The acquisition of complement clause constructions: A sentence repetition study.

Evan Kidd
Max Planck Child Study Centre, University of Manchester

Elena Lieven & Michael Tomasello
Max Planck Institute for Evolutionary Anthropology

1. Introduction

The presence of complement clause constructions in child speech marks the beginning of the child's entrance into the more complex aspects of grammar. Since complement clauses are lexically-specified, individual complement taking verbs (CTVs) must be tagged with information about the kinds of complements they can license; their acquisition suggests that children are beginning to elaborate the argument structure privileges of individual CTVs. This suggests that children must attend to the distributional properties of individual CTVs and gradually master complementation over time, an hypothesis that has been confirmed by studies of children's naturalistic speech (Bloom, Rispoli, Gartner, & Hafitz 1989, Diessel & Tomasello 2001, Limber 1973). This paper reports on an experiment investigating 3 – 5-year-old children's acquisition of sentential complements. We present results that suggest children's acquisition of complement clause constructions is intimately tied to CTV frequency.

1.1 Complement clauses

Complement clauses are sentential structures that function as an argument of a matrix verb. The complement clause can occupy the subject argument slot, as in (1), or the object slot, as in (2) (the complement clause is underlined).

(1) That Rufus was late angered his boss.
(2) Rufus could see that he had angered his boss.

In the unmarked case (2) the complementiser *that* is optional, and consequently the complementiser does not occur frequently in the naturalistic speech of children or adults (Diessel & Tomasello 2001, Thompson & Mulac 1991).

A number of syntactically distinct complement clauses can be identified. First, a distinction is made between finite and non-finite complements. In (3) the complement clause is finite, since the subordinate verb is tensed. In sentence (4) the complement clause is non-finite; the subordinate verb is tenseless.

(3) I see Elena walked to the Institute this morning.
(4) Mike enjoys chatting with his students.

Finite and non-finite complement clauses are not subject to the same grammatical constraints. Notably, non-finite clauses frequently lack an overt subject (4), and do not require subject-verb agreement (McCawley 1988). Non-finite complement clause constructions like (4) have been the subject of much acquisition research investigating children's understanding of control structures, where the focus has been on how children assign coreference (e.g., Cairns, McDaniel, Hsu, & Rapp 1994, Hsu, Cairns, Eisenberg, & Schlisselberg 1989, Sherman & Lust 1986).

Furthermore, there are different types of finite and non-finite complement clauses, of which this paper considers only one form – sentential complements. The distribution of finite and non-finite complements depends on the matrix verb; some allow either clause type, and some only allow either finite or non-finite complements, as shown in (5) – (10).

(5) *I think he run/running to the shop.
(6) I think he ran to the shop.
(7) I watched him run/running to the shop.
(8) *I watched him ran to the shop.
(9) I see him *run/running to the shop.[1]
(10) I see he ran to the shop.

The verb *think* allows finite complement clauses only, whereas *watch* only allows a non-finite complement. In contrast, *see* allows both complement types, although it only allows non-finite clauses in the restricted *–ing* case. These cases suggest that learning the grammatical restrictions on the content of the subordinate clause is a non-trivial process.

The issues surrounding children's acquisition of sentential complements like (5) – (10) have not been the topic of much experimental investigation.[2] However, there have been some naturalistic studies.

1.2 Past research

Limber (1973) and Bloom and her colleagues (1991) charted the development of complex sentence types in the naturalistic speech of children between the ages of 2- and 3-years. Both studies reported a uniform pattern for the acquisition of sentential complements. The children that they followed first began to use complement clause constructions at around age 2, including infinitival complements and sentential complements. These constructions occurred with a restricted range of verbs, consisting mainly of light verbs such as *go*, *want*, and *make*, and verbs of perception such as *see* and *look*. Following this period the children in these studies began to use a larger range of verbs in their production of sentential complements, but their productions were restricted to object complements only. A prominent feature of children's first sentential complements were that they did not appear to express two full propositions, as they might in the adult grammar. Instead, the main clause was propositionally empty, such as in the sentence *I think I want grape juice*, where the main clause is used parenthetically. In their more detailed analysis of four CTVs (*think*, *know*, *see*, *look* [*at*]), Bloom, Rispoli, Gartner, and Hafitz (1989) showed that while children were syntactically productive within the complement clause, their matrix clauses were highly formulaic; in particular, the matrix subject was restricted largely to first and second person pronouns (although this depended on the individual verb) and the matrix verb rarely inflected. In contrast, the subjects within the complement clauses were more varied, and subordinate verbs were inflected or marked for modality in more than 50% of children's utterances.

Diessel and Tomasello (2001) conducted a more comprehensive naturalistic study of seven children's acquisition of complement clause constructions. Restricting their analysis to finite complement clauses, Diessel and Tomasello analysed the development of these structures in seven children followed longitudinally, from age 1;2 to 5;2, enabling them to study children's use of a larger number of CTVs than Limber (1973) and Bloom et al. (1989). Once again, Diessel and Tomasello showed that children's first uses of complement clause constructions were highly formulaic. For instance, children's use of psychological predicates and verbs of perception such as *think*, *know*, and *see* were restricted to short, formulaic clauses such as *I think + S, I know + S, See if + S*. Since the main clauses in these sentences almost never deviated from their formulaic form (i.e., the verbs were rarely inflected, there were no modals), the authors suggested that they are more accurately interpreted as clausal operators that modify the content of the rest of the sentence. As such, these utterances do not fit the standard definition of a complement clause construction, since they do not strictly

[1] In some dialects of English the bare stem form of the verb in this construction is acceptable, particularly when co-occurring with a modal verb in the main clause, as in *I can see him run to the shop from my bedroom window*, or when occurring with a supporting discourse context.

[2] There have been a number of experimental studies that have used complement constructions to investigate hypothesised grammatical phenomena, such as *wh*-movement (e.g., de Villiers & de Villiers, 1999; de Villiers, Roeper, & Vainikka, 1990; Roeper & de Villiers, 1994).

express two propositions. The children in the study gradually began to express propositional content within the main clause as they grew older, although the rate at which children became more productive differed for each child. This developmental pattern was also argued to be true of *if*- and *wh*-complements.

Diessel and Tomasello (2001) considered three explanations for these results. First, they suggested that the children's use of formulaic complement constructions is likely to have mirrored their input, since these are highly frequent in adult language (see Thompson & Mulac 1991). Second, they suggested that children have difficulty processing two propositions at the same time, as they would have to do when producing complement constructions that deviate from the formulaic use. A third of line explanation is related to the second, where the authors suggested that children's formulaic use of CTVs is in some part due to the fact that they had yet to develop the cognitive abilities required to use many of the CTVs in their lexicon in the appropriate sense. This is because many CTVs denote mental states that young children find difficult to entertain.

The naturalistic data suggest a very uniform qualitative developmental pattern: Children initially produce highly formulaic complement constructions where the main clause is propositionally empty, and gradually develop productive competence over the main and subordinate clause over time. Furthermore, initial knowledge of complementation appears to be tied to individual CTVs. These results are consistent with current 'usage-based' approaches to syntactic development (Tomasello 2003), which suggest that children converge on the adult end state by abstracting over initially low-scope lexically-specific constructions. The present study was designed to investigate children's emerging knowledge of complementation in an experimental setting. In particular, we aimed to investigate children's knowledge of the complement clause constructions they most often say and hear: formulaic sentential complements. We investigated the extent to which the frequency of the CTV in formulaic sentential complements affects children's ability to remember and repeat grammatical complement clause constructions and correct ungrammatical complements clauses constructions. The effects of lexical frequency are well attested in studies investigating adults and children's language processing (e.g., Hare, MacRae, & Elman 2003, 2004, Trueswell 1996, Snedeker & Trueswell 2004). These studies suggest that the frequency with which a verb appears in a particular syntactic environment influences the manner in which the parser processes upcoming language. Since children's early language has been shown to be lexically-specific (see Tomasello 2000, 2003), the corresponding argument in acquisition is that the frequency of lexical co-occurrences plays a significant role in how children acquire the syntactic regularities of their language.

Following this logic, we hypothesised that children would provide more exact repetitions of sentences containing high frequency (HF) CTVs verbs than sentences containing low frequency (LF) CTVs, since the argument structure privileges of HF CTVs are likely to be better attested than the argument structure privileges of LF CTVs. For this same reason, we also hypothesised that children would provide more corrections to ungrammatical sentences containing high frequency CTVs than to those containing LF CTVs.

2. Method

2.1 Participants

Seventy-eight children aged between 2;10 and 5;9 were recruited from nurseries and primary schools in Manchester, England. Eleven children were excluded because they failed to complete both testing sessions, four were excluded because they were identified by their teacher or parent as having language or cognitive problems, and two were excluded because they were bilingual. The final sample consisted of 61 children aged between 2;10 and 5;9. The sample was divided into three age groups: 3-year-olds (n = 20, mean age: 3;4, range: 2;10 – 3;8), 4-year-olds (n = 23, mean age: 4;2, range: 3;9 – 4;7), and 5-year-olds (n = 18, mean age: 5;1, range: 4;8 – 5;9). None possessed any known language impairment or learning difficulty.

2.2 Materials

Forty test sentences were constructed using ten complement-taking verbs: five high frequency (HF) and five low frequency (LF). To identify the verbs, a corpus study was conducted using data from six children (Adam, Eve, Sarah, Peter, Nina, Naomi) in the CHILDES corpus (MacWhinney 2000). For each child, the frequency with which each verb occurred in a sentential complement construction relative to the total number of times that verb was used, was computed. The data were then pooled to obtain an overall frequency estimate. The verbs were identified and matched roughly for semantic similarity; the verb pairs chosen were *think-pretend*, *bet-say*, *hope-know*, *see-hear*, and *watch-look at* (HF verbs first). *See* and *hear* take either a finite or non-finite complement. When appearing with finite S-comps *see* was a high frequency and *hear* low frequency; when appearing with a non-finite S-comp this pattern was reversed (for details of the corpus study see Kidd, Lieven, & Tomasello, in prep.).

Each verb pair appeared in four sentence frames manipulating grammaticality: two containing a finite complement clause and the other two containing a non-finite complement clause. These sentence frames, for the verb pair *say-bet*, are shown in (11) – (14).

(11) I say/bet he is talking on the telephone.
(12) I say/bet she eats some chocolate ice-cream.
(13) *I say/bet him jumping over the fence.
(14) *I say/bet her give the present to her mum.

All test sentences began with the first person singular pronoun *I*, followed by the main verb and sentential complement. That is, the main clause was propositionally empty. Additionally, following the naturalistic data, the main verb was not marked for tense. This format was chosen because Diessel and Tomasello (2001) reported that this form marks children's earliest and most frequent productions of sentential complements. Grammaticality was manipulated by changing the finiteness of the subordinate clause. For example, since *say* and *bet* take only a finite complement clause, (13) and (14) are ungrammatical because the subordinate clauses in these sentences are non-finite.

Each sentence was preceded by its own discourse context. These were included to ensure a felicitous reading of the test sentences. Toy characters served as referents for the test sentences, providing children with concrete referents onto which they could map the language they heard. The characters were a mother and a father, three girls ('Alice', 'Jenny', and 'Sarah'), two boys ('Tim' and 'Johnny'), and two pets (a dog and a horse).

2.3 Procedure

The children were tested individually in a quite area of their nursery/school. They were told that they were going to play some games with the experimenter. The first game was called the 'parrot game'. In this task the experimenter discussed with the child how a parrot talks by copying 'exactly what people say'. The child was then asked if they could be a parrot, and repeat exactly what the experimenter said. Upon the child's approval, the experimenter read through three example sentences (two simple sentences and coordinate structure), and asked the child to repeat the sentence after they had heard it, just as a parrot would. If children made any mistakes the experimenter corrected them, and they were asked to repeat the sentence again. This continued until the child gave an exact repetition. Therefore the pre-training task served to instruct children that they were required to repeat the sentences read by the experimenter exactly.

The children were then given three example items that were in the same format as the test items to follow. That is, a discourse context was established which led to one of the toy characters producing an utterance that the child would be required to repeat. Therefore the three examples were used to train the children on the format of the test sentences. In particular, it was important that children were aware that they were required to repeat the

characters' utterances in their entirety. The training sentences therefore provided an opportunity to correct the children if they did not understand the full requirements of the task.

The test sentences immediately followed the pre-training and training sessions. An example test item proceeded as follows. First, the experimenter told the child which characters were needed for that particular 'story'. The characters were named and put into position on the table in front of a purpose built set (a house). The children were then told the story. An example is given in (15) (the test sentence is underlined):

(15) *Tim and his mum were just about to eat dinner but they couldn't find Dad. 'Maybe he's in the garden', said Mum. 'No', said Tim, 'I bet he is talking on the telephone'.*

The experimenter then said to the child 'Can you say that?' The child's response was then recorded on a response sheet. Each sentence had its own discourse context. There were two testing sessions. Half of the sentences were presented in the first session, and the other half were presented in the second session. The sessions were between one and two weeks apart. Each session lasted between 20 – 30 minutes. Some of the younger children were tested over a greater number of sessions because they did not want to sit through the full sessions. When this occurred the children in question were always re-introduced to the testing procedure using the training items. All sessions were audio-recorded. The presentation of test sentences was pseudo-randomised to avoid order effects; eight orders were used.

3. Results

Children's utterances were transcribed and coded for: (i) exact repetitions and, (ii) corrections to ungrammatical sentences. A number of analyses were conducted. The analyses of children's exact grammatical repetitions are reported, followed by the analyses of children's corrections to the ungrammatical sentences.

3.1 Repetitions of grammatical sentences

The proportion of exact repetitions of finite sentential complements by verb frequency and age is shown in Figure 1, and the proportion of exact repetitions of non-finite sentential complements by verb frequency and age is shown in Figure 2.

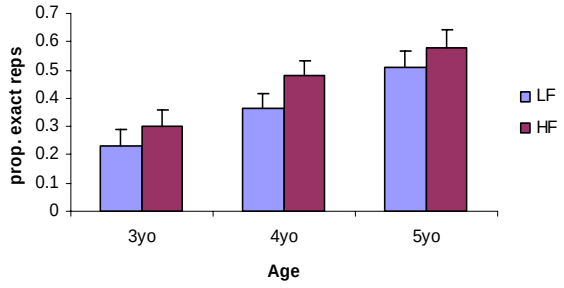

Figure 1. Exact repetitions of finite S-comps.

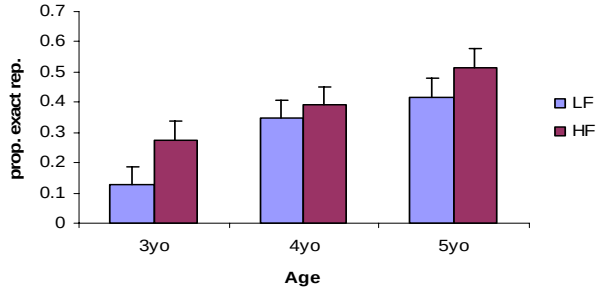

Figure 2. Exact repetitions of non-finite S-comps.

Both Figures 1 and 2 show that children in each age group provided more exact repetitions of grammatical sentences containing HF CTVs than LF CTVs. Furthermore, the number of exact repetitions increased with age. Two 3 (age: 3-, 4-, & 5-year-olds) by 2 (CTV frequency: HF vs. LF) mixed ANOVAs were conducted using the proportion of exact repetitions of finite and non-finite S-comps as the dependent measures, respectively.

For the finite S-comps there was a significant main effect of CTV frequency ($F(1, 58) = 12.751$, $p = .001$, partial $\eta^2 = .180$), showing that all children made significantly more exact

repetitions when the test sentence contained a HF CTV. There was also a significant main effect of group ($F(2, 58) = 6.304$, $p = .003$, partial $\eta^2 = .179$). Least Significant Difference (LSD) post hoc comparisons showed that both the 5yo and 4yo children produced significantly more exact repetitions than the 3yo children (both p's < .05), but did not differ from each other. There was no CTV frequency by group interaction.

For the non-finite S-comps there was a main effect of CTV frequency ($F(1, 58) = 7.377$, $p = .009$, partial $\eta^2 = .113$), once again showing that all children made significantly more exact repetitions when the test sentence contained a HF CTV. There was also a significant main effect of group ($F(2, 58) = 6.338$, $p = .003$, partial $\eta^2 = .179$). Least Significant Difference (LSD) post hoc comparisons showed that, once again, both the 5yo and 4yo children produced significantly more exact repetitions than the 3yo children (both p's < .02), but did not differ from each other. Once again there was no CTV frequency by group interaction.

3.2 Corrections to ungrammatical sentences

An analysis of children's exact repetitions of ungrammatical sentences yielded no significant effects. Children's corrections were divided into two categories: (i) corrections that preserved the semantics of the intended message and corrected the syntactic violation only ('syntax only' correction), and (ii) corrections that changed both the syntax and semantics of the sentence ('syntax + semantics' correction). These latter corrections were cases where children substituted a different main verb into the test sentence. Figure 3 shows the 'syntax only' correction for each age group by CTV frequency, and Figure 4 shows the 'syntax + semantics' corrections for each age group by CTV frequency.

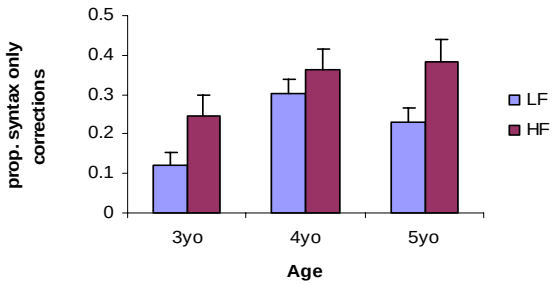

Figure 3. Proportion of 'syntax only' corrections.

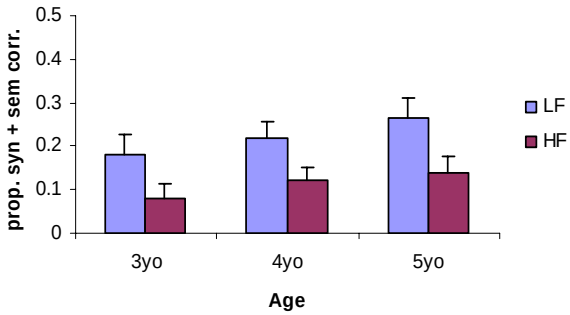

Figure 4. Proportion of 'syntax + semantics' corrections.

Figure 3 shows that children provided more 'syntax only' corrections when test sentences contained HF CTVs. Conversely, Figure 4 shows that children made more 'syntax + semantics' corrections when test sentences contained LF CTVs. The number of 'syntax only' corrections increased with age, but the number of 'syntax + semantics' corrections appeared not to. Two 3 (age: 3-, 4-, & 5-year-olds) by 2 (CTV frequency: HF vs. LF) mixed ANOVAs were conducted using the proportion of children's 'syntax only' and 'syntax + semantics' corrections as the dependent measures, respectively.

For the 'syntax only' corrections there was a main effect of CTV frequency ($F(1, 58) = 16.954$, $p < .0001$, partial $\eta^2 = .226$), suggesting that children in each age group were better able to make 'syntax only' corrections when the test sentence contained a HF CTV. There was also a significant main effect of group ($F(1, 58) = 4.737$, $p = .012$, partial $\eta^2 = .14$). Post hoc LSD tests showed that both the 5yo and 4yo children made significantly more 'syntax only' corrections than the 3yo children (both p's < .03), but did not differ from each other.

For the 'syntax + semantics' corrections there was a main effect of frequency ($F(1, 58) = 26.31$, $p < .0001$, partial $\eta^2 = .312$), suggesting that children most often used this correction strategy when test sentences contained a LF CTV. There were no other significant effects.

The results from the analyses of children's corrections suggest that, when ungrammatical test sentences contained a HF CTV, children were able to preserve the intended message of the test sentence and correct the grammatical violation only. Conversely, when ungrammatical test sentences contained a LF CTV, children most often corrected by substituting a different verb into their repetition, thereby changing the syntax *and* semantics of the original test sentence. An analysis of children's main verb substitutions showed that, across all age groups, children were over four and a half times more likely to correct an ungrammatical test sentence by substituting in a HF CTV than a LF CTV ($p < .0001$). This suggests that when children were presented with an ungrammatical sentence containing a CTV they were less familiar with (i.e., a LF CTV), they were likely to 'trade up' to better entrenched alternatives. Over half of these 'syntax + semantics' corrections involved children using the verb *think* – a verb that was identified by Diessel and Tomasello (2001) as the most frequently used CTV by children in their study.

4. Discussion

The present study aimed to investigate experimentally a feature of children's acquisition of complement clauses that has been documented several times in naturalistic studies: children's initial productions of sentential complements are highly restricted in function and form. We investigated the extent to which the frequency with which a CTV appears in a sentential complement construction contributes to children's ability to remember and repeat this structure. The results suggest that frequency plays an important role.

Hypothesis one, that children would produce more exact repetitions of grammatical sentences containing high frequency CTVs than sentences containing low frequency CTVs, was supported. Children provided more exact repetitions of finite and non-finite S-Comp constructions when the test sentence contained a HF CTV. This provides experimental evidence to support the idea that children's grammatical knowledge is organised around high frequency lexical items (Tomasello 2003).

Hypothesis two, that children would make more corrections to ungrammatical sentences containing HF CTVs, was supported by the results. In fact, the pattern of corrections exhibited a very striking effect. Children from every age group made more 'syntax only' corrections of ungrammatical sentences containing HF CTVs, but more 'syntax + semantics' corrections of ungrammatical sentences containing LF CTVs. A 'syntax only' correction is clearly the more sophisticated correction type, since by only correcting the syntactic violation children were preserving the intended message of the test sentence. In contrast, when children were making 'syntax + semantics' corrections they were changing the message because they were substituting in a different main verb, most often a HF verb. These results provide further support for the idea that knowledge of complementation is intimately tied to CTV frequency.

The lack of age by frequency interactions in the present study suggest that frequency is an important factor in the acquisition and subsequent processing of grammatical forms. There is a rich literature in adult psycholinguistics that has identified a central role for frequency information in parsing (Hare et al. 2003; MacDonald, Pearlmutter, & Seidenberg 1994, Trueswell 1996). The role of frequency is not as widely studied in acquisition research; however, a growing body of literature is now exploring the role of probabilistic information in language acquisition (Kidd 2003, Snedeker & Trueswell 2004). These approaches, in general, build upon much of the early work investigating children's sentence processing strategies (e.g., MacWhinney & Bates 1989). The focus for these researchers has been elucidating the mechanisms of construal (and by implication, acquisition). Within the usage-based approach to the acquisition of grammar (e.g., Goldberg 1995, Tomasello 2003), frequency is argued to play an important role in the conventionalisation and abstraction of grammatical forms. The central focus of the usage-based approach is the nature of children's representations. Children are argued to make form-function mappings at the utterance level. Therefore, an important

issue to address to address is how frequency information might contribute to process of establishing such form-function links.

Diessel and Tomasello (2001) suggested that the main clause in children's formulaic sentential complements acts as a clausal operator that modifies the propositional content of the subordinate clause. For instance, 'I think' in *I think she is riding away on the horse* functions like an epistemic modal, marking the degree of certainty the child has over the action being described. That is, the main clause is a marker of illocutionary force. One finding from the current study suggests children's use of the parenthetical main clause may be even less sophisticated than Diessel and Tomasello propose. By far the most common verb children substituted into their repetitions was *think*. The pervasive and often indiscriminate use of *think* in children's repetitions suggests that the function is not initially a marker of illocutionary force, but an unanalysed attention getter signifying the child's intention to speak.

This suggestion finds some continuity with research characterising children's early multi-word speech. Lieven, Pine, and Baldwin (1997) and Lieven, Pine, and Rowland (1998) suggest that children's early multi-word utterances can be categorised by 'slot and frame' patterns organised around high frequency lexical items. Initially these frames are argued to be grounded by unanalysed units, such as *get it* in *get it daddy*, or *where's* in *where's mommy gone?* On this approach, the use of formulaic complement constructions is a more sophisticated version of such item-based learning, suggesting that the beginnings of sentential recursion is an outcome of the same or similar strategy used in early language development. Children can break into the system by using forms (e.g., *I think, I hope,…*) that initially have only cursory pragmatic function. These initial form-function mappings provide the foundation from which children can flesh out the appropriate linguistic facts about complementation in English. Future research conducted at the discourse level will help identify the true pragmatic function of children's parenthetical main clauses in sentential complement constructions.

The process of abstracting over lexically-based knowledge and converging on the adult model is an important topic for future research. Construction-based approaches represent knowledge in the form of inheritance networks; the challenge to researchers working within this tradition is to explain how children establish network links between isolated construction 'islands'. The results from the present study suggest that frequency information could play an important facilitative role in the abstraction process. Appeals to frequency are hollow unless one can identify the grain at which probabilities are computed over. The results from the present study suggest that the frequency with which a CTV appears in a complement clause construction *relative* to the verb's total use predicts children's patterns of repetitions.[3] However, additional mechanisms apart from frequency must be imported into the system in order to explain the process of abstraction. Tomasello (2003) suggests that children do so by making analogies over lexically-specific constructions. Since children are argued to perform a distributional analysis of their input, it can be assumed that they will identify the basis for analogising once they identify what each lexically-specified construction has in common: a sentential constituent in the argument slot of a CTV. Of course, a central tenet of the construction approach is that form-meaning correspondences are made at the utterance level. Thus the pervasiveness of formulaic sentential complements, which are similar in surface structure (i.e., Formulaic Main Clause + Complement), provide a simplified model over which children can begin to make analogies for this construction type.

There are also other cues to which children could attend in order to make analogies. For instance, as many CTVs encode mental states, analogies could be made on the basis of verb semantics, a suggestion for which there is empirical support in the adult sentence processing literature (Hare et al. 2003; see also Pinker 1989). Another potential cue is the prosodic contour of the utterance, although there is little developmental research on children's use of prosody in acquisition. The success of the construction-based approach will depend on its ability to provide an adequate explanation of the process by which children abstract over

[3] Interestingly, when the same analyses were performed using raw frequency (i.e., total number of times a given verb appears with a complement clause) as a predictor, the frequency effect disappears.

locally-bound lexically-specific knowledge to converge onto the adult state. At present there is little research that bears on this important issue.

5. Conclusion

The present paper has shown that children's acquisition of sentential complement constructions is intimately tied to CTV frequency. Children in every age group tested provided more exact repetitions of sentences containing HF CTVs than sentences containing LF CTVs. Furthermore, children were better able to correct syntactic violations in ungrammatical test sentences when those sentences contained a HF CTV. When ungrammatical sentences contained a LF CTV, children were more likely to 'trade up' to higher frequency CTVs, providing corrections where they substituted the LF CTV for a HF CTV. Overall, these results suggest that the availability of frequent exemplars provide the basis upon which children can construct a grammar.

Acknowledgements

We would like to thank the staff and pupils at Christies Day Nursery, Our Lady Lancashire Primary School, and Fiddlers Lane Primary School for participating in this study. Thanks also to Anna Roby for testing the final few children. This research was supported by a postdoctoral fellowship from the Max Planck Institute for Evolutionary Anthropology to the first author. Correspondence should be directed to: Evan Kidd , Max Planck Child Study Centre, School of Psychological Sciences, University of Manchester, Oxford Road M13 9PL, Manchester, UK; Email: evan.j.kidd@manchester.ac.uk

References

Bloom, L. (1991). *Language development from two to three*. Cambridge, UK: Cambridge University Press.

Bloom, L., Rispoli, M., Gartner, B., & Hafitz, J. (1989). Acquisition of complementation. *Journal of Child Language, 16,* 101 – 120.

Cairns, H. S., McDaniel, D., Hsu, J., & Rapp, (1994). A longitudinal study of principles of control and pronominal reference in child English. *Language, 70,* 260 – 288.

de Villiers, J., & de Villiers, P. (1999). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 189–226). Hove, UK: Psychology Press.

de Villiers, J., Roeper, T., & Vainikka, A. (1990). The acquisition of long-distance rules. In L. Frazier & J. de Villiers (Eds.), *Language processing and language acquisition* (pp. 257 – 297). Dordrecht: Kluwer.

Diessel, H., & Tomasello, M. (2001). The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics, 12*, 97 – 141.

Goldberg, A. E. (1995). *Constructions.* Chicago: University of Chicago Press.

Hare, M., MacRae, K., & Elman, J. (2003). Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language, 48,* 281–303.

Hare, M., MacRae, K., & Elman, J. (2003). Admitting that verb sense into corpus analyses makes sense. *Language and Cognitive Processes, 19,* 181 – 224.

Hsu, J., Cairns, H. S., Eisenberg, S., & Schlisselberg, (1989). Control and coreference in early child language. *Journal of Child Language, 16,* 599 – 622.

Kidd, E. (2003). An investigation of children's sentence processing: A developmental perspective. Unpublished PhD dissertation, La Trobe University.

Kidd, E., Lieven., E. V., & Tomasello, M. (in prep.). Lexical frequency effects in the acquisition of sentential complements.

Lieven, E. V., Pine, J., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24,* 187 – 220.

Lieven, E. V., Pine, J., & Rowland, C. (1998). Comparing different models of the development of the English verb category. *Linguistics, 36,* 807 – 830.

Limber, J. (1973). The genesis of complex sentences. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 169 – 185). New York, NY: Academic Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101,* 676 – 703.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.

MacWhinney, B., & Bates, E. (Eds.) (1989). *The crosslinguistic study of sentence processing.* Cambridge, UK: Cambridge University Press.

McCawley, J. (1988). *The syntactic phenomena of English* (Vols. 1 & 2). Chicago: University of Chicago Press.

Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure.* Cambridge, MA: MIT Press.

Sherman, J. C., & Lust, B. (1986). Syntactic and lexical constraints in the acquisition of control in complement sentences. In B. Lust (Ed.), *Studies in the acquisition of anaphora*, (pp. 279 – 308). Dordrecht: Reidel.

Snedeker, J., & Trueswell, J. (2004). The developing constraints on parsing decisions: The role of lexical-biases and referential scenes in child and adult sentence processing. *Cognitive Psychology*.

Thompson, S. A., & Mulac, A. (1991). A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In E. C. Truscott & B. Heine (Eds.), *Approaches to grammaticalization* (Vol. 2) (pp. 313 – 329). Amsterdam: John Benjamins.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition.* Cambridge, MA: Harvard University Press.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language, 35,* 566 – 585.