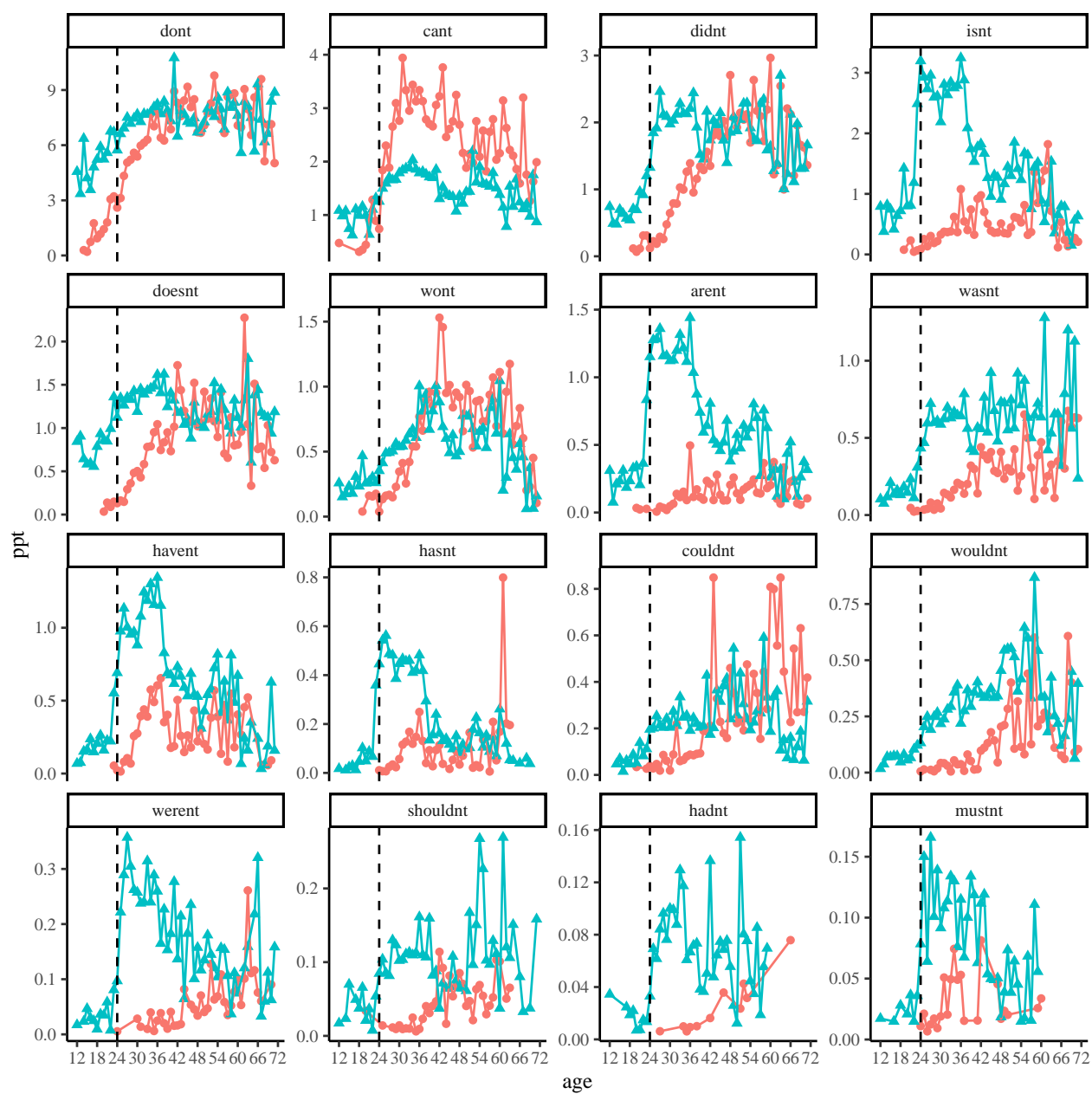


Figure 5. Relative frequency of negated auxiliary verbs in the speech of parents (green triangles) and children (red circles). The dashed line marks 24 months on the x axis.

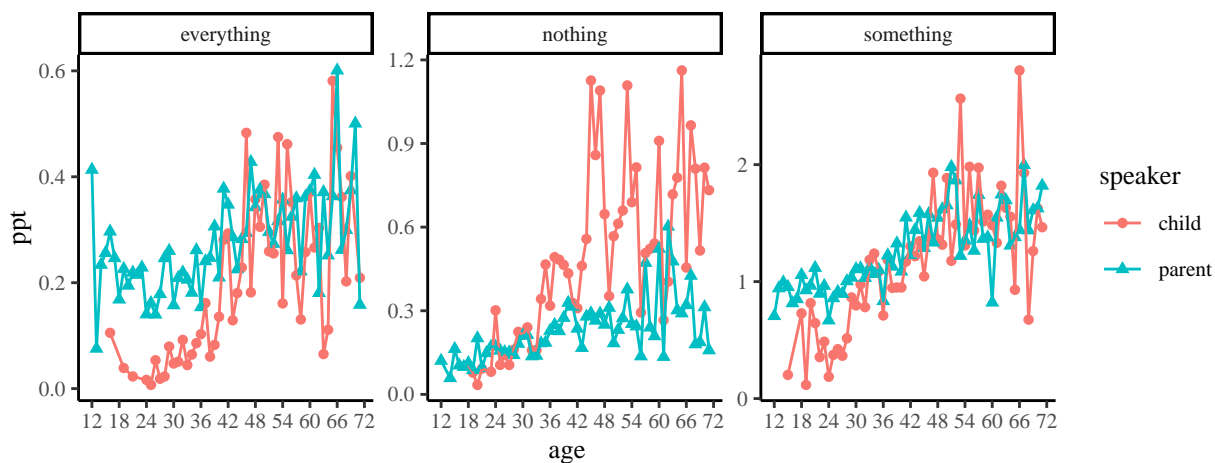


Figure 6. Relative frequency of pronouns *everything*, *something*, and *nothing*

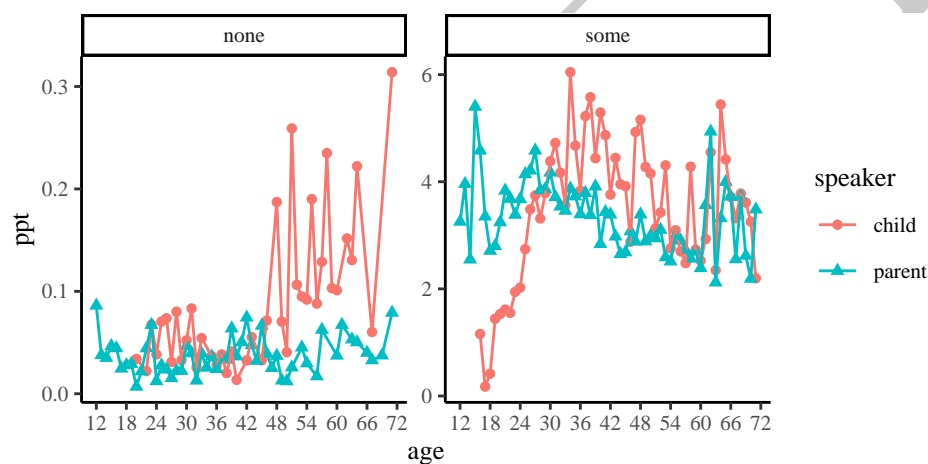


Figure 7. Relative frequency of quantifiers *none*, *some*, and *all*.

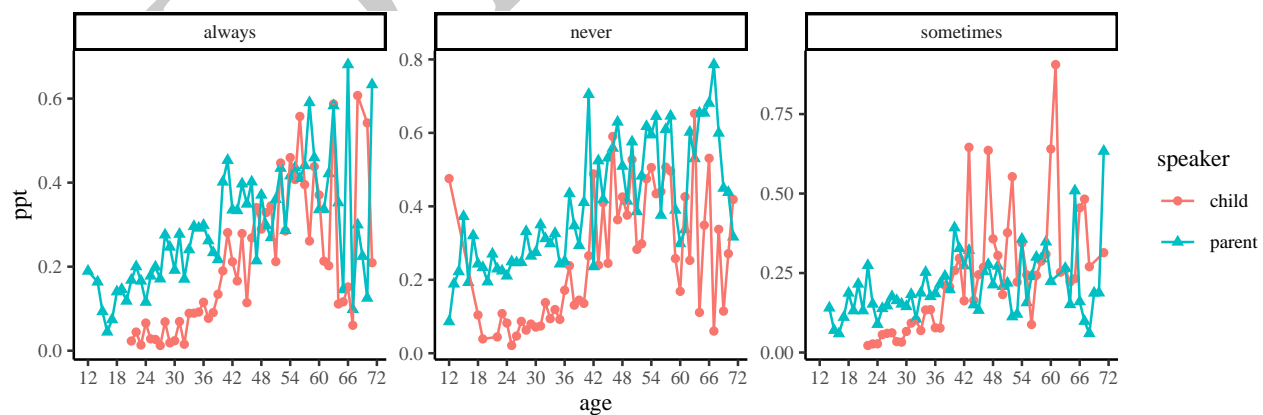


Figure 8. Relative frequency for adverbs of frequency *always*, *never*, and *sometimes* in the speech of parents and children.