Comparing different accounts of inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?'*

BEN AMBRIDGE AND CAROLINE F. ROWLAND

University of Liverpool, UK

ANNA L. THEAKSTON

University of Manchester, UK

AND

MICHAEL TOMASELLO

Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany

ABSTRACT

This study investigated different accounts of children's acquisition of non-subject wh-questions. Questions using each of 4 wh-words (*what, who, how* and *why*), and 3 auxiliaries (BE, DO and CAN) in 3sg and 3pl form were elicited from 28 children aged 3;6–4;6. Rates of non-inversion error (*Who she is hitting?*) were found not to differ by wh-word, auxiliary or number alone, but by lexical auxiliary subtype and by wh-word+lexical auxiliary combination. This finding counts against simple rule-based accounts of question acquisition that include no role for the lexical subtype of the auxiliary, and suggests that children may initially acquire wh-word+lexical auxiliary combinations from the input. For DO questions, auxiliary-doubling errors (*What*

does she does like?) were also observed, although previous research has found that such errors are virtually non-existent for positive questions. Possible reasons for this discrepancy are discussed.

INTRODUCTION

Non-subject wh-questions[1] (e.g. *Who can she see?*) have attracted much interest in the language acquisition literature for two reasons. First, questions are perhaps the only syntactic structure for which English-speaking children commonly make word-order errors (Stromswold, 1990). Second, from a generativist perspective, wh- (and yes/no) questions represent a paradigmatic case of syntactic movement, and therefore represent an ideal 'test case' for both movement-based generativist accounts and competing constructivist accounts of language acquisition.

Under the former approach, non-subject wh-questions are formed as follows. The wh-word moves from its initial position in the inflectional phrase, IP (*she can see who?*), to the specifier position of the complementizer phrase, CP[2] (*who she can see?*). The auxiliary verb[2] (*can*) then raises from its original position within IP to the head of CP ([V to] I to C movement).

(1) [CP Who$_i$ [C' [C can$_j$] [IP she t$_j$ see t$_i$?]]]

The key assumption of all movement-based accounts of question formation is that this latter process is governed by a *subject–auxiliary inversion rule* (obligatory for wh-questions) which children must acquire. Potentially problematic for such accounts is the finding that many children pass through a stage in which correctly-formed questions co-exist with questions that display various types of inversion error. The challenge for generativist accounts is to explain why, if children are acquiring an inversion rule, this rule is applied inconsistently or erroneously across different question types. Importantly, inversion errors are not distributed randomly throughout children's questions. Many studies have demonstrated that particular types of inversion error occur predominantly or solely with particular question types. It is incumbent upon both movement- and construction-based accounts alike to explain this precise patterning of inversion errors.

The structure of the remainder of this introduction is as follows. First, we briefly survey the existing data on inversion errors in non-subject question

[1] This term is used here to refer to any wh-question which requires so-called subject–auxiliary inversion; that is both argument and adjunct wh-questions.

[2] As is usual in the wh-question literature, the term auxiliary (AUX) is used here to refer to genuine auxiliary verbs (BE, HAVE), the copula (BE), modals (CAN, SHOULD etc.) and the dummy auxiliary (DO), unless specifically stated otherwise. Upper case is used to refer to a lexical auxiliary type (e.g. BE), lower case to refer to particular lexical auxiliary subtypes (e.g. *is, are*).

formation, focussing particularly on wh-questions. We then introduce some of the accounts that attempt to explain the patterning of these errors, and conclude by introducing a new experimental study designed to investigate the various accounts. It should be noted from the outset that all of the research discussed, including the present study, is specific to English, and does not address the acquisition of the full range of wh-questions that exist in other languages.

Many different studies (some of which are summarized in Appendix Table A1) have investigated inversion errors in children's question production. Generally, children's inversion errors take the form of non-inversion, raising, or double-marking errors. Examples of each of these error types are shown in Table 1.

Researchers have typically focussed on how the frequency of these different error types varies according to the wh-word, auxiliary and the presence or absence of negation, or a combination of these factors. Although the studies are often not directly comparable, and sometimes produce contradictory results, three findings consistently emerge:

(1) *What* is the first wh-word to be acquired (Bloom *et al.*, 1982; Rowland *et al.*, 2003), displays the lowest rate of non-inversion (Kuczaj & Brannick, 1979; Rowland & Pine, 2000) and the highest rates of correct use (Labov & Labov, 1978; Erreich, 1984), and overall use (Tyack & Ingram, 1977). Conversely, *why* is typically acquired later than *what, who* and *how* (Bloom *et al.*, 1982; Rowland *et al.*, 2003), and displays the highest rate of non-inversion (Kuczaj & Brannick, 1979, but see Table 1; Rowland & Pine, 2000), and the lowest rate of correct use (Labov & Labov, 1978; Erreich, 1984).

(2) DO (especially in negated form) is typically associated with more inversion errors (Santelmann *et al.*, 2002; Rowland *et al.*, 2005; Rowland, submitted, for yes/no questions only), particularly double-marking errors (for negated DO only, Maratsos & Kuczaj, 1978; Stromswold, 1990; Hattori, 2003) and lower rates of correct use (Labov & Labov 1978, but see Table 1; Valian & Casey, 2003) than all other auxiliaries (but see Stromswold, 1990, p. 201). Both copula and auxiliary BE (the two are not always distinguished) generally display very low rates of inversion error (Maratsos & Kuczaj, 1978; Rowland & Pine, 2000, copula BE excluded; Santelmann *et al.*, 2002, for auxiliary but not copula BE; Rowland *et al.*, 2005; but see Stromswold, 1990; p. 154–155) and high rates of correct use (Labov & Labov, 1978; Valian & Casey, 2003).

(3) Negated wh-questions typically attract extremely high rates of non-inversion error, often 100% for a certain period, (Bellugi, 1971; Labov & Labov, 1978; Erreich, 1984; Stromswold, 1990; Rowland & Pine, 2000)

TABLE 1. *A taxonomy of possible inversion errors in children's positive and negative wh- and yes/no questions (all types have been attested at least once in the literature, but not all are observed in the present study)*

Inversion error type			(Negated) Wh-question example	(Negated) Yes/no question example
Non-inversion Raising*			What she does(n't) like? What she (not) likes?	She does(n't) like cake?** She (not) likes cake?***
Double marking (or 'doubling')	Auxiliary doubling	AUX … AUX	What does she does like?	Does she does like cake?
		AUX+NEG … AUX+NEG	What doesn't she doesn't like?	Doesn't she doesn't like cake?
		AUX … AUX+NEG	What does she doesn't like?	Does she doesn't like cake?
		AUX+NEG … AUX	What doesn't she does like?	Doesn't she does like cake?
		Different Auxilaires	What can('t) she does(n't) like?	Can('t) she does(n't) like cake?
	Tense doubling	AUX+, Main verb +	What does(n't) she likes?	Doesn't she likes cake?
	Tense on main verb only		What do(n't) she likes?	Do(n't) she likes cake?

* Possible only for DO questions.
** Acceptable in certain contexts (e.g. as echo questions).
*** Positive form, but not negative form, acceptable in certain contexts (e.g. as echo questions).
Note. Note the terminological distinction between 'inversion error' (or 'error of inversion'), which refers to any of the error types shown here and 'non-inversion error' which is a particular type of inversion error. Similarly, note that the term 'double-marking error' or simply 'doubling error' is a catch-all term that includes all of the seven different types of error in which tense/agreement is erroneously marked twice. For the example questions shown the target questions are of the form *What does(n't) she like* and *Does(n't) she like cake?* for wh- and yes/no questions respectively.

and doubling error (Maratsos & Kuczaj, 1978; Stromswold, 1990; Guasti *et al.*, 1995; Hattori, 2003). Indeed, auxiliary-doubling errors occur almost exclusively with negative questions (Maratsos & Kuczaj, 1978; Stromswold, 1990; Guasti *et al.*, 1995).

It is also sometimes suggested that wh-questions display higher rates of inversion error than yes/no questions (e.g. Bellugi, 1971; Labov & Labov, 1978), but the picture is complicated by the fact that non-inverted yes/no questions are not always counted as errors (as they are permissible in certain contexts) and by the findings of Erreich (1984) and Stromswold (1990) that at least some children display more inversion for wh- than yes/no questions.

Of course, no single experimental study can investigate all of these issues and, indeed, no author has yet proposed a theory that attempts to explain all these diverse findings. In the present study, then, we restrict ourselves to positive wh-questions, and investigate a number of theories that seek to explain the patterning of inversion errors with reference to the identity of the wh-word, the auxiliary, or a combination thereof.

Generativist accounts

Under generativist accounts such as Stromswold (1990), DeVilliers (1991), Valian, Lasser & Mandelbaum (1992) and Santelmann *et al.* (2002) it is generally claimed that the subject–auxiliary inversion rule, as a general principle of Universal Grammar, is available to children from the very first stages of multi-word speech. Inversion errors occur when the child has not yet learned to correctly apply this rule to particular wh-words or particular auxiliaries.

Wh-word-specific approaches. DeVilliers (1991) claims that children initially mis-analyse wh-questions such that the wh-word is analysed as an unmoved element, generated in place at the start of the sentence (adjoined to IP). This mis-analysis is caused by the presence in the input of a certain uncommon type of adjunct wh-question in which the subject and auxiliary are not inverted (e.g. *How come he is leaving?*), which children then extend to all wh-questions, producing non-inversion errors (e.g. **Why he is leaving?*).

These errors cease when the child produces sentences in which the wh-word is used as a complementizer (e.g. *I know why he is leaving*), as such sentences demonstrate to the child that the particular wh-word must always appear in the specifier of CP. However, this reanalysis 'comes in piecemeal with each wh-word' (p. 171). DeVilliers (1991) claims that the misanalysis of 'adjuncts 'how' and 'why' … persists for some time' (p. 171), presumably because the mis-analysis was triggered by adjunct, as opposed to argument, wh-words in the first place. Stromswold (1990) also argues that non-inversion errors will be more common for adjunct questions because 'the wh-word may be in a pre-SPEC position' (p. 199). Thus these theories

attempt to explain the common finding that *why* and, to a lesser extent, *how* seem to resist inversion whilst *what* is almost invariably the first wh-word to display inversion.

Valian *et al.* (1992) also argue that children possess knowledge of subject–auxiliary inversion from the earliest observable stages but initially apply this rule only optionally, and 'require a great deal of evidence for each wh-word' (p. 23) that the rule is obligatory for all non-subject wh-questions. Although this account does not specifically address the question of why particular wh-words seem to display higher rates of inversion error than others, it can potentially explain the more general, common finding that the frequency of such errors does indeed differ according to the identity of the wh-word.

Auxiliary-specific approaches. The central claim of generativist, auxiliary-specific approaches to inversion errors is that, from the point at which children first begin to produce non-subject wh-questions, their 'grammars initially contain knowledge of inversion' (Santelmann *et al.*, 2002: 820). Thus children know that inversion is obligatory, and always attempt to invert when producing non-subject questions. Errors are argued to occur only for two auxiliaries that display idiosyncratic properties: copula BE and the dummy auxiliary DO (Santelmann *et al.*, 2002; see also Stromswold, 1990; Hattori, 2003). Copula BE is unique in that it is the only main verb that displays inversion in questions (compare 2 and 3).

(2) Minnie is a mouse→What is Minnie?
(3) Minnie Mouse likes cake→*What likes Minnie Mouse?

Since children learn that, in English, main verbs do not invert, they are predicted to have difficulty in forming questions with copula BE until they learn to 'override their grammatical knowledge that main verbs do not raise … for their language' (Santelmann *et al.*, 2002: 837) for this particular item.

The dummy auxiliary DO is also unique in that, unlike all other auxiliaries, it is not present in the underlying declarative sentence, unless added for emphasis or negation:

(4) Minnie Mouse likes cake
(5) Minnie Mouse *does* (not) like cake

In the case of questions, DO is added to bear TNS and AGR:

(6) What does Minnie Mouse like?

TNS and AGR cannot be carried on the main verb, as main verbs in English do not invert (see 3). Thus this process of *DO-support* in questions does not fall naturally out of children's UG knowledge of inversion, but is a historical accident of the development of English (specifically that main verbs may no longer invert). Thus children's documented difficulties in

forming questions with DO (e.g. Stromswold, 1990; Santelmann *et al.*, 2002; Hattori, 2003) are argued to be a result of the difficulty of integrating their UG knowledge of inversion with their language specific knowledge of the process of DO-support.

It is notable that none of the generativist theories considered here include a role for the lexical form of the auxiliary. Under these accounts, non-subject wh-questions are formed by an inversion rule which operates on the CATEGORICAL VARIABLES of SUBJECT and AUX, and not on particular lexical subjects or auxiliaries. Stromswold (1990) explicitly states that once children 'hear a particular auxiliary, they generalize across tense, number and person' (p. 20). This prediction is also implicit in Santelmann *et al.*'s (2002) study, in which evidence children's use of a particular auxiliary form (e.g. *is*) is used to make claims about their knowledge of the patterning of the auxiliary verb (BE) in general. More generally, all formalist accounts make the assumption that children (and adults) produce utterances by manipulating lexical or functional CATEGORIES of items rather than individual lexical items themselves (see Wilson, 2003). Thus any finding that errors of inversion differ according to lexical auxiliary subtype (e.g. Rowland & Pine, 2000) is potentially problematic for such theories, at least in their current form.

A constructivist, lexical learning (wh-word + lexical auxiliary-specific) approach Rowland & Pine (2000) argue that none of these accounts are fully compatible with the available data, and advocate instead a construction-based account under which children 'will only produce correctly inverted wh-questions when they have been able to learn the relevant wh + aux combinations necessary to produce the question from the input' (p. 177).

Under functionalist, construction-based accounts of language acquisition, children acquire a structured inventory of grammatical constructions, which become increasingly abstract as development proceeds. For example, a child might begin with a lexically-specific WANT [X] schema, which gradually develops into a [SUBJECT] [VERB] [OBJECT] transitive construction schema (e.g. Tomasello, 2003). The adult-endpoint under such an account is not a formal grammar based on principles such as X-bar theory, but a hierarchically ordered network of interrelated constructions such as that advocated by Goldberg (1995) or Croft (2001).

Under such an account, the non-subject wh-question is simply another construction schema that the child must acquire (see also Dabrowska & Lieven, 2005). Like other construction schemas, the child begins by acquiring a number of lexically-specific schemas (e.g. *Where's [X] going?*, *What's [X] doing?*), and then may generalize across these to form a more abstract schema (e.g. [Wh-word] [BE] [SUBJECT] [VERB$_{progressive}$]?). It is important to note that there is no role for syntactic movement, transformations, or a subject–auxiliary inversion rule under this account.

Non-subject wh-question constructions have an independent existence, and are not viewed as transformations of an underlying declarative utterance.

As is often the case with construction-based approaches, the difficulty lies in specifying *a priori* the nature of the early lexically-specific constructions with which the child is assumed to be operating. Rowland & Pine (2000) argue that 'the child's lexically-specific knowledge is likely to centre round wh-word + auxiliary combinations, rather than auxiliary + subject combinations' (p. 164). One reason to favour this interpretation is that the range of wh-words and auxiliaries is relatively narrow (especially, perhaps, in speech to young children), whereas the range of subjects is potentially infinite. Also, Fletcher (1985) provides evidence that virtually all of one child's earliest non-subject wh-questions fitted one of only three specific wh-word + auxiliary templates (*How do …, What are …* or *Where is …*). It must be stressed that the term 'auxiliary', for Rowland & Pine's theory, refers to a specific lexical auxiliary subtype (e.g. *is*) and not to an auxiliary type (e.g. BE).

In contrast to the generativist theories outlined above, Rowland & Pine's (2000) theory predicts that non-inversion errors will pattern by neither wh-word nor auxiliary alone but by wh-word + lexical auxiliary combination. This is because the child is hypothesized to produce correct non-subject wh-questions only 'when [she] has learnt a wh-word + auxiliary marker around which to base her question frame' (p. 164) Non-inversion errors will occur only 'when the child has not learnt the particular wh-word + auxiliary marker around which to base the question she wishes to ask' (p. 165). In such cases non-inversion errors may be the result of a 'groping pattern' (Braine, 1976) whereby the child, in the absence of knowledge of the appropriate construction frame, for example, simply concatenates the wh-word onto part of a declarative frame (e.g. *What + she likes X = *What she likes?*).

Note that since, under this account, children do not acquire a subject–auxiliary inversion rule at all, the rate of non-inversion errors will decrease over time not because an optional inversion rule becomes obligatory, but because children gradually acquire more wh-word + auxiliary combinations from the input, and slowly generalize across all members of the wh-word and auxiliary categories. When a child has learned a wh-word + auxiliary combination for every question she may wish to ask, or has formed a general [WH-WORD] [AUX] [SUBJ] [VERB]? wh-question schema, there will be no need to apply 'groping pattern' strategies, and hence she will no longer make non-inversion errors.

Rowland & Pine's (2000) claim that non-inversion errors pattern by neither wh-word nor auxiliary alone but by wh-word + lexical auxiliary combination predicts an interaction between the variables of wh-word, auxiliary type, and number (auxiliary subtype) with respect to

non-inversion errors. This is because any particular wh-word+lexical auxiliary combination produced by children in the present study constitutes a unique combination of each of these three variables.

SUMMARY

Although, as documented in Table 1, many studies have investigated children's acquisition of non-subject wh-questions, a number of short-comings mean that the findings of these studies are often difficult to interpret. The studies of Bellugi (1971), Tyack & Ingram (1977), Labov & Labov (1978), Maratsos & Kuczaj (1978), Bloom *et al.* (1982), Rowland & Pine (2000), Hattori (2003), Rowland *et al.* (2003) and Rowland (submitted) all use naturalistic data. Whilst corpus analysis enjoys the advantage of relatively high ecological validity, one shortcoming of the approach is that certain wh-word+auxiliary combinations are produced too infrequently for reliable statistical analyses to be conducted. Equally problematic is that whilst virtually all authors note the occurrence of potentially rote-learned forms (e.g. *What's that?*) in children's speech, few attempt to systematically exclude such forms from their frequency counts. Of the experimental studies reported in Table 1, three (Kuczaj & Brannick, 1979; Erreich, 1984; Guasti *et al.*, 1995) include no statistical analyses (this is also the case for the naturalistic data studies of Bellgui, 1971; Tyack & Ingram, 1977; Maratsos & Kuczaj, 1978 and Bloom *et al.*, 1982) and two (Guasti *et al.*, 1995; Valian & Casey, 2003) do not present results separately for different wh-words or auxiliaries. The only experimental study which presents data separately for different auxiliaries and conducts appropriate statistical analyses is the elicited imitation study of Santelmann *et al.* (2002). Whilst an imitation task clearly reveals something about children's grammatical competence, the results of such a study clearly do not provide unambiguous evidence concerning children's ability to spontaneously produce correctly formed questions of particular types.

The goal of the present study then was to systematically elicit a range of non-subject wh-questions using different wh-words (*what*, *who*, *how* and *why*) auxiliary types (auxiliary BE, auxiliary DO and modal CAN) and lexical auxiliary subtypes (the 3sg and 3pl present tense forms of each auxiliary type) from young children, in order to provide further data against which to evaluate the theories outlined above.

METHOD

Participants

Participants were 28 monolingual English-speaking children (12 male and 16 female) aged between 3;6 and 4;6 (mean=4;1) recruited from two primary schools in Manchester, England. In total, 43 children were tested, of whom

15 were subsequently excluded. Children were excluded if they failed to ask at least one question during the warm up period (six children, all of whom attempted to answer rather than ask questions), failed to produce at least five correctly-formed non-subject wh-questions over the course of the study (four children), gave a large number of unintelligible responses (two children) or failed to complete the study due to absence or uncooperative behaviour (three children).

Materials

A toy dog with an internal loudspeaker connected to a minidisc player was used to provide responses to children's questions (see Valian & Casey, 2003). Four discs containing the same answers presented in different order (to allow for counterbalancing for question order) were recorded, with the first author providing the (male) voice for the dog. Five animal puppets (Mickey and Minnie Mouse, a frog, a duck and a bear) and eight small toys (e.g. cup, ball, pen) were used to enact the scenarios about which questions were to be asked. A small cardboard screen was used to hide characters from the children where necessary.

Design

The experiment employed a $(4) \times (3) \times (2) \times (4)$ mixed design. The three within-subjects variables were wh-word (*what/who/how/why*) auxiliary type[3] (BE/DO/CAN) and number (3sg/3pl), corresponding to the lexical subtype of the relevant auxiliary (*is/are, does/do, can/can*) and to the pronominal subject (*she/they*). The between-subjects variable was the order in which questions were elicited. Four pseudo-random orders were generated with the stipulation that no two consecutive questions could use the same wh-word or auxiliary. A further stipulation was that the wh-word *who* did not appear in any of the first four elicited questions as subject *who* questions were used during the warm-up period.

Procedure

Twenty-four non-subject wh-questions, each corresponding to a unique combination of the three within-subjects variables (wh-word, auxiliary and

---

[3] Eight questions with copula BE were also elicited. However, as noted by an anonymous reviewer, *what*, *who* and *how* questions with copula BE (e.g. *What is she? Who is she? and How is she?*) are potentially confusing as they are pragmatically odd and could be interpreted as subject questions (indeed the latter is not a true adjunct *how* question at all). In support of this interpretation, children made large numbers of errors with these questions, but not with *why* + copula BE questions (e.g. *Why is she sad?*) that are pragmatically more felicitous. The decision was therefore taken to remove all questions involving copula BE from further analysis.

number) were elicited from each participant. The child was first shown and asked to name the toys and animal puppets, which all were easily able to do. The experimenter then introduced the child to the 'talking dog' toy, and explained that the dog would speak only to answer questions which the child had put to him. The experimenter explained that he would help the child, by telling her what she should ask.

Four warm-up trials were used to introduce the child to the game of asking the dog questions about Mickey and Minnie Mouse, in response to a prompt from the experimenter. All warm-up trials used subject *who* questions (in the past tense to avoid the use of an auxiliary). Subject *who* questions were chosen as they allow the child to ask a well-formed question by imitating a part of the experimenter's prompt (since they do not require subject–auxiliary inversion) and are common in young children's speech (Stromswold, 1995).

At the start of each warm-up trial, the experimenter placed Mickey or Minnie Mouse and three other animal puppets behind a screen, out of the view of the child. Actions were performed behind a screen, as previous research (e.g. Crain & Thornton, 1998) has demonstrated that children display a pervasive tendency to answer rather than ask questions, unless they are unaware of the correct answer. The experimenter then performed the relevant action behind the screen and said *Oh no! Somebody hit [or kicked/bit/dropped] Mickey [or Minnie] Mouse. Let's ask the dog who hit him. Can you ask the dog who hit him? Say it after me 'who hit him?' and again 'who hit him?'*. In the majority of cases the child produced the appropriate subject wh-question (in this case, *Who hit him?*). The experimenter then operated the minidisc player, in order to have the dog produce an appropriate response (e.g. *The frog hit him*), which the experimenter and/or the child then enacted with the toys.

For each of the 24 test trials the procedure was as follows. The experimenter selected the relevant character(s) (Minnie Mouse, or Mickey and Minnie Mouse) and other animals and toys as necessary, placed them behind a screen and, out of view of the child, performed the relevant action (where appropriate). For trials in which the elicited question concerned states (*Who does she like?*), or future actions (*What can she draw?*, *How can they reach the cup?*) the screen was not used. The experimenter then produced three utterances (constituting the *prompt* for that trial), in order to elicit the appropriate non-subject wh-question from the child. Table 2 shows some sample prompts used in the experiment. The full text of all prompts used can be found in the appendix (Table A2).

The first utterance made some kind of general statement, establishing the background for the question to be asked, and included the auxiliary verb (in negative form for DO questions) and the main verb and any transitive object NP that was to be used in the question form. The second and

TABLE 2. *Sample experimenter prompts illustrating the use of the appropriate pronominal subject, wh-word and lexical auxiliary form in non-inverted order*

Wh-Word	AUX Type	AUX Form (No.)	(1) Statement	(2) 'I wonder' clause	(3) 'Ask the dog' clause	Child's response (and dog's answer)
What	BE	3sg	Minnie **is drinking** something	I wonder **what she is drinking**	Ask the dog **what she is drinking**	What is she drinking? (She is drinking tea)
Who	BE	3pl	Mickey and Minnie **are drinking** something	I wonder **what they are drinking**	Ask the dog **what they are drinking**	What are they drinking? (They are drinking coke)
How	DO	3sg	Minnie **doesn't kick the ball** with her hands	I wonder **how she *does* kick the ball**	Ask the dog **how she *does* kick the ball**	How *does* she kick the ball? (She kicks the ball with her feet)
Why	CAN	3pl	Mickey and Minnie **can hear the frog**	I wonder **why they can hear the frog**	Ask the dog **why they can hear the frog**	Why can they hear the frog? (The frog is singing very loud)

third utterances contained the appropriate pronominal subject, wh-word and lexical form of the auxiliary in the main clause of the utterances *I wonder …* and *Ask the dog …* respectively. All auxiliaries were supplied in full, uncontracted form in both the experimenter's prompt and the dog's reply.

In order to provide a 'pure' test of movement-based theories, it was decided to supply the appropriate form of the auxiliary verb in each prompt. As noted above (see sentences 4–5) auxiliary DO is not normally supplied in basic canonical sentences. To ensure that provision of DO in the second and third utterances (*I wonder who she* does *like. Ask the dog who she* does *like*) was felicitous, a negated form was used in the first utterance (*Minnie doesn't like the duck*). This ensured that inclusion of *does/do* in subsequent sentences, as a contrast with *doesn't/don't*, was natural. Whilst this does represent something of a confound, it was felt that omitting DO in the prompt sentence, to provide a more natural utterance (*I wonder who Minnie likes. Ask the dog who she likes*) would result in a greater confound, since all other auxiliaries were supplied in the experimenter's prompt. Our design also allows us to determine whether children's documented problems with DO (Stromwold, 1990; Santelmann *et al.*, 2002; Hattori, 2003) are caused by the need to insert the dummy auxiliary into the basic untransformed sentence. If so, then one might expect children to perform better when this auxiliary is supplied in the untransformed sentence, thus triggering any putative subject–auxiliary inversion rule.

All questions concerned transitive actions. This was to ensure parity between the argument wh-questions (*what* and *who*), which, by their nature require a direct object, and the adjunct wh-questions (*how* and *why*). All questions used the pronoun form *she* (3sg) or *they* (3pl). The decision to use only a female sentence subject (*she*, denoting Minnie Mouse) was taken for two reasons. Firstly, since female subjects are less frequent than male subjects (as revealed by an informal search of the CHILDES database), we felt that they would attract a greater number of inversion errors. Secondly, since the talking dog character was depicted as male (and had a male voice), the referent of the pronoun *he* would have been ambiguous between a male sentence subject (e.g. Mickey Mouse) and the dog character.

Thus, for each question, the child, on two occasions, heard the pronominal subject *she*, the appropriate wh-word, and the appropriate lexical form of the relevant auxiliary, but in NON-INVERTED order. In order to form an appropriate non-subject wh-question, the child merely had to reproduce a part of the experimenter's prompt, inverting the subject and the auxiliary (i.e. applying a hypothesized subject–auxiliary inversion rule to elements already present in the untransformed utterance).

When the child had attempted to produce an appropriate question, the experimenter then operated the minidisc player, in order to have the dog

produce an appropriate response, which was then enacted with the animal toys (where appropriate). If the child failed to produce a response, the experimenter reminded the child to *Ask the dog* but did not repeat any of the prompt. The 24 questions were evenly split into four blocks, each of which was presented on a different session on one of four consecutive days (with the four warm-up trials presented immediately before the first test session on Day 1). To control for order effects, four different versions of the test were constructed, with the 24 questions elicited in a different order for each.

Scoring

Children's responses were transcribed and classified by the first author. A representative sample (approximately 10% of the data) was also transcribed and classified by an independent researcher, blind to the theories under investigation. Agreement between the two coders with respect to the response category assigned to each utterance was 96% (Kappa=0·93). In no case did one coder class a response scored as an inversion error by the other coder as a correct question, or vice versa. Each response was classified into one of five mutually exclusive response categories:

- CORRECT QUESTION A well-formed non-subject wh-question with no errors.
- NON-INVERSION ERROR A question in which the subject and auxiliary appear in non-inverted order (e.g. *Who they do like?*, *What she is drinking?*) but with no other error. Raising errors (e.g. *Who she likes?*) were not observed in the present study (probably because the experimenter always supplied the relevant form of DO, e.g. *Ask the dog who she does like*).
- AUXILIARY-DOUBLING ERROR A question in which a child produces one correct auxiliary form, plus a second auxiliary form, whether correct or incorrect (e.g. *What do they do like?*, *What do they does like?*, *Why is she can hear the frog?*), but with no other error. No other types of double-marking errors (see Table 1) were observed in the present study (again, probably because the experimenter supplied the correct auxiliary and main verb forms in the prompt).
- OMITTED AUXILIARY A question in which the auxiliary verb is either absent or unclear, but with no other error.
- OTHER ERROR Any other error, including incorrect auxiliary forms (e.g. *What do she like?*), case marking errors (*What does her want?*), subject omission (*What does like?*), and responses in which the child asked a non-target question (e.g. *Who does she like?* instead of *Why does she like the bear?*). In fact, since the experimenter supplied the correct wh-word, auxiliary and subject pronoun forms in the prompt

sentence, such errors were extremely rare, and no questions scored as *other* exhibited non-inverted word order. Zero, unclear or irrelevant responses, or trials for which the child attempted to answer rather than ask a question were also scored as *other* and constituted a large majority of the 5% of utterances classified as *other*.

Substitutions of *he* for *she* and of full NPs for pronouns (e.g. *Mickey and Minnie* for *they*) were permitted, and occurred at rates of approximately 0·5% and 6% respectively. On only a handful of occasions did a child produce more than one response for a trial. On each occasion, the second utterance was clearly an attempted correction of the first, and only this second utterance was scored.

Responses involving auxiliary *is* were also scored as to the form of the auxiliary (full or contracted) used. The other auxiliaries either do not occur in contracted form, or have a contracted form deemed not be reliably distinguishable from the full form.

RESULTS AND DISCUSSION

Table 3 shows, for each elicited question, the proportion that were scored as non-inversion errors ($M = 11$%), correct questions (76%), errors of auxiliary doubling (6%) and omission (2%), and other errors (5%). Although all statistical analyses were conducted on raw data, for clarity we report only proportional data (the two are equivalent, since every child provided exactly one response for each question type, with no missing data).

As a preliminary analysis, the data were entered into a $(4) \times (3) \times (2) \times (4)$ mixed ANOVA with wh-word, auxiliary, and number as within-subjects variables, presentation schedule as a between-subjects variable, and non-inversion errors as the dependent variable (see the following paragraph). This revealed that the variable of presentation schedule was not associated with any main effects or interactions. Subsequent analyses were therefore conducted using $(4) \times (4) \times (2)$ repeated measures ANOVAs. A significance level of $p < 0·05$ was adopted throughout the study. Although this includes no adjustment for multiple comparisons, due to the extremely low power of our analyses (a result of having only a single response from each child in each cell) a lower p value would have been likely to have caused Type 2 errors. We also wanted to be maximally fair to theories that predict high error rates for particular wh-words or auxiliaries and not to dismiss as non-significant predicted effects that were observed, but weak (perhaps due to the low power of our analyses).

The primary set of analyses used non-inversion errors as the dependent measure, as the majority of accounts under investigation here were designed largely or exclusively to explain the patterning of these errors (and not, for example, auxiliary-doubling errors). For example, Rowland & Pine (2000)

TABLE 3. *Proportion of correct and erroneous questions for each question elicited in the study*

Wh-	AUX	No.	Target Question	Correct M	Correct S.D.	Non-inversion M	Non-inversion S.D.	Auxiliary-doubling M	Auxiliary-doubling S.D.	Auxiliary omission M	Auxiliary omission S.D.	Other error M	Other error S.D.
What	BE	s	What is she drinking?	0·82	0·39	0·11	0·31	0·00	0·00	0·04	0·19	0·04	0·19
		pl	What are they drinking?	0·71	0·46	0·04	0·19	0·04	0·19	0·04	0·19	0·18	0·42
	DO	s	What does she want?	0·71	0·46	0·07	0·26	0·14	0·36	0·00	0·00	0·07	0·26
		pl	What do they want?	0·39	0·50	0·25	0·44	0·29	0·46	0·00	0·00	0·07	0·26
	CAN	s	What can she draw?	0·89	0·31	0·11	0·31	0·00	0·00	0·00	0·00	0·00	0·00
		pl	What can they draw?	0·89	0·31	0·11	0·31	0·00	0·00	0·00	0·00	0·00	0·00
Who	BE	s	Who is she touching?	0·86	0·36	0·14	0·36	0·00	0·00	0·00	0·00	0·00	0·00
		pl	Who are they touching?	0·86	0·36	0·04	0·19	0·04	0·19	0·04	0·19	0·04	0·19
	DO	s	Who does she like?	0·64	0·49	0·14	0·36	0·14	0·36	0·00	0·00	0·07	0·26
		pl	Who do they like?	0·50	0·51	0·36	0·49	0·07	0·26	0·04	0·19	0·04	0·19
	CAN	s	Who can she see?	0·89	0·31	0·11	0·31	0·00	0·00	0·00	0·00	0·00	0·00
		pl	Who can they see?	0·89	0·31	0·11	0·31	0·00	0·00	0·00	0·00	0·00	0·00
How	BE	s	How is she eating the cake?	0·79	0·42	0·11	0·31	0·04	0·19	0·04	0·19	0·04	0·19
		pl	How are they eating the cake?	0·64	0·49	0·11	0·31	0·00	0·00	0·14	0·36	0·11	0·31
	DO	s	How does she kick the ball?	0·71	0·46	0·14	0·36	0·11	0·31	0·04	0·19	0·00	0·00
		pl	How do they kick the ball?	0·75	0·44	0·11	0·31	0·07	0·26	0·00	0·00	0·07	0·26
	CAN	s	How can she reach the cup?	0·89	0·31	0·04	0·19	0·00	0·00	0·00	0·00	0·07	0·26
		pl	How can they reach the cup?	0·86	0·36	0·11	0·31	0·00	0·00	0·00	0·00	0·04	0·19
Why	BE	s	Why is she pushing the car?	0·96	0·19	0·04	0·19	0·00	0·00	0·00	0·00	0·00	0·00
		pl	Why are they pushing the car?	0·68	0·48	0·04	0·19	0·00	0·00	0·14	0·36	0·14	0·39
	DO	s	Why does she like the bear?	0·75	0·44	0·07	0·26	0·14	0·36	0·00	0·00	0·04	0·19
		pl	Why do they like the bear?	0·57	0·50	0·11	0·31	0·21	0·42	0·04	0·19	0·07	0·26
	CAN	s	Why can she hear the frog?	0·82	0·39	0·07	0·26	0·04	0·19	0·00	0·00	0·07	0·26
		pl	Why can they hear the frog?	0·86	0·36	0·11	0·31	0·00	0·00	0·00	0·00	0·04	0·19
			TOTAL	0·76		0·11		0·06		0·02		0·05	

argue that children will make non-inversion errors 'when the child has not learnt the particular wh-word + auxiliary marker around which to base the question she wishes to ask' (p. 165). Clearly, then, it is appropriate to test this prediction only on non-inversion errors, and not auxiliary-doubling errors, which, in fact, constitute evidence that the child HAS learned the relevant 'wh-word + auxiliary marker'. Recall that no errors of inversion of other types (e.g. raising or tense-doubling errors) were observed in the present study.

It is important to recognize that, when evaluating the theories under consideration here, correct use is less valid as a dependent measure than non-inversion errors (and, for some theories, auxiliary-doubling errors). A correct-use measure 'lumps together' non-inversion errors with errors that provide no evidence one way or the other with respect to inversion (i.e. errors of auxiliary omission and unclassified errors), and with auxiliary-doubling errors which, despite their error status, constitute, from a constructivist standpoint, evidence of the child having acquired the relevant wh-word + auxiliary combination (and which have unclear status under accounts such as DeVilliers', 1991 and Santelmann *et al.*, 2002).

Non-inversion errors

A $(4) \times (3) \times (2)$ ANOVA conducted for non-inversion errors yielded no significant main effects, but did reveal significant interactions of auxiliary by number ($F_{3,27} = 4 \cdot 50$, $p = 0 \cdot 02$, partial $\eta^2 = 0 \cdot 14$) and of wh-word by auxiliary by number ($F_{6,81} = 2 \cdot 23$, $p = 0 \cdot 04$, partial $\eta^2 = 0 \cdot 08$). These interactions are illustrated in Figures 1 and 2 respectively.

The finding of a null effect for wh-word does not support accounts under which non-inversion errors are argued to be more persistent for adjunct than argument wh-words (e.g. Stromswold, 1990; De Villiers, 1991) (or indeed any account under which such errors are argued to pattern simply by wh-word, e.g. Valian *et al.*, 1992). In fact, the non-significant ($p = 0 \cdot 09$) trend was in the opposite direction, with both *how* ($M = 0 \cdot 10$, *S.E.* $= 0 \cdot 03$) and *why* ($M = 0 \cdot 07$, *S.E.* $= 0 \cdot 03$) showing lower rates of non-inversion than either *what* ($M = 0 \cdot 11$, *S.E.* $= 0 \cdot 04$) or *who* ($M = 0 \cdot 15$, *S.E.* $= 0 \cdot 04$).

Similarly, the finding of a null effect for auxiliary does not support the idea that children have particular difficulty in forming questions with auxiliary DO (e.g. Stromswold, 1990; Santelmann *et al.*, 2002[4]). In this case, though, there was a marginally significant ($p = 0 \cdot 10$) trend in the

[4] In support of a further prediction of these authors (see introduction), the withdrawn questions with copula BE did display a high rate of non-inversion error. However, this was carried entirely by performance on the pragmatically odd questions (see Footnote 3). *Why* + copula BE questions, the only type that are not pragmatically odd, in fact displayed a lower rate of non-inversion error than all but one (*Why* + auxiliary BE) of the other question types elicited in the present study.
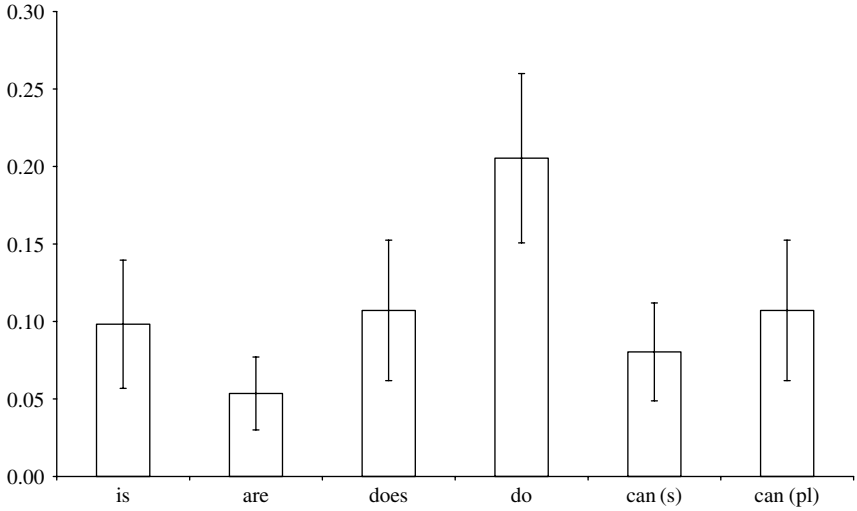
Fig. 1. Non-inversion errors by lexical auxiliary (auxiliary × number) as a proportion of all responses for each lexical auxiliary (error bars represent standard error).
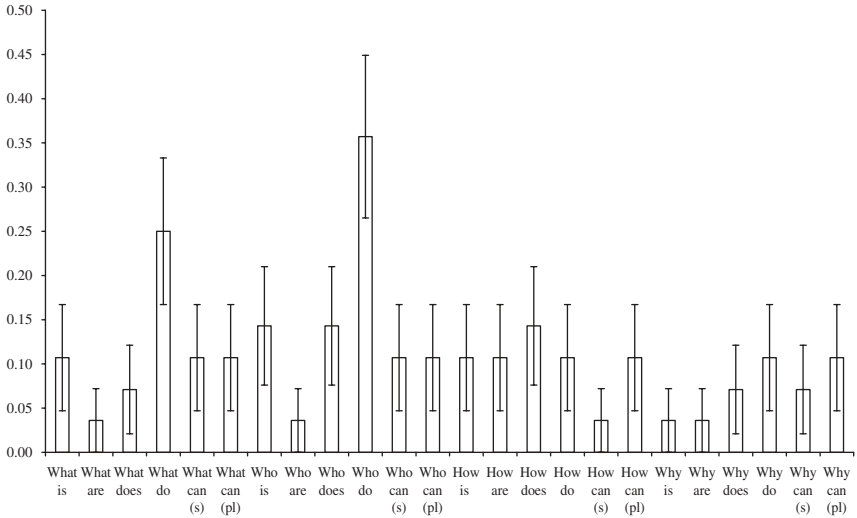

Fig. 2. Non-inversion errors by wh-word+lexical auxiliary combination (wh-word × auxiliary × number) as a proportion of all responses for each wh+lexical auxiliary combination (error bars represent standard error).

predicted direction with DO ($M=$ 0·16, $s.e.=$ 0·04) showing a higher rate of non-inversion than either BE ($M=$ 0·08, $s.e.=$ 0·03) or CAN ($M=$ 0·09, $s.e.=$ 0·03).

Whilst this might suggest that the present results 'provide some support for theories such as those proposed by Stromswold (1990) and Santelmann et al. (2002), within-number post hoc tests conducted to investigate the interaction of auxiliary by number (see Figure 1) revealed that the non-significant trend for children to produce more non-inversion errors for DO than BE or CAN was carried solely by children's performance on *do* (and not *does*) questions. That is, children produced significantly more non-inversion errors for 3pl *do* ($M=$ 0·21, $s.e.=$ 0·06) than 3pl *are* ($M=$ 0·05, $s.e.=$ 0·02, $p=$ 0·002) or 3pl *can* ($M=$ 0·11, $s.e.=$ 0·05, $p=$ 0·05), whilst the error rates for *does* ($M=$ 0·11, $s.e.=$ 0·05), *is* ($M=$ 0·10, $s.e.=$ 0·04) and 3sg *can* ($M=$ 0·08, $s.e.=$ 0·03) did not differ significantly. Thus the predicted effect for children to display difficulty with inversion for DO questions was observed only for 3pl and not 3sg questions. This finding cannot be attributed to some general difficulty which children might have in forming 3pl as opposed to 3sg questions as no main effect for number was observed (3sg $M=$ 0·10, $s.e.=$ 0·03; plural $M=$ 0·12, $s.e.=$ 0·03).

Under all the generativist accounts considered here (Stromswold, 1990; DeVilliers, 1991; Valian et al., 1992; Santelmann et al., 2002; Hattori, 2003), subject–auxiliary inversion is a movement rule that operates on auxiliary TYPES such as AUXILIARY DO and not on individual lexical auxiliary SUBTYPES such as *do* and *does*.[5] Thus none of these accounts would predict that non-inversion rates should differ between singular and plural forms of the same lexical auxiliary. In fact, within-auxiliary *post hoc* tests conducted to further investigate the auxiliary by number interaction revealed a higher non-inversion rate for 3pl *do* ($M=$ 0·20, $s.e.=$ 0·06) than 3sg *does* ($M=$ 0·10, $s.e.=$ 0·05; $p=$ 0·05), though the comparisons between 3pl *are* ($M=$ 0·05, $s.e.=$ 0·02) and 3sg *is* ($M=$ 0·10, $s.e.=$ 0·04) and between 3pl *can* ($M=$ 0·11, $s.e.=$ 0·05) and 3sg *can* ($M=$ 0·08, $s.e.=$ 0·03) did not reach statistical significance.

The finding of a significantly higher non-inversion rate for 3pl *do* than 3sg *does* is incompatible with all the generativist accounts considered here. In fact, on the surely uncontroversial assumption that children aged 3;6 and above have acquired the highly frequent lexical forms *do* and *does*, this finding would seem to be incompatible with such accounts IN PRINCIPLE. Once the different lexical forms of an auxiliary have been acquired, there

---

[5] In fact, strictly speaking, it is a rule that acts on a variable ('AUX') which stands for the lexical category of AUXILIARY VERBS. However, we accept that there are good grounds for treating copula BE and auxiliary DO as, at best, idiosyncratic members of this class.

is simply no mechanism which would allow for rates of commission error to differ for the different auxiliary subtypes (see Wilson, 2003, for similar arguments).

Rowland & Pine (2000) argue that rates of non-inversion error will pattern not by wh-word, auxiliary, or number alone, but by wh-word + lexical auxiliary combination. This account therefore predicts a three way interaction of wh-operator by auxiliary by number in the absence of any main effects. This pattern was observed in the present study. As we will argue in the general discussion, this finding is potentially problematic for all the other theories under consideration, as none currently include a role for both the wh-operator and the lexical form of the auxiliary.

Further post hoc tests were conducted to explore the three way interaction of wh-word by auxiliary by number (see Figure 2). The most interesting comparisons for our purposes are the between-number (i.e. within wh-word + auxiliary type) comparisons. Of the 12 comparisons, only two reached statistical significance. More non-inversion errors were observed for *what do* ($M = 0.25$, *S.E.* $= 0.08$) than *what does* ($M = 0.07$, *S.E.* $= 0.05$), and for *who do* ($M = 0.35$, *S.E.* $= 0.09$) than *who does* ($M = 0.14$, *S.E.* $= 0.07$), demonstrating that the observed effect of a higher non-inversion rate for *do* than *does* was carried entirely by these particular combinations (for *who does/do* and *how does/do*, the means are in the opposite direction).

Specific between-auxiliary (i.e. within wh-word + number) and between-wh-word (i.e. within auxiliary + number) comparisions are not directly relevant to our purposes here and so are not described in detail (the five significant comparisons observed all demonstrated that the two combinations that showed the highest rate of non-inversion error, *what do* and *who do*, showed significantly higher rates of this error than questions in which *do* was used with certain other wh-words, and *what* and *who* were used with certain other auxiliaries.) The point to note is simply that these interactions demonstrate that rates of non-inversion error were significantly different for different wh-word + auxiliary combinations, and did not pattern in any systematic way by wh-word, auxiliary or number alone. We did not conduct tests to compare error rates for each individual wh-word + auxiliary combination with every other as this would have required 276 separate tests ($23 + 22 + 21 \cdots + 1$). However, the comparisons that would have reached statistical significance in such a test can be determined with considerable accuracy by visual inspection of Figure 2 (paying particular attention to the error bars representing standard error).

Auxiliary-doubling errors

A $(4) \times (3) \times (2)$ ANOVA conducted for auxiliary-doubling errors found a significant main effect of auxiliary ($F_{2,27} = 12.64$, $p < 0.001$, partial $\eta^2 = 0.32$),
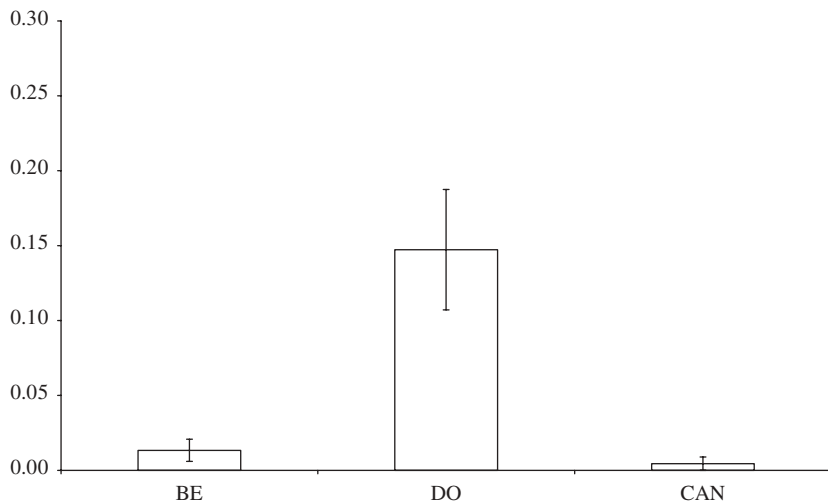
Fig. 3. Auxiliary-doubling errors by auxiliary as a proportion of all responses for each auxiliary (error bars represent standard error).

illustrated in Figure 3, but no other main effects or interactions. Significantly more auxiliary-doubling errors were observed for DO ($M = 0.15$, $S.E. = 0.04$) than BE ($M = 0.01$, $S.E. = 0.01$) or CAN ($M = 0.004$, $S.E. = 0.004$). Since these latter two figures correspond to three utterances and a single utterance respectively, it is fair to say that auxiliary-doubling errors were observed in meaningful numbers for DO only. This finding is consistent with the idea that children will have particular difficulty in forming questions with auxiliary DO, due to the language specific nature of this item (Stromswold, 1990; Santelmann *et al.*, 2002; Hattori, 2003).

GENERAL DISCUSSION

In the present study, non-subject wh-questions using each of four wh-words (*what, who, how* and *why*) and three auxiliaries (BE, DO and CAN), each in both 3sg and 3pl form (*is/are, does/do,* and *can/can*), were elicited from 28 children aged between 3;6 and 4;6. Non-inversion errors (e.g. *Who she is hitting?*) were found to pattern by lexical auxiliary subtype (auxiliary + number combination) and wh-word + lexical auxiliary subtype (wh-word + auxiliary + number combination) but not by wh-word, auxiliary or number alone. Most errors were observed for the combinations *who + do, what + do, how + does, who + does* and *who + is* with fewest errors for *what + are, who + are, how + can, why + is* and *why + are* (displaying a non-significant trend for DO to show the most non-inversion, and BE the least). Auxiliary-doubling errors (e.g. *What does she does like?*), generally

539

unattested for positive questions, were observed at a rate of 6%, almost exclusively reflecting performance on questions with auxiliary DO.

Thus it is clear that theories that attempt to explain the patterning of non-inversion errors with reference to the identity of the wh-word alone (e.g. DeVilliers, 1991; Valian *et al.*, 1992) cannot explain the current findings. Accounts based on the identity of the auxiliary (e.g. Stromswold, 1990; Santelmann *et al.*, 2002; Hattori, 2003) fare somewhat better. Although no main effect of auxiliary was observed for non-inversion errors, there was a nonsignificant trend for DO to display higher rates of non-inversion than BE or CAN (though this was carried by poor performance on the 3pl form *do*), and this effect reached statistical significance for auxiliary-doubling errors.

The finding that non-inversion errors patterned according to lexical auxiliary subtype, as evidenced by the interaction of auxiliary by number, is problematic for all the generativist accounts considered here. All generativist accounts of language acquisition (not just those under investigation here) view language as a formal system of rules that operate on CATEGORICAL VARIABLES (e.g. VERB, NOUN, AUX, SUBJECT) as opposed to individual lexical items. In an important sense, this is the whole point of the generativist approach, and lends its theories their characteristic elegance and parsimony. This means that generativist theories CAN build in a role for lexical content but only to the extent that different lexical items have different formal properties. For example, whilst DO and BE are both AUXs, they differ formally in that BE but not DO is present in the un-transformed sentence. Thus it is legitimate to predict different patterns of non-inversion and double-marking errors with these verbs (Stromswold, 1990; Santelmann *et al.*, 2002; Hattori, 2003). It would not be legitimate, though, for generativist theories to posit a difference between different forms of the same auxiliary verb (e.g. *does/do*) as these are formally identical (both represent the verb DO with appropriate TNS and AGRs marking as assigned or checked by the relevant functional projections). Yet this pattern is found in the present data (and Rowland *et al.*, 2005, report different error rates for the *is* and *are* subtypes of auxiliary BE). It cannot be argued that children had not acquired the plural form *do* as the experimenter supplied the appropriate lexical auxiliary form three times before each question was produced and indeed, for questions exhibiting non-inversion or double-marking errors, children always used the appropriate auxiliary subtype in their own productions. Although it would be possible to argue that children were using this form imitatively, without a true understanding of its function, it would seem unlikely that children aged 3;6 and above would not have acquired so frequent a lexeme as *do*. It is difficult to see then how it would be possible, in principle, for a generativist, movement-based model to account for the pattern of data observed.

Turning then to constructivist approaches, there are two theories that could potentially explain the current findings; the lexical learning account of Rowland & Pine (2000), and an alternative functionalist account presented by Van Valin (2002), grounded in the theoretical framework of Role and Reference Grammar (Van Valin & LaPolla, 1997). Under this analysis, when children produce correctly formed non-subject questions, they do so on the basis of their knowledge that in interrogative utterances, the element that is marked for tense (i.e. the auxiliary verb) appears CORE-INITIALLY (i.e. in pre-subject position). Thus children are argued initially to make non-inversion errors for auxiliaries that do not enter into clear tense opposition pairings (e.g. *can, may, might*) but not for those that are clearly tensed (e.g. '*is/are* vs *was/were*). According to Van Valin (2002: 167), this predicts that *is, are … do* [and] *does* should occur … in inverted questions, while *can …* should occur non-inverted, during the initial phase of the development of inversion'.

The findings of the present study do not provide for a test of this specific prediction, as children aged 3;6 and above can hardly be considered to be in the 'initial phase of the development of inversion', and, under Van Valin's (2002) analysis, may well have begun to analyse *can* as being tensed. The finding of a non-inversion rate of 21% for 3pl *do*, however, is potentially problematic for Van Valin's (2002) account. One possibility is that children have failed to analyse 3pl *do* as bearing tense as (unlike the 3sg form *does*, which shows a significantly lower non-inversion rate) it is homophonous with the non-finite form. If this were the case, however, it is difficult to see why the non-inversion rate for CAN would not also be high given that this item is also not clearly tense-marked. If, on the other hand, children analyse both 3pl *do* and CAN as tensed, then Van Valin's (2002) account has no way to explain the high non-inversion rate for the former item. The only way in which one could explain the current pattern of findings within Van Valin's (2002) framework would be to argue that children have begun to analyse CAN but not 3pl *do* as tensed. It is not easy to see why children should begin to analyse CAN as tensed before 3pl *do*, and, indeed, Van Valin (2002: 167) predicts precisely the opposite.

One way in which Van Valin's (2002) theory could be adapted to take account of the current findings would be to build in a significant role for lexical learning. Unlike generativist theories which, we have argued, rule out a role for the lexical subtype of a particular auxiliary in principle, there is no reason why a role and reference grammar account could not include a substantial lexical learning component. In practice, though, given the finding of a wh-word + lexical auxiliary interaction, in the absence of any main effects, it would seem that the only way to do this would be to include a role for the learning of wh-word + lexical auxiliary combinations. Whilst a role and reference grammar account could, of course, include such a component,

it is difficult to see how the predictions of such an account would differ from those of lexical learning account such as Rowland & Pine (2000), which does not require children to learn principles of illocutionary force marking such as 'in interrogative utterances, tense appears core-initially' (Van Valin, 2002: 165).

A potential objection to the explanations offered above for our findings of a null effect for wh-word and auxiliary is that some theories that seek to explain non-inversion errors in terms of the identity of the wh-word (e.g. Stromswold, 1990; DeVilliers, 1991; Valian *et al.*, 1992) or auxiliary (e.g. Van Valin, 2002) are concerned with the developmental order in which different wh-words or auxiliaries display inversion, and not with rates of non-inversion when children have begun to invert for some of their questions. Whilst it is not possible to address this objection directly without longitudinal experimental data, the findings of the present study do not sit comfortably with the claim that (obligatory) inversion appears later for *why* and *how* than *what* and *who* (DeVilliers, 1991), or that '*is, are … do* [and] *does* should occur … in inverted questions, while *can* … should occur non-inverted, during the initial phase of the development of inversion' (Van Valin, 2002: 167). If this were indeed the case, then one would have to argue that the present null findings for wh-word and auxiliary were caused by a ceiling effect, with children performing well on all questions. In fact, an overall error rate of 24% was observed, with over two thirds of these errors constituting non-inversion or auxiliary-doubling errors. Despite these relatively high error rates, there was no trend for the mean number of inversion errors to be higher for *why* and *how* than *what* and *who*, or for CAN than DO, which would surely have been the case if even a minority of children were displaying delayed inversion with adjunct wh-words, or apparently un-tensed auxiliaries, with the remainder at ceiling. The finding of an interaction of wh-word, auxiliary and number for non-inversion errors, despite the relatively low power of our analysis, demonstrates that at least a good number of children were not at ceiling. In fact, only two of the 28 children studied displayed perfect performance across the study. In conclusion then, whilst it is always difficult to draw conclusions from null findings, it seems that the finding that non-inversion errors did not pattern by wh-word or auxiliary alone is potentially theoretically interesting and, at the very least, is not solely attributable to ceiling effects.

In contrast to theories based solely on the identity of the wh-word or auxiliary, the lexical-learning theory of Rowland & Pine (2000), under which it is argued that children acquire wh-word + lexical auxiliary combinations from their input, effectively predicts the wh-word by auxiliary by number interaction observed for non-inversion errors in the present study. In one respect, though, the data are somewhat problematic for Rowland & Pine's (2000) theory. As predicted by generativist accounts such as those of

Stromswold (1990) and Santelmann *et al.* (2002), children did seem to have a certain amount of difficulty in producing correctly formed questions with DO (though this effect was significant only for auxiliary-doubling errors). Under Rowland & Pine's (2000) constructivist account, questions are not viewed as transformations of declarative utterances. Thus the theory cannot explain children's difficulty with DO in terms of its idiosyncratic properties with respect to the process of syntactic movement (Stromswold, 1990; Santelmann *et al.*, 2002; Hattori, 2003).

One factor which could potentially explain this finding, under Rowland & Pine's (2000) model, is input frequency. These authors claim that 'the wh-word+auxiliary combinations that the child uses in inverted wh-questions will be of significantly higher frequency in the child's input than the wh-word+auxiliary combinations that the child fails to use' (p. 165). Thus if it were the case that the particular wh-word+*do/does* combinations elicited in the present study tend to be rare in caregiver speech, then a simple input frequency account could potentially explain children's apparent problems with DO. In fact it would seem extremely unlikely that this is the case. Rowland & Pine's (2000) own data (see Table 5, p. 175) show that, excluding wh-word+auxiliary combinations that Adam (Brown, 1973) did not correctly produce over the course of the investigation, nine out of the ten most frequent wh-word+auxiliary combinations in Adam's mother's data use a form of DO (and *what+do*, which displayed the second highest non-inversion rate in the present study, was in fact the most frequent in this input corpus).

If the input data from Adam are at all representative of the input which the participants of the present study are likely to have received, Rowland & Pine's (2000) prediction that children will make fewer non-inversion errors for wh-word+auxiliary combinations that are frequent than infrequent in their input is at odds with the data. Clearly, then, a simple input-frequency based account cannot predict the wh-word+auxiliary combinations with which children will produce inverted and non-inverted questions, as Rowland & Pine (2003) themselves acknowledge:

> 'We do not wish to state that all children's acquisition data can be explained in terms of input frequencies (in fact, this claim would not be upheld by the data).' (p. 211)

What then is the relationship between input frequency and the pattern of children's language acquisition? One possibility, of course, is that there is no relationship between the two. Clearly such a position is untenable given that input frequency has been shown many times to be related to both children's language acquisition (e.g. Theakston, Lieven, Pine & Rowland, 2004, 2005; Rowland *et al.*, 2003) and adults' linguistic performance and representations (for a review see Bybee & Hopper, 2001).

A second possibility is that input frequency is closely related to children's language acquisition, but that researchers have yet to establish the undoubtedly complex nature of this relationship. The effects of input frequency on the acquisition of particular linguistic structures are likely to be complex and nonlinear, and to emerge from the precise interaction of the type and token frequency of particular items in the input. For example, Bybee (1995) discusses how constructions that have a high type frequency in the input (e.g. the English past tense $[VERB]+ed$ construction, in which many different verbs can participate), with any particular instantiation of that construction (e.g. *played*) occuring with only relatively low token frequency, are extremely productive. Conversely constructions that have low type frequency (e.g. the past tense vowel change construction, as in *give/gave*, in which far fewer verbs participate), with each of these types (e.g. *gave*) occurring with high type frequency, are less productive.

Applying this logic to non-subject wh-question constructions may help to explain why some wh-word + auxiliary combinations that are highly frequent in children's input nevertheless display high rates of non-inversion error. Rowland & Pine (2000) note that, for Adam (Brown, 1973), the combination *why + don't* attracted a non-inversion rate of 100% despite being one of the most frequent in the input data. Rowland & Pine (2000) further note that every recorded instance of *why + don't* in Adam's mother's speech was an instance of *Why don't you* [*X*]. The authors argue that, because the construction *Why don't you* [*X*] occurred with high token frequency in the input, but with no variation in the subject position (i.e. with low type frequency), Adam abstracted not a more productive *Why don't* [*SUBJECT*] [*X*] construction, but a less productive *Why don't you* [*X*] construction. The combination *why + don't* therefore displayed a high rate of non-inversion error because Adam attempted to form a number of *why + don't* questions with subjects other than *you*, but had not yet formed the necessary construction to do so (all but one of Adam's non-inversion errors with *why + don't* had subjects other than *you*).

A similar account could potentially explain why *what + do* displayed a higher non-inversion rate than all but one of the remaining wh-word + auxiliary combinations in the present study, despite presumably being very common in speech to young children (recall that it was the most frequent combination in the maternal data of the Adam corpus, see Rowland & Pine, 2000). It would seem likely that questions beginning *What do you* … are far more frequent in the input of our experimental participants than questions beginning *What do they* … (In the dataset analysed by Rowland & Pine, 2000, there are 89 tokens of questions beginning *What do you* … but only two beginning *What do they* …). If this is indeed the case, then it could be that the participants of the present study had acquired a *What do you* [*X*] formula, and not the more productive *What do*

[*SUBJECT*] [*X*] formula (or a *What do they* [*X*] formula) required to form a 3pl *what + do they* question.

For at least some wh-question constructions then, it could be that children are acquiring not wh-word + auxiliary but wh-word + auxiliary + subject combinations (or even in the case of highly frequent questions such as *What do you want?*, rote learned wh-word + auxiliary + subject + verb combinations). Indeed, this would not seem particularly unlikely given that much recent research (Pine & Lieven, 1993; Childers & Tomasello, 2001) has highlighted the importance of pronouns and frequent proper nouns (e.g. *Mummy* or the child's own name) in the formation of early construction schema. An important goal for future constructivist research, then, is to find a principled way to attribute to children knowledge of particular slot and frame construction schemas, the size and productivity of which would seem likely to vary depending on the particular lexical items involved, and their relative type and token frequencies. It is on these complex construction schemas, defined relative to particular lexical items and particular children (and not on simple pre-specified schemas such as wh-word + auxiliary), that theorists must base models of the acquisition of grammatical constructions, in which input frequency (in terms of the complex interaction of type and token frequency) will no doubt play a role. Focussing specifically on non-subject wh-questions, future research could seek to determine the generality and scope of particular question constructions in individual children by comparing performance on questions in which, for example, the wh-word + auxiliary combination (e.g. *What are …*) is held constant, but the subject is varied (*… you doing?* vs *… they doing?*).

A third possible characterization of the relationship between input frequency and children's language acquisition is that input frequency (though perhaps not simple input frequency) is one of a number of many different factors that influence acquisition, quite possibly in a complex and interacting fashion.

A recent study by Theakston, Lieven, Pine & Rowland (2005) highlights the importance of socio-pragmatic factors on the acquisition of particular schemas. Theakston *et al.*'s naturalistic data analysis of auxiliary acquisition found that the subject + auxiliary combinations *you're* and *you've*, although by far the most frequent in children's input, were typically acquired later than other, less frequent forms (e.g. *I'm*). One possibility discussed by these authors is that children are more eager to talk about themselves than a discourse partner, and, consequently preferentially abstract the first person frame.

With regard to non-subject wh-questions in particular, the work of Rowland & Pine (2000) and Van Valin (2002) highlights the importance of negation in inversion. Rowland & Pine show that all of Adam's wh + negated auxiliary combinations (e.g. *why can't*, *why don't*) use

exclusively non-inverted word order. Van Valin (2002) hypothesizes that the presence of the negation marker obscures tense marking, making it difficult for the child to locate the tensed element, here the auxiliary, which must be placed CORE-INITIALLY (i.e. in pre-subject position; see also Guasti *et al.*, 1995 for an alternative proposal). Whether or not this explanation is correct, it is clear from the literature that the presence of negation does, for whatever reason, hinder inversion in wh-question formation. Future work using the current paradigm should attempt to elicit negative questions to examine the pattern of non-inversion errors observed.

Both the present study, and the previous literature (particularly Labov & Labov, 1978 and Rowland & Pine, 2000) demonstrate that contracted and uncontracted auxiliary forms appear to operate quite differently with respect to non-subject wh-question formation, most likely because contracted auxiliary forms are often associated with potentially rote-learned, highly formulaic questions (e.g. *What's that?*), and subject+auxiliary combinations that use declarative word order (e.g. *she's*). Again, future research using an elicitation paradigm should attempt to elicit questions using contracted and uncontracted auxiliary forms to investigate the way in which the two forms seem to operate differently in question formation and, indeed, in auxiliary acquisition in general (Wilson, 2003; Theakston *et al.*, 2005).

Many of the findings of the present study are rather different to those of previous investigations of children's question acquisition (see Appendix A1). It is important to investigate these discrepancies in some detail, in order to determine whether they are mere artefacts of the present experimental procedure, or whether, in fact, they represent real findings that have not been uncovered by previous work.

The finding of the present study that non-inversion errors did not pattern by wh-word is at odds with the previous literature. Virtually all studies that have investigated the patterning of non-inversion errors by wh-word have found that errors are most frequent for *why* (and, to a lesser extent, *how*) and least frequent for *what* (and, to a lesser extent, *who*). It is significant that the only previous study for which this pattern was not observed (Kuczaj & Brannick, 1979) is the only one of the relevant studies that used experimental, as opposed to naturalistic data. (Although Kuczaj & Brannick found that *what* displayed the lowest rate of non-inversion, *why* and *how* both displayed lower rates than *who*.) One reason why naturalistic data studies might exaggerate children's ability at forming questions with *what* is that this wh-word tends to feature in a number of high frequency, potentially rote learned schema (e.g. *What's that?, What are you doing?*). Although virtually all authors note this possibility, few make any attempt to control for it when reporting error rates by wh-word. One notable exception is the study of Labov and Labov (1978) who report inversion rates separately for *what* questions with full and contracted forms of the auxiliary. The

difference in results is striking. *What* questions with a contracted auxiliary (including potentially rote learned questions such as *What's that?*) displayed an inversion rate of 77%, with this rate dropping to just 20% for questions with a full auxiliary form. The implication is that low error rates for *what* typically reflect children's knowledge of formulaic questions such as *What's that?*, and that when, as in the present study, children are placed in a situation in which these formulaic questions cannot be used, the advantage for *what* questions disappears. Conversely, *why* may show high error rates in naturalistic data studies because this wh-word does not feature in any rote-learned formulae, and so children are forced to apply creative 'groping' procedures when forming these questions (Brown, Cazden & Bellugi, 1969 and Rowland & Pine, 2000 suggest that the 'groping' procedure of concatenating a wh-word onto a declarative utterance may be particularly common for *why* questions, especially those produced in response to a declarative utterance from the caregiver, e.g. after the mother says *You don't throw things* the child asks *Why you don't throw things?*). When children are forced to form all their questions creatively (i.e. without the use of potentially rote-learned formulae) in an experimental study, we find that performance does not differ by wh-word.

Recall that in the present study too, children's responses with for questions with auxiliary *is* were coded for whether they used the full (*is*) or contracted (*-'s*) form of the auxiliary. Questions using a contracted form of auxiliary *is* displayed a non-inversion rate of 26·67% – somewhat higher than the overall rate across the study – and an auxiliary-doubling rate of 6·67% (though this represents only two utterances). The high non-inversion rate is perhaps seen because the subject + auxiliary combination *she's*, which appears in non-inverted questions, has been more well learned than any particular wh-word + auxiliary combination, due to its high frequency in declarative sentences. Of course, this explanation remains highly speculative at present. Future research could investigate this explanation by eliciting questions using (a) wh-words and auxiliaries that appear together with high and low frequency in interrogative utterances and (b) subjects and auxiliaries that appear together with high and low frequency in declarative utterances. The prediction from the above explanation is that high frequency wh-word + auxiliary combinations should resist non-inversion, whilst high frequency subject + auxiliary combinations should increase the rate of such errors.

Returning to the present study, there is a second reason why naturalistic data might exaggerate children's difficulty with *why* questions, particularly relative to *what* questions. Examination of the data from Rowland & Pine (2000: Table 3, p. 172) reveals that all the *what* questions in Adam's data occurred with positive auxiliaries, whereas six of the twelve *why* + lexical auxiliary question types that displayed non-inversion errors included

negated auxiliaries. Although the reasons for this are currently not well understood (but see Guasti *et al.*, 1995 and Van Valin, 2002, for proposals), we are aware of no study that has failed to find a higher rate of non-inversion (and double-marking) errors for negative than positive questions (see Table A1), regardless of the identity of the wh-word or, indeed, whether wh- or yes/no questions are studied. It could well be the case, then, that children do not have any particular difficulty with *why* as opposed to *what* questions *per se*. Rather, children have difficulty in forming negative questions which, probably for pragmatic purposes, tend to occur far more often with *why* than *what* (compare the naturalness of the sentences *Why can't I have a drink?* and *What can't I have?*). If this is true, then it is to be expected that when an experimental scenario is set up such that only positive questions are elicited, the difference in non-inversion rates between *what* and *why* will disappear.

A number of previous studies (Stromswold, 1990, for double-marking errors only; Santelmann *et al.*, 2002; Rowland *et al.*, 2005 [for auxiliary BE, but not modals]) have found that errors of inversion are more frequent for questions with DO than other auxiliaries (excluding copula BE). Considering first solely non-inversion errors, the present study found only a non-significant trend in this direction, carried by the 3pl form. In contrast Labov & Labov (1978) report higher rates of correctly inverted questions for DO than CAN (though the rate for auxiliary BE was higher again), and Stromswold (1990) reports perfect performance on positive DO questions.

These discrepancies may well result from the type of question studied. The study that provides the strongest support for the view that non-inversion errors will be most frequent for DO questions is the experimental study of Santelmann *et al.* (2002), which investigated solely yes/no questions, whereas the present study, and that of Labov & Labov (1978), investigated solely wh-questions. Rowland (in press) found that when yes/no and wh-questions are analysed separately, errors of commission are more frequent for DO than the modal auxiliaries for yes/no questions (although non-inverted yes/no questions were not counted as errors) but not for wh-questions.

Considering solely wh-questions then, the studies of Labov & Labov (1978) and Rowland (in press) suggest a split between DO and the modal auxiliaries, for which errors are common, and auxiliary BE, for which errors are less frequent. Although no main effect of auxiliary was observed in the present study, the distribution of non-inversion errors was by no means inconsistent with this pattern.

Considering now auxiliary-doubling errors, the findings of the present study, where such errors were observed at a rate of 15% for DO questions, are certainly inconsistent with those of previous research. Although, in line

with the findings of Maratsos & Kuczaj (1978), Stromwold (1990) and Hattori (2003), doubling errors were significantly more frequent for DO than BE or CAN, these authors (as well as Guasti *et al.*, 1995) observed virtually no doubling errors of the type AUX … AUX; the only type observed in the present study (see also Hurford, 1975; Kucazj, 1976). These authors report that, for positive questions, nearly all doubling errors were errors of tense-doubling (e.g. *What does she likes?*), with auxiliary-doubling errors restricted to negative questions (all but a handful of the form AUX … AUX + NEG; e.g. *What does she doesn't like?*). The studies of Maratsos & Kuczaj (1978), Stromswold (1990) and Hattori (2003) involve analysis of naturalistic data, so it is possible that auxiliary-doubling errors for positive questions were present in children's speech but were simply missed. The study of Guasti *et al.* (1995) however, used an experimental procedure similar to that of the present study. It would seem fair to conclude then that, outside the context of the present experiment, children almost never make auxiliary-doubling errors with positive auxiliaries.

It would seem likely, therefore that some factor in the design of the present study somehow caused children to produce auxiliary-doubling errors in contexts in which they would not normally do so. One possibility is that the experimenter's supplying DO in the 'untransformed' prompt utterances (e.g. *Minnie doesn't like the frog. I wonder who she* **does** *like. Ask the dog who she* **does** *like*) confused the children, since DO is not normally present in such utterances. Children may have thought that, since the experimenter was supplying a form of DO in a context in which it would not normally be present, in order to provide a contrast with *doesn't*, they too should highlight this contrast by providing an additional DO (giving, for example, *Who does she* **does** *like?*).

If, as would seem likely, the significantly higher rate of auxiliary-doubling errors with DO than CAN or BE observed in the present study was an artefact of the experimental design, then the findings of this study provide very little support for the claims of Stromswold (1990), Santlemann *et al.* (2002) and Hattori (2003) that children will have difficultly in forming questions with DO, due to its language-specific properties (recall that the higher non-inversion rate for DO than CAN or BE did not reach statistical significance).

In conclusion, the results of the present study suggest that generativist, movement-based accounts cannot account for the pattern of correctly formed and non-inverted non-subject wh-questions observed. The inter-actions of auxiliary by number and wh-word by auxiliary by number are simply not compatible with theories that explain the formation of wh-questions in terms of the application of rules that are blind to lexical content operating on categorical variables, and, indeed constitute evidence

against formalist approaches to language acquisition in general. These findings are, in principle, more compatible with a construction-based account of language acquisition, under which children first acquire lexically specific constructions based around highly frequent markers (e.g. *Why is she [X]ing Y?*), and then later form more abstract constructions. However, the present study adds to a growing body of work which demonstrates that constructivist theorists must move beyond simple input frequency, and incorporate a whole range of linguistic and sociopragmatic factors into their explanations of particular phenomena (Slobin, 1997; Rowland *et al.*, 2005; Theakston *et al.*, 2005). Future work should combine experimental and naturalistic methods in order to determine both the precise pattern of data that a theory of non-subject wh-question acquisition must explain and the various factors which influence this acquisition process.

REFERENCES

Bellugi, U. (1971). Simplification in children's language. In R. Huxley & E. Ingram (eds), *Language acquisition: models and methods*. New York: Academic Press.

Bloom, L., Merkin, S. & Wooten, J. (1982). Wh-questions: linguistic factors that contribute to the sequence of acquisition. *Child Development* **53**, 1084–92.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development* **41**(164).

Brown, R. (1973). *A first language*. Cambridge, MA: Harvard University Press.

Brown, R., Cazden, C. & Bellugi, U. (1969). The child's grammar from I to III. In J. P. Hill (ed.), *Minnesota symposium on child psychology 2*. Minneapolis: University of Minnesota Press.

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes* **10**(5), 425–55.

Bybee, J. & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins.

Childers, J. B. & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* **37**(6), 739–48.

Crain, S. & Thornton, R. (1998). *Investigations in universal grammar*. Cambridge, MA: MIT.

Croft, W. (2001). *Radical construction grammar: syntactic theory in typological perspective*. Oxford: Oxford University Press.

Dabrowska, E. & Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* **16**(3), 437–74.

DeVilliers, J. (1991). Why question? In T. L. Maxfield & B. Plunkett (eds), *Papers in the acquisition of wh: proceedings of the UMASS Roundtable, May 1990*. Amherst, MA: University of Massachusetts Occasional Papers.

Erreich, A. (1984). Learning how to ask: patterns of inversion in yes–no and wh-questions. *Journal of Child Language* **11**, 592–7.

Fletcher, P. (1985). *A child's learning of English*. Oxford: Blackwell.

Goldberg, A. E. (1995). *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Guasti, M., Thornton, R. & Wexler, K. (1995). Negation in children's questions: the case of English. *Proceedings of the 19th Annual Boston University Conference on Language Development*, 228–39.

Hattori. (2003). Why do children say *did you went?*: the role of do-support. *Supplement to the Proceedings of the 28th Boston University Conference on Language Development*. (http://www.bu.ed/linguistics/APPLIED/BUCLD/supp.html)

Hurford, J. (1975). A child and the English question formation rule. *Journal of Child Language* **2**, 299–301.

Kuczaj, S. (1976). Arguments against Hurford's 'Aux copying rule'. *Journal of Child Language* **3**, 423–7.

Kuczaj, S. A. & Brannick, N. (1979). Children's use of the wh question modal auxiliary placement rule. *Journal of Experimental Child Psychology* **28**, 43–67.

Labov, W. & Labov, T. (1978). Learning the syntax of questions. In R. Campbell & P. Smith (eds), *Recent advances in the psychology of language*. New York: Plenum Press.

Maratsos, M. & Kuczaj, S. (1978). Against the transformationalist account: a simpler analysis of auxiliary overmarking. *Journal of Child Language* **5**, 337–45.

Pine, J. M. & Lieven, E. V. M. (1993). Reanalysing rote-learned phrases: individual differences in the transition to multi-word speech. *Journal of Child Language* **20**(3), 551–71.

Rowland, C. F. (in press). Explaining errors in children's questions. *Cognition*.

Rowland, C. F. & Pine, J. M. (2000). Subject–auxiliary inversion errors and wh-question acquisition: 'What children do know?' *Journal of Child Language* **27**(1), 157–81.

Rowland, C. F. & Pine, J. M. (2003). The development of inversion in wh-questions: a reply to Van Valin. *Journal of Child Language* **30**(1), 197–212.

Rowland, C. F., Pine, J. M., Lieven, E. V. M. & Theakston, A. L. (2003). Determinants of acquisition order in wh questions: re-evaluating the role of caregiver speech. *Journal of Child Language*, 609–35.

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.

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**, 813–42.

Slobin, D. (1997). The universal, the typological and the particular in acquisition. In D. Slobin (ed.), *The crosslinguistic study of language acquisition; volume 5, Expanding the contents*. New Jersey: Erlbaum.

Stromswold, K. (1990). *Learnability and the acquisition of auxiliaries*. Unpublished Ph.D. dissertation, MIT.

Stromswold, K. (1995). The acquisition of subject and non-subject wh-questions. *Language Acquisition* **4**(1), 5–48.

Theakston, A. L., Lieven, E. V. M., Pine, J. M. & Rowland, C. F. (2004). Semantic generality, input frequency and the acquisition of syntax. *Journal of Child Language* **31**, 61–99.

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**, 247–77.

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

Tyack, D. & Ingram, D. (1977). Children's production and comprehension of questions. *Journal of Child Language* **4**, 211–24.

Valian, V. & Casey, L. (2003). Young children's acquisition of wh-questions: the role of structured input. *Journal of Child Language* **30**, 117–43.

Valian, V., Lasser, I. & Mandelbaum, D. (1992). *Children's early questions*. Paper presented at the 17th Annual Boston University Conference on Language Development, Boston, MA.

Van Valin, R. (2002). The development of subject–auxiliary inversion in English wh-questions: an alternative analysis. *Journal of Child Language* **29**(1), 161–75.

Van Valin, R. & LaPolla, R. (1997). *Syntax: structure, meaning and function*. Cambridge: C.U.P.

Wilson, S. (2003). Lexically specific constructions in the acquisition of inflection in English. *Journal of Child Language* **30**, 1–41.

TABLE A1. *Summary of previous research investigating children's acquisition of questions*

Study	Methodology/Participants	Dependent measure	Results
Tyack & Ingram (1977)	Parental diaries of 22 children aged 2;0–3;11, split into 4 age groups.	Frequency of questions using each wh-word as % of all questions, excluding 'identical repetitions' (p. 215) but presumably including both correct use and errors	• *What* = 14·5%, *where* = 11·8%, *why* = 6·9%, *how* = 2·1%, *who/when*/others ⩽ 1·5% (*Yes/no* questions = 60%) (p. 215) No statistical analyses conducted
Labov & Labov (1978)	Parental diary: 'A near-total record of … all [one child's] wh-questions – from the very first occurrence to the acquisition of the adult system' (p. 2)	% of questions using each wh-word that display inversion (from a random sample consisting of 1/6 of the corpus).	• *What* full form of AUX = 20% *What* contracted AUX = 77% *Why* full form of AUX = 6% *Why* contracted AUX = 3% (p. 15). *Who* = 87%, *How* = 89% (p. 21) • *Is, am* = 59%, *are* = 65%, DO = 48%, CAN = 35% (No statistics conducted for AUX) • Significant effects for wh-word, contraction (full = 64%, contracted = 52%), negation (neg = 11%, pos = 57%) & tense (past = 69%, present *-ing* = 40%) (p. 21) • Inversion 'learned for yes–no but not wh-questions' (p. 5)
Erreich (1984)	Elicited production (omitting AUX in prompt sentence for wh-questions (e.g. *Ask Anne her mommy's name*). 18 children aged 2;5–3;0	% of scorable questions (those containing an AUX) that display inversion, including both elicited and spontaneous questions	• *What* = 76%, *Where* = 72%, *Why* = 23%, *When* = 8%, *How* = 28%, *Who* = 94% (p. 588). • 'In affirmative questions children … used inversion randomly for yes–no questions [49% inversion] and preferred inversion in wh-questions [69%]' (p. 587)

			• All negative questions (5 yes/no, 11 wh-) in non-inverted form (p. 587) (no statistics conducted for any analysis).
Kuczaj & Brannick (1979), Study 1	Elicited imitation of wh- questions with a modal auxiliary. 20 children aged 3;2–4:10 divided into 2 age groups.	'Mean number of imitations in which the modal aux occupied a position different from that it had occupied in the model sentence [out of 5 trials per auxiliary]' (p. 48)	• *What*=0·4, *When*=1·1, *Where*=0·5, *Who*=1·3, *Why*=0·8, *How*=1·1, *How long*=1·2, *How come*=1·9 (p. 48) (no statistics conducted)
Bellugi (1971)	Adam corpus (Brown, 1973), plus informal elicited production experiment with 10 pos. and 10 neg. questions. Prompts use indirect questions with different subject (e.g. *Ask the Old Lady why she can't sit down*)	Total number of questions using modal auxiliary verbs that were inverted/non-inverted. Particular wh-words and modal auxiliaries used not reported.	• (Corpus) Positive *yes/no questions*= 198 inverted, 8 non-inverted. Negative *yes/no* questions=3 inverted, 0 non-inverted • (Corpus) Positive wh-questions=68 inverted, 34 non-inverted. Negative wh-questions=5 inverted, 16 non-inverted • (*Experiment*) All positive questions inverted, all negative questions non-inverted (no statistics conducted for any analysis).
Maratsos & Kuczaj (1978); Hattori (2003)	Naturalistic data from 2 children, ages not supplied (Maratsos & Kuczaj) and from Adam, Eve and Sarah from the Brown (1973) corpus (Hattori).	% of *yes/no* questions using each auxiliary that constituted tense doubling errors (e.g. *What did you got?*). Note that only one auxiliary-doubling error (e.g. *What is this is?*) was produced (by K.R). M&K note that Menyuk (1969) reports 2 such errors (Footnote 2) but conclude 'nearly anything can happen once' (p. 342)	• *Does*=13·3% (child D.C) 5·9% (K.R) *Did*=17·5% (D.C), 15·6% (K.R) *is*=0% (D.C.), 2·6% (K.R) *are & modals*=0% (both children) (Maratsos & Kuczaj, as presented in Hattori) • DO=14·6% (Adam), 33·3% (Eve), 18·2% (Sarah) BE=2·7% (Sarah), 0% (Adam, Eve) MODAL=0% (all 3 children) (Hattori, p. 10) (no statistics conducted for any analysis).

Study	Methodology/Participants	Dependent measure	Results
Bloom, Merkin & Wootten (1982)	Naturalistic data from 7 children aged 1:10–3:0	Rank order of emergence for productivity each wh-word, 'defined as the use of at least 3 different questions with a particular wh-form by at least 5 of the 7 children' (p. 1085), presumably including both correct use and errors	● Sequence (for questions containing aux+main verb)=where, what (2:2), who (2;4), how (2;9), why (2;11) (p. 1086). (No statistical analysis conducted on this sequence.)
Rowland et al. (2003)	Naturalistic data from the CHILDES Manchester corpus (Theakston et al., 2001). 12 children aged 1;8–3;0	'The 3rd use acquisition criteria adopted by Bloom et al. was used' (p. 617).	● '*What* or *where* were the first acquired wh- words for all 12 children … Four children acquired *how* or *why* before *who*'. (p. 617) (No statistical analysis conducted on this sequence).
Rowland et al. (2005)	Diary data from one child aged 2;7–2;11 (plus data from Manchester corpus-details as above – not reported here).	Errors of inversion (non-inversion, raising or double marking) as a % of all questions that use a form of each auxiliary.	● Statistically significant effects: DO (15·39%)>auxiliary BE (2·38%) Modals (23·53%)>auxiliary BE (2·38%) Copula BE (1·22%)=auxiliary BE (2·38%) Copula BE (1·22%)<Modals (23·53%) (Presumably DO>Copula BE, and DO=Modals, but not reported) (p. 400 & Table 8) For correctly formed questions: Auxiliary *is* (92%)>auxiliary *are* (75%)
Rowland (in press)	Naturalistic data from 10 children	Commission errors (i.e. all errors excluding subject or auxiliary omission) as a % of all questions that use a form of each auxiliary (DO or a modal). Wh and *yes/no* questions counted. NB: non-inversion-errors were not possible for yes/no questions.	● Main effect of question type (more errors for *wh* than *yes/no* questions). ● Interaction of question type × auxiliary: *Yes/no*:: DO (11·29%)>modals (1·17%) *Wh*: DO (8·91%)=modals (18·56%) (Table 4). Interaction holds excluding *why* questions and negatives (Table 5)

Valian & Casey (2003)	Elicited imitation of *when* & *where* questions with auxiliaries *do/does*, *can*, *will* and copula *is/are*. 29 children aged 2;6–3:2	% of questions using each wh-word that display inversion (before training intervention)	● *Can*+COP *is*+ COP *are*=50%, *Will*+*does*+*do*=40%. 'During the pre-intervention session, children performed better with *can* and *be* than with *will* and *do*' (p. 136) (It is unclear whether this difference is statistically significant. Inversion rates for individual auxiliaries not reported)
Rowland & Pine (2000)	Analysis of Adam corpus (Brown, 1973) between ages 2;11 and 3;8.	% of questions using each wh-word/ lexical auxiliary that display inversion. All analyses conducted on types not tokens: 'Tokens were defined as the same type if the wh-word, aux, subject, verb and, if present, PP were identical' (p. 167)	● Statistically significant effects: *Why* (8·3%)+*How* (85·4%)<*What* (78·6%)+*Who* (100%) (i.e. adjunctss<arguments) BUT ● *How* (adjunct)=*What*+*Who* (arguments) ● 'Only *why*, not *how*, shows a greater delay in inversion that *what*' (p. 170) ● Inversion rates for different lexical auxiliaries (not compared statistically): *can*=2/16, *can't*=0/13, *does*=21/21, *doesn't*=0/3, *do*=52/52, *don't*=0/6, aux *is*=5/6, -aux '*is*=2/6, aux *are*=15/15, aux -'*are*=3/3 (copula BE excluded).
Stromswold (1990)	Corpus analysis of 14 children with ages ranging from 0;11 to 2:10 (start of study) and 5;11 to 7;10 (end of study)	Inversion rate (%) 'for every question that contained an auxiliary and a subject NP' (p. 146), including only children who used at least 4 different wh-words for wh-word analysis (*N*=9), and who used both Cop and Aux BE for auxiliary analysis (*N*=11)	● 'mean inversion rates [for children who used ≥4 different wh-words]: how=98·5%, what=91·9%, when=73·2%, where=94·6%, which=89·5%, who=100%, why=80·0%' (p. 152). (Sig. main effect but no post hocs) ● (Wh and yes/no questions) DO=98·9%, modals=91·8%, Cop BE=86·0%, Aux BE=88·4% (p. 154, 155; Sig. main effect, but no post hocs).

Study	Methodology/Participants	Dependent measure	Results
		'Highest possible' (i.e. including doubtful utterances; p. 231) rate (%) of double-tensed questions	• Positive AUX=93·4%, NEG AUX=71·3% • Double-tensed questions=0·4% (164/40,600 questions; p. 231). Excluding 'stuttered or unclear utterances…93% of the double-tensed questions involved do-support'.
Santelmann *et al.* (2002)	Elicited imitation of yes/no questions (and declaratives) with aux *is*, *does*, *can*, copula *is* and can+copula *is* (e.g. *Can Donald Duck be a teacher?*). 65 children aged 2;1–5;3 divided into 3 age groups.	'Word order mismatches ([responses that] involved a clear change in the position of the auxiliary or the finiteness marking relative to the subject) as % of items' (Table 5, p. 833)	• Aux *is*=6%, *does*=14%, *can*=5%, copula *is*=11%, copula *is*+*can*=5%. (No statistical comparisons conducted for these figures, but statistical analysis shows that performance on questions is impaired relative to declaratives only for *does* and copula *is*).
Guasti *et al.* (1995)	Elicited production of negative wh- and yes/no questions 'with … do-support … [and] modals … [and] positive questions and [negative] declaratives … as controls' (p. 230) (data for auxiliary be were collected but excluded). 10 children aged 3;8–4;7	(1) % Aux … aux doubling errors (e.g. What does he does like?) (2) Aux+neg … aux+neg doubling errors (e.g. What doesn't he doesn't like?) (3) Aux … aux+neg doubling errors (e.g. What does he doesn't like?) (4) Non-inversion/raising errors (e.g. What he doesn't like; What he likes?) (5) Full form of not structure (What does he not like?) (6) Adult (correct questions)	(Results are presented collapsed across wh- and yes/no questions, and across different wh-words and auxiliaries. No statistical analyses are reported). • (1) Appears to be 0%, though data for positive questions not reported systematically • (2) 8·3% • (3) 39·7% • (4) 21·7% • (5) 10·7% • (6) 19·8% (calculated from Table 1, p. 231)

Wh-	AUX	No.	Prompt
What	BE	3sg	Minnie is drinking something. I wonder what she is drinking. Ask the dog what she is drinking?
		3pl	Mickie & Minnie are drinking something. I wonder what they are drinking. Ask the dog what they are drinking?
	DO	3sg	Minnie doesn't want any biscuits. I wonder what she does want. Ask the dog what she does want?
		3pl	Mickey & Minnie don't want any biscuits. I wonder what they do want. Ask the dog what they do want?
	CAN	3sg	Minnie can draw a picture of something. I wonder what she can draw. Ask the dog what she can draw?
		3pl	Mickey & Minnie can draw a picture of something. I wonder what they can draw. Ask the dog what they can draw?
Who	BE	3sg	Minnie is touching someone. I wonder who she is touching. Ask the dog who she is touching?
		3pl	Mickey & Minnie are touching someone. I wonder who they are touching. Ask the dog who they are touching?
	DO	3sg	Minnie doesn't like the duck. I wonder who she does like. Ask the dog who she does like?
		3pl	Mickey and Minnie don't like the duck. I wonder who they do like. Ask the dog who they do like?
	CAN	3sg	Minnie can see someone back there. I wonder who she can see. Ask the dog who she can see?
		3pl	Mickey & Minnie can see someone back there. I wonder who they can see. Ask the dog who they can see?
How	BE	3sg	Minnie is eating the cake – but I don't know how she is eating the cake. Ask the dog how she is eating the cake
		3pl	Mickey & Minnie are eating the cake – but I don't know how they are eating the cake. Ask the dog how they are eating the cake?
	DO	3sg	Minnie doesn't kick the ball with her hands. I wonder how she does kick the ball. Ask the dog how she does kick the ball?
		3pl	Mickey & Minnie don't kick the ball with their hands. I wonder how they do kick the ball. Ask the dog how they do kick the ball?
	CAN	3sg	Minnie is quite short but can reach the cup. I wonder how she can reach the cup. Ask the dog how she can reach the cup?
		3pl	Mickie and Minnie are quite short but can reach the cup. I wonder how they can reach the cup. Ask the dog how they can reach the cup?
Why	BE	3sg	Minnie is pushing the car. I wonder why she is pushing the car. Ask the dog why she is pushing the car?
		3pl	Mickie & Minnie are pushing the car. I wonder why they are pushing the car. Ask the dog why they are pushing the car?
	DO	3sg	Minnie doesn't like the frog. I wonder why she does like the bear. Ask the dog why she does like the bear?
		3pl	Mickey & Minnie don't like the frog. But wonder why they do like the bear. Ask the dog why they do like the bear?
	CAN	3sg	Minnie can hear the bear. I wonder why she can hear the bear. Ask the dog why she can hear the frog?
		3pl	Mickey and Minnie can hear the bear. I wonder why they can hear the bear. Ask the dog why they can hear the frog?