6-2004

A Corpus Study on the Item-based Nature of Early Grammar Acquisition

Adam Hodges
Carnegie Mellon University

Valerie Krugler
Stanford University

Deborah Law
University of Colorado Boulder

A Corpus Study on the Item-based Nature of Early Grammar Acquisition

Adam Hodges[1], Valerie Krugler[2], and Deborah Law[1]

[1]*University of Colorado,* [2]*Stanford University*

This paper explores the item-based nature of child language acquisition by examining data from the CHILDES database (MacWhinney 2000). Two studies are explicated: the first uses pooled data from several children, and the second follows a single child longitudinally. The results show that the learning of the complex construction consisting of a main clause followed by an infinitival compliment, e.g. *I want to play*, center around a single verb, *want*, even though other candidate verbs exist in the children's vocabulary. We provide empirical evidence to show that children initially learn grammar via item-based units and gradually break down complex *constructions as units* into smaller pieces in a process that leads towards the organization of language into the abstract categories consistent with a fully competent adult grammar.

1 Introduction

How do children acquire the grammar of their language? Chomsky (1957) theorizes that a child is born with an innately equipped language module (cf. Chomsky 1981, 1986a, 1986b). Following Chomsky, Pinker posits *a priori* grammatical categories, such as verbs and nouns, in his proposals for semantic bootstrapping and linking rules (1984, 1989; cited in Hoff 2001: 250, cf. Pinker 1994). Other nativists such as Bower and Wexler (1992, cited in Tomasello 2000b: 160) posit a more gradual turning on of genes and syntactic structures. Innatist proposals such as these base their starting point on the idea that the nature of children's grammar is simply a miniature version of a fully competent adult. Yet, from the outside perspective of an adult, child speech could be *perceived* as grammatical while operating quite differently during the developmental process of learning a first language.

Tomasello (1992) points out the highly item-based nature of child's speech. Rather than requiring the presence of innate categories or abstract structures, "children can also produce 'grammatical' language by simply reproducing the specific linguistic items and expressions (e.g. specific words and phrases) of adult speech, which are, by definition, grammatical" without productively using those items (Tomasello 2000b: 156). For example, Diessel and Tomasello (2001) show that in early patterns of child usage, complement clauses are generally introduced by formulaic frames centered around only a few verbs, which they term constructional islands, or verb islands. Their data demonstrate an early reliance on item-based patterns, with more novel uses occurring as experience with the language increases.

Thus, with regard to early syntactic development, it is important to ask whether children productively create utterances based on an abstract grammar or mainly reproduce item-based units as an initial step in the language learning process. If they possess an abstract grammar, they should be able to productively manipulate constituent categories, such as verbs. If they are merely reproducing constructions *as units*, then such productivity should be lacking in early stages and gradually increase over time.

This paper explores these ideas with data from the CHILDES database (MacWhinney 2000) in two separate studies to illustrate trends associated with a usage-based understanding of early syntactic development. We focus on variations of the complex construction consisting of a main clause followed by an infinitival complement, e.g. *I want to play*, *he needs to eat*, etc. The first study (section 2) uses pooled data from several children, while the second study (section 3) focuses on a longitudinal study of a single child. We provide evidence consistent with a usage-based understanding of child language development. That is, (1) the children in this study demonstrate early usage of this construction centered around a single verb, *want,* even though other verbs are present in their vocabulary. (2) The input language provides a preponderance of exemplars centered around the verb *want* but nevertheless shows a more level usage of different verbs in the main clause slot, a usage pattern that the children move toward in the course of development. (3) Instances of this construction where the *same* subject is present in both the main clause and infinitival complement (*I want to play* versus *I want **him** to play*) are learned first; the ability to switch subjects from the main clause to infinitival clause comes much later in development, and when it does, it initially follows the same pattern demonstrated with single subject uses of the construction, relying on the verb *want* in the main clause before productively adding other verbs to that slot. A general discussion on the findings from both studies is provided in section 4.

2   Initial Study with Pooled Data from Seven Corpora

The first study uses pooled data from several children and focuses specifically on first person usages of the main-clause-plus-infinitival-complement construction, e.g. *I want to play*, *I need to eat*, etc. We make two primary claims based on the results. (1) Before 3 years of age, children rely on a single verb, *want*, even though other candidate verbs exist in their vocabulary, i.e. *have, got, like, need*. (2) After 3 years of age, children move away from their reliance on the single verb *want* in the main clause, and add a second verb *have* to the possibilities. This shift demonstrates increased productivity for this construction.

2.1   Methodology

The pooled corpus consists of data selected from the CHILDES database (MacWhinney 2000), and incorporates data from seven corpora, Bates (1988; Carlson-Luden 1979), Belfast (Henry 1995; Wilson and Henry 1998), Bloom (1970; Bloom and Lightbown 1974; Bloom et al 1975), Bloom (1973), Bohannon (Bohannon and Marquis 1977; Stine and Bohannon 1983), Clark (1978a, 1978b, 1979, 1982a, 1982b), and Hall (Hall et al 1984; Hall et al 1981; Hall and Tirre 1979). All children in this combined corpus are between the ages of 1 year 1 month and 4 years 5 months. The corpus includes 304 data sessions from 67 children, a total of 86,010 child utterances. The original transcripts include varying degrees of additional notation with respect to non-speech events, pauses, unclear utterances, etc. We preprocessed the data with a normalization script to remove most of these inconsistencies.

To determine a frequently used frame to focus the study, sets of n-grams were generated for each age group, each age group represents approximately a six-month age

range.   N-grams are collocations of terms within an utterance.   Bigrams (2 term collocates), trigrams, 4-grams, and 5-grams were generated to find frequent collocates in each group. One of the most frequent collocations across age groups above 2;0 is *I want to*, which upon further investigation generalized to *I (don't) V VP-INF* (first person singular pronoun followed by an optional *don't*, followed by a verb from a limited list, with an infinitive verb phrase complement.) This generalized frame was chosen as the focus for this study because of its relative frequency and complexity. Regular expression search patterns were used to identify instances of this construction, 225 instances total. The regular expression patterns took into account another inconsistency in the data sample: the tendency for some transcriptions to contain more casual expressions of colloquial speech (e.g. 'hafta', 'want ta'), while transcribed to a more formal style of speech (e.g. 'have to', 'want to'.) The results of the regular expression searches were hand-checked to verify that all utterances considered aligned with the desired construction. Eighteen utterances were discarded during the hand-checking.

2.2   Results

The construction: I (don't) V VP-INF, and candidate verbs. In the complex construction, *I (don't) V VP-INF*, the pooled data reveal a set of five verbs that are used in the main clause by the children: *want, have, got, like, need*. Importantly, each of these verbs is found in the vocabulary of each age group studied, even though the frequencies vary markedly in their appearance in the construction studied. (N.B. No attested examples of this construction occur in the children below 2 years old. The complete set of data is given in appendix 1.)

Tables 1 and 2 show the frequency of occurrence of the five verbs found in this construction—*want, have, got, like, need*. As Table 1 shows, *want* is used 84% of the time by the children in the 2;0-2;5 age group, 86% of the time by the 2;6-2;11 age group, 46% of the time by the 3;0-3;5 age group, 47% of the time by the 3;6-3;11 age group, and 54% of the time by the 4;0-4;5 age group.

The use of the verb *have* is broken down as follows: 11% for the 2;0-2;5 age group, 4% for the 2;6-2;11 age group, 48% for the 3;0-3;5 age group, 47% for the 3;6-3;11 age group, and 42% for the 4;0-4;5 age group.

The remaining verbs—*got, like, need*—make up a small percentage of use across all age groups, as shown in Table 1.

Table 2 splits the numbers into two age groups: under 3 years old vs. 3 years old and above. We highlight this split since age 3 is an obvious dividing line in the children's usage (as evidenced in Table 1.) Children below 3 years old use *want* 85% of the time in the main clause of this construction, while children 3 years old and above use *want* 48% of the time. In children 3 years old and above, the usage of *have* increases from 8% to 46%. The remaining verbs—got, like, need—make up a small remainder of the usage in both age groups.

Figures 1 through 4 illustrate the major trends from Tables 1 and 2. Figure 1 shows the usage of *want* across the age groups, as shown by the numbers in Table 1. Figure 2 illustrates the usage of *want* below vs. above 3 years, as detailed in Table 2. Figure 3 illustrates the trend that children over 3 years use verbs other than *want* more frequently in the main clause of this construction; and Figure 4 shows the breakdown of the verbs used by children over 3 years.

Table 1. Distribution of the verbs used in the main clause by children across the age groups

VERB	AGE											
	1;6-1;11		2;0-2;5		2;6-2;11		3;0-3;5		3;6-3;11		4;0+	
	#	%	#	%	#	%	#	%	#	%	#	%
want	0	0%	51	84%	60	86%	23	46%	8	47%	14	54%
have	0	0%	7	11%	3	4%	24	48%	8	47%	11	42%
got	0	0%	2	3%	6	9%	0	0%	0	0%	0	0%
like	0	0%	1	2%	0	0%	2	4%	0	0%	1	0%
need	0	0%	0	0%	1	1%	1	2%	1	6%	1	4%
TOTAL	**0**	**0%**	**61**	**100%**	**70**	**100%**	**49**	**100%**	**17**	**100%**	**27**	**100%**

Table 2. Distribution of the verbs used in the main clause by children below and above 3 years

VERB	AGE			
	Under 3 years old		3 years old and above	
	#	%	#	%
want	111	85%	45	48%
have	10	8%	43	46%
got	8	6%	0	0%
like	1	1%	3	3%
need	1	1%	3	3%
TOTAL	**131**	**100%**	**94**	**100%**

Figure 1. Use of 'want' in the main clause across age groups

A Corpus Study on the Item-based Nature of Early Grammar Acquisition   5
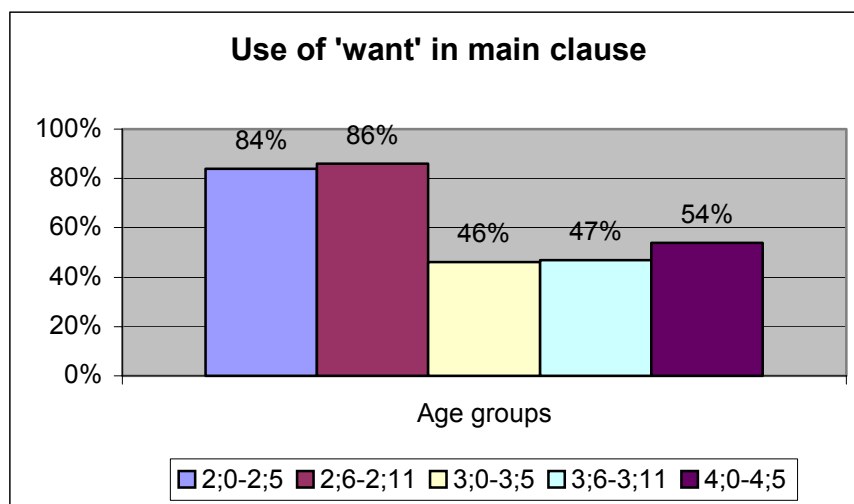
Figure 2. Use of 'want' in the main clause by children below 3 years vs. children above 3 years

Use of 'want' in main clause

85%

48%

Age

☐ Under 3 years   ■ Over 3 years
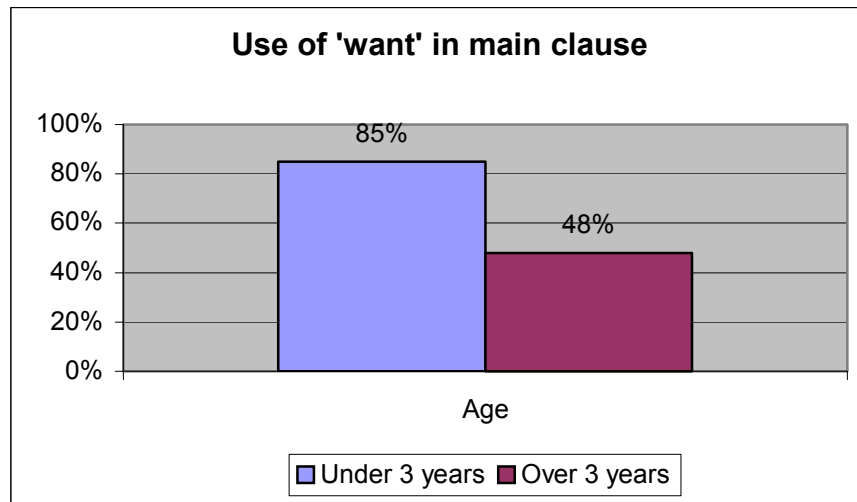
Figure 3. Use of verbs other than 'want' by children below 3 years vs. children above 3 years

Use of verbs other than 'want'

52%

15%

Age

☐ Under 3 years   ■ Over 3 years
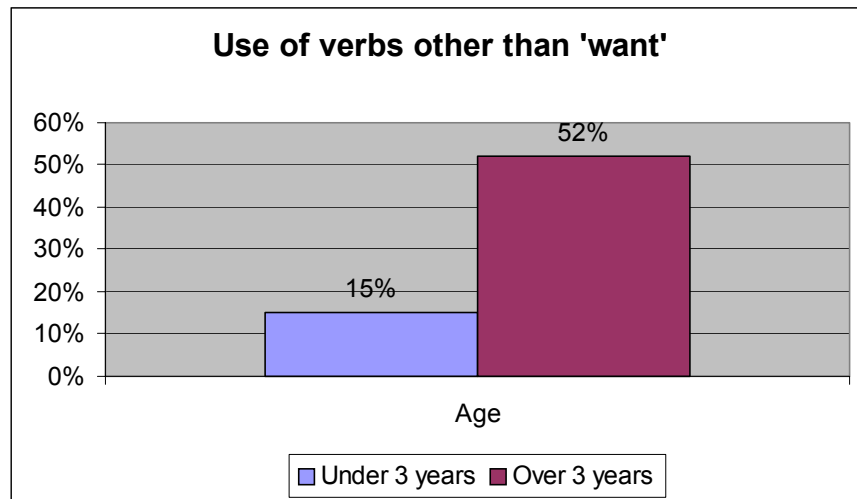
Figure 4. Distribution of the verbs used in the main clause by children above 3 years old

Verbs used by children above 3 years

48%   46%

0%   3%   3%

Verbs

☐ want ■ have ☐ got ☐ like ■ need
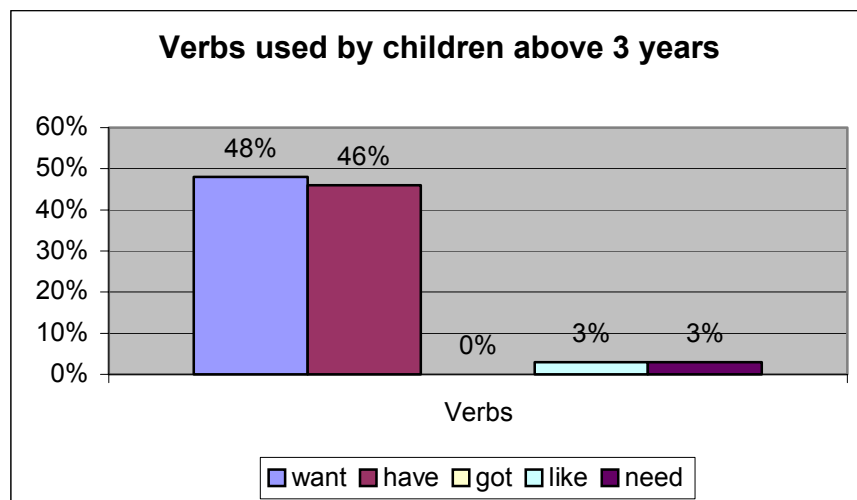
2.3   Conclusions

These data lead to the following conclusions.

1) Before 3 years of age, children rely on a single verb, *want*, even though other candidate verbs exist in their vocabulary, i.e. *have, got, like, need*.

2) After 3 years of age, children move away from their reliance on the single verb *want* in the main clause, and add a second verb *have* to the possibilities. This shift demonstrates increased productivity for this construction.

These conclusions suggest that children initially learn the main clause in this construction as a single unit. As their language develops, they gradually break down this entrenched unit into smaller pieces. Past age three, they are able to apply other verbs to the construction, notably the verb *have,* in addition to *want*. This progression illustrates a gradual move toward the full competence characterized by adult grammar, a competence that demonstrates the ability to substitute constituents within larger clauses.

3   Follow-up Longitudinal Study of a Single Child

While our initial study uses pooled data and looks solely at utterances of child output, the purpose of our second study is to examine this particular construction again but in a slightly different light. (1) The construction we search for still involves a main clause followed by an infinitival compliment, but our search pattern captures usages of this construction across the paradigm rather than simply focusing on first person uses (i.e. *I want to play* and *you want to play* are both captured). (2) In addition to looking at uses of this construction where the subject is the same in both the main clause and infinitival clause, we carry out an additional search to find all uses of this construction where a noun is present between the main clause verb and infinitival marker, which demonstrates a switch in person between the subject of the main clause and subject of the infinitival clause. (3) We follow the development of a single child, rather than pooling data from several children, and also look at the adult input received by that child.

By looking at variations of this construction and focusing on the development of a single child and the input language he receives, we provide another illustration of the item-based nature of the early stages of grammar acquisition. A single child was chosen for this study to avoid the problems of conflating the development of several children. An inherent problem in this type of approach, and indeed in any work done on child language development, involves the lack of sufficient longitudinal data. Tomasello and colleagues are attempting to overcome this problem by collecting longitudinal data that represents a consistent and detailed look at the development of individual children. However, for the purposes of this study, we chose one of the best longitudinal corpora available on the CHILDES database (MacWhinney 2000), the corpus of Adam collected by Brown (1973). Until sufficient longitudinal data from several children are available, we feel the pooled approach in section 2 and the individual approach in this section provide a starting point and evidence for a general trend worth further examination in the future.

A Corpus Study on the Item-based Nature of Early Grammar Acquisition   7

We make the following claims based on the results of this section. (1) The child under study, Adam, demonstrates the general pattern illustrated in section 2 where initial uses of the main-clause-plus-infinitival-complement construction are centered around the verb *want* even though other verbs are in his vocabulary. (2) The input language provides a preponderance of exemplars centered around the verb *want* but nevertheless shows a more level usage of different verbs in the main clause slot, a usage pattern that Adam moves towards in his development. (3) Adam's usage of this construction relies heavily on same subject uses; the ability to switch subjects from the main clause to infinitival clause comes much later in development, and when it does, it initially follows the same pattern demonstrated with single subject uses of the construction, relying on the verb *want* in the main clause before productively adding other verbs to that slot.

3.1   Methodology

Our data comes from a longitudinal study of the child Adam conducted by Brown (1973) and available as a part-of-speech tagged corpus on the CHILDES database (MacWhinney 2000). "Adam was the child of a minister and an elementary school teacher. His family was middle class and well educated. Though he was Black, he was not a speaker of American Black English, but of Standard American. There are 55 files in the Adam corpus and his age ranges from 2;3 to 4;10" (MacWhinney 2000, based on Brown 1973). The data indicate the age of Adam at time of recording. The total number of utterances from Adam in the corpus is 46,722, and the total number of utterances from caretakers is 114,081. While the documentation in the database states that data for Adam ends at age 4;10, file number 52 lists the age of collection as 5;2. Due to this discrepancy, we excluded file number 52 from the results.

We used a Perl script (appendix 2) to search the files (which are marked by age) for the presence of the following construction: *(NOUN) (don't) V VP-INF*. Regular expression search patterns were used to identity instances of this construction using the five potential candidate verbs identified in section 2—*want, have, like, need*, and *got*. The first four of these verbs were found in the speech of Adam and his caretakers.

The regular expression patterns took into account an inconsistency in the data— the tendency for some transcriptions to contain more casual expressions of colloquial speech (e.g. 'hafta', 'want ta'), while transcribed to a more formal style of speech (e.g. 'have to', 'want to'). The following is the regular expression used to search for the verb 'want' in its different forms: /^%mor:.*pro\|.*v\|wan(t|ts|ted|ting)(|~-)inf\|t(o|a) v\|/. Each tagged line starts with **%mor:,** and each tag is in front of the word separated with a pipe. Infinitives are separated from the main clause verb by ~-. Counts of each verb occurring in the main clause were made for both the child and his caretakers. The Perl script in appendix 2 includes the additional regular expressions used to search for the other candidate verbs in this construction. Samples of the output data are provided in appendix 3.

An additional search was made for a variation of this construction: *(NOUN) (don't) V NOUN VP-INF*. That is, the same phrase but with a noun inserted between the main verb and infinitive marker to indicate a switch of subject from the main clause to the infinitival clause, e.g. *I want **him** to go*. The following is the regular expression used to search for the verb 'want' in this search: /^%mor:.*pro\|.*v\|wan(t|ts|ted|ting).*inf\|t(o|a) v\|/.

8   Colorado Research in Linguistics, Volume 17 (2004)

3.2   Results

The construction NOUN (don't) V VP-INF, and candidate verbs. In the complex construction *NOUN (don't) V VP-INF*, the Adam corpus reveals four verbs that are used in the main clause by both Adam and his caretakers: *want, have, like,* and *need*. Importantly, each of these verbs is found in Adam's vocabulary throughout his development, even though the frequency of their appearance in the construction varies throughout that development. Thus, it is not their lack in the lexicon that keeps them from appearing productively in this construction.

Table 3 shows the distribution of these verbs throughout Adam's development. He starts out relying on the verb *want* and gradually decreases this reliance as the construction begins to be used more productively with other verbs. From 2;3 to 2;11, Adam uses *want* 95% of the time in this construction. From 3;0 to 3;11, the reliance on *want* decreases to 81% of the time as the second most used verb, *like*, increases to 13%. Then from 4;0 to 4;10, *want* is used only 69% of the time, while *have* is used 18% of the time, *like* 10% of the time, and *need* 2% of the time in this construction.

Table 4 shows the corresponding input for this construction over the same time periods. The input language uses the verb *want* to a fairly large extent, providing the key exemplar that Adam is picking up on; however, the other candidate verbs are also present to a large degree, providing the end result exemplar that Adam's development moves towards.

Table 3. Distribution of the verbs used in the main clause of the target construction PRO (don't) V VP-INF by Adam across ages

	2;3-2;11		3;0-3;11		4;0-4;10	
	#	%	#	%	#	%
want	56	95%	218	81%	175	69%
have	1	2%	9	3%	46	18%
like	2	3%	35	13%	26	10%
need	0	0%	7	3%	5	2%
TOTAL 621	59	100%	269	100%	252	100%

Table 4. Distribution of the verbs used in the main clause of the target construction PRO (don't) V VP-INF from the input language across ages

	2;3-2;11		3;0-3;11		4;0-4;10	
	#	%	#	%	#	%
want	77	63%	69	36%	43	39%
have	21	17%	44	23%	34	31%
like	21	17%	50	26%	30	27%
need	4	3%	29	15%	4	4%
TOTAL 434	123	100%	192	100%	111	100%

8

A Corpus Study on the Item-based Nature of Early Grammar Acquisition   9

NOUN (don't) V NOUN VP-INF.   Tables 5 and 6 show the usage of the more complicated form of the main-clause-plus-infinitival-complement construction.   In this form of the construction, an additional noun is placed between the main clause verb and infinitival marker, which requires the speaker to productively manipulate the constituents that comprise the construction to indicate different subjects for the main and infinitival clauses.

The data in Table 5 show that Adam does not have this ability before the age of 3 years, as no instances of this form of the construction occurred between 2;3 and 2;11. This is in contrast to the 59 instances of the simpler version of this construction that Adam used during the same time period shown in Table 1. As Adam starts to use this more complicated form of the construction after 3 years, he shows a strong reliance on the verb *want*, in 86% and 87% of the occurrences from 3;0 to 3;11 and 4;0 to 4;10, respectively.

Table 6 shows the presence of this form of the construction in the input language across the same time periods. It is present at all stages of development, and provides a dominant number of exemplars using the verb *want*—91% during the age range of 2;3 to 2;11, 81% from 3;0 to 3;11, and 71% from 4;0 to 4;10. The presence of these additional uses of *want* adds to the number of exemplars of the general construction provided to the child during development.

Table 5. Distribution of the verbs used in the main clause of the target construction PRO (don't) V PRO VP-INF by Adam across ages

	2;3-2;11		**3;0-3;11**		**4;0-4;10**	
	#	%	#	%	#	%
want	0	0%	24	86%	13	87%
have	0	0%	1	4%	1	7%
like	0	0%	2	7%	1	7%
need	0	0%	1	4%	0	0%
TOTAL	**0**	**0%**	**28**	**100%**	**15**	**100%**
47						

Table 6. Distribution of the verbs used in the main clause of the target construction PRO (don't) V PRO VP-INF from the input language across ages

	2;3-2;11		3;0-3;11		4;0-4;10	
	#	%	#	%	#	%
want	29	91%	21	81%	10	71%
have	0	0%	3	12%	2	14%
like	2	6%	2	8%	1	7%
need	1	3%	0	0%	1	7%
TOTAL	**32**	**0%**	**26**	**100%**	**14**	**100%**
73						

3.3   Conclusions

The trends illustrated in these results lead to the following conclusions.

1) Adam's learning of this construction treats it as an item-based unit anchored around the verb *want*. As the child's language skills develop, he is gradually able to treat the constituents of this complex construction as such and more productively use other verbs, and even switch subjects between the main clause and infinitival clause.

2) The high frequency occurrences of the verb *want* in the input language provide the most notable exemplars picked up on by the child. That is, the child begins learning the construction as a unit centered around the most frequently heard verb used in this construction by adults—*want*.

These findings suggest that children initially treat this construction as a unit. As their language develops, they gradually break down this unit into smaller pieces. This progression illustrates a gradual move toward the full competence characterized by adult grammar, a competence that demonstrates the ability to productively manipulate various components within complex constructions.

4   General Discussion

A usage-based account of acquisition posits that language is built from the bottom up, from item-based units rather than innately present abstract categories. Tomasello (2000b) states that "children imitatively learn concrete linguistic expressions from the language they hear around them, and then—using their general cognitive and social-cognitive skills—categorize, schematize and creatively combine these individually learned expressions and structures to reach adult linguistic competence" (Tomasello 2000b: 156). The findings in this paper provide further evidence for this type of acquisition model.

4.1   The learning of language via item-based units

The importance of larger-than-word units in language is an important idea stressed in cognitive linguistics. Langacker (1987) explains, "With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit" in grammatical organization (Langacker 1987: 59). That is, language is composed of a 'structured inventory of constructions.' With this idea in mind, we posit that children are apt to learn language by pulling in such units into their own usage. Therefore, their first attempts at producing their own utterances rely heavily on these constructions as units (see Langacker 1987: 57 for a detailed discussion of units; also, see Langacker 1987: 50 for a detailed discussion of entrenchment in grammar, and Tomasello 2000a: 72 for its relation to language acquisition; cf. Langacker 1991, 1998). The data from this paper show that children initially rely on exemplars centered around the verb *want* when learning the complex construction that involves a main-clause-plus-infinitival-complement.

This finding is backed up by previous research. For example, Peters (1986) provides the examples of *Iwant* and *Idontknow* (cited in Hoff, 2001: 210) as multiword phrases memorized as units in the lexicon of children. Hoff (2001: 210) refers to these as "rote-learned wholes" and "jargon + word combinations." Lieven et al (1997) show that initial multi-word utterances of children are based on specific lexical patterns (cited in Tomasello 2000: 157). And Bolinger (1976) suggests that language learners, whether children or adults, often learn via collocations (cited in Johnson, 2000: 193). Bolinger (1977) presents examples of "collocation mixing" to show the process of internal analysis that then takes place in children as they gradually break up larger phrases into words (cited in Johnson 2000: 192).

Similar to the case of the construction presented here, *wh*-word development has been shown to have formulaic beginnings. Johnson (2000: 187) cites Bellugi (1965), Brown (1968, 1973), Klima and Bellugi (1966), and Johnson (2000), as evidence for such a process. Once a *wh*-frame is rote-learned, then the child abstracts the *wh*-word to new situations in what Johnson's (2000) data show is a gradual process. The end-result is competence in using the *wh*-words themselves in appropriate, novel situations. And in the case of negation, Hoff (2001: 219) cites evidence that children often start out using whole forms such as *can't* and *don't* as negative markers before analyzing and breaking out the conflated auxiliaries (*can, do*) and negative marker (*not*).

The data presented in this paper clearly show a similar pattern. The children produce the complex construction studied by relying on the verb *want*, a verb that is most frequently used in the adult input, as attested in the study of Adam. Rather than needing to master the complex grammatical abstractions involved in such a construction, the child can simply reproduce the most typical exemplar. Once this complex construction is learned as an item-based unit, then the child begins to gradually abstract out the finer, individual components, and use the construction more productively—first adding other verbs to the main clause verb slot and then switching subjects between the main clause and infinitival clause.

4.2   From item-based units to abstract grammar

The data in this study show a shift from a reliance on the use of the single verb *want* in the main clause toward an ability to more productively substitute other verbs into the appropriate slot. This progression is demonstrated in the results of both section 2 and section 3. Section 2 shows that for the pooled data, there is a marked shift in productive ability with this construction after 3 years of age. The study of Adam in section 3 demonstrates the same progression toward productivity over his development. Clearly, the abstract category of *verb* is not being productively applied in the early usages demonstrated in the data, suggesting that such an innate category is lacking.

In explaining his Verb Island hypothesis, Tomasello (2000b: 157) claims that children have "lexically specific syntactic categories" rather than abstract syntactic categories. This is similarly demonstrated in our data, which point to the lexically specific nature of the construction studied in early grammatical development. The near exclusive reliance on the collocation of *want* with the surrounding constituents of the construction, suggests that children do not yet have a more abstract schema that treats this construction as being composed of categories (whether present or underlying) such as NOUN + VERB + (NOUN) + INF-marker + VERB. Out of the five verbs used by

children in the initial study and the four verbs used by Adam in this type of construction, all are present in the vocabulary during each age range examined. However, the children fail to substitute these plausible candidates into the main clause in this construction at earlier ages. Productivity gradually increases as they get older and develop more experience with language in general and this construction in particular.

Children in the initial study begin to move away from the reliance on the verb *want* after age 3. However, they do not immediately exhibit free use of any of the candidate verbs, but rather move from reliance on one verb to reliance on two verbs—*want* and *have*. Similarly, from 3;0 to 3;11, Adam begins to replace his reliance on the single verb *want* by adding another verb to his repertoire—*like*. From 4;0 to 4;10, the verb *have* begins to play more of a role in Adam's productive ability. This demonstrates that productivity increases gradually, and the movement towards increased productivity involves a step-by-step process. This process can be thought of as one that involves learning separate *constructions as units*—first a construction with the verb *want*, then a construction with the verb *like*, then a construction with the verb *have*, and so on. Rather than a complete, all-or-nothing shift towards knowledge of abstract categories, the child gradually builds up a repertoire of constructions until the point is reached when generalizations can be made about this set as a whole. Then the child learns to abstract out the individual constituents and manipulate those constituents more productively. The end result is the fully productive, abstract grammar of an adult.

The data presented in this study are similar to other findings. Tomasello (2000b: 157) cites cross-linguistic studies which imply that the most frequent verb forms in the input language are first learned as unanalyzed items by children; then the inflections are gradually abstracted out, applied to other verbs and eventually abstracted to the rest of the forms in the verb paradigm (Pizutto and Caselli 1992, 1994; Rubino and Pine 1998; Berman and Armon-Lotem 1995; Berman 1982; MacWhinney 1978; Behrens 1998; Allen 1996; Gathecole 1999; Stoll 1998). Experimental studies cited by Tomasello (2000b: 158-159) demonstrate this as well. Studies by Akhtar (1999), Akhtar and Tomasello (1997), Brooks and Tomasello (1999), Berman (1993), Dodson and Tomasello (1998), Maratsos et al (1987), Olguin and Tomasello (1993), Pinker et al (1987), Tomasello and Brooks (1998) imply a gradual progression in children's ability to extend novel verbs to constructions different from those in which the verbs were learned, such that children younger than 3 lack the productivity that children older than 3 are able to demonstrate. Furthermore, samples of children's overgeneralizations as compiled by Bowerman (1982, 1988) and Pinker (1989) predominantly occur after age 3, presumably once more abstract categories have been extracted from item-based units (cited in Tomasello 2000b: 158). As Tomasello states, "…it is clear that young children are productive with their early language in only limited ways. They begin by learning to use specific pieces of language and only gradually create more abstract linguistic categories and schemas" (Tomasello 2000b: 159).

Finally, the exemplars provided in the input language are an important component in the process of learning grammar. The most frequently heard exemplars of the complex construction, as attested in the Adam study, are centered around the verb *want*. The child picks up on this frequently heard exemplar and bases early uses of the construction on this verb. The frequency of the verb *want* in the input data is even greater when the more complex form of the construction that involves a switch of subject between the main

clause and infinitival clause is examined (section 3). While a young child is unable to pick out every constituent of that construction, he hears *want* enough so that his early production of the simpler version of the construction is based upon that 'verb island,' or 'constructional island' (Diessel and Tomasello 2001, cf. elsewhere).

5   Summary

This study provides empirical evidence, pertaining to the complex construction comprised of a main clause plus infinitival complement, for the idea that children initially learn grammar via item-based units, and then gradually break down complex *constructions as units* into smaller pieces in a process that leads towards the organization of language into abstract categories consistent with a fully competent adult grammar. This is a pattern "consistent with a more constructivist or usage-based model in which young children begin language acquisition by imitatively learning linguistic items directly from adult language, only later discerning the kinds of patterns that enable them to construct more abstract linguistic categories and schemas" (Tomasello 2000b: 158). The evidence provided here creates a difficult obstacle for a high-innateness model of language acquisition, which relies on *a priori* abstract categories in the mind of the young language learner. If these categories were present from the beginning, then they would be expected to result in much greater productivity than what these data show.

14   Colorado Research in Linguistics, Volume 17 (2004)

References

AKHTAR, N. and M. TOMASELLO. 1997. "Young children's productivity with word order and verb morphology." *Developmental psychology,* 33: 952-965.

AKHTAR, N. 1999. "Acquiring basic word order: evidence for data-driven learning of syntactic structure." *Journal of child language,* 26: 261-278.

ALLEN, S. 1996. *Aspects of argument structure acquisition in Inukitut,* John Benjamins.

BATES, E. AND I. BRETHERTON AND L. SNYDER. 1988. *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge, MA: Cambridge University Press.

BEHRENS, H. 1998. "How difficult are complex verbs? Evidence from German, Dutch, and English." *Linguistics,* 36: 679-712.

BELLUGI, U. 1965. "The development of interrogative structures in children's speech." In K. Riegel, ed., *The development of language functions,* Report No. 8, 103-137. Ann Arbor, MI: Language Development Program.

BERMAN, R.A. and S. ARMON-LOTEM. 1995. "How grammatical are early verbs?" *Annales Littéraires de l'Université de Franche-Comté,* 631 : 17-56.

BERMAN, R. 1982. "Verb-pattern alternation: the interface of morphology, syntax, and semantics in Hebrew child language." *Journal of child language,* 9: 169-191.

—. 1993. "Marking verb transitivity in Hebrew-speaking children." *Journal of child language,* 20: 641-670.

BLOOM, L. AND L. HOOD AND P. LIGHTBOWN. 1974. "Imitation in language development: If, when and why." *Cognitive Psychology, 6*, 380–420.

BLOOM, L. AND P. LIGHTBOWN AND L. HOOD. 1975. "Structure and variation in child language." *Monofigures of the Society for Research in Child Development, 40*, (Serial No. 160.)

BLOOM, L. 1970. *Language development: Form and function in emerging grammars.* Cambridge, MA: MIT Press.

—. 1973. *One word at a time: The use of single-word utterances before syntax.* The Hague: Mouton.

BOHANNON, J.N. AND A.L. MARQUIS. 1977. "Children's control of adult speech." *Child Development, 48*, 1002–1008.

BOLINGER, DWIGHT. 1976. "Meaning and memory." *Forum linguisticum, 1,* 1-14.

—. 1977. "Idioms have relations". *Forum linguisticum, 2,* 157-169.

BOWER, H. and WEXLER, K. 1992. "Bi-unique relations and the maturation of grammatical principles." *Natural language linguistic theory,* 10: 147-187.

BOWERMAN, M. 1982. "Reorganizational processes in lexical and syntactic development." In L. Gleitman and E. Wanner, eds., *Language acquisition: The state of the art,* 319-346. Cambridge, England: Cambridge University Press.

—. 1988. "The 'no negative evidence' problem: How do children avoid constructing an overgeneral grammar?" In J.A. Hawkins, ed., *Explaining language universals,* 73-101. Basil Blackwell.

BROOKS, P. and M. TOMASELLO. 1999. "Young children learn to produce passives with nonce verbs." *Developmental psychology,* 35: 29-44.

BROWN, ROGER. 1968. "The development of Wh questions in child speech." *Journal of verbal learning and verbal behavior, 7,* 279-290.

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

CARLSON-LUDEN, V. 1979. *Causal understanding in the 10-month-old.* Unpublished doctoral dissertation. University of Colorado at Boulder.

CHOMSKY, NOAM. 1957. *Syntactic structures.* The Hague: Mouton.

—. 1981. *Lectures on government and binding.* Dordrecht: Foris.

—. 1986a. *Knowledge of language: Its nature, origin and use.* New York: Praeger.

—. 1986b. *Barriers.* Cambridge: MIT Press.

CLARK, E.V. 1978a. "Awareness of language: Some evidence from what children say and do." In R. J. A. Sinclair & W. Levelt, eds., *The child's conception of language.* Berlin: Springer Verlag.

—. 1978b. "Discovering what words can do." In W. J. D. Farkas & K. Todrys, eds., *Papers from the parasession on the lexicon.* Chicago: Chicago Linguistic Society.

—. 1979. "Building a vocabulary: Words for objects, actions and relations." In P. Fletcher & M. Garman, eds., *Language acquisition: Studies in first language development.* New York: Cambridge University Press.

—. 1982a. "Language change during language acquisition." In M. E. Lamb & A. L. Brown, eds., *Advances in child development: Vol. 2.* Hillsdale, NJ: Lawrence Erlbaum Associates.

—. 1982b. "The young word maker: A case study of innovation in the child's lexicon." In E. Wanner & L. R. Gleitman, eds., *Language acquisition: The state of the art.* Cambridge, MA: Cambridge University Press.

DIESSEL, HOLGER and MICHAEL TOMASELLO. 2001. "The acquisition of finite complement clauses in English: A corpus-based analysis." *Cognitive linguistics,* 12-2: 97-141.

DODSON, K. and M. TOMASELLO. 1998. "Acquiring the transitive construction in English: The role of animacy and pronouns." *Journal of child language,* xx: xxx-.

GATHECOLE, V. et al. 1999. "The early acquisition of Spanish verbal morphology : across-the-board or piecemeal knowledge ?" *International journal of bilingualism,* 3: 138-182.

HALL, W.S. AND W.C. TIRRE. 1979. *The communicative environment of young children: Social class, ethnic and situational differences.* Champaign, IL: University of Illinois.

HALL, W.S. AND W.E. NAGY AND G. NOTTENBURG. 1981. *Situational variation in the use of in-ternal state words.* Champaign, IL: University of Illinois.

HALL, W.S. AND W.E. NAGY AND R. LINN. 1984. *Spoken words: Effects of situation and social group on oral word usage and frequency.* Hillsdale, NJ: Erlbaum.

HENRY, A. 1995. *Belfast English and Standard English: Dialect variation and parameter setting.* New York: Oxford University Press.

HOFF, ERIKA. 2001. *Language development,* Belmont, CA: Wadsworth/Thomson Learning.

JOHNSON, CAROLYN E. 2000. "What You See Is What You Get: The importance of transcription for interpreting children's morphosyntactic development." In L. Menn and N.B. Ratner, eds., *Methods for studying language production.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

KLIMA, E. and U. BELLUGI. 1966. "Syntactic regularities in the speech of children." In J. Lyons and R. Wales, eds., *Psycholinguistics papers*, 183-208. Edinburgh: University Press. (Reprinted in A. Bar-Adon and W. Leopold, eds., *Child Language: A book of readings,* 412-424, Englewood Cliffs, NJ: Prentice Hall.)

LANGACKER, RONALD. 1987. *Foundations of cognitive grammar: Volume I.* Stanford: Stanford University Press.

—. 1991. *Foundations in cognitive grammar: Volume 2*. Stanford: Stanford University Press.

—. 1998. "Conceptualization, symbolization, and grammar." In Michael Tomasello, ed., *The new psychology of language: Cognitive and functional approaches to language structure*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

LIEVEN, E. 1997. "Variation in a cross-linguistic context." In D. Slobin, ed., *The cross-linguistic study of language acquisition,* Vol. 5: 233-287. Erlbaum.

MACWHINNEY, BRIAN. 1978. "The acquisition of morphophonology." *Monogr. Soc. Res. Child Dev.*, No. 43.

—. 2000. *The CHILDES Project: Tools for analyzing talk*. 3rd Edition. Vol. 2: The Database. Mahwah, NJ: Lawrence Erlbaum Associates.

MARATSOS, M. et al. 1987. "A study in novel word learning: The productivity of the causative." In B. MacWhinney, ed., *Mechanisms of language acquisition,* 89-114, Erlbaum.

OLGUIN, R. and M. TOMASELLO. 1993. "Twenty-five month old children do not have a grammatical category of verb." *Cognitive development,* 8: 245-272.

PETERS, ANNE. 1986. "Early syntax." In P. Fletcher and M. Garman, eds., *Language acquisition,* 2nd edition: 307-325. Cambridge, England: Cambridge University Press.

PINKER, STEVEN et al. 1987. "Productivity and constraints in the acquisition of the passive." *Cognition,* 26: 195-267.

PINKER, STEVEN. 1984. *Language learnability and language development.* Cambridge, MA: Harvard University Press.

—. 1989. *Learnability and cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.

—. 1994. *The language instinct: How the mind creates language.* New York: Morrow.

PIZUTTO, E. and C. CASELLI. 1992. "The acquisition of Italian morphology." *Journal of child language,* 19: 491-557.

—. 1994. "The acquisition of Italian verb morphology in a cross-linguistic perspective." In Y. Levy, ed., *Other children, other languages*: 137-188. Erlbaum.

RUBIN, R. and J. PINE. 1998. "Subject-verb agreement in Brazilian Portuguese: what low error rates hide." *Journal of child language*, 25: 35-60.

STINE, E.L. AND J.N. BOHANNON III. 1983. "Imitations, interactions, and language acquisition." *Journal of Child Language*, 10, 589–603.

STOLL, S. 1998. "The acquisition of Russian aspect." *First language,* 18: 351-378.

TOMASELLO, MICHAEL. 1992. *First verbs: A case study in early grammatical development,* Cambridge, England: Cambridge University Press.

—. 2000a. "First steps toward a usage-based theory of language acquisition." *Cognitive linguistics,* 11-1/2: 61-82.

—. 2000b. "The item-based nature of children's early syntactic development." *Trends in cognitive science*, vol. 4, no. 4: 156-163.

WILSON, J. AND A. HENRY. 1998. "Parameter setting within a socially realistic linguistics." *Language in Society, 27,* 1–21

Appendix 1: Data from initial study

All instances of the *I (don't) V VP-INF* construction used by children in this study

2;0   would you # I have to put the pen back in my pocketbook.

2;1   <ssht> [?] i have to put um +...

2;1   I I don't want to go anywhere.

2;1   Mummy Mummy I wanna [: want to] help Daddy cause I'm a good girl.

2;1   and I want to come too.

2;1   and i want to play wuh wi with <shome apes> [?] some

2;1   but i have to put a bench for him.

2;1   dat's i wanna p i wanna i wanna hang this up wif dis.

2;1   duh floor is dirty # so i have to um bwush it wif my

2;1   get a screwdriver # I have ta find one.

2;1   i i wanted to put him in in.

2;1   i wanna be dey cook i wanna i wanna cook dem.

2;1   i wanted to uhh i want to +...

2;1   n+no # i wanna give um no # i'll go do it.

2;1   no # af no no first I have to invite all my

2;1   no # i wanna play play+dough.

2;1   no # i want to put him next to +...

2;1   no no # i wanna put some i wanna put some juice in dem

2;1   no no no # i wanna give duh froggie dat banana.

2;1   nope # I got ta get another one.

2;1   somebody can sit # um <le> [?] oh # i have to wear uh hat.

2;1   this i wanna i don't wanna see Mister Wogers.

2;1   with these tools # <I want ta put> I got ta fix em.

2;1   yeah # uh um i wanna go play now.

2;1   yeah I like beer cos I like to xxx.

2;2   i <wanta> i wanna xxx xxx in the water.

2;2   no # i wanna go uh see uh baby.

2;2   this my writing # I want ta write.

2;2   uh mommy and <uh> i <wanna> i wanna put on uh that.

2;3   <n> <n> i wanna go but uh and see it uh turn around that

2;3   I wanna eat it and like that # i wanna broken and all page

2;3   aw # i <wanna> i wanna fo.

2;3   he's doing work in the garage # i wanna go walk up and see the xxx

2;3   i think i wanna go bed.

2;3   i think i wanna go binky bed.

2;3   i want <ba> i wanna bite some.

2;3   i wanna <go> i want <tell> i wanna <tell> i wanna's

2;3   i wanna play+dough # i wanna play+dough in the room.

2;3   i want it # i wanna play+dough.

2;3   tape # i wanna play tape that go in a box.

2;3   xxx xxx what in a box &i i wanna climbing &i &i in mommy's room?

2;3   yeah # i wanna carry.

2;3   yeah # i wanna put <that> i wanna that record off.

2;4   <because> to do that # because i wanna pah pah pah pah!

2;4   <i> <i> <i> i wanna get the book.

2;4   i <wanna> i wanna punch out this.

2;4   i <wanna> i want to write ana's name # i want to write ana's

2;4   kay # i want to get +...

2;4   yeah # xxx i wanna big kitty.

2;5   <uh> because i wanna go in the water # daddy # i wanna go out.

2;5   because i wanna get down.

2;5   because the wind blow down there to the grass because i wanna go

2;5   i <wand> i <wanna> i wanna put dis on dat fing # and

2;5   i want do <wu> # i want to wun on duh stickers now.

2;5   i want # i want ta play airplane in the sky with mine # okay?

2;5   i want to play with daddy machine # i want to <play> i wanna

2;5   in the sky because i wanna get down.

2;5   no # i <wand> i wanna wear pants and uh short # okay?

2;5   no # i wanna pour some more.

2;5   yes # i wanna go play.

2;6   <I want ta write with> I want ta write with it.

2;6   <and> <uh> and so i have to fix it.

2;6   <and> i want to wun awound.

2;6   <mommy> I want ta put this mommy on the +...

2;6   <uh> <uh> i want to play them.

2;6   <uhh> i want to go outside.

2;6   I want ta give I want ta give you a piece # a # paper # of

2;6   help me walk <hm+hm> <hm> i jump by myself # i wanna jump on

2;6   i want see <duh> how duh clock work # i wanna <have> have

2;6   it broke # i have to fix it with other.

2;6   no # i wanna go in <uh> in <uh> in <uh> <i> in <uh>

2;6   no # i wanna go wif mommy and daddy.

2;6   no # i want to go on dere.

2;6   no I xxx # I want ta be quiet in the bus.

2;6   xxx # I want ta write.

2;7   <I want a ya> <I want ta put a> I want ta put this -.

2;7   <and> and i wanna # i wanna fings <dat> &da you go

2;7   <uh> <uh> i want to go in duh store and buy some string

2;7   <uhh> <uhh> <uhh> <uhh> <um> i wanna put it

2;7   <uhh> want <get> want <you> <um> i wanna get

2;7   ah # i want to go and see.

2;7   dere # i <wan> i wanna put it outside.

2;7   i <wand> i <wantuh> i want ta put it in the sky # now.

2;7   i wanna <fi> <uh> <um> i wanna put it together # go

2;7   no # I got ta walk in here.

2;7   no I need to put my lipstick on first.

2;7   no I want ta write on your paper # ok.

2;7   no i wanna go to minky+bed # dat means i want to go to house.

2;7   now # I want to write.

2;7   now i can go up there # i wanna climb up here.

2;7   oh # i <have> i have to wing duh phone first # i have to wing

2;7   oh ## better hurry # I got ta get some food.

2;7   ok # I want to write.

2;7   what's that I wanna [: want to] see # penguins show me # see

2;7   yeah # see if dis broke # i wanna put some words <tuh> too

2;7   yeah ## I got ta get the doctor.

2;8   cos I wanna [: want to] put it in here.

2;8   <I want see> <I want ta> I want ta wear that

2;8   but i but i said i wanna go to your house.

2;8   but i wanna i i wanna read it back here.

2;8   but i want to # uhh+oh # it fits.

2;8   i i wanna +...

2;8   i i wanna put it right dere next to your yips.

2;8   i wan i wanna jump on your waterbed.

2;8   i wan i wanna put it right here.

2;8   i want i wanna put it near here.

2;8   i want i wanna xxx in dere.

2;8   i wanna play wif i wanna play wif i wanna play wif dis.

2;8   like that # I got ta go there and can't fit.

2;8   momma # i want to go <sit> [?] up here.

2;8   no # i wanna put it on here.

2;8   uhh # i wanna go to room again.

2;8   uhh # i wanna put it in the house.

2;8   uhh i i wanna put it in the cup.

2;8   uhh uhh uhh i want to watch.

2;8   we i wanna play uh wecord in my room.

2;8   xxx I want to pick it.

2;8   yeah # I want ta wear you ring.

2;9   and also i wanna play dat fing.

2;9   and here's a book dat i want to to make.

2;9   and i can take it off because i wanna put it on here.

2;9   because i wanna um leave it like &da in case dis goes like

2;9   but i wanna put another one on &i and two fings.

2;9   don't # i i want i wanna you put it down.

2;9   i want duh duh ladder and i wanna show i wanna get dose #

2;9   now # i wanna fing wight here # okay ?

2;9   there's a car in there # I got ta go see it.

2;9   this was broken and I got ta fix it.

2;9   where when i wanna get off i just can jump off and then +...

3;0   <like dat> [?] first # i hafta get uh game.

3;0   <yeah> [?] i wanna play some playdough on some paper.

3;0   No I wanna weed to you.

3;0   and this is the bus because I like to go to the grandma's house.

3;0   another children could s i wanna put duh bed in +...

3;0   but i don't wanna do i don't wanna go in our my car!

3;0   but i wanna play wif duh families.

3;0   first i hafta get um this xxx this out # and this.

3;0   fo i wanna put <it> [?] back on dere.

3;0   i think i wanna get this.

3;0   i wanna put it in i wanna put this string in there.

3;0   no # i i have to lie dem down <for a> [?] second.

3;0   no # i'll put it on i want to put it here.

3;0   now i have to put dem back in.

3;0   um but first i hadded to get like dis +...

3;0   um i have to uhh i wanna play wif uh little game.

3;0   yeah # i i wanna go pee.

3;1   because i haf to go to another place.

3;1   I I I oh I have to look for the sun # sun # nope [: no].

3;1   I want to # I want to come!

3;1   because i wanna <get> [?] s'm some more potato chips.

3;1   but but now i hafta put the legs.

3;1   cause I had them # I had to knock them down.

3;1   em # I xxx I had to xxx them down.

3;1   ha ha ha # xxx every day # and I have to &s uh # do all my homework

3;1   hey # i hafta get some chairs.

3;1   hey # i have to c take him somewhere.

3;1   hey # i wanna leave it closed.

3;1   i think i hafta pound it harder and harder.

3;1   i think i have to give this to Amanda uhh uhh

3;1   i wan i wanna get dese off.

3;1   i wanna play i wanna play with something else.

3;1   mummy I want to play!

3;1   now i have to dry this up and i hafta go to another place.

3;1   oh # no # i haf i have to go on dis motor+cycle.

3;1   oh # no # oh # no # i hafta put dis in.

3;1   uh I have to keep them down here.

3;1   uhh no # dese are i wanna cover up duh elephant because

3;1   um # I want to be the fire brigade.

3;1   um i think i haf to wait.

3;1   well I want to go.

3;1   yeah # i hafta climb up here.

3;1   yeah # i hafta poke it uh little more.

3;1   yep # i have to go on dis motor+cycle # because we need to drive s

3;2   no # i like to get um uhh no # i have popcorn # but

3;2   yeah # and i wanna i wanna see it make duh sound.

3;5   I was standing # I don't need to look at the # these animals.

3;5   and then I want to go to my mummy.

3;5   get him back on my # on your knee # I want to xx.

3;6   I have to wash my hands but I need to go to the toilet as well.

3;6   a wee thing but I have to put all of them in.

3;6   and I want to put makeup on to me.

3;6   and you keep it closed # cause I have to get my bike with me.

3;6   get sand in # I have to get a new bucket and get sand in # to the

3;6   now I have to catch another one.

3;6   yeah but I have to get your breakfast ready.

3;7   I don't want one on there # I want to go on the floor.

3;7   no I don't want to catch a fish.

3;7	the wee ones aren't # where I had to keep them.
3;7	xxxx I # I don't want to play with xxx.
3;9	I will feed him # I # I want to bring my xx back xx xx.
3;9	Mummy I want to keep my tracker coat on.
3;9	and I want to play with that # xx lift him again.
3;9	can I put it back in # I want to put it back in.
3;9	oh # I have to get this in!
4;0	I know what I want to play with xxx.
4;0	I'm going to be # I'm go # I want to put all these back in cos I'm
4;0	Mummy I don't wanna [: want to] go on holidays.
4;0	and I have to get a trailor.
4;0	can't find what I have to put on.
4;0	cause I wanna [: want to] do #.
4;0	now if you open your bags and I wanna [: want to] see what's in
4;0	oh I # I have to close this up I need everything in it.
4;0	well # that's all I have to give him.
4;0	xxx I have to xxx.
4;0	xxx I wanna [: want to] go out and play xxx.
4;0	yep I have to get a new one.
4;1	so I have to be uh # this one.
4;2	do I have to catch anyone I like to?
4;2	no I want to watch a video with David.
4;2	yes and I want to play that.
4;4	I like it and I want to keep it.
4;4	cause I like # cause I like calling it Neil # I want to call it
4;4	cause I wanted to build a city so I breaked it.
4;4	cause the man couldn't get down the stairs # I have to lift him.
4;4	no cause # cause I don't want to # to do it the day [: today].
4;4	um # I wanna [: want to] have a look xxx.
4;4	well I need to build you a wee boat.
4;4	why do I have to go to granny Kenny's?
4;4	yip # if the two dies I have to get # I have to call him Grape and
4;5	well I don't want to go.

Appendix 2: Study 2 Perl script to gather data for the search of (NOUN) (don't) V VP-INF

```perl
#!/usr/local/bin/perl
use strict "vars";
my $filecount=55;
my $name='ADAM';
my $childid='CHI';
my $currentline;
my $previousline=undef;
my %childlines;
my %otherslines;
my %totalverbs;
my %verbcounts;
my $child='CHILD';
my $others='OTHERS';
open(RESULT,">>results$name.txt")||die "Couldn't open <results$name.txt>: $!\n";
print RESULT "CHILD: $name\n";
print RESULT "SEARCH PATTERN: PRO (don't) V VP-INF\n";
for(my $i=1;$i<$filecount+1;$i++)
{
        open(IN,"$i$name.txt")||die "Couldn't open <$i$name.txt>: $!\n";
        %childlines=();
        %otherslines=();
        %totalverbs=();
        while($currentline=<IN>)
        {
                chomp($currentline);
                findandprintage($currentline);
                matchregex($currentline,$previousline,$childid);
                totalverbcounts($currentline,$previousline,$childid);
                $previousline=$currentline;
        }
        %verbcounts=();
        counteachverb(%childlines);
        printverbcount($child,%verbcounts);
        %verbcounts=();
        counteachverb(%otherslines);
        printverbcount($others,%verbcounts);
        printdata($child,%childlines);
        printdata($others,%otherslines);
        printverbtotal($child,%totalverbs);
}
sub findandprintage
{
        my ($line)=@_;
        if($line=~/^\@ID:.*\|(\d;\d.(\d|))/)
        {
                print RESULT "\n\n.....................................\n\n";
                print RESULT "AGE: $1\n";
        }
}
sub matchregex
{
```

A Corpus Study on the Item-based Nature of Early Grammar Acquisition   25

```perl
my ($current,$previous,$id)=@_;
if($current=~/^%mor:.*v\|wan(t|ts|ted) inf\|t(o|a) v\|/)
{
        if($previous=~/^*$id/)
        {
                $childlines{$previous}='want';
        }
        else
        {
                $otherslines{$previous}='want';
        }
}
elsif($current=~/^%mor:.*v\|ha(ve|s|d) inf\|t(o|a) v\|/)
{
        if($previous=~/^*$id/)
        {
                $childlines{$previous}='have';
        }
        else

        {
                $otherslines{$previous}='have';
        }
}
elsif($current=~/^%mor:.*v\|g(et|ets|ot) inf\|t(o|a) v\|/)
{
        if($previous=~/^*$id/)
        {
                $childlines{$previous}='get';
        }
        else

        {
                $otherslines{$previous}='get';
        }
}
elsif($current=~/^%mor:.*v\|nee(d|ds|ded) inf\|t(o|a) v\|/)
{
        if($previous=~/^*$id/)
        {
                $childlines{$previous}='need';
        }
        else
        {
                $otherslines{$previous}='need';
        }
}
elsif($current=~/^%mor:.*v\|lik(e|es|ed) inf\|t(o|a) v\|/)
{
        if($previous=~/^*$id/)
        {
                $childlines{$previous}='like';
        }
        else
        {
                $otherslines{$previous}='like';
        }
}
```

26   Colorado Research in Linguistics, Volume 17 (2004)

```
}
sub totalverbcounts
{
        my ($current,$previous,$id)=@_;
        if($current=~/v\|wan(t|ts|ted)/)
        {
                if($previous=~/^*$id/)
                {
                        $totalverbs{want}++;
                }
        }
        if($current=~/v\|ha(ve|s|d)/)
        {
                if($previous=~/^*$id/)
                {
                        $totalverbs{have}++;
                }
        }
        if($current=~/v\|g(et|ets|ot)/)
        {
                if($previous=~/^*$id/)
                {
                        $totalverbs{get}++;
                }
        }
        if($current=~/v\|nee(d|ds|ded)/)
        {
                if($previous=~/^*$id/)
                {
                        $totalverbs{need}++;
                }
        }
        if($current=~/v\|lik(e|es|ed)/)
        {
                if($previous=~/^*$id/)
                {
                        $totalverbs{like}++;
                }
        }
}
sub counteachverb
{
        my $value;
        my (%verbhash)=@_;
        my @values=sort(values(%verbhash));
        foreach $value (@values)
        {
                if($value=~/want/)
                {
                        $verbcounts{want}++;
                }
                elsif($value=~/have/)
                {
                        $verbcounts{have}++;
                }
                elsif($value=~/get/)
```

A Corpus Study on the Item-based Nature of Early Grammar Acquisition   27

```perl
                {
                        $verbcounts{get}++;
                }
                elsif($value=~/need/)
                {
                        $verbcounts{need}++;
                }
                elsif($value=~/like/)
                {
                        $verbcounts{like}++;
                }
        }
}
sub printverbcount
{
        my $element;
        my($person,%verb)=@_;
        print RESULT "\n$person\n\n";
        my @array=sort(keys(%verb));
        foreach $element (@array)
        {
                print RESULT "\t$element: $verb{$element}\n";
        }
}
sub printdata
{
        my $key;
        my ($person,%hash)=@_;
        my @total=sort(keys(%hash));
        print RESULT "\n$person DATA\n\n";
        foreach $key (@total)
        {
                print RESULT "\t$key\n";
        }
}
sub printverbtotal
{
        my $element;
        my($person,%totalverb)=@_;
        print RESULT "\nTOTAL VERB COUNT FOR $person\n\n";
        my @array=sort(keys(%totalverb));
        foreach $element (@array)
        {
                print RESULT "\t$element: $totalverb{$element}\n";
        }
}
```

Appendix 3: Sample data output from the Perl script in appendix 2

CHILD: ADAM
SEARCH PATTERN: PRO (don't) V VP-INF

....

AGE: 2;3.4

CHILD

       want: 1

OTHERS

       have: 1
       like: 5
       want: 8

CHILD DATA

       *CHI: I [?] wan(t) (t)a [?] box .

OTHERS DATA

       *MOT: Adam # want to close the box ?
       *MOT: I guess she might like to see that .
       *MOT: do you want to read a book ?
       *MOT: do you want to see what I have ?
       *MOT: do you want to write on here ?
       *MOT: why do you like to throw your book ?
       *MOT: would you like to have your books on the bookshelf too ?
       *MOT: wouldn't you like to pick these up ?
       *MOT: you like to walk ?
       *URS: do you want to bring over the high stool for me to sit on ?
       *URS: do you want to paste it ?
       *URS: do you want to play with them ?
       *URS: do you want to see this ?
       *URS: you'll have to pick them up .

TOTAL VERB COUNT FOR CHILD

       get: 48
       like: 18
       want: 3

....

AGE: 4;10

CHILD

       need: 1

want: 12

OTHERS

have: 2
want: 5

CHILD DATA

*CHI: I don't wan(t) (t)a do it .
*CHI: I don't wan(t) (t)a go under the bottom of the sea .
*CHI: I don't want to be a monkey .
*CHI: I don't want to be staying here .
*CHI: I wan(t) (t)a be my own self .
*CHI: I wan(t) (t)a go with you .
*CHI: I wan(t) (t)a make a kite .
*CHI: I wan(t) (t)a see it go higher .
*CHI: I want to go too .
*CHI: I want to try one out .
*CHI: Ursula # I wan(t) (t)a make my kite .
*CHI: in case you want to go someplace ?
*CHI: we don't need to use de tops .

OTHERS DATA

*MOT: Paul # we might have to go upstairs # hmm ?
*MOT: you want to put that one on ?
*URS: I think they have to be the double kind .
*URS: d(o) you want to use another color ?
*URS: d(o) you want to use the newspaper ?
*URS: in case you want to carry it with you .
*URS: you want to look at the little book .

TOTAL VERB COUNT FOR CHILD

get: 43
have: 34
like: 23
need: 9
want: 34

....

AGE: 4;10

CHILD

have: 8
like: 1
want: 8

OTHERS

have: 5
like: 1
want: 4

CHILD DATA

*CHI: Daddy # you don't have to be strong .
*CHI: I don't have to need it off .
*CHI: I wan(t) (t)a make a bow and arrow .
*CHI: I want to get some +...
*CHI: I want to hold it .
*CHI: after you finish # so nobody won't have to clean the place up .
*CHI: have to tear it out because how can we do it ?
*CHI: here # we going to have to build one with another string on it .
*CHI: hey # I wan(t) (t)a make a chair and a hammer .
*CHI: hey # you wan(t) (t)a make it a little smaller again ?
*CHI: if you want to try one # wait .
*CHI: ok # I'm gonna have to pick things up .
*CHI: see if the flowers would like to watch me .
*CHI: we gonna have to need a small bowl .
*CHI: you don't have to push him .
*CHI: you want to keep the hand+buckler@c ?
*CHI: you want to take only one home ?

OTHERS DATA

*MOT: I don't want to have a dog .
*MOT: now I don't have to get any birthday presents .
*URS: alright # but first I have to get it all collected .
*URS: d(o) you want to ask your [?] mother [?] for that ?
*URS: d(o) you want to look at them ?
*URS: maybe you'll have to show me some things # alright ?
*URS: then we have to make a sprinkling can .
*URS: we may have to put another drop of ink on there .
*URS: would you like to bring your chair over here ?
*URS: you want to make some too ?

TOTAL VERB COUNT FOR CHILD

get: 68
have: 51
like: 32
need: 9
want: 36