
ELSEVIER
Volume 104, Number 1, July 2007
ISSN 0010-0277
pp. 1–162
COGNITION

Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Cognition 104 (2007) 106–134
COGNITION
www.elsevier.com/locate/COGNIT

Explaining errors in children's questions ☆

Caroline F. Rowland *

School of Psychology, University of Liverpool, Bedford Street South, Liverpool, L69 7ZA, UK

Received 4 October 2005; revised 22 May 2006; accepted 25 May 2006

Abstract

The ability to explain the occurrence of errors in children's speech is an essential component of successful theories of language acquisition. The present study tested some generativist and constructivist predictions about error on the questions produced by ten English-learning children between 2 and 5 years of age. The analyses demonstrated that, as predicted by some generativist theories [e.g. Santelmann, L., Berk, S., Austin, J., Somashekar, S. & Lust. B. (2002). Continuity and development in the acquisition of inversion in yes/no questions: dissociating movement and inflection, *Journal of Child Language*, *29*, 813–842], questions with auxiliary DO attracted higher error rates than those with modal auxiliaries. However, in wh-questions, questions with modals and DO attracted equally high error rates, and these findings could not be explained in terms of problems forming questions with why or negated auxiliaries. It was concluded that the data might be better explained in terms of a constructivist account that suggests that entrenched item-based constructions may be protected from error in children's speech, and that errors occur when children resort to other operations to produce questions [e.g. Dąbrowska, E. (2000). From formula to schema: the acquisition of English questions. *Cognitive Liguistics*, *11*, 83–102; Rowland, C. F. & Pine, J. M. (2000). Subject-auxiliary inversion errors and wh-question acquisition: What children do know? *Journal of Child Language*, *27*, 157–181; Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press]. However, further work on constructivist theory development is required to allow researchers to make predictions about the nature of these operations.

☆ This manuscript was accepted under the editorship of Jacques Mehler.
* Tel.: +44 151 794 1120; fax: +44 (151) 794 2945.
E-mail address: crowland@liverpool.ac.uk

0010-0277/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2006.05.011

Keywords: Errors; Questions; Language acquisition; Usage-based model; Language-specific knowledge

1. Introduction

Although a substantial proportion of children's utterances are correct from the beginning of multi-word speech, there are significant pockets of systematic error in certain parts of the system. Many of these errors have been intensively studied and have been influential in theory development (e.g. case errors and past tense over-generalisation errors; Maratsos, 2000; Marcus, 1995; Rispoli, 1998, 1999; Schütze, 1997) but the errors that children make in learning to form questions have attracted less interest in recent years. This may be because early descriptions of children's questions over-estimated the prevalence of such errors, but it still remains true that children make a significant number of errors in their early questions.

In English, most errors in children's questions are errors of auxiliary omission (e.g. *what he doing?* instead of *what is he doing?*, see Rowland, Pine, Lieven, & Theakston, 2005). However, children also make a number of commission errors. The most well known of these are subject-auxiliary inversion errors, which are perhaps the most common word order error in English acquisition and indicate that children struggle to master the rules governing auxiliary placement in questions (producing for example *why he can go?* instead of *why can he go?*). Children also make a substantial number of other errors including double auxiliary errors (e.g. *what can he can do?*), double tensing errors (e.g. *what does he likes?*) and raising errors (e.g. *what he likes?*; see e.g. Hurford, 1975; Kuczaj, 1976a; Radford, 1990; Rowland et al., 2005). There are also reports of errors indicating that children fail to apply correct nominative case to the syntactic subject or fail to apply subject-auxiliary agreement correctly (e.g. *where can me go?, where does the boys go?*; see Radford, 1990). Finally, although most of the work has focussed on wh-questions, similar types of error can be found in children's early yes–no questions (Derwing & Smyth, 1988; Klee, 1985; Valian, Lasser, & Mandelbaum, 1992).

Two key factors about these errors make them an important test case for the investigation of current theories of language acquisition. First, they seem to demonstrate a failure on the part of the child to master the grammatical rules governing question formation, yet they occur at the same time as correct questions (see e.g. Rowland & Pine, 2000; Stromswold, 1990). It is clear that children who can demonstrate mastery of rules such as movement and case and agreement marking in some questions are not freely applying their knowledge across all questions. Second, the patterning of correct use and error, and the nature of the errors produced, indicate areas in which children have particular difficulty mastering correct production. Successful theories of language acquisition must be able to account for these factors. The aim of this paper is to investigate the pattern of errors in ten children's naturalistic data to test the solutions proposed by some current accounts of question formation.

2. Modern generativist solutions

The central theme of many recent accounts of question acquisition is that movement is an innate principle present in Universal Grammar (UG). According to this approach, relatively little exposure to the English language is required for children to register that English is a language that allows movement operations such as subject-auxiliary inversion. As a result, children quickly grasp that in wh-questions, tense and agreement must be obligatorily marked on the auxiliary, which is placed in the pre-subject (inverted) position (e.g. *where are you going?*). They also quickly learn that such inversion is usual (although optional) in English yes–no questions (e.g. *are you going to the park?*).

The presence of movement in UG explains why children are capable of producing correct questions from very early on in the acquisition process. However, theorists must also explain why children make a number of systematic errors in their early questions. On the face of it, the presence of errors in child grammars "challenges a Strong Continuity Hypothesis of UG as a model of the child's language faculty" (Santelmann, Berk, Austin, Somashekar, & Lust, 2002, p. 815). Errors are hard to explain in terms of a theory which posits very early parameter setting with little scope for extended learning or maturation. However, errors can be explained in ways that are compatible with the idea of movement as part of UG. Broadly, three such explanations have been suggested. One is that errors occur because of performance limits on production (cf. Bloom's explanation of subject omission, 1990). Another is that errors occur because some aspects of UG (e.g. functional categories) have yet to mature (e.g. Radford, 1990). However, both these approaches have difficulty explaining the precise pattern of error in question acquisition. For example, the performance limitation theory makes predictions about the patterning of auxiliary omission across different forms of the same auxiliary subtype that are not supported by the data (see Theakston, Lieven, Pine, & Rowland, 2005). The maturation theory has problems explaining the co-occurrence of correct use and error, as well as the fact that cross-linguistic studies provide clear evidence for the presence of structures that rely on functional categories from the beginning of multi-word speech (see Lust, 1999, for a synopsis and further arguments).

2.1. Language-specific knowledge: A solution

A more successful explanation is that children make errors because they have to coordinate their innate knowledge of movement with the task of learning when and how movement applies in their particular language. On this view, the problem lies outside UG, and more in the realm of interpreting innate knowledge in the light of the idiosyncrasies of individual languages. One suggestion is that subject-auxiliary inversion errors stem from problems analyzing the correct placement and role of the wh-word. For example, DeVilliers (1991) argues that children initially misanalyse the wh-word as being in topic position of the inflectional phrase, so that inversion only becomes possible after the correct reanalysis of the word into the specifier position of the complementizer phrase (see Valian et al., 1992, for a similar idea). However, errors do not pattern purely in terms of the wh-word and many studies have reported a role for other elements such as auxiliary choice or the presence of negation (e.g. Bellugi, 1971; Kuczaj & Maratsos, 1983; Labov & Labov, 1978; Rowland & Pine, 2000, 2003).

A more recent idea is that children make errors while mastering how question formation rules apply to particular auxiliaries. Santelmann et al. (2002; see also Stromswold, 1990) present such a theory. They, too, start from the generativist assumption that "knowledge of movement allowing inversion is a fundamental component of Universal Grammar and that this UG is continuously available to the child" (Santelmann et al., p. 815). Errors occur because the English learning child still has to learn to integrate the formal features of inflection (tense, number and person) with their knowledge of phrase structure rules and their knowledge of how inversion applies in their language. Although all these components may be specified in UG, how they are to be integrated and how they are realized through the lexicon and morphology depend on the nature of the language to be learnt. Difficulties in question formation stem from integrating the innate components, including inversion, in the correct language-specific way.

The pattern of error is explained in terms of the level of complexity required to integrate UG with the language-specific rules governing question formation. In English, the formation of yes–no questions with auxiliary BE, HAVE[1] and modal auxiliaries (e.g. *can, can't, won't*) differs from the formation of the corresponding declarative in only one major way - the requirement for inversion (e.g. *he can eat it* vs. *can he eat it?*). All other requirements (e.g. the marking of tense) are identical. Thus, since the inversion component is specified in UG, there should be equal numbers of errors in declarative and yes–no questions with these auxiliaries. This prediction is upheld by the results of Santelmann et al.'s (2002) matched sentence imitation experiment and is consistent with much of the data from naturalistic studies, which report very few errors of commission in either declaratives or yes–no questions with forms of auxiliary BE, HAVE and modal auxiliaries (e.g. Bellugi, 1971; Guasti, Thornton, & Wexler, 1995; Stromswold, 1990).

The situation is different for questions requiring copula BE and auxiliary DO. The requirement to insert auxiliary DO into a question to carry tense and agreement is the "jerry-rigged result" (Stromswold, 1990, p. 246) of the property of English that requires the raising of tense and agreement but prohibits the raising of main verbs[2]. To produce a correct question with DO, children must learn to integrate UG with this English-specific restriction on movement, a complex task that will lead to error. Similar problems occur with copula BE. Copula BE is the only English main verb that commonly undergoes inversion (e.g. *is he hot?* is grammatical, *do he be hot?* is not)[3]. Children struggle to integrate this fact with the evidence that, as a general rule,

[1] Capital letters indicate reference to the auxiliary subtype. For example auxiliary BE refers to all forms of the auxiliary (*am*, *is*, *are*, *was*, *were*).

[2] In English, questions with main verb inversion are ungrammatical, unlike in many other languages (e.g. *walks he to the park?*).

[3] Main verb HAVE can also undergo inversion in English (e.g. *have you a penny?*) but these questions are very infrequent in both adult and child speech.

"main verbs do not raise in their grammar for their language" (Santelmann et al., 2002, p. 837). These questions will also be more error prone compared to their declarative counterparts.

The predictions about questions with copula BE and auxiliary DO are borne out by Santelmann et al.'s (2002) own study and the literature on naturalistic data. Santelmann et al., reported that, as they predicted, children performed significantly worse in their attempts to imitate yes–no questions with DO and copula BE than in their attempts to imitate matched declaratives. The naturalistic data also seems to show that children make errors with these auxiliaries in yes–no questions but not (or to a lesser extent) in declaratives (e.g. Stromswold, 1990).

In sum, the literature on yes–no questions and declaratives patterns according to the predictions of the theory. However, the picture for wh-questions is less clear. In support of Santelmann et al. (2002), to our knowledge, there are no reports of substantial numbers of errors of commission with either declarative or wh-question forms containing auxiliary BE or HAVE, and there are a number of studies that report high rates of errors in both yes–no and wh-questions with DO (e.g. Rowland et al., 2005; Stromswold, 1990). However, the data on wh-questions with modals is more problematic as both Van Valin (2002) and Rowland et al. (2005) have reported high rates of error. Rowland et al. (2005) have demonstrated that although error rates in wh-questions with DO are higher than in wh-questions with other primary auxiliaries, error rates are in fact equally high in wh-questions with modal auxiliaries. Van Valin (2002) has argued that children make substantial numbers of errors with auxiliaries that are not overtly marked for tense (e.g. modals), and are much more accurate in the placement of overtly tensed auxiliaries such as *does* and *did*. Thus, contrary to Santelmann et al.'s prediction, both Rowland's and Van Valin's data suggest that wh-questions with modals may attract high rates of error. However, no study has explicitly compared error rates across yes–no questions and wh-questions. The first aim of the present study was to use naturalistic data from ten English-learning children to investigate error rates in yes–no and wh-questions with modal auxiliaries and auxiliary DO in order to ascertain whether alternative explanations are required to account for the pattern of error in wh-questions.

3. Usage-based theory

The constructivist theory known as the usage-based theory (Tomasello, 2000, 2003) proposes that children are not acquiring a generative grammar but are learning constructions, which are seen as the basic units of language (Croft, 2001; Goldberg, 1995; Langacker, 1991). Constructions are structural units that express particular meanings (e.g. the transitive, the resultative) and, because they are defined at least partly by the functions they express, they can be learnt via the powerful species specific learning mechanisms that children possess (e.g. pattern-finding skills, intention reading; Tomasello, 2003).

The learning of grammatical constructions involves two central processes. First, the child builds up an inventory of lexically based frames (i.e. item-based constructions) that have occurred with high frequency in the input, and that express particular meanings. These frames are not merely rote-learned forms but are semi-abstract constructions that are associated with a particular communicative function. For example, the child's early intransitive utterances may be restricted to a number of lexically based frames such as *I'm V-ing* or *you're V-ing* (see Wilson, 2003). Thus, children's early grammatical knowledge consists "not of an abstract and coherent formal grammar but rather of a loosely organized inventory of item-based construction islands" (Cameron-Faulkner, Lieven, & Tomasello, 2003, p. 866).

The second process of acquisition involves generalization mechanisms that derive more complex abstract constructions from the simpler lexical constructions. In generalizing, the child utilizes basic learning mechanisms such as analogy and schematization to abstract across previously learnt lexical frames that share semantic, pragmatic and/or distributional information. In this way, more abstract linguistic schemas develop (see Tomasello, 2003). For example, the abstraction of commonalities across lexically specific frames such as *I'm V-ing* and *you're V-ing* will result in the development of semi-abstract constructions such as *NP Aux V-ing* and eventually to the acquisition of the adult like intransitive construction (see Wilson, 2003).

The assumptions behind this theory are that children can learn lexical patterns, can reproduce them accurately and can use them as the basis of later generalization. These are based on three separate strands of evidence. First, there is evidence that high frequency formulaic patterns are common in English usage (e.g. Sinclair, 1991; Van Lancker-Sidtis & Rallon, 2004). Second, it is clear that children and adults can use frequently occurring surface distributional cues to parse the language accurately; for example, to discover word boundaries (e.g. Saffran, Aslin, & Newport, 1996), gender like subclasses (e.g. Brooks, Braine, Catalano, Brody, & Sudhalter, 1993), and even syntactic categories (e.g. Braine, 1987; Wilson, 2000; Wilson, Gerken, & Nicol, 2000). Third, the presence of a converging invariant cue is an important factor in this learning as there is often no evidence of accurate categorisation when converging cues are absent. In particular, Mintz (2002, see also Mintz, Newport & Bever, 2002) reported that adults can acquire grammatical categories in an artificial language based on frequently occurring frames (frames such as "*the X is* ..." where X could be one of several nouns)[4]. It is important to note that much more than rote learning is involved in this process. The lexical frame consists not only of certain lexical units (e.g. *where's*, *what can*) but also of a subset of lexical forms that can combine with these units. For example, a child who has learnt how to form *where has* questions is also learning what can and cannot combine with *where has* (e.g. *gone* rather than *go*, *he* rather than *him*) and the relative position of these forms (*he gone* rather than *gone he*).

Within this theory, children's questions are not derived by the manipulation of structures using operations such as MOVE or invert. Instead, the word order is

[4] The extend to which statistical learning can explain the acquisition of language is debated (see e.g. Peña, Bonatti, Nespor, & Mehler, 2002). However, it is the role of statistical learning in the language acquisition process, not the evidence that children and adults are capable of acquiring such frames and memorising them accurately, that is controversial.

specified directly in the lexically based frame; early questions will consist of lexical frames such as *what's the NP* and *where's the NP V-ing* (see Dąbrowska, 2000; Dąbrowska & Lieven, 2005; Rowland & Pine, 2000, 2003). Importantly, the use of these frames has implications for error. The use of formulae in both child and adult speech reduces the working memory demands of the processing task, resulting in fewer errors, less hesitation phenomena and faster reaction times because "it is easier for us to look something up than to compute it" (Bresnan, 1999; also see Ellis, 2002; Kuiper, 1996). In addition, the use of the frame is likely to restrict the opportunity for error. For example, a child who wishes to ask the question *where's the doggie going?* is much less likely to make an error if she has already heard a substantial number of *where's X?*[5] questions in her input and has learnt that to form these types of questions you must conjoin a *where's* pivot with a particular set of nouns and verb types. The child who can use a frame to form a question is much less likely to make a grammatical error in such a question than a child without a frame.

Thus, like generativism, the usage-based theory predicts low rates of error overall because the child is capable of producing a variety of correct questions by the application of entrenched lexical frames. Since both theories predict relatively low rates of error, it is not possible to distinguish between them by testing predictions about overall error rates, only by testing predictions about the pattern of error. Whereas theories such as that of Santelmann et al. (2002) predict that errors will occur in questions that require language-specific knowledge, the usage-based theory focuses on a failure to acquire lexically specific knowledge and on initially conservative and limited generalization as the source of error. A child who knows how to form one particular question type (e.g. *what can* + *X*), will not necessarily generalize this knowledge to other question types immediately (e.g. *what does* + *X*). This means that if the child wishes to ask a question that expresses a particular meaning, but has no lexical template appropriate to express this meaning, she will be required to apply other generalization strategies such as combining frames with interrogative fragments (e.g. *what* + *can he do* (*it*) = *what can he do?*) or retrieving a related but sometimes inappropriate frame (e.g. *what's* + *I do* = *what's I do?*). Sometimes these strategies will result in the formation of correct questions. However, since the child is not guided by a lexical frame that can reduce the demands of the processing task, or by complex abstract knowledge of question formation rules, the theory predicts that questions produced in this way are likely to contain more grammatical errors than those formed by the application of a lexical frame. The second aim of the present work was to test this prediction on the naturalistic data of ten language-learning English children.

Although the use of entrenched frames for question formation will help protect from error, errors within frame-based questions are not ruled out by the theory because the children must creatively combine an appropriate pivot with the appropriate elements in the variable slot. However, if a child makes an error in a frame-based question, that error should be more likely to occur in the variable slot than

[5] *X* is used to refer to the material that can occupy the variable slot in the frame.

in the frame; as deciding which elements to place in the variable slot should be more difficult than simply retrieving the appropriate pivot to express the intended meaning. Any errors produced by children in frame-based questions should tend to occur in the variable slot (e.g. *what did he did?*) rather than the pivot part of the frame (e.g. *what do he do?*). The present paper also tested this prediction.

3.1. Defining lexical frames

The usage-based theory hinges on the ability of the acquisition mechanism to extract semantic-distributional patterns from the input. The composition of the input, in terms of variability and regularity, is thus central to its predictions. Drawing on work on the type of information available from the statistical dependencies in the speech stream (e.g. Gómez, 2002; Onnis, Monaghan, Christiansen, & Chater, 2004; Onnis, Christiansen, Chater, & Gómez, under review), and following theorists such as Bybee (see e.g. Bybee & Hopper, 2001), a combination of high token frequency and low type frequency is said to lead to the entrenchment of word combinations. Lexical items that tend to occur together frequently in the child's input are likely to become entrenched as a unit (or pivot). For example, the subject pronoun *I* and the contracted auxiliary form *am* occur together in English speech extremely frequently. Thus, they are likely to be learnt as a unit – *I'm*[6]. Conversely, low token frequency and high type frequency are more likely to lead to productivity. For example, if *I'm* is followed by a range of different verbs, the child will learn to combine it with a range of verbs – resulting in an *I'm V-ing* lexical frame.

Thus, lexical frames will often consist of an entrenched unit (pivot) combined with a number of lexical items (variable). In declaratives, for example, children may learn pronoun based frames such as *he's V-ing*, *you're V-ing* and *I'm V-ing*, because closed class forms such as *he*, *you* and *I* occur frequently as subjects, they occur most often with contracted copula BE, and they occur with a number of variable verbs. In wh-questions, the most likely pivot is the wh-word + auxiliary combination. This is because there are only a small number of possible wh-word and auxiliaries that can be combined (and only a few of these will occur with high frequency in the input), which means that there will be much less variation within the wh-word + auxiliary unit than between the unit and the remainder of the question (see Rowland & Pine, 2000 for an expanded rationale for this decision). For yes–no questions, similar ideas suggest that the auxiliary + subject combination will be a pivot. First, the small number of auxiliary forms makes the auxiliary a likely pivot. Second, there is evidence that subject positions are filled by a limited range of pronominal and nominal forms, which suggests that subjects may function as pivots in children's speech. Third, Cameron-Faulkner et al. (2003) have demonstrated an apparent lack of lexical variation in the auxiliary + subject combinations that occurred in maternal input; with nine core auxiliary + subject combinations

[6] As Wilson (2003) points out, this does not mean that *I'm* is equivalent to *I* as some researchers have suggested. Although *I'm* will function as a unit, the child will have learnt the functional differences between *I'm* and *I*.

accounting for 71% of yes–no questions addressed to the children they studied. In summary, the present paper defines lexical frames as pivot + variable patterns. For wh-questions, the pivot of the frame is defined as the wh-word + auxiliary combination. For yes–no questions, the pivot is defined as the auxiliary + subject combination.

To summarize, the first analyses investigated the idea that the language acquisition mechanism works by mapping detailed innate knowledge of abstract linguistic structure onto the language being heard, with errors occurring where the mapping is less straightforward (e.g. Santelmann et al., 2002). The analyses tested the prediction that children make errors when they have to apply the rules governing the use of auxiliary DO by investigating whether auxiliary DO attracts higher error rates than modal auxiliaries in both yes–no and wh-questions. The second analyses assessed the constructivist idea that the acquisition mechanism is a semantic-distributional analyzer, which builds up an inventory of constructions and the functions they express, builds links between constructions that share semantic-distributional properties and abstracts from more lexical to more general constructions. Under this account, questions that can be produced with entrenched lexical frames will be protected from error, and so errors will occur more often in questions that must be formed using generalization strategies. First, we tested the prediction that questions that the child can produce using previously learnt entrenched frames will be less likely to contain grammatical errors than those the child has to produce by other, perhaps more complex, operations. Second, we tested the idea that the child is learning frames which are of high frequency in the input, which predicts significantly lower error rates in questions based on frames of high input frequency. Third, we tested the prediction that errors in frame-based questions should tend to occur in the variable slot, not the pivot part, of the frame.

4. Method

4.1. Participants

The participants were ten children, nine whose corpora are available on the CHILDES database (MacWhinney, 2000) and one whose data were collected by the author (Lara). The children from the CHILDES database were Anne, Aran and Becky from the Manchester corpus (Theakston, Lieven, Pine, & Rowland, 2001), Adam and Sarah from the Brown corpus (Brown, 1973), Nina from the Suppes corpus (Suppes, 1974), Peter from the Bloom corpus (Bloom, Lightbown, & Hood, 1975), Abe from the Kuczaj corpus (Kuczaj, 1976b), and Ross from the MacWhinney corpus (MacWhinney, 2000). The children were chosen because they all produced a large number of yes–no and wh-questions with auxiliary DO and modal auxiliaries. Further details about all the children except Lara can be found on the CHILDES database.

Lara was the first-born monolingual English daughter of two university graduates, who was born and brought up in Nottinghamshire, England. Lara was audio-recorded for approximately one to two hours every week. During recording,

Table 1
Age and MLU of children during the error period, details of the transcripts used and the number of questions produced

Name	Age range	MLU at start and end	Transcripts in error period	Number of questions
Abe	2;6.29–5;0.4	5.55–8.40[b]	19–209	1712
Adam	2;6.3–4;10.23	2.69–5.16	7–55[a]	1673
Anne	2;1.20–2;7.28	2.70–3.35	11–29	151
Aran	2;4.27–2;10.7	3.05–3.66	17–32	69
Becky	2;6.5–2;11.15	3.38–3.20	14–34[a]	653
Lara	2;7.23–3;3.22	2.95–3.46	1–66	2576
Nina	2;9.26–3;2.12	3.51–4.91	34–51	302
Peter	2;8.14–3.1.21	3.48–3.38[c]	17–20[a]	109
Ross	2;7.18–4;4.4	4.96–5.85	22–51	643
Sarah	3;6.16–5;1.6	3.12–3.51	65–139[a]	539

[a] Indicates that errors were produced on the last recording.
[b] Abe's MLU is large because of the way the data are transcribed, with more than one utterance often transcribed on one line.
[c] Indicates that MLU was calculated in words because morphological coding was not available.

Lara engaged in everyday play activities with her regular caregivers (mother, father, grandmothers and grandfather). The recorded data were supplemented by data from a written diary of the wh-questions that Lara produced from age 2;7.21 to 3;3.30. Lara's caregivers were the diary-keepers. No notes were made when the child was at nursery (for 2 part days a week) so it is estimated that the diary contains approximately 80% of the wh-questions that were produced by Lara during this period.

All the children had language acquisition within the normal range and none were identified as having language problems. A number of other children were considered but were excluded because they did not produce enough questions of the relevant type. Data were included only from the period during which the children made errors of commission in questions with DO or modal auxiliaries. This was to reduce the chance of development and later correct use masking any effects. The error period began at the first transcript in which the child made an error with a DO or modal question form and ended after the last recorded transcript with such an error. The only exception was Ross, who produced 26 errors between the age of 2;7 and 4;4 but then also produced one error – *does donuts also have TV* – at 6;7. The decision was taken not to include the transcripts from 4;4 to 6;7 because of the 2-year gap in error production. Table 1 details the children's age and MLU range corresponding to this period, the numbers of the transcripts that were used in the analysis and the number of questions produced.

4.2. Transcription

All data were orthographically transcribed. Further details about transcription can be found in the CHILDES database manual (MacWhinney, 2000). Lara's data were transcribed according to the conventions used for the Manchester corpus children.

4.3. Corpora

All yes–no questions and object and adjunct wh-questions that contained a form of auxiliary DO or a modal auxiliary were extracted from the transcripts. Because of the importance of clearly distinguishing questions with DO from those with modals, questions with omitted auxiliaries were excluded. In questions with omitted auxiliaries it is difficult to tell whether the omitted element is a form of DO or a modal (e.g. *what you do?* could be an attempt to produce *what do you do?* or *what can you do?*).

The analysis was restricted to matrix questions, excluding utterances such as echo questions and single word questions (e.g. *what?*). Any utterances with parts marked as unclear (e.g. marked with a question mark or *xxx* according to CHAT convention) were also removed. Subject wh-questions were excluded because they do not require inversion. Non-inverted yes–no questions were also excluded because it is often difficult to determine from the transcript whether a true question was asked or whether the utterance had been given a question mark in error. Finally, only full questions were included (i.e. yes–no questions which had a subject, auxiliary and main verb present, and wh-questions with a wh-word, subject, auxiliary and main verb present).

4.4. Error coding

The questions produced by all eight children were coded as follows:

4.4.1. Correct questions

For wh-questions, the choice and placement of the wh-word, auxiliary, main verb and subject had to be correct. For yes–no questions, the choice and placement of auxiliary, subject and main verb had to be correct. Questions with omissions and errors not pertinent to the grammatical rules that apply specifically to questions (e.g. determiner omission) were included.

4.4.2. Double marking errors

These errors included doubling of the auxiliary (e.g. *where does he does go?*, *does he does go?*), errors in which tense and agreement were correct but were marked on both auxiliary and main verb (e.g. *where does he goes?*, *does he goes?*) and errors in which an auxiliary was present but tense and agreement were marked only on the main verb (e.g. *where do he goes?*, *do he goes?*). It also included errors with two different auxiliaries present (e.g. *does he can go?*).

4.4.3. Non-inversion errors

Subject-auxiliary inversion errors (e.g. *where he does go?*), which could only occur in wh-questions.

4.4.4. Agreement errors

Errors in which an auxiliary was present but did not agree with the subject (e.g. *where does you go?*, *do he go?*).

4.4.5. Case errors

Errors in which the subject had incorrect non-nominative case (e.g. *does her like it?*, *where does her go?*) or where the object had non-accusative case (e.g. *do you want he to sit down?*).

4.4.6. Auxiliary-verb mismatch errors

Errors where the form of the main verb and auxiliary were incompatible, suggesting that either the wrong auxiliary or the wrong verb form had been used (e.g. *does he going to the shops?*, *did you be there?*).

4.4.7. Miscellaneous

Errors that could not be categorized according to the scheme above or where the type of error could not be reliably identified (e.g. *what for do you need the sudocrem?*).

Information on coding reliability can be found in Rowland et al. (2005). The level of agreement between coders was 97.5%.

5. Results

The children produced a wide range of different question types. Six children used either five or six of the six possible forms of DO (*do*, *don't*, *does*, *doesn't*, *did*, *didn't*) and four children used four of the forms. Thirteen modals were used (*can*, *can't*, *could*, *couldn't*, *shall*, *should*, *shouldn't*, *will*, *won't*, *would*, *wouldn't*, *may*, *might*; mean number across children = 7.6, range = 4 - 12). *Can* was the most frequent, accounting for a mean of 60% of questions with modals (SD = 23.40). Other modals each accounted for between 0.02% and 12% of all modal questions on average.

Table 2 demonstrates the mean number of correct questions and errors, and the mean percentage of errors accounted for by each error type. Overall, most questions were correct, errors accounted for only 7% of the children's questions on average (SD = 3.98). Nearly half the errors produced were double marking errors, nearly a quarter were non-inversion errors and 14% were auxiliary-verb mismatch errors. There were extremely few errors that could not be categorised according to the coding scheme (0.31%).

Table 2
Mean number of correct questions and errors

	Correct	Errors					
		Double marking	Non-invert.	Agree.	Case	Auxiliary-verb mismatch	Misc.
Mean No. (SD)	799.40 (821.75)	17.80 (14.99)	14.00 (25.86)	4.50 (7.01)	1.10 (1.97)	5.70 (6.29)	0.20 (0.63)
Mean % of total errors (SD)		47.21 (20.17)	21.26 (16.49)	12.32 (11.83)	4.98 (8.77)	13.92 (11.89)	0.31 (0.97)

Table 3
Mean number of correct questions and errors and % of questions that were errors for questions with auxiliary DO and modal auxiliaries

	Questions with modal auxiliaries			Questions with DO		
	No. correct	No. errors	% Error	No. correct	No. errors	% Error
Mean (SD)	395.50 (480.76)	13.30 (21.83)	4.08 (3.68)	403.90 (382.10)	30.00 (26.80)	9.87 (6.05)

5.1. Testing Santelmann et al.'s (2002) theory

Analysis 1 examined the overall error rate for questions with DO and modal auxiliaries to investigate whether there were higher error rates in questions with DO than in those with modals. Table 3 reports the results. A repeated measures one-way ANOVA that compared error rates on the two types of auxiliary revealed that the children produced significantly fewer errors in questions with modal auxiliaries than in those with DO, $F_{(1,9)} = 10.97$, $p = 0.009$, $\eta_p^2 = .55$[7]. This confirmed the finding from the literature that questions with DO attract higher error rates than questions with modal auxiliaries. However, on average, the children produced 603 yes–no questions but only 240 wh-questions. Thus, the error rates overwhelmingly represent the children's performance with yes–no questions.

The second analysis investigated error rates in yes–no and wh-questions separately. Table 4 shows the mean number of questions and percentage of questions that were errors in yes–no and wh-questions with auxiliary DO and modal auxiliaries. There were very different patterns of error across yes–no and wh-questions. In particular, there were big differences in error rates between yes–no and wh-questions with modals. Yes–no questions with modals attracted the lowest mean error rate (1.17%) and wh-questions with modals attracted the highest mean error rate (18.56%).

A two-way repeated measure ANOVA was performed with auxiliary type and question type as independent variables and error rate as the dependent variable. The results revealed a main effect for question type, $F_{(1,9)} = 5.91$, $p = .04$, $\eta_p^2 = .40$, with wh-questions attracting higher error rates than yes–no questions. There was no main effect for auxiliary type, $F_{(1,9)} = .008$, $p = .93$, but there was a highly significant interaction between question and auxiliary type, $F_{(1,9)} = 8.75$, $p = .016$, $\eta_p^2 = .49$.

Post-hoc within-question t-tests revealed that, within yes–no questions, auxiliary DO attracted a significantly higher error rate than the modal auxiliaries; $t = 3.59$, $df = 9$, $p = .006$, $\eta_p^2 = .59$. For wh-questions, error rates were equally high for modals as for DO, $t = 1.82$, $df = 9$, $p = .10$. The prediction that DO will attract higher rates of error than modals holds only for yes–no questions not wh-questions.

Post-hoc within-auxiliary t-tests revealed no significant differences in the error rates for auxiliary DO across the two question types, $t = 1.50$, $df = 9$, $p = .17$. However, for modal auxiliaries, error rates were significantly higher in wh-questions than

[7] For all analyses, the significance value $p = .05$.

Table 4
Mean number of questions and % questions that were errors in yes–no and wh-questions with auxiliary DO and modal auxiliaries

	Yes–no questions				Wh-questions			
	With modal		With DO		With modal		With DO	
	No. qs.	% Errors	No. qs.	% Errors	No. qs.	% Errors	No. qs.	% Errors
Mean (SD)	354.60 (445.31)	1.17 (2.28)	248.50 (248.17)	11.29 (8.34)	54.20 (54.82)	18.56 (18.75)	185.40 (165.90)	8.91 (5.60)

Table 5
Mean number of questions and % questions that were errors in (a) wh-questions excluding why questions and (b) positive yes–no and wh-questions

	Yes–no questions				Wh-questions			
	With modal		With DO		With modal		With DO	
	No. qs.	% Errors	No. qs.	% Errors	No. qs.	% Errors	No. qs.	% Errors
(a) Mean with *why* questions excluded (SD)					45.90 (47.95)	16.19 (16.55)	142.20 (125.07)	7.23 (5.77)
(b) Mean for positive questions only (SD)	349.80 (444.19)	1.07 (1.96)	241.30 (240.18)	10.66 (8.60)	47.50 (48.99)	16.15 (16.69)	171.10 (150.93)	7.96 (5.80)

yes–no questions, $t = 2.79$, $df = 9$, $p = .02$, $\eta_p^2 = .46$. Thus, wh-questions with modal auxiliaries seem to attract unexpectedly high rates of error.

However, there are other explanations for high error rates in wh-questions with modals that might explain our data. First, there is some evidence that error rates, especially for errors of inversion, are particularly high in children's *why* questions (see e.g. Labov & Labov, 1978)[8]. Although problems with *why* cannot be the only reason for errors as errors also occur on other wh-words, if a large proportion of wh-questions with modals occur with the wh-word *why*, these errors may stem from problems forming *why* questions, not problems with modals per se.

If errors in wh-questions with modals are due to problems with *why*, the error rate should reduce substantially when *why* questions are excluded. Table 5 demonstrates the percentage of errors produced in wh-questions when *why* questions

[8] It is unclear why this may be. Some have argued that questions with *why* can be formed without movement of the wh-word into the specifier position of the complementizer phrase (spec CP), and it is the presence of the wh-word in spec CP that triggers inversion (DeVilliers, 1991). Another idea is that children might have been misled into thinking *why* behaves as it does in some Romance languages. Alternatively, it may be that questions requiring *why* are of relatively low frequency in the child's input, thus, the child has not yet learned how to form them properly (see Rowland & Pine, 2000, 2003).

were excluded and shows that the error rate in wh-questions with modals remained high. A two way repeated measures ANOVA comparing error rates across question types (yes–no questions/wh-questions excluding *why*) and auxiliary types (DO/modals) confirmed this. There was an interaction between auxiliary type and question type, $F_{(1,9)} = 9.98$, $p = .01$, $\eta_p^2 = .53$. There were no main effects for question type, $F_{(1,9)} = 4.24$, $p = .07$, or for auxiliary type, $F_{(1,9)} = .06$, $p = .82$. Post-hoc within-question *t*-tests showed no significant difference between error rates in wh-questions with DO and modal auxiliaries, $t = 1.90$, $df = 9$, $p = .09$. Post-hoc within-auxiliary *t*-tests showed that for modal auxiliaries, error rates were significantly higher in wh-questions than yes–no questions, $t = 2.72$, $df = 9$, $p = .02$, $\eta_p^2 = .45$. The effect was reversed for auxiliary DO, with significantly higher error rates in yes–no questions than wh-questions, $t = 3.01$, $df = 9$, $p = .02$, $\eta_p^2 = .50$.

Another possible explanation for the high error rate in wh-questions with modals is that children have difficulty forming negative questions (e.g. *what can't he do?*). Guasti et al. (1995) have reported the results of an elicitation experiment in which children produced auxiliary doubling errors (e.g. *can she can't go underneath?*) in negative, not positive, questions. Bellugi (1971) reported similar findings in spontaneous speech. Although not all commission errors can be explained in terms of problems forming negated questions, it may be that high error rates in wh-questions with modals can be attributed to problems with negation.

Table 5 demonstrates the mean error rates in yes–no and wh-questions with positive DO and modal auxiliaries only (i.e. when questions with negated auxiliaries were excluded). The table shows that the error rate in wh-questions with modals remained high. A two way repeated measures ANOVA comparing error rates across question and auxiliary type confirmed this. There was a main effect for question type, with wh-questions attracting significantly more errors than yes–no questions, $F_{(1,9)} = 5.93$, $p = 0.04$, $\eta_p^2 = .40$. There was no main effect for auxiliary type, $F_{(1,9)} = .08$, $p = .79$, but even with negative questions excluded, there was an interaction between auxiliary type and question type, $F_{(1,9)} = 7.86$, $p = 0.02$, $\eta_p^2 = .47$. Once again, post-hoc within-question *t*-tests showed that auxiliary DO attracted a significantly higher error rate than the modal auxiliaries in yes–no questions; $t = 3.27$, $df = 9$, $p = .01$, $\eta_p^2 = .54$; but for wh-questions, error rates were equally high for modals as for auxiliary DO,$t = 1.67$, $df = 9$, $p = .13$. Post-hoc within-auxiliary *t*-tests again revealed that, for auxiliary DO, there was no significant difference in error rate across the two question types, $t = 1.72$, $df = 9$, $p = .12$. However, for modal auxiliaries, error rates were significantly higher in wh-questions than yes–no questions, $t = 2.73$, $df = 9$, $p = .02$, $\eta_p^2 = .45$.

To summarize, the prediction that questions with DO will attract higher rates of error than questions with modals holds for yes–no questions only, not for wh-questions. In wh-questions, questions with modals and DO attracted equally high error rates, even when questions with *why* and negative auxiliaries were excluded.

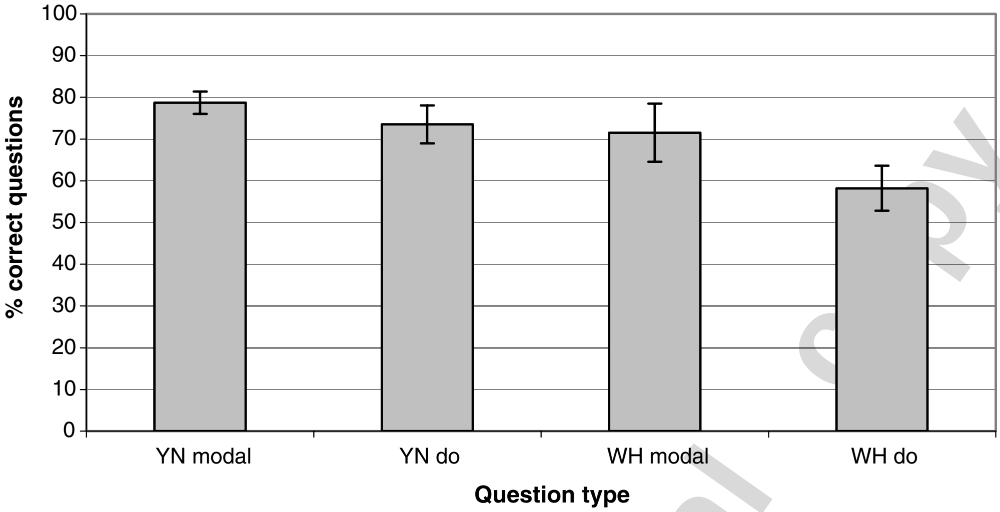

Fig. 1. Mean % correct questions (tokens) accounted for by the three most frequent frames for each question type (error bars indicate standard error).

5.2. Testing the usage based model

5.2.1. The importance of lexical frames

The first analysis assessed how many correct utterances could be explained in terms of limited lexical learning using the traditional measure of lexical specificity, which is to calculate how many utterances can be explained in terms of a small number of frequently occurring lexical frames, usually three (see e.g. Lieven, Pine, & Baldwin, 1997; Pine & Lieven, 1997). The three most frequent frames (wh-word + auxiliary + variable or auxiliary + subject + variable) produced by each child for each of the four question types were identified, and the percentage of the children's questions that could have been produced using these frames was assessed. Fig. 1 indicates the results, showing that the majority of correct questions could have been produced by the application of just three frequently occurring lexical frames.

5.2.2. Predicting error rates from frames

The usage-based model predicts that error rates will be lower in questions that could have been produced using a prior-learnt lexical frame, and that children will be more likely to make errors when the absence of these frames forces them to generate the question using other operations. Prior-learned frames were defined as wh-word + auxiliary + variable or auxiliary + subject + variable combinations that had been used to produce a correct question at least once by the child before the first error occurred (i.e. in the pre-error period). This decision was taken on the basis that the earliest learnt questions are those that are most likely to be entrenched as frames for subsequent production[9]. For example, a child was attributed a *what can* + *X*

[9] For this analysis, the frequent frames analysis used above could not be applied because using the same data to define frames and to test the effect of such frames introduces a statistical confound.

frame only if *what can* had appeared in at least one correct question before the error period began. Because few questions were produced in the pre-error period, it was not possible to use a more robust criterion. However, the measure reliably identified similar frames to the traditional measure used for the previous analysis. On average 8.1 frames were identified per child (SD = 5.88), 78% of these frames also occurred in the error period (and thus contributed to the analysis), and 76% of these were identical to those used for the frequent frames analysis above.

Three children produced no questions with auxiliary DO or modal auxiliaries before the first error – Abe, Adam and Lara. For these children, frames were defined using data from the early stages of the error period. Abe produced only five errors before age 2;10.7 so his data from age 2;4.24 to 2;10.7 were used to define frames. Adam produced only two errors before 3;0.10: *what did you did?* (produced once) and *what do you doing?* (produced six times). Thus, data from this period were used to determine the frames Adam had learnt. For Lara, data were used from 2;7.23 to 2;7.26 (only four errors were produced in this period). For all three children, both correct questions and errors produced during the frame definition period were excluded from the subsequent analysis[10].

The correct questions produced by the children during the error period were then divided into questions that could have been produced correctly by the application of an entrenched frame (frame based questions) and those that could not have been generated in this way (non-frame based questions). For example, Sarah's correct question *can it bounce?* was coded as non-frame based because she was not attributed with a *can it* + *X* frame (i.e. she did not produce any *can it* questions before the error period).

The errors produced by the children were similarly divided into two sets. The first set contained errors that could have been produced correctly using an entrenched frame but were not. These were errors that were not consistent with the prediction that children only make errors when they do not have a frame on which to base their question. For example, the target of Adam's error, *does you tear dese off dese papers?* (*do you tear dese off dese papers?*) could have been produced correctly by the application of his previously learnt *do you* + *X* frame. The second set consisted of errors that could not have been produced correctly using an entrenched frame. These were consistent with the idea that children only make errors in non-frame based questions. For example, Becky's error *does lambs like apples?* was consistent with the prediction because the auxiliary + subject combination required for correct production (*do lambs* + *X*) was not one of the entrenched frames that she had learnt. Thus, as predicted by the theory, Becky could be said to have produced this error because she had no previously learnt lexical frame available on which to base this question.

[10] It will inevitably be the case that children will learn new frames during the error period. Thus, restricting the definition of frames to those learnt before the period will underestimate the amount of data that can be attributed to a frame, and will reduce the power of the analysis and the possibility of finding a significant difference. However, it is not possible to use data to define frames which are subsequently used to test a prediction on the same data; this would invalidate the statistical analyses.

Errors in which the target form was not clearly a form of DO or a modal auxiliary were excluded from the analysis (e.g. Anne's *can he drinking it?*).

Table 6 indicates the mean percentage error rate for yes–no and wh-questions that could and could not have been based on a previously learnt frame, and the mean number of questions on which the error rates were based. Meaningful mean error rates for frame-based wh-questions with modals could not be calculated because eight children produced no frames for these questions in the pre-error period. Thus, the data for each auxiliary type have been combined to increase the reliability of the analysis.

For both frame and non-frame based questions, correct questions are more likely than errors (error rates are never higher than 11%). Thus, it is clearly possible to produce significant numbers of correct questions even without a frame. However, the prediction of the usage-based account relates not to the number of correct questions produced but to the issue of whether non-frame questions are more likely to contain errors. To assess this, a two-way repeated measures ANOVA was performed with question type and frame/non-frame as independent variables and the percentage error as the dependent variable. There was no effect of question type, $F_{(1,7)} = .71$, $p = .43$, and no interaction, $F_{(1,6)} = .95$, $p = .36$. However, there was a main effect of frame – questions that could have been produced correctly using a frame had significantly lower error rates than those that could not, consistent with the constructivist prediction, $F_{(1,6)} = 15.18$, $p = .006$, $\eta_p^2 = .68$. In summary, error rates were significantly lower in questions that could have been produced using a prior-learnt entrenched frame.

The second analysis tested the prediction that question types that had occurred with high frequency in the input would be picked up as frames by children and so would be protected from error. High frequency input-based frames were defined as wh-word + auxiliary + variable or auxiliary + subject + variable combinations that had occurred three or more times in the adult input. Adult input was taken from the first five transcripts for Adam, Anne, Aran, Becky, Lara, and Nina. In order to ensure there were enough adult utterances for reliable analysis, the first 40 transcripts were used for Abe and Sarah, and the first ten for Ross. All the transcripts (20) were used for Peter. For Abe and Ross, the father's speech was used as the input because the father spoke more than the mother on the transcripts. For all other children, maternal utterances were used as the input.

Table 6
Mean number of questions and % error rate in questions that could and could not have been produced correctly using entrenched frames

	Questions that could have been based on frames		Questions that could not have been based on frames	
	Yes–no questions	Wh-questions	Yes–no questions	Wh-questions
Mean % error (SD)	1.22 (1.75)	1.27 (2.35)	6.58 (3.58)	11.24 (9.82)
Mean No. Qs (SD)	316.70 (451.38)	29.90 (43.15)	269.60 (298.08)	202.90 (198.18)

Table 7
Mean number of questions and % error rate in questions that could and could not have been produced correctly using a frame of high frequency in the input

		Questions that could have been based on frames		Questions that could not have based on frames	
		Yes–no questions	Wh-questions	Yes–no questions	Wh-questions
Mean % error (SD)	Questions with modals	0.64 (1.57)	11.42[a] (18.51)	2.74 (5.17)	18.93 (19.83)
	Questions with DO	1.95 (3.08)	7.26 (6.47)	23.47 (28.08)	8.43 (5.95)
Mean No. Qs (SD)	Questions with modals	247.30 (428.86)	7.70 (10.73)	107.00 (123.67)	46.40 (47.82)
	Questions with DO	169.40 (123.67)	98.10 (86.82)	74.90 (80.44)	85.10 (91.46)

[a] Figure is based on data from seven children as three parents produced no frames for wh-questions with modals.

On average, 20 lexical input-based frames (SD = 5.64) were identified for each adult. These overlapped substantially with the types of questions that occurred in the children's data: on average, 72% of the frames identified for each parent were also produced by that parent's child. On average, 53% of the combinations identified as frames from each child's pre-error period were also identified as frames in their parent's speech.

As for the previous analysis, the children's correct questions and errors were coded as being consistent or inconsistent with one of the input frames. Table 7 indicates the mean percentage error rate for frame and non-frame based questions, and the mean number of questions on which the error rates were based. Again, most questions were correct for both frame and non-frame based questions (error rates were never higher than 23%). A 3-way repeated measures ANOVA was performed with question type, auxiliary type and input frame/non-frame as independent variables and the percentage error as the dependent variable. There was no effect of question type, $F_{(1,6)} = 1.24$, $p = .31$, and no main effect of auxiliary, $F_{(1,6)} = .19$, $p = .68$. There were no significant interactions except for a marginally significant interaction between question type and auxiliary, $F_{(1,6)} = 4.41$, $p = .08$. Once again, however, there was a main effect of frame. Questions that could have been produced correctly using a frame occurred with significantly lower error rates than those that could not, $F_{(1,9)} = 7.03$, $p = .04$, $\eta_p^2 = .54$. In summary, error rates were significantly lower in the questions that could have been based on high frequency lexical frames.

There are two alternative explanations for the effects of frame reported above. First, it might be that errors are less likely in frame-based questions simply because these questions are learnt earlier, are less complex and require less cognitive load for production. If this were the case, we would expect to see differences in complexity between correct questions and errors. This was investigated by comparing correct questions and errors using the traditional measure of sentence complexity in children's speech – mean length of utterance (MLU)[11]. The results showed numerical, though not significant, differences in MLU in the wrong direction: errors were in fact shorter than correct questions (mean MLU of errors = 5.54, SD = .57; mean MLU of correct questions = 5.62, SD = .65, $t = .66$, $df = 9$, $p = .53$). It is not the case that errors are more likely on longer, more complex questions.

However, it might be that the children make an error because the target correct question is complex. If this were the case, we would predict that the correct target of the error would be longer in terms of MLU than the error. For each error, the correct target of the error was identified (e.g. for *what he did do?* the target would be *what did he do?*) and the MLU of the targets was then compared to the MLU of the error itself. Results showed that errors were, in fact, significantly longer than the target of the error (mean MLU of target = 5.23, SD = .57, mean MLU of error = 5.54, SD = .57, $t = 2.33$, $df = 9$, $p = .04$). Thus, it is not the case that errors are simply a mechanism by which children reduce the complexity of their questions.

The second potential explanation concerns word frequency, and in particular the frequency of the words that make up the pivot section of the frames (i.e. the frequency with which children produce particular lexical auxiliaries and subjects or wh-words in their speech). In adult and child speech, high frequency words are less likely to attract error; for example, Stemberger (2002) has demonstrated that over-tensing errors in adult speech are more frequent with low frequency verbs than with high frequency verbs. It is also probable that the wh-word + auxiliary and auxiliary + subject combinations we have identified as pivots for lexical frames are made up of words that are of high frequency in the language. Thus, we might be seeing an effect of the frequency of the pivot combination (i.e. of frames versus non-frames) on error rates simply because children make fewer errors with high frequency auxiliaries, subjects and wh-words.

Logistic regression was used to assess whether the input frequency of the lexical frame predicted error over and above the frequency of the individual words in the pivot.[12] For all questions produced by the children, the frequency of the lexical frame in the child's input was calculated. Then, for each wh-word, subject and auxiliary produced by a child, word token frequency in the input speech as a whole was calculated, to yield an individual frequency count for each word based on all uses in all contexts.[13] A logistic regression was performed with three predictor variables: wh-word or subject frequency, auxiliary frequency and lexical frame frequency, and with correct question/error (correct questions = 1, error = 0) as the dependent variable.

[11] MLU was calculated in words.

[12] The data from the pre-error period was not substantial enough to support performing this analysis on the frequency of the frames defined on the child's speech.

[13] A frequency measure based on the child's own input rather than an independent measure of word frequency was used in order to assess correctly the frequency of words that were commonly produced by the child but that may not be frequent in the language as a whole (e.g. the name of a sibling).

Table 8
Mean number of errors in the pivot and variable slot (SD)

	Questions that could have been based on frame		Questions that could not have been based on frame	
	Mean no. errors in pivot (SD)	Mean no. errors in variable slot (SD)	Mean no. errors in pivot (SD)	Mean no. errors in variable slot (SD)
Frames based on pre-error period	0.4 (0.97)	2.6 (3.37)	20.5 (32.28)	11.9 (12.39)
Frames based on input data	4.6 (4.55)	6.1 (6.31)	17.3 (29.84)	8.4 (8.00)

The predictor variables were transformed to correct for skew and a Huber-White sandwich estimator was used to correct for non-independence.[14]

Overall, the three predictors explained 8% of the variance (Wald χ^2 (3) = 164.06, $N = 8326$, $p < .001$). The frequency of the lexical frame remained a significant predictor even when the frequency of the individual words in the pivot was taken into account, Odds Ratio (OR) = 1.45, 95% confidence interval = 1.25 – 1.70, $p < .001$.[15] Wh-word or subject frequency was also a significant predictor, OR = 1.23, 95% confidence interval = 1.12 – 1.35, $p < .001$. However, auxiliary frequency was not an independent significant predictor, OR = 1.10, 95% confidence interval = .81 – 1.47, $p = .75$. Even when the frequency of the words in the pivot was taken into account, children made significantly fewer errors with pivot combinations that had occurred with high frequency in their input.

5.2.3. The nature of frame-based errors

This analysis investigated whether errors in frame-based questions tended to occur in the variable slot of the frame to test the prediction that errors in frame-based questions will result from the incorrect choice of elements in the variable slot. Table 8 indicates the mean number of errors in the pivot and the variable slot for questions that could and could not have been produced using a frame. This analysis must be considered cautiously as there were very few errors that occurred in frame-based questions. However, there was some evidence that errors in frame-based questions tended to occur in the variable rather than the pivot slot as predicted, though this was significant only for frames based on the pre-error period not for input-based frames (Wilcoxon $Z = 2.03$, $N = 10$, $p = .04$, Wilcoxon $Z = 1.13$, $N = 10$, $p = .26$ respectively). For questions that could not have been based on a frame, the data patterned in the opposite direction, with larger numbers of errors occurring in the pivot rather than the variable slot. However, this difference was not significant for either

[14] The regression used a Huber-White sandwich estimate to correct for non-independence rather than simply entering the child as a predictor, which would not have corrected for the covariance differences.

[15] The Odds Ratio (OR) is the ratio of the probability of occurrence of an event to the probability of the event not occurring. In our data, an OR greater than 1 means that the more frequent the frame, the more likely a correct question was to occur.

child data or input based frames (Wilcoxon $Z = 1.01$, $N = 10$, $p = .31$, Wilcoxon $Z = .83$, $N = 10$, $p = .41$, respectively).

5.3. The interaction between frames and question and auxiliary type

The final analysis assessed the relationship between the results reported in the first half of this paper – that errors pattern according to an interaction between question and auxiliary type – and the results reported in the second half – that frames attract fewer errors than non-frames. A logistic regression was performed in order to assess the independent contributions of the three predictors: question type (yes-no/wh), auxiliary type (DO/modal) and the input frequency of the lexical frame, on the pattern of correct use and error. As for the regression in the previous section, the lexical frame frequency predictor was log transformed and a Huber-White sandwich estimator was applied to correct for non-independence.

Overall, the three predictors explained 10% of the variance (Wald χ^2 (3) = 220.09, $N = 8326$, $p < .001$). Lexical frame frequency was a significant independent predictor, with fewer errors on the more frequent frames, OR = 1.70, 95% confidence interval = 1.38 – 2.09, $p < .001$. Question type was also a significant independent predictor, with more errors in wh-questions than yes–no questions, OR = .39, 95% confidence interval = .22 – .71, $p = .002$. However, auxiliary type was not an independent significant predictor, OR = .66, 95% confidence interval = .32 – 1.36, $p = .26$. The results suggest that the frequency of the lexical frame and the type of question are independent predictors of error.

To summarize, the children made significantly fewer errors in questions that could have been based on frames, whether the frames were defined on question types that are learnt early by the child or on question types that have occurred more frequently in the child's input. These effects cannot be attributed to the length of the question or to the frequency of the words making up the pivot section of the frame and may go part way towards explaining why children make more errors with some types of question (e.g. wh-questions with modals) than others (e.g. yes–no questions with modals).

6. Discussion

The present paper investigated error rates in yes–no and wh-questions with auxiliary DO and modal auxiliaries. The results demonstrated that error rates differed not only according to auxiliary type but also according to question type. Although overall error rates were consistent with the idea that children make more errors with auxiliary DO than modal auxiliaries, these rates disproportionately reflected the children's performance with the more frequently produced yes–no questions, and hid very different rates across question types. Distinguishing between the two question types revealed that the prediction held only for yes–no questions. In wh-questions, error rates were equally high with modals as with auxiliary DO. More importantly, for modal auxiliaries, error rates were significantly higher in wh-questions than in yes–no questions. This was not the case for auxiliary DO.

The difficulty with modal auxiliaries could not be attributed to problems forming questions with *why*; the error rate for wh-questions with modals remained high even when *why* questions were excluded. Nor could the findings be explained in terms of the children producing a high number of errors with negative questions because the results held for positive questions as well. It is important to note that the analysis does not show that children have no problems with *why* questions or with negated auxiliaries or with auxiliary DO. However, we cannot explain errors in questions simply in these terms.

The results suggest a clear structural difference between yes-no questions and wh-questions, which seem to prompt different reactions to the requirements on the movement of different auxiliaries. Explaining this pattern of error requires a principled explanation of why modals might be hard to acquire in wh-questions. It may be possible to achieve this by incorporating a greater role for the lexical learning of particular auxiliaries into generativist theories or by citing performance limitations or maturation. Another potential solution could exploit the structural differences between the two question types. However, no across-the-board explanation will suffice. Theories that posit that extra constraints on wh-questions in general (e.g. because of the additional syntax associated with wh-movement; Santelmann et al., 2002) or on questions with certain types of wh-words (e.g. DeVilliers, 1991) may have difficulty explaining why such constraints cause more commission errors in wh-questions with modals and DO than in those with other auxiliaries (BE and HAVE, see Rowland et al., 2005). Theories that look to problems with modals (e.g. the complex semantics of many modals) will have to incorporate an explanation of why these problems manifest themselves in wh-questions but not yes–no questions (or, in fact, in declaratives; Guasti et al., 1995). Finally, theories that posit that negation causes error must explain why errors occur, albeit at lower rates, with positive auxiliaries too.

In summary, errors in children's questions seem to indicate a complex interaction between question type and auxiliary type. In addition, as predicted by the usage-based account, the children were significantly more likely to produce correct questions if they could apply a lexical frame. Error rates were significantly lower in questions that could have been derived from previously learnt entrenched frames or from frames that had occurred with high frequency in the input. This finding could not be explained in terms of the length of the utterance or the overall frequency of the words in the pivot.

There was also some evidence that, for frames defined on the child's data, errors in frame-based questions reflected problems filling the variable slot of the frame (not the pivot). Many of the errors produced in frame-based questions were errors in which the child had used an entrenched pivot correctly but made a mistake in the variable slot – usually by choosing a tensed or agreeing verb form (e.g. Sarah's *did I did six*?, Nina's *does she has ears*? and Ross's *why did they ran away*?).

Finally, the results demonstrated that frame frequency was an independent predictor of error when question and auxiliary type were taken into account, suggesting that the interaction between question and auxiliary type reported in the first part of the paper may be partially explicable in terms of the frequency of individual frames. However, there was also an independent effect of question type in this analysis, indicating that wh-questions attract more, and yes–no questions attract fewer, errors than we would expect from a pure frame-based account. One possibility is that the frequency of the question type over all (wh-questions are less frequent than yes–no questions) may interact with the frequency of the frame; this possibility requires further analysis.

In sum, rates of error in children's questions seem to be explicable in terms of the speed at which children learn, and the frequency with which they hear, particular lexical question types. These results are consistent with the usage-based idea that a significant number of children's correct questions are based on an entrenched lexical frame, and the notion that errors tend to occur more often if no such frame is available. However, two provisos must be made. First, the analyses above focussed only on object and adjunct wh-questions. An additional layer of complexity needs to be added to the theory when we consider how children acquire a wider range of question structures, such as subject wh-questions (e.g. *who made that?*) and more complex questions. In particular, the theory must address how children might distinguish between structurally ambiguous uses of particular pivots. For example, *who do you* can be analysed as a frame for a simple object question or for a complex question raised from a subordinate clause (e.g. *who do you think is tall?*). Predictions about whether such pivots are applied to both structures will depend not only on the frequency of the pivot in the different structures in the input but on the amount of overlap between the lexical forms that are used in the two structures and on the level of semantic and conceptual similarity between the structures. The version of the theory presented in this paper does not address this issue but it will need to be considered in further developments (see Abbot-Smith & Behrens, 2006, for more on the role of similarity in the acquisition of semantically and/or syntactically related constructions).

Second, the children studied here were clearly capable of producing correct questions even without frames. Error rates in non-frame based questions were never higher than 30% and the regression analysis demonstrated that input frequency only accounted for 8% of the variance between correct use and error. Some of the correct non-frame questions may be based on frames that were not picked up by our frame definition schemes, a problem that may be resolved by more sophisticated definitions of a frame. For example, 19 of Becky's 20 correct *where does* questions were of the form *where does X go*, in which the subject form was variable but the verb *go* was constant. These examples indicate that the definition of frame should specify the form of the verb in certain cases (e.g. *where does X go?*). In other examples, it seemed as if the subject should be incorporated into the frame; 99 of Adam's 104 correct *how do* questions were of the form *how do you* + *X* and his one error with *how do* occurred with a different subject – *how do this go?* (see Lieven, Behrens, Speares, & Tomasello, 2003, for further work on frame definition).

However, it must be the case that children can generate correct questions even when they do not have a frame on which to base their question. In addition, the results of the regressions suggested that question type (yes–no or wh) and the frequency of the wh-word or subject of the question also had an independent effect on whether the child produced a correct question or an error. These results are not a problem per se for the usage-based theory, which predicts that children learn via a process of generalisation across frames (and within and between question types) as well as by lexically specific learning, and is consistent with the idea that linguistic representations are stored at a variety of grain sizes. However, it is clearly important not to over-emphasise the importance of lexically based frames and to focus attention too on how the pattern of data predicted by a constructivist generalisation mechanism would differ from that predicted by generativist theory. Some constructivist studies are starting to illustrate how generalisation might work. For example, Freudenthal, Pine, and Gobet (2005) have demonstrated that a symbolic computational model with an utterance final bias and a chunking mechanism (which treats frequent multi-word phrases as one unit) can learn to generate both correct utterances and errors productively by paying attention to the distributional characteristics of the input. Other work demonstrates how the acquisition of a particular construction can either support or hinder the later acquisition of related constructions (e.g. Abbot-Smith & Behrens, 2006).

Such work is necessary not only to explain the presence of non-frame based correct questions but also to illuminate the mechanisms behind the different types of error that we see in children's questions. Some of the error types reported in the literature may be straightforwardly explained in terms of inappropriately applied generalisation strategies. For example, agreement errors in wh-questions tend to occur more often with high frequency wh-word + auxiliary combinations (e.g. *what's* + *X*) in which the child slots a non-agreeing subject (e.g. *what's the cows eating*), then with low frequency combinations (e.g. *what does* + *X*). Similarly, Brown (1968) has demonstrated that many non-inversion errors (e.g. *why he can go to the park?*) are direct transformations of an antecedent utterance produced by the adult interlocutor, (*he can go to the park*), suggesting that the child is concatenating a wh-word with a recently heard declarative. However, other error types such as the 'positive + negative' auxiliary doubling errors reported by Guasti et al (1995, e.g. *can he can't eat it?*) may provide more problematic, and almost certainly require a theory that attributes the child with productive knowledge of the relationship between positive and negative forms of the auxiliary.

It is also important to note that many of the errors identified here have also been reported in the literature on adult speech (e.g. agreement errors and over-tensing errors; Bock, Eberhard, & Cutting, 2004; Stemberger, 2002). This does not mean that child and adult errors must stem from identical sources: child errors are more prevalent, more robust, and, unlike adult errors, are often unaffected by overt correction and show a gradual decline over the acquisition process. However, the similarities between child and adult errors may reveal pockets of difficulty that remain cognitively demanding even in adulthood. The most obvious explanations for the similarity between adult and child error concern frequency and complexity. Structures that are semantically, syntactically or phonologically complex and/or of low frequency are likely to be late acquired and to be more prone to error in production.

In addition, it may be that the utterance types that have never been learnt as lexical frames but have to be constructed on-line each time are harder for both children and adults to produce. These explanations are compatible both with ideas from cognitive linguistics about adult language (e.g. Bybee & Scheibman, 1999, who argue that lexical templates may underlie much of adult production) and with some models of sentence production (e.g. Gerken, 1991, who proposes that the use of templates in production can help reduce cognitive load). Further study of the relationship between complexity, frequency and the pattern of errors in child and adult speech is likely to be central to usage-based research on the language acquisition and production processes.

Acknowledgments

Thanks are due to Kirsten Abbot-Smith, Ben Ambridge, Colin Bannard, Ewa Dąbrowska, Bruno Estigarriba, Evan Kidd, Elena Lieven, Danielle Matthews, Julian Pine, George Rowland, Anna Theakston, and Mike Tomasello for their comments on a previous draft and for useful suggestions about relevant literature. Thanks to Jeremy Miles for help with the regression analyses and to Nick Ellis and Ginny Gathercole for advice on statistical analysis. Parts of this paper were presented at the Child Language Seminar, University of the West of England, Bristol, UK, 2004, the 6th Annual Gregynog Conference on Child Language, Nant Gwrtheyrn, UK, 2004, the 30th Boston University Conference on Language Development, Boston, USA, 2005 and the Latsis Colloquium of the University of Geneva, Geneva, Switzerland, 2006. The research was funded by the Economic and Social Research Council, Grant Numbers: RES-000-22-0241 and R000236393.

References

Abbot-Smith, K., & Behrens, H. (2006). How known constructions influence the acquisition of other constructions: The German passive and future constructions. *Cognitive Science, 30*(4).

Bellugi, U. (1971). Simplification in children's language. In R. Huxley & E. Ingram (Eds.), *Language acquisition: Models and methods* (pp. 95–117). London: Academic Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21*, 491–504.

Bloom, L., Lightbown, P., & Hood, L. (1975). Structure and variation in child language. *Monographs of the Society for Research in Child Development*, 40 (Serial No. 160).

Bock, K., Eberhard, K. M., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language, 51*, 251–278.

Braine, M. D. S. (1987). What is learned in acquiring word classes: A step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 65–87). Hillsdale, NJ: Lawrence Erlbaum.

Bresnan, J. (1999). *Linguistic theory at the turn of the century.* Plenary presentation at the 12th World Congress of Applied Linguistics, Tokyo.

Brooks, P. J., Braine, M. D. S., Catalano, L., Brody, R., & Sudhalter, V. (1993). Acquisition of gender-like subclasses in an artificial language: The contribution of phonological markers to learning. *Journal of Memory and Language, 32*, 76–95.

Brown, R. (1968). The development of wh questions in child speech. *Journal of Verbal Learning and Verbal Behavior, 7*, 279–290.
Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
Bybee, J., & Hopper, P. (Eds.). (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamin.
Bybee, J., & Scheibman, J. (1999). The effect on degree of constituency: The reduction of don't in English. *Linguistics, 37*, 575–596.
Cameron-Faulkner, T., Lieven, E. V. M., & Tomasello, M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*, 843–873.
Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.
Dąbrowska, E. (2000). From formula to schema: The acquisition of English questions. *Cognitive Linguistics, 11*, 83–102.
Dąbrowska, E., & Lieven, E. V. M. (2005). Developing question constructions: Lexical specificity and usage-based operations. *Cognitive Linguistics, 11*, 437–474.
Derwing, T., & Smyth, R. (1988). Auxiliary placement revisited. *Journal of the Atlantic Provinces Association, 10*, 1–15.
DeVilliers, J. (1991). Why questions? In T. Maxfield & B. Plunkett (Eds.), *Papers in the acquisition of WH* (pp. 155–171). Amherst, MA: University of Massachusetts Occasional Papers.
Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143–188.
Freudenthal, D., Pine, J. M., & Gobet, F. (2005). On the resolution of ambiguities in the extraction of syntactic categories through chunking. *Cognitive Systems Research, 6*, 17–25.
Gerken, L. (1991). The metrical basis for children's subjectless sentences. *Journal of Memory and Language, 30*, 431–451.
Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431–436.
Guasti, M.T., Thornton, R. & Wexler, K. (1995). Children's negative questions: The case of English. In D. MacLaughlin & S. McEwan (Eds.). *Proceedings of the Boston University Conference on Language Development*, *Vol. 19*, pp. 228–239.
Hurford, J. (1975). A child and the English question formation rule. *Journal of Child Language, 2*, 299–301.
Klee, T. (1985). Role of inversion in children's question development. *Journal of Speech and Hearing Research, 28*, 225–232.
Kuczaj, S. (1976a). Arguments against Hurford's Aux copying Rule. *Journal of Child Language, 3*, 423–427.
Kuczaj, S. (1976b). *-ing, -s and -ed: A study of the acquisition of certain verb inflections*. Unpublished doctoral dissertation, University of Minnesota.
Kuczaj, S., & Maratsos, M. (1983). The initial verbs of yes–no questions: A different kind of general grammatical category. *Developmental Psychology, 19*, 440–444.
Kuiper, K. (1996). *Smooth talkers: The linguistic performance of auctioneers and sportscasters*. Mahwah, NJ: Erlbaum.
Labov, W., & Labov, T. (1978). Learning the syntax of questions. In R. Campbell & P. Smith (Eds.), *Recent advances in the psychology of language* (pp. 1–44). New York: Plenum Press.
Langacker, R. W. (1991). *Foundations of cognitive grammar, Descriptive application* (Vol. 2). Stanford, California: Stanford University Press.
Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Positional learning and early grammatical development. *Journal of Child Language, 24*, 187–219.
Lieven, E., Behrens, H., Speares, J., & Tomasello, M. (2003). Early syntactic creativity: A usage-based approach. *Journal of Child Language, 30*, 333–370.
Lust, B. (1999). Universal grammar: The strong continuity hypothesis in first language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 111–155). San Diego: Academic Press.

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

Maratsos, M. (2000). More overregularisations after all: New data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen and Xu. *Journal of Child Language, 27*, 183–212.

Marcus, G. F. (1995). Children's overregularization of English plurals: A quantitative analysis. *Journal of Child Language, 22*, 447–460.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition, 30*, 678–686.

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science, 26*, 393–425.

Onnis, L., Monaghan, P., Christiansen, M. H., & Chater, N. (2004). Variability is the spice of learning, and a crucial ingredient for detecting and generalizing in nonadjacent dependencies. In *Proceedings of the 26th Annual Conference of the Cognitive Science Society*.

Onnis, L., Christiansen, M. H., Chater, N. & Gómez, R. (under review). Variability is an important ingredient in learning. *Journal of Experimental Psychology: Learning, Memory and Cognition.*

Peña, M., Bonatti, L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science, 298*, 604–607.

Pine, J. M., & Lieven, E. V. M. (1997). Slot and frame patterns and the development of the determiner category. *Applied Psycholinguistics, 18*, 123–138.

Radford, A. (1990). *Syntactic theory and the acquisition of English syntax: The nature of early child grammars in English*. Oxford: Blackwell.

Rispoli, M. (1998). Patterns of pronoun case errors. *Journal of Child Language, 25*, 533–554.

Rispoli, M. (1999). Case and agreement in English language development. *Journal of Child Language, 26*, 357–372.

Rowland, C. F., & Pine, J. M. (2000). Subject-auxiliary inversion errors and wh-question acquisition: What children do know?. *Journal of Child Language 27*, 157–181.

Rowland, C. F., & Pine, J. M. (2003). The development of inversion in wh-questions: A reply to Van Valin. *Journal of Child Language, 30*, 192–212.

Rowland, C. F., Pine, J. M., Lieven, E. V. M., & Theakston, A. L. (2005). The incidence of error in young children's wh-questions. *Journal of Speech, Language and Hearing Research, 48*, 384–404.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928.

Santelmann, L., Berk, S., Austin, J., Somashekar, S., & Lust, B. (2002). Continuity and development in the acquisition of inversion in yes/no questions: Dissociating movement and inflection. *Journal of Child Language, 29*, 813–842.

Schütze, C. T. (1997). *INFL in child and adult language: Agreement, case and licensing.* Ph.D. dissertation, MIT, Cambridge, MA (Distributed by MIT Working Papers in Linguistics).

Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Stemberger, J. P. (2002). Overtensing and the effect of regularity. *Cognitive Science, 26*, 737–766.

Stromswold, K. (1990). *Learnability and the acquisition of auxiliaries.* Dissertation, MIT (Distributed by MIT Working Papers in Linguistics).

Suppes, P. (1974). The semantics of children's language. *American Psychologist, 29*, 103–114.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2001). The role of performance relations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*, 127–152.

Theakston, A. L., Lieven, E. V. M., Pine, J. M., & Rowland, C. F. (2005). The acquisition of auxiliary syntax: BE and HAVE. *Cognitive Linguistics, 16*, 247–277.

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209–253.

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

Valian, V., Lasser, I. & Mandelbaum, D. (1992). *Children's early questions.* Manuscript, New York: Hunter College.

Van Lancker-Sidtis, D., & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. *Language and Communication, 24*, 207–240.
Van Valin, R. D. (2002). The development of subject-auxiliary inversion in English wh-questions: An alternative analysis. *Journal of Child Language, 29*, 161–175.
Wilson, R. (2000). *Category induction in second language learning: What artificial grammars can tell us*. Unpublished M.A. dissertation. University of Arizona.
Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language, 30*, 75–115.
Wilson, R., Gerken, L. A. & Nicol, J. (2000). *Artificial grammar research extends to natural language: Implicit learning of categories*. New Orleans, LA: Annual Meeting of the Psychonomics Society.