Language Acquisition

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hlac20

Children's early negative auxiliaries are true auxiliaries

André Eliatamby & Virginia Valian

To cite this article: André Eliatamby & Virginia Valian (08 Jul 2024): Children's early negative auxiliaries are true auxiliaries, Language Acquisition, DOI: 10.1080/10489223.2024.2356668

To link to this article: https://doi.org/10.1080/10489223.2024.2356668

LANGUAGE ACQUISITION
https://doi.org/10.1080/10489223.2024.2356668

Routledge
Taylor & Francis Group

Children's early negative auxiliaries are true auxiliaries

André Eliatamby[a] and Virginia Valian[b]

[a]CUNY The Graduate Center; [b]CUNY – Hunter College and Graduate Center

ABSTRACT

This study investigates young children's acquisition of functional categories through their use of negative words and negative auxiliaries in particular. Drawing from CHILDES, we analyze twelve months of spontaneous speech by 14 children (youngest age 1;9, oldest age 3;1) and their mothers, in order to assess whether children's earliest negative productions are morphological combinations and reflect possession of abstract syntactic categories or are instead input-driven formulae. In five analyses we show that (i) two-year-olds use a wide and overlapping range of negative and positive auxiliaries; (ii) the range of the negative auxiliaries children produce is strongly correlated with the range of the positive auxiliaries they produce; (iii) children's most common negative auxiliary, *don't*, is used grammatically with respect to the syntactic category being negated and with respect to overt markings of tense; (iv) children's subject agreement errors with *don't* are mirrored by subject agreement errors with *do, have*, and *haven't*; and (v) children omit auxiliaries with *not* at rates that cannot be attributed to properties of their input. Our findings support the hypothesis that children's earliest negations are syntactically adult-like and reflect the possession of abstract syntactic categories. By age 2, English-learning children productively combine auxiliary, negation, and tense categories and syntactically differentiate different negative morphemes.

ARTICLE HISTORY
Received 09 March 2023
Accepted 14 May 2024

1. Introduction

The acquisition of functional categories provides a window into children's early syntactic representations. Unlike lexical categories, functional categories do not directly map onto visible conceptual categories. Unlike nouns and verbs, for example, determiners lack direct physical correlates. They are semantically rich, but their properties, such as definiteness, are not easily targeted or illustrated. Consequently, how and when children acquire functional categories can shed light on the properties of the mechanism children employ in language acquisition.

We examine two-year-old English learners' use of negators, which are functional categories, with a focus on negative auxiliaries. We target negation not only because negation is itself a functional syntactic element, but because many languages, like English, have negators that interact with other functional categories. "Standard" English has two distinct types of morphemes that carry negative polarity: the words *no* and *not*, and the clitic *n't*. While *no* and *not* are both free morphemes, *n't* is a bound morpheme that combines with auxiliaries, copulas, or modals to create negative auxiliaries like *don't, didn't, doesn't, can't, couldn't, isn't*.

That negative auxiliaries are true auxiliaries can be seen when comparing their distributional behavior to their positive counterparts. First, like auxiliaries in positive contexts, negative auxiliaries cannot occur after other auxiliaries:

CONTACT André Eliatamby aeliatamby@gradcenter.cuny.edu Linguistics, CUNY The Graduate Center, Linguistics, 365 Fifth Avenue, New York, NY 10016

This article has been corrected with minor changes. These changes do not impact the academic content of the article. Please see Correction (http://doi.org/10.1080/10489223.2024.2380569).

Supplemental data for this article can be accessed online at https://doi.org/10.1080/10489223.2024.2356668

1. *The tigers **are** don't wear a hat
2. *The tigers **are** do wear a hat

Second, negative auxiliaries have the same selectional requirements as their positive counterparts. For example, *don't, doesn't*, and *didn't* must occur with a verb phrase (VP) containing an untensed verb, just like their associated positive counterparts *do, did*, and *does*:

3.	The tiger didn't wear a hat	(VP with untensed verb)
4.	The tiger did wear a hat	(VP with untensed verb)
5.	*The tiger doesn't wears a hat	(VP with present-tense inflected verb)
6.	*The tiger does wears a hat	(VP with present-tense inflected verb)
7.	*The tiger doesn't red	(AP = adjective phrase)
8.	*The tiger does red	(AP)
9.	*They don't my toys	(DP = determiner phrase)
10.	*They do my toys	(DP)

Similarly, *isn't, wasn't,* and *aren't* are only licensed in copula constructions that lack a main tensed verb, like their counterparts *is* and *was*:

11. *The tiger isn't wears a hat
12. *The tiger is wears a hat
13. The tiger isn't red
14. The tiger is red
15. Those aren't my toys
16. Those are my toys
17. The tiger isn't wearing a hat
18. The tiger is wearing a hat

Third, tense is expressed on the auxiliary in both positive and negative contexts:

19.	The tiger did/didn't want a hat	(past tense)
20.	The tiger does/doesn't want a hat	(present tense)
21.	The tiger was/wasn't wearing a hat	(past tense)
22.	The tiger is/isn't wearing a hat	(present tense)

These properties suggest that, in adult English, negative auxiliaries are morphological combinations of *n't* with an auxiliary, modal, or copula. As a result, to use negative auxiliaries in an adult-like manner, a speaker must have acquired three distinct functional elements—the negation category, the broad category containing auxiliaries, modals, and copulas, and a representation of tense (tense category)—and must also possess the ability to combine these functional elements to create a morphologically complex negator.

We focus on two-year-olds' productions of negators since children start producing multi-word or telegraphic utterances around this age (or earlier). It has traditionally been argued that functional words are absent at the start of multi-word speech (e.g., Bowerman 1973, Brown 1973). More recent research, however, has shown that despite an absence in production, children as young as 12 months encode and represent at least an underspecified functional category. Children acquiring a variety of languages make use of functional elements to segment speech (Shi & Lepage 2008, Bernard & Gervain 2012) to help in the acquisition of content words (Bernal et al. 2007, Yuan & Fisher 2009), and are sensitive to the presence of functional words and elements in the speech they hear (Shady & Gerken 1999, Shi et al. 1999).

With respect to specific categories, two-year-olds have a productive determiner category, in that their use of determiner-noun bigrams reflects knowledge of the fact that fragments like "*the ball*" encode two distinct abstract categories that can combine in novel ways (Valian et al. 2009, Yang 2013; but see Pine et al. 2013). The use of negative auxiliaries at the start of multi-word speech can provide

further evidence about the encoding and productive use of specific functional elements. Specifically, if two-year-olds behave like adults with respect to the tense-bearing and selectional properties of negative auxiliaries, that will suggest that functional elements, including negation, auxiliaries, and tense, are acquired early.

1.1. *Previous research into early negative auxiliaries*

1.1.1. *Monomorphemic negation*

Research on negation has come to contradictory conclusions about children's early productions. Starting with Klima & Bellugi (1966), many investigators have concluded that English-learning children's use of negators prior to age 3 is non-adult-like (Kuczaj & Marastos 1975, Stromswold 1990, Capdevila i Batet & Llinàs i Grau 1995, Cameron-Faulkner et al. 2007, Thornton & Tesan 2013, Thornton & Rombough 2015, *inter alia*). These investigators argue that children's initial negators are limited to *no, not, don't*, and *can't*, and that early uses of *don't* and *can't* are misanalyzed morphological "wholes" rather than combinations of auxiliaries and *n't*.

Evidence putatively supporting non-adult negation is: (i) children only produce *don't* and *can't* before age 3 (Klima & Bellugi 1966, Cameron-Faulkner et al. 2007), (ii) *do, can*, and other auxiliaries do not occur in positive contexts during the same period (Kuczaj & Marastos 1975, Capdevila i Batet & Llinàs i Grau 1995), (iii) *don't* and *can't* occur ungrammatically with other tensed elements of a sentence (Thornton & Tesan 2013, Thornton & Rombough 2015); and (iv) *don't* occurs with a third person singular subject (Bellugi 1967). Only after age 3, these investigators argue, do children start producing true negative auxiliaries that reflect morphological combination.

Grammatical accounts of misanalyses of *don't* and *can't* propose that children have a non-target grammar. Klima & Bellugi (1966), for example, claim that children initially lack the phrase structure rule needed to morphologically compose auxiliaries with *n't*. They group *don't* and *can't* with *no* and *not* as forming an initial set of monomorphemic negators. Capdevila i Batet & Llinàs i Grau (1995) argue that children in the monomorphemic stage lack negation as a functional category; syntactic negation only becomes available after biological maturation (Borer & Wexler 1987, Tsimpli 1992). Thornton & Tesan (2013) and Thornton & Rombough (2015) also link the misanalysis of *don't* and *can't* to the syntactic status of the negators, but consider the misanalysis period to be continuous with adult language (cf. Pinker 1984). Thornton and colleagues (Thornton & Tesan 2013, Thornton & Rombough 2015) propose that English-learning children's first negators reflect an adverbial grammar,[1] following proposals that adult languages vary according to whether negative elements are adverbs or syntactic heads of fully specified negation phrases (Zanuttini 1991, Zeijlstra 2004).

Usage-based proposals offer a different account of children's non-target productions: children do not possess any abstract categories until sometime after age 3. Children's early negative sentences are item-specific formulae that reflect the input they receive (Cameron-Faulkner et al. 2007). For example, early uses of negation reflect the possession of formulae such as *don't* + *X, can't* + *X, no* + *X*, and *not* + *X* derived from common input sequences rather than possession of abstract auxiliary, negation, and tense representations. Consequently, children's early negators are monomorphemic, their sentences are non-compositional, and any non-target errors are the result of production limitations or the distributional patterns in the input. Cameron-Faulkner et al. (2007) argue that negator use develops along a *no-not-n't* trajectory driven by input frequency, that is, negators used frequently by caregivers are the first to emerge in child speech.

Thus, under both grammatical and usage-based explanations, children's early representations of *don't* and *can't* do not combine *n't* with a tense-bearing auxiliary, and consequently are not true negative auxiliaries. Although arising from very different first principles, both explanations predict that important aspects of productive functional word use are still under development from age 2 to age 3.

[1]This proposal assumes that only syntactic heads are available for morphological composition.

1.1.2. *Early Competence*

In contrast, some researchers have adopted what we call the Early Competence Hypothesis, adducing evidence that children do in fact combine auxiliaries with *n't* during age 2 and that those combinations reflect possession of abstract syntactic knowledge. Schütze (2010), for example, argues that children's productions of *don't* are overwhelmingly grammatical with respect to tense and category selection. The five children in Schütze's analysis do not make the errors shown in (1) and (7). If early *don't* is monomorphemic, and consequently untensed, one would expect children to make such errors. Using a large-scale cross-sectional study of 571 children, Jasbi et al. (2021) show that children start producing negative and positive auxiliaries at around the same time, with positive auxiliaries being used at higher rates than negative auxiliaries. This suggests that the auxiliary system is productive when children first start producing negative auxiliaries.

Errors that children do make include using *don't* with a 3rd-person singular subject (Bellugi 1967, Guasti & Rizzi 2002, Schütze 2010, Thornton & Rombough 2015), as in

23. He don't want to play.

Given that *don't* is otherwise used grammatically, Schütze (2010) argues that such errors are the result of tense and agreement features being underspecified in children's grammars, rather than *don't* being morphologically simple. In support of that claim, Jasbi et al. (2021) find that children produce first-person singular subjects with *do* in positive contexts at similar rates as they produce them with *don't*. This is consistent with *don't* being the negative form of *do*. Together, those findings suggest that children's early negative auxiliaries are morphological compositions, even if they are not entirely adult-like.

1.2. *Small initial set of negators: Limited sampling and long-tailed distribution*

Fewer investigations have examined the set of negators children produce. If children at age 2 have an adult-like capacity to generate negative auxiliaries, their putative limited range (*not, no, don't* and *can't*) requires further explanation. We hypothesize that the limited range is a mirage, a sampling artifact due to the interaction of the low underlying frequency distribution of negator use and relatively sparse naturalistic speech corpora. Most research claiming monomorphemic uses of *don't* has involved corpus analysis of naturalistic speech (save for two elicited production studies by Thornton & Tesan 2013 and Thornton & Rombough 2015). Sparse taping will undercount infrequent productions in a way that makes them appear rare or non-existent.

Using a Poisson distribution to model target capture, Tomasello & Stahl (2004) show that for a construction that occurs 10 times a day, a 1-hour per week collection window has slightly less than a 2% chance of capturing more than three instances of the target construction. In other words, a construction that is actually being produced 70 times a week is most likely to be captured between zero and 3 times under a weekly collection scenario. For a construction that occurs 5 times a day (35 times a week), a 1-hour per week collection has less than a 0.2% chance to capture more than three instances.

We combine that observation with the fact that negator production follows a Zipfian (or some other long-tailed) distribution, where a few negator types are produced highly frequently and the majority occur extremely infrequently. The result is that most negator types have a lower chance of being picked up in a 1-week sample than the few most frequent negators. If negator production is generally infrequent, a 1-week collection will only capture the few top ranked negators. Since each subsequent sample from a long-tailed distribution increases the chance of less frequent elements being selected, it follows that if one is repeatedly sampling a child's speech every few weeks, the less frequent items are more likely to turn up in later draws. This gives the illusion that the infrequent types are produced later and that development occurs in stages.

If our hypothesis is correct, previous findings of a limited set of negators are the result of a sampling artefact: children are not limited to producing *don't* and *can't* and they can and do produce true negative auxiliaries by age 2. This would mean that tense, auxiliary, and negation categories are in place in child English by 24 months.

1.3. *Errors with "not"*

Researchers have also noted that children systematically produce errors when using *not* (Klima & Bellugi 1966, Harris & Wexler 1996, Thornton & Rombough 2015), producing sentences like (24):

24. I not want that

Not differs from *n't* in adult English in that it is neither an auxiliary nor a tense bearing element. Errors with *not* thus do not bear on children's knowledge of negative auxiliaries. Early Competence proposals do not predict error-free production with all negators. Children might be productive and adult-like with negative auxiliaries while being non-target with *not*. Errors with *not* are, however, informative as to how abstract children's early syntactic representations are. Sentences like (24) are ungrammatical and thus expected to be absent (or extremely infrequent) in the child's input.

For usage-based theories of development, the prevalence of children's errors with *not* requires explanation. Where can such errors come from if children's input lacks them? Previous research has neglected this point. For example, Cameron-Faulkner et al. (2007) claim that children treat *not* as a "general purpose" negator because it occurs in a variety of constructions. However, the fact that *not* is general purpose does not explain why children omit the auxiliary that precedes it. Further, if children have learned from their input that *not* occurs in a variety of syntactic contexts, this means they have already made certain syntactic generalizations that differentiate these contexts.

1.4. *Current study*

In this study we control for the statistical artefacts that are likely in naturalistic corpus data and assess whether children's earliest negative auxiliaries are compositional and are syntactically distinct from early uses of *not*. We use the Manchester (Theakston et al. 2001) and Manchester-Dense (Lieven et al. 2009) corpora in order to satisfy the three conditions shown in the following list. Together, the two corpora provide naturalistic data for 14 child–mother pairs over a 12-month period for each child, with collection sessions of around 1 hour in length, occurring roughly between age 2 and age 3.

- *First*, we divide the data into two developmental halves of roughly six months each. That allows us to look at developmental changes while mitigating the time course properties of infrequently occurring phenomena.
- *Second*, we use corpora that maximize consistency across samples in the length of each collection session and the type of session being observed (naturalistic, structured-play, elicited production task, etc.).
- *Third*, we examine longitudinal corpora to ensure uniformity in collection methodology across the developmental period of interest.

Analysis 1 examines the range and distribution of positive and negative auxiliaries used by children and their mothers to determine the relationship between positive and negative auxiliary use and whether 2-year-olds use a limited set of negators. Analysis 2 assesses children's use of *don't, can't*, and *not* to determine whether they are tense-bearing and select for right the syntactic categories. Analysis 3 focuses on the subject-person agreement errors, examining the prevalence of agreement errors with *don't, do, haven't, and have*. Analyses 4 and 5 assess usage-based accounts of negator development by examining the extent to which children's ungrammatical uses of *not* are reflected in the input.

2. General method

2.1. *Corpora selection*

We analyzed 12 mother–child dyads from the Manchester corpus (Theakston et al. 2001; 595,000 utterances) and 2 mother–child dyads from the Manchester Dense corpus (Lieven et al. 2009; 496,000 utterances), both from CHILDES (MacWhinney 2001). Both corpora provide transcripts of 1-hour naturalistic play sessions between mother and child (roughly) over a 12- to 14-month period for each child, with ages across all children ranging from 21 to 37 months. There are some differences in the rate of collection across this period, as shown in Appendix Table 1.[2] In the Manchester corpus, children were recorded on average every ten days. The Manchester Dense corpus included two six-week intensive collection periods (at the start and end of the collection period) in which children were recorded for five hours each week. For the rest of the collection period, children were recorded for five hours in one week at the start of each month. Despite the variation in collection rate, the uniformity of transcription length and sampling time period between these two corpora make them suitable for combined analysis.

2.2. *Procedure*

We performed our analyses using a custom Python program, making use of the morphological tagging provided in the CHILDES files. We categorized utterances as negative based on the presence of an element with the **neg** tag or of an element transcribed as "*no*" with the **co** tag.[3] For negative utterances, we extracted the negator, and for positive utterances we extracted elements tagged as auxiliaries (tag: **aux**), modals (tag: **mod**), and copulas (tag: **cop**).

Utterances were categorized into three types based on the terminating element of the transcription line. Lines ending in a "." were categorized as non-interrogatives, lines ending in a "?" were categorized as interrogatives, and lines ending with a "+ . . . " or "+//." were categorized as continuations. Continuations involve sentence fragments where a speaker has been interrupted by another speaker. Some are continued later in the discourse, others are not, and their true type might be an interrogative or non-interrogative. Our program only categorizes continuations without subcategorizing them.

We split the corpus into two developmental halves. Utterances from age 29 months or younger were classified as being in the first half and those from age 30 months or older were classified as being in the second half. As Appendix Table 2 shows, there is some variation in the age ranges of children in the corpus, meaning some children provide more data for the first half and some provide more for the second. Overall, however, the corpus division is fairly well balanced.

2.3. *Exclusions*

We excluded all utterances with an unintelligible element, all one-word utterances using a negator (e.g., "*no*," "*not*"), and all utterances involving the repetition of the same negator (e.g., "*no no*"). We removed all utterances containing more than one negator type (e.g., "*I can't remember when it was not actually*," "*no you're not naughty but kicking the dog is naughty, isn't it?*") to simplify the automated analysis. Finally, although the following elements were tagged as negators in CHILDES, we excluded them as non-standard[4]: '*s, nos, <don't, <can't, no(t), nots, (h)a(v)en't, (h)a(ve)n't, '(t), <not, you@d, i(s) n't, n't@d*, and *we*. In total, we excluded 5% of potential negative utterances, leaving 83,376 remaining negative utterances, representing 7.6% of the entire corpus.

[2]All appendices are provided in the supplementary material, available at https://doi.org/10.1080/10489223.2024.2356668.
[3]*No* is tagged as a communicative speech act element by CHILDES
[4]These items were tagged as a negator in CHILDES, either erroneously by the corpus creators or by the part-of-speech tagger that generates morphological information for CHILDES corpora. We chose to ignore them rather than attempt to infer what the transcriber intended.

3. Analysis 1

Analysis 1 examines the negators and positive auxiliaries that 2-year-olds produce. If children's first negative auxiliaries are compositional, they will produce a range of negative and positive auxiliaries. Furthermore, children's positive and negative auxiliary use will correlate: if all the components needed to generate negative auxiliaries are in place, the only limit on negative auxiliary use will be knowledge of particular auxiliary words. If, in contrast, children only have access to *not, no, don't*, and *can't*, only those forms will appear initially, in the first half. If children lack auxiliaries more generally, positive auxiliaries will initially be absent or limited in use and there will be no connection between negative and positive auxiliary use in the first half.

3.1. *Method*

For each developmental half we calculated mean production statistics for mothers and children (averaged by speaker); individual data by child and mother were also examined in order to ensure that the average statistics are not due to the contributions of a few children.

- Averaged data (by individual speaker) for children and mothers
  - the mean rate of occurrence of each negator
  - the mean rate of occurrence of each bare auxiliary in positive contexts
- Individual data
  - the number of children and number of mothers using each negator
  - the number of children and number of mothers using each positive auxiliary
  - the number of unique negators used by each child
  - the number of unique positive auxiliaries used by each child
  - the correlation between the range of negative auxiliaries used and positive auxiliaries used by children

3.2. *Results*

3.2.1. *Negators (including negative auxiliaries) as a group*

As a group, children use 18 different negators in the first half and 22 in the second half, while mothers use 27 different negators across both halves. Examples of children's utterances are shown in Table 1.

Figures 1, 2, and 3, show the mean rate of negator use for children and mothers averaged by individual and the number who use each negator in the corpus. The distributions of mothers' and children's negator use are long-tailed, with the majority of negators occurring at low rates for both mothers and children—less than 2 uses per 1,000 utterances.

The range of negator use is not due to one or two prolific children. In the first half, 9/18 negators—including *doesn't, isn't, didn't, haven't*, and *won't*—are used by at least half the children, and 14/18 negators are used by at least a quarter of the children. In the second half, 13/22 negators are used by half of the children, with 18/22 being used by at least a quarter of the children. For mothers, the majority of negators (19/27) are used by at least half.

In the first half, children as a group primarily use *not* and *no*, followed by *don't* and *can't*; in the second half, they primarily use *not, don't*, and *can't*. Mothers primarily use *don't*, followed by *not* and *isn't*.

3.2.2. *Negators, individual children*

The distribution of individual children's negator use is similar to the collective distribution in both halves (Appendix Figures 1 and 2). Children's productions of negative forms follow a power-law-like distribution.

Table 1. Sample of children's negative utterances.[a]

Utterance	Child	Age
"I no like sweetcorn"	Joel	2;0
"not train set"	Dominic	2;1
"Thomas not scaring me"	Carl	2;2
"I don't know"	Fraser	2;3
"no Mama in there"	Ruth	2;4
"lives in that one, doesn't he?"	Joel	2;5
"I can't do them, Mum"	Becky	2;5
"I haven't got boots"	Fraser	2;5
"I couldn't see any"	Eleanor	2;5
"it not nice car"	Dominic	2;6
"the piggy doesn't stand up"	Becky	2;6
"I don't think so"	Aran	2;6
"not three babies"	Gail	2;7
"because I don't like to get up"	Carl	2;7
"cause I just won't need Teddy"	Eleanor	2;0
"it's not finish yet"	Fraser	2;8
"him willn't get in there"	Nicole	2;9
"you mustn't"	Aran	2;10
"don't know"	Eleanor	2;11
"I haven't done it yet"	Dominic	2;11
"you not doing it anymore"	Ruth	3;0
"you can't go outside"	Nicole	3;0

[a]Age is in Year;Month format.

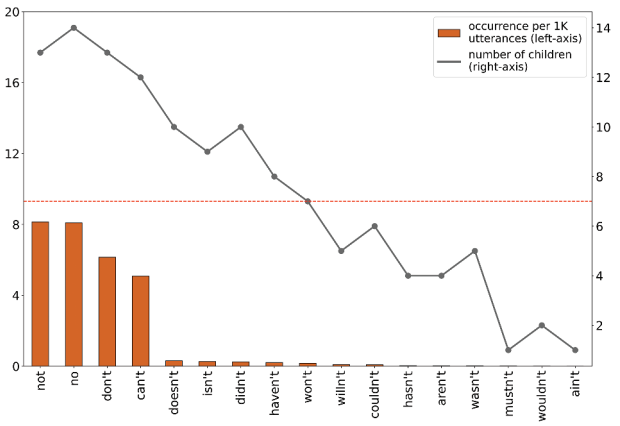

Figure 1. Children's mean rate of negator use, first half. The x-axis lists each negator produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of children who used the negator. The right-hand y-axis, the reference scale for the line plot, shows the number of children who used the negator.

In the first half of the corpora, the number of unique negator types used by each child ranged from 1 (n = 1) to 14 (n = 1). Nine out of the 14 children used nine or more negators. Two children used fewer than 5 negators: Ruth, who only used *no*, and Nicole, who used *don't, didn't, no*, and *not* (Appendix Table 3). Apart from Ruth, every child used *no, not*, and *don't*. The use of other negators varied without showing a clear developmental pattern. In the second half, the number of negator types ranged from six (n = 1) to 19 (n = 1), with 11 out of 14 children using nine or more negators (Appendix Table 4).

Children with smaller negator repertoires did not necessarily use a strict subset of the negators used by children with larger repertoires. For example, in the first half, Nicole produced four negators, among them *didn't*; John produced seven negators, but not *didn't*. Similarly, Warren produced six negators, among them *couldn't*; Anne produced 13 negators, but not *couldn't*. As in the first half, children in the second half with a smaller repertoire sometimes included negators not used by children

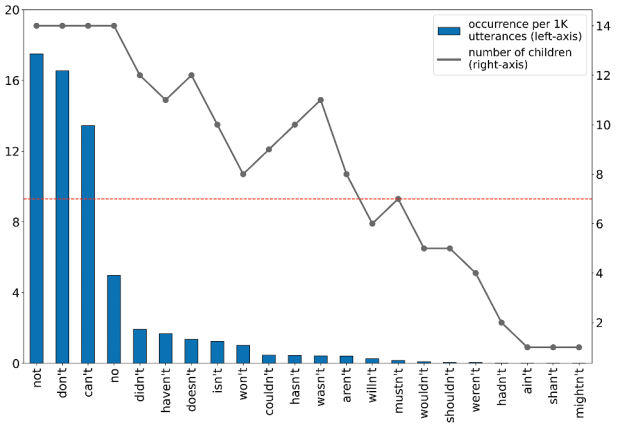

Figure 2. Children's mean rate of negator use, second half. The x-axis lists each negator produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of children who used the negator. The right-hand y-axis, the reference scale for the line plot, shows the number of children who used the negator.

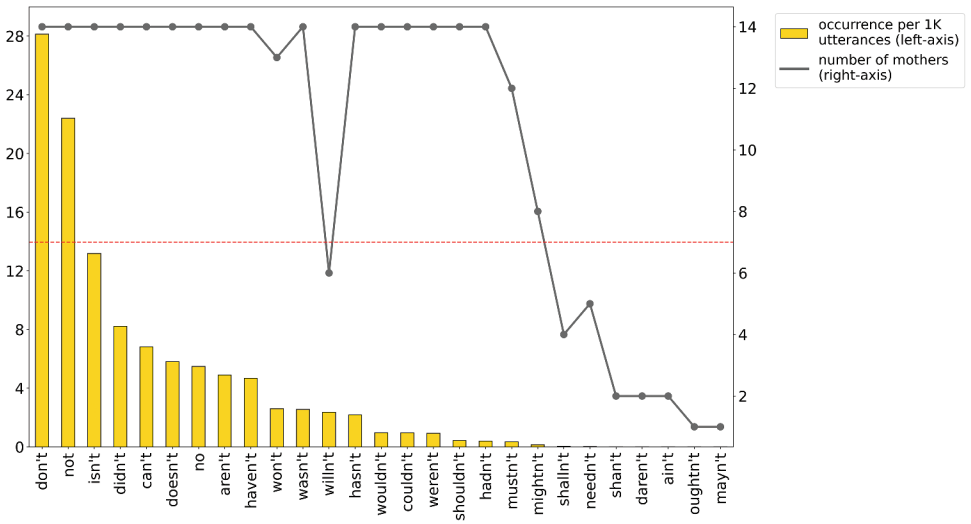

Figure 3. Mothers' mean rate of negator use, both halves. The x-axis lists each negator produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of mothers who used the negator. The right-hand y-axis, the reference scale for the line plot, shows the number of mothers who used the negator.

with a larger repertoire. *Hadn't*, for example is used by Nicole (12 negators), and not by Fraser (19), Aran (16), Liz (14), or Anne (13).

3.2.2.1. ***Child–mother differences.*** Individual children's negator preferences do not uniformly follow the preferences of their mothers. Table 2 shows the mean frequency rank of each negator averaged across children and mothers for the first half. *Can't* was, on average, the third most frequently used negator by children and the sixth most frequently used negator by mothers. *No* is similarly preferred more by children than mothers. *Isn't* is the third or fourth most common negator used by mothers, but the seventh most frequently used negator for children. Only *don't* and *not* were in the top four negators used by all children and their mothers.

Table 2. Average frequency rank for each negator across mothers and children in the first half of the corpus.[a]

Negator	CHI	MOT	Negator	CHI	MOT
ain't	14.0	19.0	mustn't	9.0	16.1
aren't	8.8	7.7	needn't	-	19.0
can't	3.2	6.1	no	2.4	6.9
couldn't	7.5	13.9	not	2.1	2.0
daren't	-	19.0	shalln't	-	18.0
didn't	6.3	5.2	shan't	-	18.0
doesn't	6.3	6.2	shouldn't	-	15.9
don't	2.2	1.3	wasn't	10.2	11.2
hadn't	-	16.8	weren't	-	14.2
hasn't	10.8	11.6	willn't	6.8	7.7
haven't	7.4	8.3	won't	5.7	11.9
isn't	7.0	3.6	wouldn't	11.0	14.9
mightn't	-	17.8			

[a]For each child and each mother, negators were ranked according to frequency of occurrence, and the mean rank then calculated for children and mothers. For example, "not" was, on average, the second most frequently used negator for both mothers and children.

The group data are mirrored by individual negator frequency ranks. *Can't* was one of the four most frequently used negators for all 12 children who produced it, but only in the top 4 for 3 of the corresponding mothers. Similarly, *no* was among the top four negators for all children, but only among the top 4 for 4 mothers. In contrast, *isn't* was one of the 4 most frequently used negators for 12 out of 14 mothers, but not for any child.

Similar patterns hold for the second half of the corpus (Table 3). Only *don't* and *not* were in the top 4 most frequently used negators used by all children and mothers. Again, *isn't* was 1 of the 4 most frequently used negators for 11 out of 14 mothers, but not for any child. *Can't* was highly preferred by all children, but only in the top 4 most used negators of 6 out of 14 mothers.

3.2.3. *Positive auxiliaries as a group*

Figures 4, 5, and 6 show the mean rate of auxiliary use in positive contexts. Children use 22 different positive auxiliary types in the first half and 23 in the second half. Mothers use 23 different positive auxiliaries across both halves. Like the negator distribution, the distribution of positive auxiliaries is long tailed, with the majority of auxilaries being produced at a rate of less than two occurences per 1,000 utterances. As the line plots show, despite the low rate of occurrence, the majority of auxiliaries are used by the majority of children and mothers.

Table 3. Average frequency rank for each negator across mothers and children in the second half of the corpus.

Negator	CHI	MOT	negator	CHI	MOT
ain't	17.0	19.0	mustn't	12.1	15.6
aren't	9.8	8.6	needn't	–	18.0
can't	2.3	5.2	no	4.7	8.1
couldn't	11.3	14.4	not	1.9	2.0
daren't	–	19.0	oughtn't	–	20.0
didn't	6.1	4.8	shalln't	–	19.3
doesn't	7.2	6.6	shan't	15.0	19.0
don't	1.9	1.3	shouldn't	14.8	15.9
hadn't	13.0	16.8	wasn't	9.3	10.2
hasn't	9.7	11.9	weren't	13.5	14.4
haven't	6.5	7.6	willn't	10.0	8.2
isn't	7.8	4.1	won't	6.0	9.5
mayn't	–	21.0	wouldn't	13.8	14.2
mightn't	18.0	18.1			

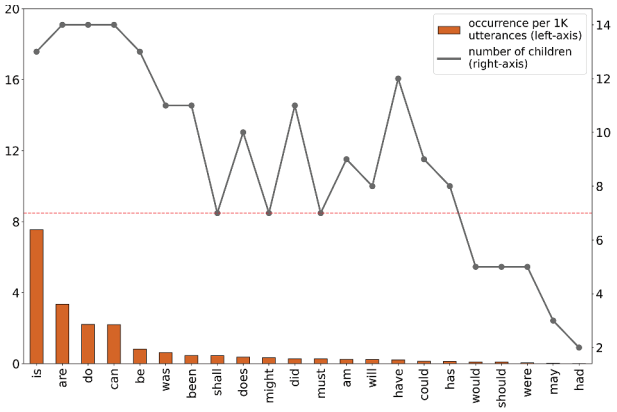

Figure 4. Children's mean rate of positive auxiliary use, first half. The x-axis lists each auxiliary produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of children who used the auxiliary. The right-hand y-axis, the reference scale for the line plot, shows the number of children who used the auxiliary.

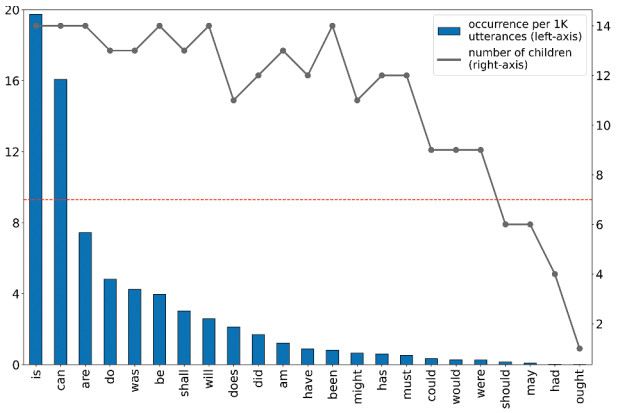

Figure 5. Children's mean rate of positive auxiliary use, second half. The x-axis lists each auxiliary produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of children who used the auxiliary. The right-hand y-axis, the reference scale for the line plot, shows the number of children who used the auxiliary.

In the first half, every child uses *do, can*, and *are* in a positive context. This includes Ruth, who produced no negative auxiliaries in the first half, and Nicole, who produced only *don't* and *didn't*. Although *are* is used by every child, only 4 out of 14 children produced *aren't*.

3.2.4. *Positive auxiliaries, individual children*

The distribution of individual children's positive auxiliary use looks similar to the group distribution in both halves (Appendix Figures 3 and 4). Children's productions of positive auxiliaries follow a power-law-like distribution.

The number of unique bare positive auxiliaries used by each child in the first half ranged from four (n = 1) to 22 (n = 2) (Appendix Table 5). Only three children used fewer than 10 different types of positive auxiliaries, John (9 different auxiliaries), Nicole (8), and Ruth (4). These children also had smaller ranges of negative auxiliary use. Smaller auxiliary repertoires were not necessarily proper subsets of larger repertoires. *Am* was used by Nicole (8 unique auxiliaries) but not by Warren (13) while *did* was used by John (8) but not Aran (13). In the second half, the number of unique positive

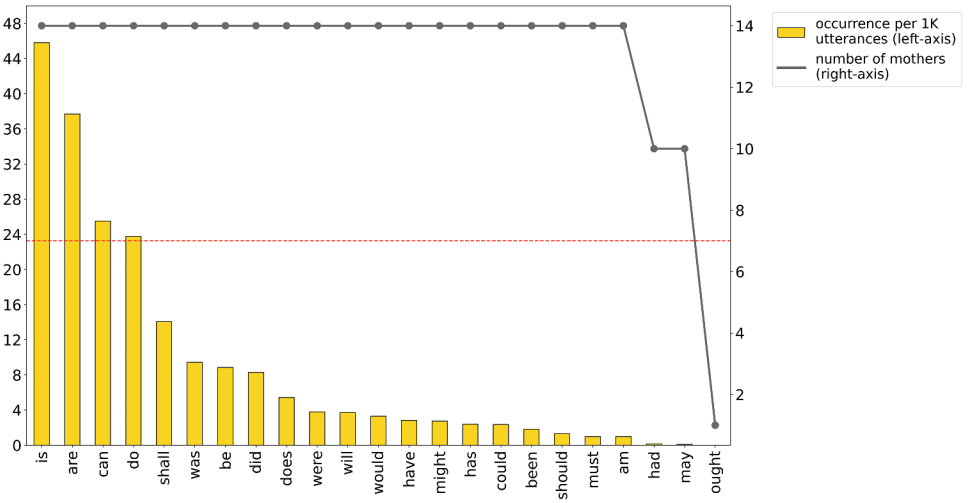

Figure 6. Mothers' positive auxiliary use, both halves. The x-axis lists each auxiliary produced. The left-hand y-axis, the reference scale for the bar-plot, shows the mean number of uses per 1,000 utterances averaged across the number of mothers who used the negator. The right-hand y-axis, the reference scale for the line plot, shows the number of mothers who used the auxiliary.

auxiliaries used by each child ranged from 11 (n = 1) to 22 (n = 4). Each child used *can, is*, and *are*, while *do* and *was* were used by 13/14 children.

3.2.5. *Positive auxiliaries vs. negative auxiliaries*

Figure 7 shows, for each child in each half, the relationship between the number of different positive auxiliaries used (on the x-axis) and the number of negative auxiliary types used (*n't* negators) (on the y-axis). The positive correlation was .89 in the first half (Pearson $r(12) = 0.89$, $p < 0.001$) and .95 in the second ($r(13) = 0.95$, $p < 0.001$). No child used more negative than positive auxiliaries. Ruth, who did not use an *n't* negator in the first half, used four different positive auxiliaries.

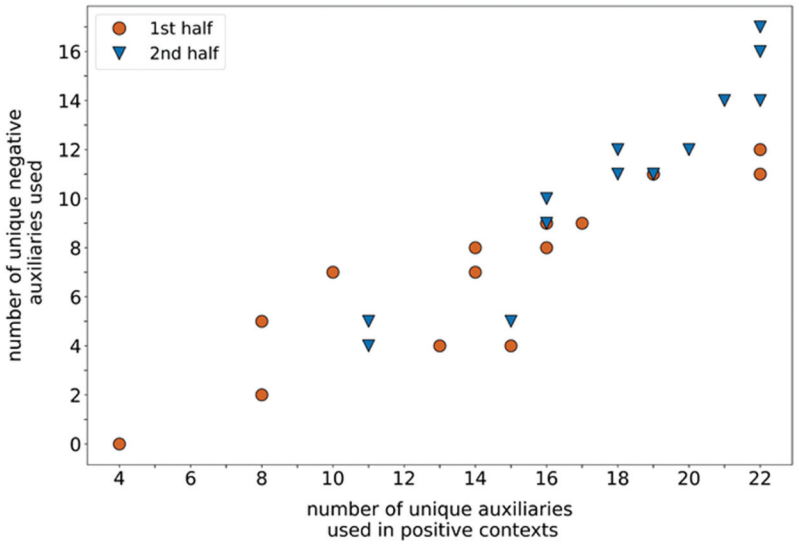

Figure 7. The number of unique positive and negative auxiliaries used by each child. Each circle represents a child in the first half. Each triangle represents a child in the second half. The x-axis shows the number of unique auxiliaries a child uses in positive sentences. The y-axis represents the number of unique negative auxiliaries a child uses.

3.3. *Discussion*

Children use a wide range of negators prior to age 3. Contrary to some earlier reports (Klima & Bellugi 1966, Bloom 1970, *inter alia*), children are not limited to *don't, can't, not*, and *no*. The range is apparent when looking at children as a group and individually. In addition, different children use different collections of negators. Negator use follows a long-tailed distributional pattern for both children and mothers. Most negators occur infrequently in child speech (both before and after 2;6) and in parental speech. Individual children's negator productions also follow a long-tailed distributional pattern. Beyond *no, not*, and *don't*, the order in which different negators first appear in the corpus is developmentally varied.

As with negative auxiliaries, so for positive auxiliaries. Children use a range of positive auxiliaries during age 2; the distribution of their auxiliary use is long-tailed, as is their mothers'. While the token frequency per 1,000 utterances of positive auxiliaries is generally lower than the token frequency of *don't, can't, not*, and *no*, all children use a wider range of positive than negative auxiliaries. Their use is also correlated: children who use more positive auxiliaries also use more negative auxiliaries. Furthermore, each child used *can* and *do* in the first half, including those who didn't produce *don't* and those who didn't produce *can't.*

In sum, children use a wide range of negative and positive auxiliaries and the distribution of their use is similar to their mothers'. The data suggest that the underlying mechanism generating negative auxiliaries is the same for both sets of speakers. Thus, the results provide evidence against two key components of Misanalysis proposals. Two-year-olds are neither limited to a small subset of negative forms nor do they lack positive auxiliaries.

The long-tailed nature of the negator distributions confirms our conjecture that previous claims of few negators were an artefact of small sample size. Negation occurs infrequently. Mothers use a negator in about 10% of their utterances. Negation is even less frequent in children, particularly below age 2;6. Cross-sectional studies and longitudinal studies with infrequent data collections are unlikely to pick up the auxiliaries that children use. Long tails also explain why children can appear to lack auxiliaries in positive contexts: auxiliaries in positive contexts are also infrequent. When utterances are split into two halves, we get a better window into children's underlying competence with these infrequent forms.

Children have different negator preferences than their mothers, a fact that is more parsimonious with theories that attribute input-independent syntactic knowledge to children than those which consider early language as predominantly input-driven. While children and mothers share a preference for *don't* and *not*, children have a stronger preference for *no* and *can't* than their mothers, while mothers have a stronger preference for *isn't* and *didn't* than their children. Children's productions are not obviously driven by the most common negators their mothers are using.

Why do children display the particular preferences they do? One possibility is that a range restricted to *no, not, don't*, and *can't* occurs very early and for a very brief period of time. All but one child uses a negator outside the restricted range in the first half. Thus, if a brief monomorphemic stage characterized by few negators exists, children must exit it very early, before age 2.

Why do children's preferences differ from their mothers'? One possible source is pragmatic: children and mothers communicate different things about themselves and the world. In the second half, when children are using a wide array of negative and positive auxiliaries, children continue to favor *can't* while mothers continue to favor *isn't*. Pragmatic differences in how different negators are used seem the most likely cause of differences in preferences.

Thus far, our analysis shows that children use a wide range of negative and positive auxiliaries throughout age two. That is initial evidence for the Early Competence Hypothesis. More convincing evidence requires an analysis of the distribution of negative elements. If the Early Competence Hypothesis is correct, children will treat negative auxiliaries as tense-bearing elements that have the same selectional properties as their positive counterparts. Analysis 2 examines children's uses of *don't* and *can't* with respect to these two properties.

4. Analysis 2

English sentences explicitly encode tense. As noted in the Introduction, in English sentences with an auxiliary and a verb, only the auxiliary is tensed. If children's early uses of *don't* and *can't* are monomorphemic, they will use *don't* and *can't* with other tense-bearing elements in constructions that are ungrammatical for adults: with verbal inflections (examples (25) and (26)), with auxiliaries (27), and with auxiliary clitics (28).

25.	*They don't liked it	(past tense inflection on verb)
26.	*She can't jumps there	(present tense inflection on verb)
27.	*They did don't like it	(additional past-tense inflected auxiliary)
28.	*It's can't fit there	(auxiliary clitic attached to subject)

In this respect, monomorphemic expressions of *don't* and *can't* should pattern like *not*, which is untensed in adult English, and **requires** a preceding auxiliary:

29.	*They not like it/They did not like it
30.	*I not made it/I have not made it/I did not make it

Finally, if *don't* and *can't* are monomorphemic, children should treat both as general purposes negators that can occur in contexts that are ungrammatical in adult English: before an Adjectival Phrase (31), before a Determiner Phrase (32 and 33), and before non-finite VPs (34 and 35).

31.	*It's don't red	(AP)
32.	*That's don't the ball	(DP)
33.	*That's can't the ball	(DP)
34.	*They're don't working	(non-finite VP)
35.	*They're can't working	(non-finite VP)

Conversely, if early *don't* and *can't* are composed of a tense-bearing auxiliary with *n't*, they will be the only tensed elements of the sentence. Tensed verbal inflections, other auxiliaries, and clitics will be absent in sentences with *don't* and *can't*. They should almost exclusively precede verbs, and in this respect pattern like "genuine" negative auxiliaries like *didn't, doesn't, won't*, and *willn't*.[5]

4.2. *Method*

To assess the generality of negator use, we examined the morphological tag of the element immediately following the negator. We ignored all instances where the negator was the last element of the sentence. Although the following items were not tagged as lexical verbs in the morphological lines, they were treated as verbs in this analysis: *like* (tagged as a communicator), *help* (tagged as a communicator), *seem* (tagged as an auxiliary), *look* (tagged as an auxiliary), *stay* (tagged as an auxiliary), *have_to* (tagged as modal), and *wanna* (tagged as modal).

To assess the presence of overt tense markers, we searched for instances of auxiliaries, modals, auxiliary clitics, and verb inflections. Instances of auxiliaries and auxiliary clitics were identified by searching for elements with the **mod**, **aux**, or **cop** morphological tag immediately preceding the negator. Instances of tense-marked verbs were identified by searching for verbs that were tagged as having a past tense (tag: **PAST**) or present tense (tag: **PRES**) morpheme immediately following the negator.

[5]We call these "genuine" negative auxiliaries because Misanalysis proposals have typically considered only *don't* and *can't* to be monomorphemic due to their prevalence in children's speech.

We included only non-interrogative utterances in order to simplify the analysis of category selection. Children's negative utterances in this corpus are overwhelmingly non-interrogative (~95%).

We do not exclude possible routines from analysis, such as "*I don't know*" or "*I can't do xxx*." Given the large universe of possible routines, it is difficult to determine what a routine is, so we opted for inclusion. We performed another analysis (that we do not report) removing utterances with the above two locutions; that analysis gave the same results as what we report in the following sections.

4.3. *Results*

4.3.1. *Category Selection*

4.3.1.1. *Not.* Children use *not* with multiple morphological categories. Figure 8 shows the distribution of morphological categories immediately following *not* averaging across children in the first half. No single category follows *not* more than 24% of the time. This pattern holds for individual speakers as well. No child uses any particular category more than 40% of the time immediately following *not*. The most biased child, Ruth, follows *not* with a verb 37% of the time.

4.3.1.2. *Don't.* Children overwhelmingly use *don't* with lexical verbs. Across the entire corpus children produce 6,568 non-interrogative utterances with *don't*. On average, 97% of the uses in the first half immediately precede a lexical verb (N = 13, min = 93%, max = 100%, SD = 2%) and 96% immediately precede a verb in the second half (N = 14, min = 92%, max = 100%, SD = 2%). Mothers use *don't* preceding a lexical verb 93% of the time (N = 14, min = 87%, max = 97%, SD = 3%). Logit-transformed paired *t*-tests show that children place lexical verbs after *don't* significantly more often than mothers do in both the first half ($t(12) = 3.97$, $p = 0.002$) and the second half ($t(13) = 3.259$, $p = 0.006$). Mothers' lower use of *don't* before lexical verbs is primarily due to their use of negative imperatives: mothers follow *don't* with a pronoun and with *be*, both of which are grammatical in imperative constructions.

4.3.1.3. *Don't vs. didn't and doesn't.* *Don't* patterns like *didn't* and *doesn't*. Across the entire corpus there is no difference in the percentage of children's productions of *don't* that precede a lexical verb and the productions of *didn't* that precede a lexical verb ($M_{don't} = 96\%$, $M_{didn't} = 90\%$, paired $t(9) = -1.30$, $p = 0.23$). The same is true for *doesn't* ($M_{don't} = 96\%$, $M_{doesn't} = 94\%$, paired $t(12) = -1.23$, $p = 0.25$).

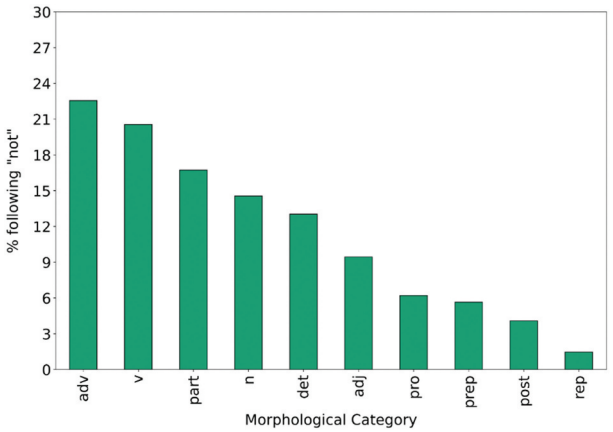

Figure 8. Children's top 10 morphological categories following "not." The x-axis shows the morphological category of the item immediately following "not." The category labels follow the CHILDES morphological tagging convention: adv = adverb; v = verb; part = participle; n = noun; det = determiner; adj = adjective; pro = pronoun; prep = preposition; post = post-nominal modifier; rep = a repetition of the negator.

4.3.1.4. *Can't.* Children overwhelmingly use *can't* with lexical verbs. Across the entire corpus children produce 3,563 non-interrogative utterances with *can't*. On average, 94% of the uses in the first half immediately precede a lexical verb (N = 12, min = 80%, max = 100%, SD = 6%)[6] and 97% immediately precede a verb in the second half (N = 14, min = 91%, max = 100%, SD = 2%). Mothers use *can't* preceding a lexical verb 90% of the time (N = 14, min = 82%, max = 100%, SD = 5%). Logit-transformed paired *t*-tests show no difference between children and their mothers in the first half ($t(10) = 1.44$, $p = 0.18$). In the second half, children are more likely to follow *can't* with a lexical verb than their mothers ($t(13) = 2.38$, $p = 0.03$). This difference is driven by mothers using *can't* before an adverb (e.g., "*I can't really see*") and before *be* (e.g., "*she can't be very hungry*") more often than their children.

4.3.1.5. *Can't vs. Won't and Willn't.* For this comparison we looked at utterances of *won't* and *willn't* together, since some children only use one variation. *Can't* patterns like *won't/willn't*. Across the entire corpus there is no difference in the percentage of children's productions of *can't* that precede a lexical verb and the productions of *won't/willn't* that precede a lexical verb ($M_{can't} = 96\%$, $M_{won't/willn't} = 94\%$, paired $t(13) = -1.38$, $p = 0.19$).

4.3.2. *Additional tense*

4.3.2.1. *Don't.* As predicted, *don't* does not co-occur with other tense-bearing elements. Two hundred uses of *don't* contain a 3rd person singular subject, and only seven of these (3.5%) occur with an additional auxiliary or auxiliary clitic. One hundred and seventy-one (171) utterances with 3rd person singular subject also have a verb immediately following *don't*. Only five of these (3%) contain present or past tense marking on the verb. In uses of *don't* with other subject types and without an identified subject, no utterances contain an additional auxiliary or auxiliary clitic and five out of the 5,349 that precede a verb (0.09%) contain tense inflection on the verb. Similar percentages are found for uses of *didn't* and *doesn't*. Four hundred and fifty-five sentences with *didn't* and *doesn't* contain a 3rd person singular subject, eight of which (1.7%) contain an additional auxiliary or clitic. Three hundred and eighty nine sentences (389) with a 3rd person singular subject also have a verb immediately following *doesn't* or *didn't* and 11 of these (2.8%) have tense inflection on the verb. In sentences with other subject types or without an identified subject, only 2 (0.4%) contain an additional auxiliary or clitic, while 17 (4%) contain tense inflection on the verb. Notably, only four out of the 14 children produce errors with *don't*. These children are all relatively advanced, using 10 or more negators in the first half. Errors with *don't* (as well as *didn't* and *doesn't*) also occur predominantly in the second half (Appendix Tables 6 and 7). For mothers, out of the 11,605 non-interrogative uses of *don't* in this data set, none occur with an additional auxiliary or auxiliary clitic and none precede verbs with past or present tense inflection.

4.3.2.2. *Can't.* *Can't* does not co-occur with other tense-bearing elements. *Can't* is used 554 times with a 3rd person singular subject, with only five of these (0.9%) occurring with an additional auxiliary or auxiliary clitic. Four hundred and eight (408) sentences with a 3rd person singular subject also have a verb immediately following *can't*, and just three of these (0.7%) contain present or past tense inflection on the verb. In uses of *can't* without a subject or with other subject types, two out of 3,218 (0.06%) contain an additional auxiliary or auxiliary clitic and 13 (0.4%) contain tense inflection on the verb. Uses of *won't* and *willn't* follow a similar pattern. Only two out of the 134 (1.5%) sentences with a third person singular subject contain an additional auxiliary or auxiliary clitic, and none of the 89 sentences where *won't* and *willn't* are followed by a verb contained inflection on the verb. In uses of *won't* and *willn't* without a subject or with other subject types, just one out of 264 uses contains an inflection on the verb

[6]Liz, the child with the lowest percentage of verbs following *can't*, produced only 10 utterances with *can't*. The 2 non-adult-like utterances were "*can't bricks*" and "*can't Mummy*." While the former might be a category selection error, the latter seems like a vocative use of "*Mummy*."

(0.3%) and just one contains an extra auxiliary. Children's errors with *can't* are shown in the Appendix (Table 8). Again, these sorts of construction are non-existent in the mothers' speech.

4.3.2.3. *Not.* Children frequently use *not* with an additional tensed element, particularly a preceding auxiliary or auxiliary clitic. Children produce 2,168 non-interrogative utterances using *not* with a 3rd person singular subject. Of these, 1,576 (72%) occur with an auxiliary or auxiliary clitic. Two hundred and twenty-eight (228) uses of *not* are also immediately followed by a verb, and 23 of these (10%) have tense marking on the verb. Our inspection of these 23 utterances suggests that many involve participles or statives that have been misclassified as lexical verbs. For example, Fraser says "it not got some milk in"; this does not seem to be a true instance of tense information being expressed as an inflection on a lexical verb. Looking at sentences with other subject types and without subjects, 668 out of 4036 uses of *not* (17%) contained a preceding auxiliary or auxiliary clitic. Thirty-six out of the 714 uses that also precede a verb contain tense inflection on the verb (5%). As with 3rd person singular subjects, many of these do not look like true tense inflections. All 14 children produced at least one of use of *not* with a tensed element (a sample is shown in Analysis 4). Mothers produced 3,504 non-interrogative uses of *not* with a 3rd person singular subject. Ninety three percent (93%) occur with an auxiliary or auxiliary clitic. Of the 155 sentences where *not* also precedes a verb, eight (5%) had tense marking on the verb. Looking at sentences with other subjects and sentences without a subject, 2,571 of mothers' 5,388 uses of *not* include an auxiliary or auxiliary clitic (51%). Of the 343 utterances where *not* precedes a verb, 44 have tense marking on the verb. Our inspection of these utterances suggests that, as with the children's data, many involve participles or statives that have been misclassified as lexical verbs. For example, John's mother says "You're not stuck" and Gail's mother says "I've not got these on yet". Again, these do not seem to reflect tense being expressed on a verb.

4.4. *Discussion*

Children's *don't* and *can't* behave like tensed auxiliaries that select for a verbal element, following the regularities of the adult grammar and providing evidence in favor of the Early Competence Hypothesis. Children use *not* before a range of different morphological categories, but they overwhelmingly use *don't* and *can't* before verbs. Replicating Schütze's (2010) analysis of *don't* and extending the analysis to *can't*, we find no evidence that *don't* and *can't* are ever used as a general purpose negator during age 2. Children almost never use *don't* or *can't* with other tense-bearing elements, confirming their multimorphemic status. *Don't* patterns like *didn't* and *doesn't*, and *can't* patterns like *won't* and *willn't*. The rare instances of children producing additional tensed elements with a negative auxiliary occur almost exclusively in the second half of the corpus, occur with *didn't* and *doesn't* in addition to *don't* and *can't*, and are produced by more advanced children (perhaps because children produce many more tensed elements in the second half, along with errors of tense or auxiliary doubling, per Mayer et al. 1978, Hiramatsu 2003, Stromswold 1990, Woods 2016). The data at hand suggests that *don't* and *can't* are multimorphemic during age 2. The errors with *don't* and *can't* are infrequent enough to be production errors rather than reflecting a syntactic generalization. Children's earliest uses of *don't* and *can't* reflect the properties of true auxiliaries.

While *not* is a general purpose negator, children use *not* with a tensed element less frequently than their mothers do (36% vs. 68% across all uses of *not*). Since English sentences require tense to be expressed, we suspect that the "omissions" by mothers are in fact sentence fragments, like "*not that one*." Children on the other hand, produce utterances like "*I not want that*" and omit tense in positive contexts. As we noted in the Introduction, the prevalence of these errors either must be related to sequences in children's input and how children make use of these sequences, or must reflect an early generalization involving either *not* or the auxiliary elements that are omitted. We examine *not* in more detail in Analyses 4 and 5.

Two previous studies (Thornton & Tesan 2013, Thornton & Rombough 2015) have findings that differ greatly from ours in the proportion of tense errors two- and three-year-olds produce with *don't* when used with 3rd person singular subjects: 21% in Thorton & Tesan (2013, Appendix C, p. 408) and

36% in Thorton & Rombough (2015:Tables 5 & 6:145-146), compared to our ~3%. We note that these two studies involved elicitation tasks, and as such do not lend themselves to direct comparison with naturalistic corpora.[7] Furthermore, those studies still report a difference between *not* and *don't*, with *not* occurring with a tensed element between 43% (Thornton & Rombough 2015) and 61% of the time (Thornton & Tesan 2013). This difference is consistent with *don't* being tensed. Compared to *not, don't* resists co-occurrence with a tensed element.

In sum, the results of Analysis 2 suggest that children obey the distributional regularities governing negative auxiliaries in the adult grammar, distinguishing between *don't* and *can't* on the one hand and *not* on the other hand. This finding is evidence in favor of the Early Competence Hypothesis: children treat *don't* and *can't* as negative auxiliaries composed of an auxiliary and *n't*, and treat *not* as a negator with a more general distribution.

Children also produce non-target sentences like (36), where *don't* is used with a third-person singular subject, as noted in the literature (Bellugi 1967, Guasti & Rizzi 2002, Schütze 2010, Thornton & Rombough 2015).

36. *He don't want to play.

Such productions could be seen as evidence that children's uses of *don't* are not compositional (Bellugi 1967); if *don't* is monomorphemic, it lacks the necessary person and number features that determine how it is expressed with third-person singular subjects. The results of Analysis 2 suggest morphological simplicity is not the cause of subject agreement errors. Children's uses of *don't* are overwhelmingly grammatical with respect to tense and category selection, suggesting they are tense bearing auxiliaries. In Analysis 3 we examine the possibility that subject agreement errors extend beyond *don't* and reflect overgeneralizations of the most common morphological form.

5. Analysis 3

Analysis 3 examines the prevalence of subject-agreement errors with *don't* and whether those errors generalize to *do, have*, and *haven't*. Children produce non-target sentences where *don't* is used with a third-person singular subject, as in (36) (Bellugi 1967, Guasti & Rizzi 2002, Schütze 2010, Thornton & Rombough 2015). While such errors are consistent with *don't* being monomorphemic, they are also consistent with children not having fully acquired the underlying agreement paradigm of *do*. *Do* is expressed as *does* when it occurs with a third-person singular subject. If children haven't fully acquired this alternation, they will produce *do* with third person singular subjects in positive contexts, as in (37)

37. *It do go there.
38. *Do he want to play?

and (38):
Errors like (37) and (38) would be evidence that errors like (36) are in fact compositional: agreement errors with *do* persist when combined with *n't*.

Have and *haven't* follow the same paradigm as *do* and *don't*. When occurring with a third-person singular subject, *have* alternates with *has* and *haven't* alternates with *hasn't*. If children overgeneralize

39. *She haven't finished
40. *It have done it
41. *Have she finished?

[7] In particular, across both studies children gave a range of responses to the prompt in question involving a number of different negators. When looking at all responses given to the prompt, children produced *don't* with an additional tensed item only 4% of the time in both Thornton & Tesan (2013) and Thornton & Rombough (2015).

this type of agreement paradigm, they will produce errors like (39), (40), and (41):
Analysis 3 examines the proportion of third person singular subjects that have a non-target auxiliary for *don't/doesn't, do/does, haven't/hasn't, and have/has* in children's and mothers' productions. When a child uses a third-person singular subject and an auxiliary, how likely are they to use the correct morphological form of that auxiliary? If children's third-person subject agreement errors generalize beyond *don't*, we expect to see non-target uses of *do, haven't*, and *have* that parallel agreement errors with *don't.*

5.1. *Method*

We examined the features of subjects in their canonical (i.e., target) position for both positive and negative sentences. For non-interrogative sentences, the presence of a subject in a clause was determined using the computational procedure outlined in the Appendix (Determining Subjects section). This procedure looked for acceptable determiner phrases (DPs) that followed a clause boundary, and appeared immediately preceding the negator/auxiliary and some acceptable filler. Only utterances where such a DP was found were treated as having a subject. For interrogatives, we looked for the element immediately following auxiliary. An element tagged as a pronoun, noun, or demonstrative was treated as the subject. For elements that were tagged as determiners, if the next element was a noun, that element was treated as the subject.

We included interrogatives in this analysis because over half of children's positive auxiliary productions were interrogatives.

5.2. *Results*

5.2.1. *Don't vs. do*

On average children produced a subject agreement error in 55% (min = 6%, max = 100%, SD = 35%) of their utterances with *don't* and in 21% (min = 0%, max = 100%, SD = 36%) of their utterances with *do* (Table 4). Logit-transformed paired *t*-tests found this difference to be significant ($t(12) = 2.94$, $p = 0.12$). Every child produced a non-target use of *don't*. Eight out of the 11 children produced a non-target use of *do.*

Table 4. Children's productions of "don't" vs. "doesn't" and "do" vs. "does" with a third-person singular subject.

child	Total number of uses of "**don't**" and "**doesn't**" with third person singular subject[a]	Proportion of these uses with "**don't**"[b]	Total number of uses of "**do**" and "**does**" with third person singular subject[c]	Proportion of these uses with "**do**"[c]
Anne	16	0.06	21	0
Aran	34	0.91	15	0.2
Becky	49	0.1	116	0
Carl	10	0.9	2	1
Dominic	51	0.49	14	0.21
Eleanor	144	0.09	124	0.05
Fraser	168	0.27	624	0.04
Gail	10	0.6	17	0.06
Joel	6	0.33	7	0
John	7	1	–	–
Liz	7	0.57	10	0
Nicole	34	0.91	12	0.17
Ruth	5	1	1	1
Warren	15	0.53	2	0
Sum	**556**		**965**	
Mean		**0.55**		**0.21**
Std		**0.35**		**0.36**

[a]Combined count of "don't" and "doesn't" productions with a third-person singular subject.
[b]Proportion of those total uses where "don't" was used. For example, Dominic produced 51 uses of a third-person singular subject with either "don't" or "doesn't", and 49% of those uses were with "don't" (i.e. ungrammatical).
[c]Same data for "do" vs. "does."

Table 5. Children's productions of "haven't" vs. "hasn't" and "has" vs. "hasn't" with a third-person singular subject.

child	Total number of uses of "**haven't**" and "**hasn't**" with third person singular subject[a]	Proportion of these uses with "**haven't**"[b]	Total number of uses of "**have**" and "**has**" with third person singular subject[c]	Proportion of these uses with "**have**"[c]
Anne	13	0.23	3	0
Aran	4	0.75	17	0.12
Becky	11	0.09	20	0.05
Carl	–	–	1	0
Dominic	12	0.08	16	0.19
Eleanor	43	0.16	40	0.1
Fraser	25	0.16	84	0.02
Gail	3	0.67	7	0.14
Joel	3	0.33	4	0.25
John	–	–	5	0.8
Liz	2	0.5	3	0.67
Nicole	5	1	2	0
Warren	1	1	3	0.33
Sum	**122**		**205**	
Mean		**0.45**		**0.21**
Std		**0.35**		**0.26**

[a]Combined counts of "haven't" and "hasn't" productions with a third-person singular subject.
[b]Proportion of those total uses where "haven't" was used. For example, Eleanor produced 36 uses of a third-person singular subject with either "haven't" or "hasn't", and 16% of those uses were with "haven't" (i.e., ungrammatical).
[c]Same data for "has" vs. "hasn't."

5.2.2. *Haven't vs. have*

On average children produced a subject agreement error in 45% (min = 9%, max = 100%, SD = 36%) of their utterances with *haven't* and in 21% (min = 0%, max = 80%, SD = 26%) of their utterances with *have* (Table 5). Logit-transformed paired *t*-tests found this difference to be marginal ($t(10) = 2.25$, $p = 0.05$). Eleven children produced *haven't* or *hasn't* (or both) with a third-person singular subject; every one of these children produced a non-target use of *haven't*. Thirteen children produced *have or has* with a third-person singular subject, eight produce at least one non-target use of *have.*

5.2.3. *Don't vs. haven't*

Logit-transformed paired *t*-tests showed no difference between proportion of non-target uses of *don't* and non-target uses of *haven't* ($t(10) = -1.18$, $p = 0.26$). Every child produced a non-target use of *don't*. Each of the 11 children who produced *haven't* produced it with a non-target subject.

5.2.3.1. *Mothers.* Mother's auxiliary use with third-person singular subject is always grammatical (96-99%) of the time.

5.3. *Discussion*

Children use third-person singular subjects with *don't, do, haven't*, and *have*. There is no difference in the proportion of non-agreeing uses of *don't* and non-agreeing uses of *haven't*. The data are sparse, with some children producing only a handful of uses of *do, have*, and *haven't*. Nonetheless, the majority of children who produce these auxiliaries produce them with ungrammatical subject agreement. The distribution of these non-target uses across different children suggests that these errors reflect a general process.

While children in our study produce subject agreement errors with *do* less often than with *don't*, our results are in the same direction as previous findings. Jasbi et al.'s (2021) cross-sectional aggregated analysis of 571 children showed that with third-person singular subject, *do* is preferred over *does* at the around the same rate as *don't* is preferred over *doesn't*. The specific proportions in our sample might be an anomaly.

These data are best accounted for by the Early Competence Hypothesis. If *don't* is a combination of *do* and *n't*, we expect agreement errors with *do* to transfer to *don't*. The same holds for *have* and *haven't*. It is hard to see how the errors could be the result of *don't* being monomorphemic. A separate cause would be needed to explain the agreement errors with *do* and *have*. We would also need to treat *haven't* as morphologically simple, with no independent reason to think it is. As Analysis 1 shows, *haven't* is about the 7th or 8th most popular negator for mothers and children.

We have suggested that these agreement errors are overgeneralizations of the most common expression of the relevant auxiliary paradigm. Schütze (2010) alternatively proposes that non-agreeing *don't* in child speech is the result of underspecified tense and agreement features in children's syntactic representations. The consequence of this under-specification is that *n't* has no lexical item to attach to, which necessitates an uninflected "dummy" *do* being inserted to host *n't*. This account is specific to *don't*, however; additional components are needed to account for errors with *do, have*, and *haven't*. It is beyond the scope of this paper to probe the exact cause of the subject agreement errors children make, but treating these errors as overgeneralizations provides a parsimonious explanation for why they occur with auxiliaries other than *don't*.

6. Analysis 4

The previous analyses show that children use a wide range of negators during age 2, and that their use of *don't* and *can't* is highly grammatical with respect to category selection and the co-occurrence of sentential tense markers. This grammaticality is consistent with two-year-olds possessing an auxiliary category, a negation category, and a representation of tense, and combining them to produce negative auxiliaries—the Early Competence Hypothesis.

In contrast to the Early Competence Hypothesis, usage-based proposals characterize early child speech as item-specific formulae, reflective not of abstract categorization but rather of the input children are exposed to. Highly grammatical uses of *don't* and *can't* are consistent with usage-based proposals: if children's production of negative utterances is driven by what they hear, their negative auxiliaries will appear grammatical.

Analyses 4 and 5 evaluate the usage-based account of negative productions by examining children's ungrammatical utterances when *not* is used in verbal constructions. In adult English *not* + *V* sentences with a subject are only grammatical when *not* is preceded by a tense-bearing modal or dummy auxiliary:

42. I will not buy a new one
43. I did not buy a new one
44. *I not buy a new one

Previous research has found that children produce sentences like (44) (Klima & Bellugi 1966, Harris & Wexler 1996, Cameron-Faulkner et al. 2007, Thornton & Rombough 2015). Our own Analysis 2 found that children use *not* with a preceding auxiliary only 38% of the time, and mothers 68% of the time. Analysis 2 was not, however, restricted to verbal utterances that had a subject. We perform that analysis here. If negation in general is input-driven, we expect children to use sentences like (44) at the same rates as their mothers. However, if children use sentences like (44) more frequently than their mothers, this suggests they have made certain generalizations about the forms involved in negation.

6.1. *Method*

We examined the morphological tag of the item immediately preceding and immediately following *not* for sentences with *not* as the negator. We restricted our analysis to declarative sentences with subjects in their canonical (i.e., target) position. The presence of a subject in a clause was determined using the same computational procedure as Analysis 3 (outlined in the Appendix, Determining Subjects section).

We classified *not* + V sentences as those where the item immediately following *not* was classified as a verb (tag:**v**). Tense-bearing elements immediately preceding *not* were those classified as auxiliaries (**aux**), modals (**mod**), copulas (**cop**), and possessives (**poss**). We included copulas and possessives even though these items do not canonically precede *not*, because the morphological tagger appears to misidentify modals and auxiliaries preceding *not* as one of those two categories.

Our inspection of sentences suggests that some *not* + V constructions attributed to mothers are more likely children's utterances. For example, it is unlikely that a mother said "me not care." We do not have corpus-wide statistics on attributions, but our inspection suggests they are relatively infrequent. In light of this we have decided not to try to correct for misattributions.

6.2. *Results*

Overall children produced 421 declarative sentences with an explicit subject where *not* preceded a verb. Mothers produced 259 such sentences. Figure 9 shows the proportion of such sentences where a tense-bearing element preceding *not* is missing. A mixed effects logistic regression assessed the probability that such a sentence was produced as a function of speaker (child or mother), with individual children treated as random effects using a random intercept structure. The model's total explanatory power is substantial (conditional $R^2 = 0.54$) and the part related to the fixed effects alone (marginal R^2) is 0.44. The model's intercept, corresponding to speaker = child, is at 0.83 (95%CI [0.30, 1.37], p = 0.002). Within this model, the effect of speaker = mother is statistically significant and negative (beta = -3.67, 95%CI [-4.27, -3.08], $p < .001$; Std. beta = -3.67, 95%CI [-4.27, -3.08]). A greater proportion of children's uses were non-target compared to mothers' uses (70% vs. 5.5%). All 14 children produced at least one non-target SUBJECT + *not* + V sentence, while only half of the mothers did.

Example sentences are shown in Table 6. As previously noted, some non-target-looking sentences attributed to mothers are likely misattributions, but are not frequent enough to detract from our main point.

6.3. *Discussion*

Children are much more likely to produce sentences like "*I not buy a new one*" than their mothers. In fact, the majority of children's sentences with *not* and a verb lack a modal or auxiliary element. Since parents do not model this structure, its presence in children's productions requires explanation

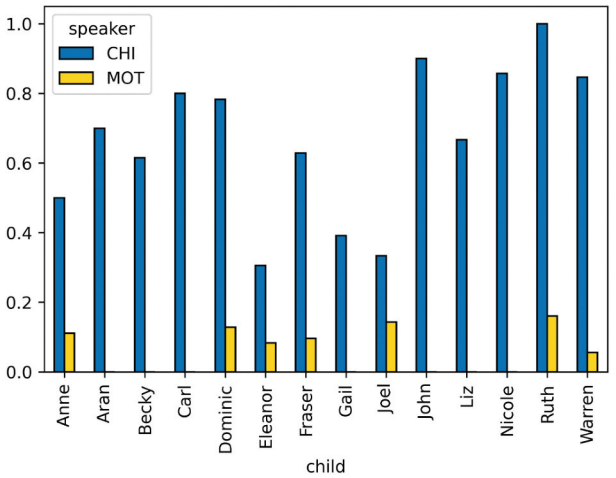

Figure 9. Proportion of *not* + V utterances with a subject that lack an auxiliary preceding *not*. The x-axis shows the proportion, the y-axis shows each child and their mother.

Table 6. The subject + *not* + V sentences produced by children and their mothers.

Child	Speaker	Utterance	Age
Carl	Child	"I not got some shopping."	2;3
Fraser	Child	"I not finish yet."	3;0
John	Child	"rain not stop."	2;9
Anne	Child	"me not care."	2;4
Nicole	Child	"and me not pay."	3;0
Fraser	Child	"I not put it in."	2;3
Ruth	Child	"Pete not go home."	2;8
Anne	Child	"I not need this."	2;2
Eleanor	Child	"I not baby."	2;4
Ruth	Child	"I not want it wet."	2;11
Anne	Mother	"me not care."	2;4
Dominic	Mother	"hat not come off."	2;1
Dominic	Mother	"Mummy not make bridge."	2;2
Ruth	Mother	"you not know."	2;5
Dominic	Mother	"Mummy not buy new one."	2;4
Fraser	Mother	"you not find your Lego pieces."	2;9
Fraser	Mother	"you not know."	2;8
Dominic	Mother	"I not listen to it now."	2;10

beyond children replaying fragments of their input. This is a general issue facing input-driven, item specific proposals of language development: why do children produce non-target forms at all?

The Early Competence Hypothesis can explain the data pattern in three ways, each of which assumes that by age 2 children possess knowledge of abstract syntactic categories. In one explanation, following Harris & Wexler (1996), *not* + V sentences reflect a grammatical stage where children treat tense as optional (the optional infinitive stage). Children's ungrammaticality in this instance reflects not an incorrect understanding of negation but non-adult-like knowledge of tense expression. In a second explanation, children omit the auxiliary/modal for performance reasons. Under this possibility, children have a target-like grammar but produce non-target utterances. A third possibility is that children's grammars allow *not* to co-occur with verbs without an auxiliary. This might be because children treat *not* as tensed. We leave open for future research which possibilities are viable. What is important is that all the possibilities involve children having abstract syntactic representations of negation, an auxiliary category, or tense.

One usage-based proposal about systematic errors in child speech contends that children make use of morphosyntactic orderings found in interrogative constructions when producing declaratives. Theakston et al. (2003), for example, argue that children's ungrammatical production of verbal constructions with a non-finite matrix clause (e.g., "*it go there*" instead of "*it goes there*") is the result of uninflected fragments of questions in their input (e.g., "*Where does it go?*"). The idea is that children make use of common frames or fragments regardless of the actual type of construction the fragment appears in. If we extend that proposal to negation, negative interrogatives produced by the mother, like "*do you not like it?*," could be the source of children's ungrammatical *not* + V declaratives. We examine this proposal in Analysis 5.

7. Analysis 5

In this analysis we assess whether the prevalence of children's ungrammatical SUBJECT + *not* + V productions like (45) are linked to the prevalence of grammatical SUBJECT + *not* + V fragments in the negative interrogatives their mothers' use, as in (46):

45. *__You not find__ your lego pieces.
46. Did __you not find__ your lego pieces?

Theakston et al. (2003) propose that systematic errors in children's declaratives are related to the frequency of fragments in the interrogatives that children hear. Errors like (45) are therefore due to the

prevalence of sentences like (46) in children's output, since both (45) and (46) share a SUBJECT + *not* + V fragment.

This proposal predicts that other interrogative constructions that occur as frequently as (46) will lead to children producing other errors in declarative sentences. We test this prediction by examining the frequency of *don't* + SUBJECT + V fragments in children's input, as in (47) and (48):

47. **Don't you want** that?
48. Why **don't we put** that away?

Analysis 2 showed that children rarely make selection errors with *don't* (96-97% of their uses preceded a verb), that is, children do not produce declaratives like (49) where *don't* is fronted:

49. *Don't I want that.

If fragment transfer is the source of children's non-target productions, interrogatives that contain a *don't* + SUBJECT + V fragment must be less frequent than those that contain a SUBJECT + *not* + V fragment. In other words, given that children produce errors like (45) but not like (49), the frequency of interrogatives like (46) must be higher than the frequency of interrogatives like (47) and (48).

7.1. *Method*

We analyzed all negative utterances in the corpus, searching for *don't* + SUBJECT + V patterns and SUBJECT + *not* + V patterns. Anything tagged as a pronoun, noun, or determiner was treated as a subject.

7.2. *Results*

Across the 14 mothers, on average 3.3% of utterances involving *not* contained a SUBJECT + *not* + V fragment (min = 0.3%, max = 16%, SD = 4%), while 4.6% of mothers' utterances involving *don't* occurred in a *don't* + SUBJECT + V fragment (min = 0.8%, max = 16%, SD = 4%). A logit-transformed paired *t*-test found no difference between the two sets of proportions ($t(13) = 1.42$, $p = 0.18$).

Another relevant comparison looks at the proportion of use relative to all negative utterances. On average 0.8% of mothers' negative utterances contained a SUBJECT + *not* + V fragment (min = 0.08%, max = 4.7%, SD = 1.3%) and 1.3% contained a *don't* + SUBJECT + V fragment (min = 0.2%, max = 4.3%, SD = 1.2). A logit-transformed paired *t*-test found this difference to be significant ($t(13) = 2.17$, p = 0.049): mothers produced a **greater** proportion of *don't* + SUBJECT + V fragments than SUBJECT + *not* + V fragments.

7.3. *Discussion*

Children's input contains at least as many *don't* + SUBJECT + V fragments as SUBJECT + *not* + V fragments, if not more. If children make use of fragments across all discourse contexts, they should make a comparable number of *don't* + SUBJECT + V ordering errors in declarative contexts. But they do not, as Analysis 2 showed. We conclude that these fragments cannot be a source of children's ungrammatical *not* + V errors. The usage-based hypothesis that children make use of frequent structures regardless of the larger syntactic contexts in which those structures appear cannot explain children's use of SUBJECT + *not* + V sequences. The input does not model those sequences more than it models *don't* + SUBJECT + V sequences.

Together with Analysis 4, the results from Analysis 5 suggest that children's early negative constructions are not input-driven formulae. Rather, they reflect abstract syntactic knowledge. By

age 2, children make generalizations involving the syntactic distribution of negative morphemes, going beyond the input.

8. General discussion

Our five analyses show that children syntactically differentiate between *not* and negative auxiliaries and that their early negative auxiliaries are adult-like. To summarize: (i) two-year-olds use a wide and overlapping range of negative and positive auxiliaries; (ii) the range of children's early negative auxiliaries is strongly correlated with the range of their positive auxiliaries; (iii) children's most common negative auxiliary, *don't*, is almost always used grammatically with respect to the syntactic category being negated and with respect to other overt markings of tense; (iv) children's agreement errors with *don't* extend to other auxiliaries like *do, have*, and *haven't*; and (v) two-year-olds' use of negative words and morphemes is not input-driven. Together, this evidence supports the hypothesis that, by age 2, English-learning children represent the syntactic categories of tense, negation, and auxiliary/modal, productively combine all three categories in target-like ways, and differentiate syntactically between different kinds of negation.

8.1. *Long-tailed distributions and infrequent data*

Our results validate the arguments of Tomasello & Stahl (2004) concerning the statistical artefacts of infrequently occurring constructions. Both negators and positive auxiliaries are infrequent in child speech, and their productions follow power-law like distributions. In conjunction with sparse taping these properties have led to an undercounting of infrequent productions and made certain markers of children's syntactic competence appear rare or non-existent. In this study we overcame these issues by aggregating data into developmental periods and by picking corpora with uniform collection methods. The large-scale aggregated analyses of Jasbi et al. (2021) provides another possible strategy for avoiding undersampling problems.

Even those strategies may not be enough. In Analysis 3 we show that the subject-agreement errors most commonly associated with *don't* generalize to *haven't* and *has* but find an asymmetry between errors in positive and negative contexts. Even when aggregating the data, the size of the corpus is not sufficient to probe this difference. Denser taping of children between two and three is required to better understand auxiliary development and its associated phenomena.

8.2. *Abstract vs. item-specific representations in early child speech*

Our results speak to a central issue in language acquisition research concerning the nature of children's early syntactic representations. Usage-based proposals about language development suggest that linguistic representations are built up from concrete particulars, with the earliest representations being memorized fragments or low-level local generalizations from the input the child is exposed to (Tomasello 2005, Ibbotson 2013, Ambridge et al. 2015, *inter alia*). Under this view, children's initial linguistic productions are tied to the properties of their input and do not reflect true syntactic combinations. Instead, abstract knowledge about syntactic operations and categories only fully emerges later in development. In contrast, nativist proposals (Fisher 2002, Lidz et al. 2003, Valian, et al. 2009, Shi 2014, *inter alia*), have argued that children's earliest representations involve abstract syntactic structures and categories that are put to use in the acquisition process to derive further generalizations about the syntactic properties of the language they are learning.

The data here provide further support for nativist proposals. Two-year-olds' ungrammatical uses of *not* and their use of third-person singular subjects with *do, don't, have*, and *haven't* are evidence that they systematically produce forms that extend beyond their input. This shows that by age 2, children have developed syntactic generalizations involving negation and certain auxiliaries, albeit non-target ones. In light of this, children's *grammatical* use of negative auxiliaries with respect to tense and category selection is

best understood as the result of a combinatorial syntax involving the negation, auxiliaries, and tense categories used in adult English.

8.3. *Functional category acquisition*

As we mentioned in our introduction, although many functional items are largely absent in children's earliest productions, there is a substantial body of evidence showing that children encode functional categories before producing multiword utterances. Furthermore, specific functional categories, like determiners, are used productively at the onset of multi-word speech. Our findings provide further evidence that functional categories are encoded in early childhood and that specific functional categories have been acquired at the onset of multiword speech. The traditional perspective that functional categories are absent from early speech does not seem tenable.

A consequence of our findings is that negation, tense, and auxiliary interaction needs to be categorized under the class of early-acquired phenomena. Languages vary in how negation, tense, and auxiliary-like categories are represented and interact. Dutch and German, for example, have negators which act like adverbs; they do not attach to other morphemes and do not act as intervenors for operations like verb movement (Zeijlstra 2004). In Italian, negative morphemes act as syntactic intervenors which can prevent clitic and verb movement, and they precede auxiliaries and modals (Zanuttini 2001). In Turkish, negation, tense, and modality are (mostly) marked as inflections on the verb (Aksu-Koç 1988). Despite this cross-linguistic variation, which testifies to the many different hypotheses that children could entertain, English-speaking children seem to have already zeroed in on the English configuration at least by age 2. This suggests an earlier representation of the relevant syntactic properties of English before they start producing multiword utterances.

The precise nature of the cues that enable children to acquire the English-like configuration remains to be discovered. Given the typological differences between English, German, Italian, and Turkish, this seems like a non-trivial learning problem. Minimally, it involves determining children's starting inventory of functional items and their initial hypotheses about how those items are expressed. One possibility is that children are aware that functional categories pair with content categories and figure out the exact specification of those functional items during development (Clahsen 1990, Demuth 1994). Another is that a number of functional categories, or at least certain classes of functional categories, are innately provided (Valian 2009, Friedmann et al. 2021). Friedmann et al. (2021), for example, propose that children's syntactic representations are proper subsets of adult syntactic representations, complete with the functional categories of the relevant subset. That English-learning children have achieved specification of some functional categories by age 2 provides a good starting point to determine what intermediary stages exist and how input is processed. We leave this and related questions to future research.

9. Summary and Conclusion

This corpus study investigated young children's acquisition of functional categories through their use of negative words and negative auxiliaries. We assessed three proposals regarding children's early use of negation: (i) that it reflects the possession of abstract, adult-like syntactic categories—the Early Competence Hypothesis, (ii) that it reflects the use of an abstract but non-adult-like negation category, and (iii) that it reflects item-specific formulae driven by input. The data support the first proposal: children's earliest negative auxiliaries are adult-like, reflecting the acquisition of negation, auxiliary, and tense categories by age 2. We conclude that at least some functional morphosyntactic categories are being acquired at the earliest stages of language development.

Acknowledgments

We would like to thank Martin Chodorow, the members of the Hunter College Language Acquisition Research Center, and audiences at BUCLD46, LSA 94, and RALFe 2018 for their helpful suggestions and feedback at various stages of this project.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

André Eliatamby http://orcid.org/0009-0003-7720-7229

References

Aksu-Koç, Ayhan. 1988. *The acquisition of aspect and modality: The case of past reference in Turkish*. Cambridge, UK: Cambridge University Press.

Ambridge, Ben, Evan Kidd, Caroline F. Rowland, & Anna L. Theakston. 2015. The ubiquity of frequency effects in first language acquisition. *Journal of Child Language* 42(2). 239–273.

Bellugi, Ursula. 1967. The acquisition of the system of negation in children's speech. Cambridge MA: Harvard Graduate School of Education dissertation.

Bernal, Savita, Jeffrey Lidz, Séverine Millotte, & Anne Christophe. 2007. Syntax constrains the acquisition of verb meaning. *Language Learning and Development* 3(4). 325–341.

Bernard, Carline & Judit Gervain. 2012. Prosodic cues to word order: What level of representation? *Frontiers in Psychology* 3. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2012.00451/full

Bloom, Lois. 1970. *Language Development: Form and Function in Emerging Grammars*. Cambridge, MA: MIT Press.

Borer, Hagit & Kenneth Wexler. 1987. The maturation of syntax. In Thomas Roeper and Edwin Williams (eds.), *Parameter setting*, 123–172. Dordrecht: Springer Netherlands.

Bowerman, Melissa. 1973. *Early syntactic development: A cross-linguistic study with special reference to Finnish*, vol. 11. Cambridge, UK: Cambridge University Press.

Brown, Roger. 1973. *A first language: The early stages*. Cambridge, MA: Harvard University Press.

Cameron-Faulkner, Thea, Elena Lieven, & Anna L. Theakston. 2007. What part of no do children not understand? A usage-based account of multiword negation. *Journal of Child Language* 34(2). 251–282.

Capdevila i Batet, Montserrat & Mireia Llinàs i Grau. 1995. The acquisition of negation in English. *Atlantis* 17(1/2). 27–44.

Clahsen, Harald. 1990. Constraints on parameter setting: A grammatical analysis of some acquisition in stages in German child language. *Language Acquisition* 1(4). 361–391.

Demuth, Katherine. 1994. On the underspecification of functional categories in early grammars. In Barbara Lust, Margaret Suñer, and John Whitman (eds.), *Syntactic theory and first language acquisition: Heads, projections, and learnability*, 1:119–134. Hillsdale, NJ: Erlbaum.

Fisher, Cynthia. 2002. Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science* 5(1). 55–64.

Friedmann, Naama, Adriana Belletti, & Luigi Rizzi. 2021. Growing trees: The acquisition of the left periphery. *Glossa: A Journal of General Linguistics* 6(1). https://doi.org/10.16995/glossa.5877

Guasti, Maria Teresa & Luigi Rizzi. 2002. Agreement and Tense as Distinct Syntactic Positions: Evidence from Acquisition. In Guglielmo Cinque (eds.), *Functional Structure in DP and IP: The Cartography of Syntactic Structures* 1, 167–194. Oxford: Oxford University Press.

Harris, Tony & Ken Wexler. 1996. The optional-infinitive stage in child English evidence from negation. In Harald Clahsen (ed.), *Generative perspectives on language acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons*, 1–42. Amsterdam: John Benjamins Publishing Company.

Hiramatsu, K. 2003. Children's judgments of negative questions. *Language Acquisition* 11(2). 99–126.

Ibbotson, Paul. 2013. The scope of usage-based theory. *Frontiers in Psychology* 4. https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2013.00255

Jasbi, Masoud, Annika McDermott-Hinman, Kathryn Davidson, & Susan Carey. 2021. Parents' and children's production of English negation. In Danielle Dionne and Lee-Ann Vidal Covas (eds.), *45th Boston University Conference on Language Development*, 360–373. Somerville MA: Cascadilla Press.

Klima, Edward & Ursula Bellugi. 1966. Syntactic regularities in the speech of children. In John Lyons and Roger Wales (eds.), *Psycholinguistic papers*, vol. 3765, 183–219. Edinburgh: Edinburgh University Press.
Kuczaj, Stan A. & Michael P. Maratsos. 1975. What children can say before they will. *Merrill-Palmer Quarterly of Behavior and Development* 21(2). 89–111.
Lidz, Jeffrey, Sandra Waxman, & Jennifer Freedman. 2003. What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition* 89(3). 295–303.
Lieven, Elena, Dorothé Salomo, & Michael Tomasello. 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. *Cognitive Linguistics* 20(3). 481–507.
MacWhinney, Brian. 2001. *The CHILDES Project: Tools for analyzing talk*, 3rd edition. Mahwah, NJ: Lawrence Erlbaum Associates.
Mayer, J. W., Erreich, A., & Valian, V. (1978). Transformations, basic operations and language acquisition. *Cognition* 6(1). 1–13.
Pine, Julian M., Daniel Freudenthal, Grzegorz Krajewski, & Fernand Gobet. 2013. Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition* 127(3). 345–360.
Pinker, S. 1984. *Language Learnability and Language Development*. Harvard University Press.
Schütze, Carson T. 2010. The status of nonagreeing don't and theories of root infinitives. *Language Acquisition* 17(4). 235–271.
Shady, Michele & Louann Gerken. 1999. Grammatical and caregiver cues in early sentence comprehension. *Journal of Child Language* 26(1). 163–175.
Shi, Rushen. 2014. Functional morphemes and early language acquisition. *Child Development Perspectives* 8(1). 6–11.
Shi, Rushen & Mélanie Lepage. 2008. The effect of functional morphemes on word segmentation in preverbal infants. *Developmental Science* 11(3). 407–413.
Shi, Rushen, Janet F. Werker, & James L. Morgan. 1999. Newborn infants' sensitivity to perceptual cues to lexical and grammatical words. *Cognition* 72(2). B11–B21.
Stromswold, Karin J. 1990. Learnability and the acquisition of auxiliaries. Cambridge, MA: MIT dissertation.
Theakston, Anna L., Elena V. M. Lieven, Julian M. Pine, & Caroline F. Rowland. 2001. The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language* 28(01). 127–152.
Theakston, Anna L., Elena V. M. Lieven, & Michael Tomasello. 2003. The role of the input in the acquisition of third person singular verbs in English. *Journal of Speech, Language, and Hearing Research* 46(4). 863–877.
Thornton, Rosalind & Kelly Rombough. 2015. The syntax-PF interface in children's negative sentences. *Language Acquisition* 22(2). 132–157.
Thornton, Rosalind & Graciela Tesan. 2013. Sentential negation in early child English. *Journal of Linguistics* 49(2). 367–411.
Tomasello, Michael. 2005. *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, Michael & Daniel Stahl. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31(1). 101–121.
Tsimpli, Ianthi-Maria. 1992. *Functional categories and maturation: The prefunctional stage of language acquisition*. London: University College London dissertation.
Valian, Virginia. 2009. Abstract linguistic representations and innateness: The development of determiners. In W. D. Lewis, S. Karimi, H. Harley, and S. Farrar (eds.), *Time and again: Theoretical perspectives on formal linguistics in honor of D. Terence Langendoen*, 189–206. Amsterdam: John Benjamins Publishing.
Valian, Virginia, Stephanie Solt, & John Stewart. 2009. Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language* 36(4). 743–778.
Woods, R. 2016. Investigating the syntax of speech acts: Embedding illocutionary force. York, UK: University of York Dissertation.
Yang, Charles. 2013. Ontogeny and phylogeny of language. *Proceedings of the National Academy of Sciences* 110(16). 6324–6327.
Yuan, Sylvia & Cynthia Fisher. 2009. "Really? She Blicked the baby?" Two-Year-Olds Learn Combinatorial Facts About Verbs by Listening. *Psychological Science* 20(5). 619–26.
Zanuttini, Raffaella. 1991. *Syntactic properties of sentential negation. A comparative study of romance languages*. Philadelphia: University of Pennsylvania dissertation.
Zanuttini, Raffaella. 2001. Sentential Negation. In Mark Baltin and Chris Collins (eds.), *The Handbook of Contemporary Syntactic Theory*, 511–35. Oxford, UK: Blackwell Publishers Ltd.
Zeijlstra, Hedde. 2004. Sentential negation and negative concord. Amsterdam: University of Amsterdam dissertation.