The acquisition of questions with long-distance dependencies

EWA DĄBROWSKA, CAROLINE ROWLAND and ANNA THEAKSTON*

Abstract

A number of researchers have claimed that questions and other constructions with long distance dependencies (LDDs) are acquired relatively early, by age 4 or even earlier, in spite of their complexity. Analysis of LDD questions in the input available to children suggests that they are extremely stereotypical, raising the possibility that children learn lexically specific templates such as WH do you think S-GAP? *rather than general rules of the kind postulated in traditional linguistic accounts of this construction. We describe three elicited imitation experiments with children aged from 4;6 to 6;9 and adult controls. Participants were asked to repeat prototypical questions (i.e., questions which match the hypothesised template), unprototypical questions (which depart from it in several respects) and declarative counterparts of both types of interrogative sentences. The children performed significantly better on the prototypical variants of both constructions, even when both variants contained exactly the same lexical material, while adults showed prototypicality effects for LDD questions only. These results suggest that a general declarative complementation construction emerges quite late in development (after age 6), and that even adults rely on lexically specific templates for LDD questions.*

* Correspondence addresses: Ewa Dąbrowska, School of English Literature, Language and Linguistics, University of Sheffield, Sheffield, S10 2TN, UK, e-mail: ⟨e.dabrowska@shef.ac.uk⟩; Caroline Rowland, School of Psychology, University of Liverpool, Liverpool, L69 7ZA, UK, e-mail: ⟨crowland@liverpool.ac.uk⟩; Anna Theakston, School of Psychological Sciences, University of Manchester, M13 9PL, UK, ⟨anna.theakston@manchester.ac.uk⟩. The research described here was supported by grant number AH/F001924/1 from the Arts and Humanities Research Council awarded to the first author. Acknowledgements: We would like to thank Claire Anderson, Hayley Lee, and Sarah Street for help with the data collection and coding, and Devin Casenhiser for his comments on an earlier version of this paper.

Cognitive Linguistics 20–3 (2009), 571–597 0936–5907/09/0020–0571
DOI 10.1515/COGL.2009.025 © Walter de Gruyter
Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Keywords: *language acquisition; long-distance dependencies; questions; complementation; usage-based approaches.*

1. Introduction

Questions and other constructions with long-distance dependencies (henceforth LDDs) have been the object of a considerable amount of research in the generative tradition (see for example, Cheng and Corver 2006; Chomsky 1977 and the references in (1) below). More recently, they have also attracted the attention of cognitive linguists (see e.g., Ambridge and Goldberg 2008; Dąbrowska 2004, 2008, in prep.; Goldberg 2006; Verhagen 2005, 2006). An interesting property of these constructions is that they contain a dependency between a WH word in the main clause and a gap in a subordinate clause, as shown in the examples in (1). In principle, there can be any number of clauses intervening between the WH word and the gap, so such dependencies are often referred to as 'unbounded'.

(1) (a) What$_1$ will John claim that you did ____$_1$? (Culicover 1997: 184)
(b) Which problem$_1$ does John know (that) Mary solved ____$_1$? (Ouhalla 1994: 72)
(c) Who$_1$ did Mary hope that Tom would tell Bill that he should visit ____$_1$? (Chomsky 1977: 74)
(d) Which problem$_1$ do you think (that) Jane believes (that) Bill claims (that) Mary solved ____$_1$1? (Ouhalla 1994: 71)

In contrast to these constructed examples, spontaneously produced LDD questions virtually never contain more than one finite subordinate clause (Dąbrowska in prep., for example, did not find a single instance of a dependency over more than one clause boundary in her sample of 423 LDD questions with finite complement clauses extracted from the spoken part of the British National Corpus). They are also extremely stereotypical (much more so than the corresponding declaratives[1]): the main clause subject is usually *you*, the verb *say* or *think*, and the auxiliary nearly always *do*; moreover, the main clause rarely contains any additional elements (Dąbrowska 2004, in prep.; Verhagen 2005). In the spo-

1. Declaratives with verb complement clauses are also quite stereotypical, but much less so. In the BNC data analysed by Dąbrowska (in prep.), the two most frequent declarative formulas, *I think S* and *I mean S*, accounted for only 35 percent of all utterances; and there was considerably more type variation in every syntactic position in the main clause. Verhagen (2005) reports very similar results for Dutch.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

ken BNC data analysed by Dąbrowska, 67 percent of the LDD questions had the form *WH do you think S-GAP?* or *WH did you say S-GAP?*, where S-GAP is a subordinate clause with a missing constituent. Most of the remaining questions were minimal variations on these patterns: that is to say, they contained a different matrix subject *or* a different verb *or* a different auxiliary *or* an additional element like an adverbial or complementizer; only 4 percent departed from the prototype in more than one respect. The extreme stereotypicality of spontaneously produced LDD questions has led some researchers (e.g., Dąbrowska 2008; Verhagen 2005) to hypothesise that while English speakers have a general complementation construction for declaratives, their knowledge about LDD questions may be most appropriately captured by means of two lexically specific templates, *WH do you think S-GAP?* and *WH did you say S-GAP?*, rather than by an abstract schema.

Researchers working within the generative tradition, in contrast, maintain that our knowledge about such constructions is expressed in terms of very general principles which apply not just to LDD questions but also to other related constructions such as indirect questions and relatives, and that, in spite of their complexity, the relevant principles are acquired by age four or even earlier (de Villiers 1995; de Villiers et al. 1990; Philip and de Villiers 1993; Thornton and Crain 1994). Thornton and Crain (1994, Experiment 1) elicited long-distance WH questions from children aged from 3;0 to 4;8, and found that seven out of the fifteen children in the study (all aged 3;11 or above) were able to produce adjunct questions (e.g., *What way do you think the Smurf went to the donut store?*), and nine out of the fifteen (all aged 3;8 or above) produced argument questions (e.g., *What do you think is in the box?*). Interestingly, children sometimes produced questions like (2) and (3), with a WH word at the beginning of both the main clause *and* the subordinate clause.

(2) What do you think what is in the box?
(3) What way do you think how he put out the fire?

Thornton and Crain regard such 'medial WH' questions as evidence for the cyclic application of movement (and hence appear to have included them in their count of productive uses of the construction, although this is not entirely clear from their description of the experiment). Note, however, that such utterances could also be produced by simply juxtaposing two independent questions (*what do you think? + what is in the box?*) or an independent question and an indirect question (*what way do you think? + how he put out the fire?*).

Thornton and Crain (1994) also tested children's comprehension of questions with long distance dependencies. The experiment involved

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

playing a guessing game with Kermit the Frog. The child was asked to hide one of two objects in a particular location, and Kermit had to guess which object the child hid. Kermit always guessed incorrectly, and the experimenter then asked questions about what Kermit had said and about the true identity of the hidden object. An example of the protocol used in the experiment is given in (4).

(4)	(i)	Kermit:	I think you probably hid the baby under there.
	(ii)	Experimenter:	What did he say's under there?
	(iii)	Child	The baby!
	(iv)	Experimenter:	What is under there?
	(v)	Child:	A bear!
			(Thornton and Crain 1994: 243)

The children (aged from 3;0 to 4;1) did not make any errors, which led Thornton and Crain to conclude that they understand questions with LDDs, and hence have the relevant grammatical knowledge, by age 3;0. The fact that they were not able to produce them until 8–11 months later, the researchers argue, is attributable to performance factors.

This conclusion, however, is premature. In order to give the correct response to the experimenter's first question, the child needs to process only the main clause in the prompt (*What did he say?*). The pragmatically most appropriate answer is *the baby*, since this is the only item of new information in (4i): all the other elements in the sentence express information that is shared by the discourse participants. Therefore, the fact that the children responded appropriately does not tell us very much about their knowledge of constructions with long distance dependencies. Similar criticisms can also be made of the other comprehension studies mentioned earlier (de Villiers 1995; de Villiers et al. 1990; Philip and de Villiers 1993): the fact that children sometimes interpret LDD questions as questions about the subordinate clause doesn't necessarily mean that they have the complex syntactic representations attributed to them by generative linguists.

However, the elicitation study does suggest that children can produce LDD questions from about 3;8 to 3;11. This is corroborated by analyses of children's spontaneous speech (Dąbrowska 2004; Thornton 2008). It should be pointed out, however, that children's LDD questions are even more stereotypical than adults, which raises the possibility that they are produced using lexically specific templates. Consider the first twelve LDD questions from the Abe corpus (Kuczaj 1976) in the CHILDES database:

(5) (a) which snake did he say was in the United States? (3;8)
(b) what do you think's under here? (3;10)

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

(c) what do you think's under here? (3;10)
(d) he's hopping to dinosaur+land what do you think the kangaroo's gon (t)a think? (3;11)
(e) I didn't know I saw him and he said ⟨hi # Abe⟩ ["] and he hitted me on the back Mommy # what do you think this is? (3;11)
(f) I know what put this here why do you think this doesn't work? (3;11)
(g) look at all those dinosaurs where do you think they're going? (3;11)
(h) no I don't # because this is gon (t)a be a dinosaur thing Mommy # look what I found where do you think the other one is? (3;11)
(j) well # how long do you think it would have to take to that crane? (3;11)
(k) what # why do you think his sword is pointing that way? (3;11)
(l) when do you think we're going to fix it? (3;11)
(m) where do you think the other army man can be? (3;11)

The Abe transcripts contain a total of 44 questions with a dependency between a WH word in the main clause and a gap in a finite subordinate clause, all recorded between the ages of 3;8 and 4;9[2]. Forty-two of these have the form *WH do you think S-GAP?* and one (interestingly, the very first LDD question in the corpus) has the form *WH did NP say S-GAP*? The remaining question (*what do you think that I was singing?*, produced at age 4;0) is a minimal variation on the *WH do you think S-GAP?* template.

To summarise, Thornton and Crain (1994) have demonstrated that four-year-olds can produce questions of the form *WH do you think S-GAP?* This, however, does not necessarily mean that they have mastered the LDD question construction, let alone other constructions involving 'movement': it is also possible that they are simply using a lexically specific template. We know that lexically specific formulas or templates play an important role in early language acquisition (Dąbrowska 2004; Lieven et al. 1997; Tomasello 2003). But do children hear enough LDD questions in order to learn such templates?

2. In addition, at age 2;10 Abe imitated an adult LDD question (*know what do you think it was?*)

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Questions with long-distance dependencies are a relatively infrequent construction: in the spoken part of the British National Corpus, for example, they occur with a frequency of about 42 per million words (Dąbrowska in prep.). More relevant for our purposes, however, is their frequency in language addressed to children. Since child-directed speech contains a relatively high proportion of questions, it is possible that LDD questions are also more frequent in this type of discourse.

To determine how often children hear LDD questions, we used CLAN software (MacWhinney 1995) to extract all instances of the construction from the Manchester corpus (Theakston et al. 2001), which consists of transcriptions of 402 hours of spontaneous interaction between 12 two-year-old children and their families. We found 325 tokens, all produced by adults. Since the corpus contains 1 450 000 adult words in total, the normalised frequency of LDD questions in the input is 225 per million words—about five times higher than in adult discourse[3].

This means that children hear an LDD question approximately once every 70 minutes on average—about as often as they hear words like *bus, old*, and *able.* LDD questions in the input are also extremely stereotypical. In 99 percent of the instances we analysed, the auxiliary in the main clause was some form of *do;* in 91 percent, the main clause subject was *you*; in 96 percent, the main verb was *think* or *say;* and only 2 percent of the main clauses contained additional elements such as complementizers or a direct object[4]. In fact, 85 percent of the questions fit the lexically specific template *WH do you think S-GAP?*; the second most frequent template, *WH did NP say S-GAP?* accounted for a further 9 percent. Most of the remaining questions involve minimal modifications of these templates (e.g., they contained a different subject *or* a different auxiliary *or* an overt complementizer).

Thus, it does not seem unreasonable to suppose that the children in the Thornton and Crain (1994) studies could have been relying on lexically specific templates. To test this hypothesis, we compared children's performance on prototypical questions (i.e., those which match one of the two high-frequency templates found in the input) and unprototypical questions (those which depart from the template in several respects). In the experiments described below we used a simpler task than Thornton and Crain (sentence repetition rather than production) and tested older chil-

3. Note that the figures given in the text are for LDD questions involving a dependency between a WH word in the main clause and a gap in the finite subordinate clause.
4. Some of the questions also contained utterance-initial elements such as *and, so, um*, and *well*. These were not counted as optional elements in the main clause.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

dren (aged from 4 to 6). Our prediction was that children will perform more accurately on prototypical questions than on unprototypical questions.

We used sentence repetition rather than elicited production because this method is thought to provide a more direct reflection of children's underlying competence while also allowing more control over what they say (Lust et al. 1998). To be able to repeat a complex sentence correctly, a child must be able to reconstruct its grammatical structure; and the errors that children make on the task often provide useful clues about their interpretation of the sentence (Slobin and Welsh 1973; Santelmann et al. 2002).

2. Study 1

In this experiment, we tested children's ability to repeat three types of questions: prototypical, unprototypical, and deeply embedded questions with long distance dependencies. Prototypical questions instantiated the *WH do you think S-GAP?* template. Unprototypical and deeply embedded questions departed from the template in various ways: the former had different lexical content in the main clause, and the latter contained a dependency spanning two clause boundaries (see Table 1 for examples). Such questions also depart from the LDD question prototype, though in a different way from the 'unprototypical' questions: they contain an additional element (the extra complement clause) between the main clause and the clause containing the gap.

According to generativist accounts of question formation, WH movement applies cyclically, so in principle, once children have learned to form simple questions and sentences with complement clauses, they ought to be able to produce LDD questions with any number of clauses intervening between the filler and the gap. However, questions with very long dependencies (spanning two or more clause boundaries) are difficult to process (Frazier and Clifton 1989; Hawkins 1999; Kluender and Kutas

Table 1. *Examples of sentences used in Study 1*

Condition	Example
Prototypical LDD questions	*What do you think the boys will really like?*
Unprototypical LDD questions	*What does the man really hope they will like?*
Deeply embedded LDD questions	*What do you think he said they will like?*
Prototypical declarative	*I think the boys will really like their shoes.*
Unprototypical declarative	*The man really hopes they will like their shoes.*
Deeply embedded declarative	*I think he said they will like their shoes.*

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

1993), and therefore one might expect poorer performance on such sentences. Generativist accounts, therefore, predict equally good performance on prototypical and unprototypical questions (since they are equally complex syntactically), and possibly some difficulties with deeply embedded questions. On the other hand, accounts which assume that children rely on lexically specific templates like *WH do you think S-GAP?* predict that children will do well on questions that match the templates but have problems with both unprototypical and deeply embedded questions.

2.1. *Method*

2.1.1. *Participants.* Thirty-four monolingual English-speaking children aged from 4;6 to 5;3 (mean 4;10) participated in the experiment. There were 14 boys and 20 girls. The children were recruited from a primary school in the North-West of England. Three additional children (all boys) were tested but not included in the final sample either because they failed to complete the task (2 children) or because of experimenter error (1 child).

2.1.2. *Design.* The experiment employed a 3 × 2 within-subjects design with two independent variables: prototypicality (with three levels: prototypical, unprototypical , and deeply embedded sentences) and construction type (with two levels: declaratives and questions). The dependent measures were, for each child, the number of correctly repeated sentences.

2.1.3. *Materials.* Prototypical LDD questions consisted of a WH word (*what, who*, or *where*) followed by *do you think* followed by a subordinate clause containing a lexical subject and an adverb (see Table 1). Unprototypical questions consisted of a WH word followed by the auxiliary *does*, a lexical NP, an adverb, a different matrix verb (*hope, expect*, or *believe*), and a subordinate clause with a pronominal subject. Deeply embedded questions consisted of a WH word followed by *do you think* followed by two complement clauses, the first consisting of the pronoun *he* or *she* and the verb *said*, and the second containing a pronominal subject, a verb, and another word. Since these sentences differ in lexical content, we used declarative sentences as controls. The declaratives contained the same lexical material as the interrogatives, except that they lacked an auxiliary and the WH word was replaced with a noun phrase or prepositional phrase as appropriate. A full list of the test sentences is given in Appendix 1.

All the test sentences were nine words long. There were three sentences in each condition, giving a total of 18 test sentences, nine interrogatives

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

and nine declaratives. In addition, there were eight simpler practice sentences (four interrogatives and four declaratives).

Associated with each sentence was a picture depicting the people and objects mentioned in it. The picture was shown to the children before they heard the test sentence, in order to maintain their interest and help them remember the sentences.

The child's interlocutor during the experiment was Dobbin-the-magic-pony, a stuffed toy with a loudspeaker hidden inside. After the child attempted to imitate a sentence, the experimenter used a remote control to play a pre-recorded comment, so that Dobbin appeared to be responding appropriately to the child's utterances.

2.1.4. *Procedure.* The children were tested individually in a quiet room in their school by a female experimenter. The testing sessions typically lasted about ten minutes, and were audio-recorded for later checking. At the beginning of the session the experimenter introduced Dobbin-the-magic-pony, and explained to the child that they were going to play a game:

We are going to play a copying game, OK? I will ask a question about a picture, and you have to ask Dobbin exactly the same question. It's very important that you say exactly the same thing. Can you do this?

When the child agreed to play the game, the experimenter produced the first practice question, showed the child the corresponding picture, and asked him/her to repeat it. If the child imitated the question correctly, the experimenter used the remote control to play a pre-recorded answer using the loudspeaker hidden inside Dobbin; if not, the experimenter repeated the question until the child was able to imitate correctly.

Experimenter:	*What is the boy doing?*
Child:	*What is the boy doing?*
Dobbin:	*He's trying to scare his sister.*

After four practice questions, the experimenter proceeded with the test items. During the test, Dobbin responded after the child produced a complete sentence, whether or not it was the same as the model. If the child was unable to repeat the entire sentence after a single presentation, the experimenter repeated it a second time.

The declaratives were tested using a similar method. First, the experimenter explained the rules of the game:

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

In this game, I will say something about the picture, and you have to repeat exactly the same thing. Do you want to try? Remember, you have to repeat exactly the same thing.

Again, the instructions were followed by four practice items and the test items themselves. The procedure was exactly the same, except that this time Dobbin commented on the statement the child repeated:

Experimenter:	*The girl is pushing the boy.*
Child:	*The girl is pushing the boy.*
Dobbin:	*That's naughty!*

The order of the two 'games' was counterbalanced across children. Within each block, the sentences were presented in a different random order for each child.

2.1.5. *Scoring.* The dependent variable was the number of correctly repeated sentences. Responses containing false starts (as in 6) and self-corrections (7) were coded as correct if there were no other errors.

(6) ⟨*what does the man*⟩ *[/] what does the man really hope they will like?*
(7) *what do you think they [//] the boys will really like?*

2.2. *Results*

Information about the mean number of correctly repeated sentences in each condition (out of a maximum score of 3) is given in Table 2. To determine whether the children performed differently across sentence types a 2 (construction type) × 3 (prototypicality) ANOVA was carried out. The results showed significant main effects of construction, $F(1, 33) = 5.63$, $p = 0.024$, $\eta_p^2 = 0.15$ and prototypicality, $F(2, 66) = 66.04$, $p < 0.001$, $\eta_p^2 = 0.67$. The interaction between construction and prototypicality approached significance: $F(2, 66) = 2.99$, $p = 0.057$, $\eta_p^2 = 0.08$.

Table 2. *Mean number (standard deviation) of correctly repeated sentences*

Condition	Questions mean (*SD*)		Declaratives mean (*SD*)	
Prototypical	1.74	(1.11)	2.21	(0.91)
Unprototypical	0.82	(0.90)	1.03	(1.00)
Deeply embedded	0.44	(0.75)	0.38	(0.70)

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

The results show that the children were significantly more accurate at repeating declaratives than questions. Planned repeated t-tests were performed to investigate the significant effect of prototypicality. As predicted by the lexically specific template hypothesis, performance on prototypical questions was significantly better than on unprototypical or deeply embedded questions (prototypical question v. unprototypical question: $t(33) = 5.34$, $p < 0.001$; prototypical question v. deeply embedded question, $t(33) = 6.77$, $p < 0.001$). There was also a difference between unprototypical and deeply embedded questions, $t(33) = 2.13$, $p = 0.04$; note, however, that this difference is no longer significant after the Bonferroni correction for multiple comparisons.

However, a similar pattern of results was found for declarative sentences. The children imitated prototypical declaratives better than both unprototypical declaratives and deeply embedded declaratives (prototypical declarative v. unprototypical declarative, $t(33) = 7.59$, $p < 0.001$; prototypical declarative v. deeply embedded declarative: $t(33) = 9.79$, $p < 0.001$), and unprototypical declaratives better than deeply embedded declaratives ($t(33) = 3.85$, $p = 0.001$).

There was also a marginally significant interaction so planned *t*-tests were performed to investigate this further. Children were significantly more accurate at imitating prototypical declaratives than prototypical questions; $t(33) = 2.69$, $p = 0.01$. However, there were no significant differences in their accuracy at imitating the two unprototypical constructions, $t(33) = 1.49$, $p = 0.15$, or between the two deeply embedded constructions, $t(33) = 0.42$, $p = 0.68$.

2.3. *Discussion*

Study 1 showed that children find prototypical declaratives the easiest construction type to imitate. Prototypical questions were significantly harder for the children to imitate. However, both were significantly easier than unprototypical declaratives and questions, both of which children found difficult (on average they imitated only 1 out of 3 correctly). Multiple embedded constructions were the most difficult, with very few correct imitations.

The experiment thus revealed a strong prototypicality effect for *both* interrogatives and declaratives. There are two possible explanations for these results. First, it is possible that children use lexically specific templates to produce both declaratives and questions at this age. This interpretation is supported by work that suggests that declaratives with verb complement clauses produced spontaneously by young children are quite formulaic (Bloom et al. 1989; Diessel 2004); and an experimental study

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

by Kidd et al. (2006) which found that children aged from 2;10 to 5;8 repeated complex declaratives more accurately when the matrix clause contained a verb which frequently occurs with sentential complements rather than in other syntactic constructions. Thus, it is possible that children have lexically specific templates for both constructions.

An equally plausible alternative interpretation is simply that the children found it easier to repeat sentences with high frequency verbs, possibly because they are more familiar. Study 2 was designed to discriminate between these two possibilities.

3. Study 2

Study 2, like study 1, was designed to compare children's performance on prototypical and unprototypical questions with long distance dependencies and their declarative counterparts. In contrast to study 1, however, the prototypical and unprototypical variants of both constructions contained exactly the same lexical material. If the prototypicality effects observed in study 1 were due merely to the lexical properties of the test sentences, they should disappear once these are controlled for. If, on the other hand, children have lexically specific templates for both constructions, we should find prototypicality effects for questions and for declaratives.

A second issue investigated was the age at which children develop verb-general knowledge about the two constructions. According to the constructivist view of language acquisition, development proceeds from lexically specific formulas to more abstract patterns, and is not necessarily synchronous: that is to say, abstract patterns may emerge at different times for different constructions. We know that verb-general knowledge about basic argument-structure constructions emerges in the third year of life or even earlier (Tomasello 2003; Goldberg 2006). The results of study 1 and the Kidd et al. study suggest that even as late as 5, children's knowledge about complementation may still be expressed in terms of verb-specific patterns such as *NP think S, NP say S.* Since abstraction is thought to be largely driven by high type frequency (Bybee 2001; Tomasello 2003), and since LDD questions are much more stereotypical than declaratives with verb complement clauses (Verhagen 2005; Dąbrowska in prep.), a fully general LDD construction should emerge even later, or not at all: it is possible that even adults rely on lexically specific patterns for this complex structure (cf. Verhagen 2005; Dąbrowska 2008). Thus, we expect to find an interaction between prototypicality, construction and age: specifically, prototypicality effects for declaratives, but not for questions, should disappear, or at least diminish, in older children.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

3.1. *Method*

3.1.1. *Participants.* Thirty-seven monolingual English-speaking children participated in the study, 18 five-year-olds (aged 4;8-5;9, mean 5;3, 10 girls and 8 boys) and 19 six-year-olds (aged 6;0-6;9, mean 6;5, 6 girls and 13 boys).

The children were recruited from a primary school in the North-West of England. One five-year-old did not attempt to repeat any of the sentences and was excluded from the analysis.

3.1.2. *Design.* The experiment employed a $2 \times 2 \times 2$ mixed-subjects design with two within-subject independent variables: prototypicality (with two levels: prototypical and unprototypical) and construction type (declaratives and questions) and one between subjects independent variable of age (five-year-olds and six-year-olds). The dependent measures were, for each child, the number of correctly repeated sentences.

3.1.3. *Materials.* Four types of sentences were used in the experiment: prototypical and unprototypical LDD questions, and their declarative counterparts. Prototypical questions had the hypothesised LDD question formula (*what do you think* or *what did you say*) in the main clause and a subordinate clause consisting of a heavy NP (4 words), an adverb and a verb which can take sentential complements (either *hope* or *expect*). The unprototypical questions contained exactly the same lexical material, but the content words which appeared in the main clause in the prototypical variant now appeared in the subordinate clause, and vice versa (see Table 3).

In the declarative counterparts of LDD questions, the WH word was replaced with a pro-form (*so* or *it*). To keep the number of words the same in both constructions, the auxiliary *will* was added to the subordinate clause. There were four items in each condition, giving a total of 16 sentences (plus six fillers and three sentences used for warm-ups). A full list of the test sentences is given in Appendix 2.

Table 3. *Examples of sentences used in Study 2*

Condition	Example
Prototypical question	What do you think the funny old man really hopes?
Unprototypical question	What does the funny old man really hope you think?
Prototypical declarative	I think the funny old man will really hope so.
Unprototypical declarative	The funny old man really hopes I will think so.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

3.1.4. *Procedure* Examination of the recordings for Study 1 revealed that the experimenter tended to articulate the non-prototypical LDD questions more slowly and more clearly than the prototypical variants, and we were concerned that this might influence the results. Therefore, the model sentences used in Study 2 were pre-recorded by a research assistant who was unaware of the purpose of the study and who was instructed to take great care to use the same speed and prosody throughout. During the experiment, the recorded sentences were played (in random order) using a loudspeaker hidden inside Dobbin. Each model sentence was followed by a beep, and the child was asked to repeat the sentence after the beep. Once the child had imitated the sentence, the experimenter provided an appropriate response (an answer to the question or a comment, as in Study 1). The reason for the beep was to introduce a short delay, and thus make the task slightly more difficult, since we were testing older children.

Apart from these two changes, the procedure was the same as in Study 1.

3.2. *Results*

3.2.1. *Target responses.* Table 4 shows the mean number of correctly repeated sentences (out of a maximum of four correct responses). These data were analysed using a $2 \times 2 \times 2$ ANOVA with the within participants factors of construction (declarative, question) and prototypicality (prototypical, unprototypical) and the between-participants factor of age (5-year-olds, 6-year-olds). The analysis revealed no significant main effects and no interactions. The main effect of prototypicality neared significance ($F(1, 34) = 3.50$, $p = 0.070$, $\eta_p^2 = 0.09$), with more correct responses for prototypical than unprototypical sentences. No other effects neared significance.

As is evident from these figures, the children found the task extremely difficult. In fact, 14 out of 17 five-year-olds and 13 out of 19 six-year-olds did not repeat a single sentence correctly: thus the differences reported

Table 4. *Mean number (standard deviation) of correctly repeated sentences (study 2, strict scoring)*

Condition	5-year-olds (*SD*)		6-year-olds (*SD*)	
Prototypical question	0.53	(1.18)	0.21	0.535
Unprototypical question	0.41	(1.18)	0.16	0.375
Prototypical declarative	0.29	(0.69)	0.47	1.020
Unprototypical declarative	0.24	(0.56)	0.11	0.459

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

above are attributable entirely to data from the remaining 9 children. The difficulty could be to the delay or the greater complexity of the experimental sentences. It is also possible that imitating a "disembodied" computer-produced voice (rather than simply imitating the person sitting next to them) requires greater concentration. It should also be pointed out that many of the errors that the children produced—for instance, omission of the adverb or one of the adjectives—are clearly uninformative with regard to their knowledge about question formation and complementation. We therefore reanalysed the data using a more focussed scoring method in which the child was given credit for a sentence if the only error(s) involved (a) omission or placement of the adverb, or substitution of a different adverb and/or (b) omission of the determiner or the adjective(s) inside the heavy NP, or substitution of a different adjective. Thus, under the new scoring system, all the responses given in (8) were coded as correct imitations of the target sentence *What does the pretty little girl really expect you said?*

(8) (a) What does the pretty little girl expect you said? [omission of adverb]
(b) What does the little girl expect you said? [omission of adverb and adjective]
(c) What does the really pretty girl expect you said? [misplacement of adverb and omission of adjective]
(d) What does the small little girl probably expect you said? [substitution of adverb and substitution of adjective]

The number of correct responses using this scoring method are given in Table 5. An ANOVA on these figures revealed a significant main effect of construction, $F(1,34) = 6.47$, $p = 0.016$, $\eta_p^2 = 0.16$, with the children correctly imitating more question than declaratives, and a significant main effect of prototypicality, $F(1, 34) = 5.82$, $p = 0.021$, $\eta_p^2 = 0.15$, with performance better on prototypical than unprototypical sentences, as predicted by the lexically specific template hypothesis. The main effect of age was not significant. However, there was a significant interaction

Table 5. *Mean number (standard deviation) of correctly repeated sentences (study 2, focused scoring)*

Condition	5-year-olds (*SD*)		6-year-olds (*SD*)	
Prototypical question	1.35	(1.46)	1.47	(1.22)
Unprototypical question	1.24	(1.25)	1.00	(0.94)
Prototypical declarative	0.71	(1.05)	1.53	(1.17)
Unprototypical declarative	0.47	(0.87)	1.00	(1.20)

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

between construction and age, $F(1, 34) = 7.51$, $p = 0.010$, $\eta_p^2 = 0.18$. The five-year-olds performed better on questions than on declaratives (prototypical declarative v. prototypical question: $t(16) = 32.67$, $p = 0.02$; unprototypical declarative v. unprototypical question: $t(16) = 2.89$, $p = 0.01$). However, by six years, the children's accurate imitation of declaratives had improved and the children were equally good on both constructions (prototypical declarative v. prototypical question: $t(18) = 0.21$, *ns*; unprototypical declaratives v. unprototypical question: $t(18) = 0.00$, *ns*).[5]

The younger children's better performance on interrogative utterances is surprising, since questions are less frequent than declaratives and commonly regarded as more complex syntactically: in a generative framework, for example, WH questions require WH movement and T to C movement (subject-auxiliary inversion). Further research will be necessary to determine whether this is a genuine developmental effect, and, if so, to explore the reasons for it. It is possible that the difference is attributable to our test materials. As explained in the Method section, the subordinate clause in the declaratives, but not in interrogatives, contained the auxiliary *will*, and the children sometimes omitted it, placed it in the main clause, or substituted a different auxiliary in its place. Such errors were relatively infrequent, with a mean frequency in the five-year-old group of 0.35 instances per child for both prototypical and unprototypical declaratives,[6] so they cannot fully account for the interrogative advantage in this age group. However, it is possible that the fact the subordinate clause in the declarative sentences always referred to an event which occurred at a different time than the event described in the main clause was an additional source of difficulty and thus contributed to errors elsewhere in the sentence. (Note that in Study 1, where all sentences had *will* in the subordinate clause, we observed the opposite pattern: children performed better on declaratives than on interrogatives.) Whatever the reason for the differences in performance on the two constructions in the five-year-

5. To ensure that the results reported in the preceding section were not simply an artefact of the scoring method we also recoded the data from Study 1 using the new method and conducted a second ANOVA on these figures. The results were similar to those reported in the main text, except that the main effects are slightly larger (for construction, $F(1, 33) = 11.49$, $p = 0.002$, $\eta_p^2 = 0.26$; for prototypicality, $F(2, 66) = 82.15$, $p < 0.001$, $\eta_p^2 = 0.71$), and the construction x prototypicality interaction is now significant ($F(2, 66) = 4.28$, $p = 0.018$, $\eta_p^2 = 0.12$).
6. This is the number of responses in which children omitted, misplaced, or replaced *will* and made no other errors (apart from those allowed in the focused scoring system). In other words, the scores for declaratives given in Table 5 would increase by 0.35 if children were given credit for these responses.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

old group, it is important to stress that they do not affect our findings about prototypicality, since the critical comparisons were between prototypical and unprototypical interrogatives (neither of which contained *will*) and prototypical and unprototypical declaratives (which both contained *will*).

3.2.2. *Errors.* The children made a variety of errors, mostly omissions or substitutions of lexical material. Most of the errors were quite unsystematic; three types of incorrect response, however, recurred in a number of children and provide some additional clues about the source of their difficulties with the constructions under investigation. We discuss these in more detail in this section.

Monoclausal responses. The children sometimes produced a simple clause instead of a complex one. These errors can be divided into three types:

– main clause only (e.g., *What does the tall woman expect?* for *What does the tall young woman probably expect you think?*)
– subordinate clause only (e.g., *What does the little boy hope?* for *What did you say the scared little boy probably hopes?*)
– amalgam of main and subordinate clause (e.g., *What does the funny old man think?* for *What does the funny old man really hope you think?*)

The frequencies of these three types of errors in each condition, collapsing the data across children, are given in Table 6.

Table 6. *Frequency of monoclausal responses in Study 2*

	Prototypical questions	Unprototypical questions	Prototypical declaratives	Unprototypical declaratives
Main clause only	0	7	0	8
Subordinate clause only	39	0	35	0
Amalgam of main and sub. cl.	1	8	0	3

Subordinate clause only responses were the most common, accounting for 74 percent of all monoclausal responses, and show a striking pattern: they occur only with the prototypical variant of each construction. This makes sense: the main clause in the prototypical question and prototypical declarative condition contains a light verb functioning as an epistemic marker (Thompson 2002; Verhagen 2005); thus, in imitating the subordinate clause only, the child repeats the gist of the stimulus sentence,

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

showing that s/he has understood it. Main clause only responses, in contrast, occurred only with unprototypical variants. Such responses are more difficult to interpret. In this case, it is the main clause that contains the semantically 'heavier' verb, so one could argue that the main clause contains the main thrust of the question. However, such responses could also arise if the child simply repeated the first (and hence most salient) clause in the stimulus sentence, without processing the subordinate clause at all. Finally, amalgams of the main and subordinate clause clearly involve a change in meaning, and thus indicate that the child had not understood the stimulus sentence. Such errors occur overwhelmingly (over 90 percent of the time) in unprototypical variants, especially unprototypical questions. In short, while monoclausal responses are considerably more frequent with prototypical variants of both constructions, the distribution of errors strongly suggests that it is the unprototypical variants which the children find more difficult to understand.

Medial WH questions. Like the children tested by Thornton and Crain (1994), the children in this study sometimes inserted an additional WH word at the beginning of the subordinate clause: for instance, one child produced (10) in response to the prompt in (9):

(9) *what does the tall young woman probably expect you think?*
(10) *what does the tall young woman probably expect what you think?*

There were nine errors of this kind, made by six children. Interestingly, all of them occurred in the unprototypical question condition. Since this is clearly an immature form, the fact that such errors occurred only in unprototypical questions suggests that they were causing the children more difficulty than the prototypical variant.

Lexical substitution errors involving verbs. Another common error involved replacing the verb in the main or subordinate clause with another verb, usually another complement-taking verb used in the experiment. In the unprototypical variants replacing the main clause verb with *think* or *say* makes the sentence more similar to the prototype. In the prototypical variant, on the other hand, the main clause verb is *think* or *say*, so replacing it with another verb makes the sentence less prototypical. Thus, reliance on lexically specific templates would make children prone to make the first type of error but not the second one. There were 14 verb substitutions in the main clauses of unprototypical questions, 11 of which involved replacing the main clause verb with *think* or *say*, and only 4 in prototypical questions, two of which involved replacing one template verb (*say*) with another (*think*). In declaratives, there were 17 main clause verb substitution errors in unprototypical variants, 12 of which involved

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Table 7. *Frequency of lexical substitution errors in Study 2*

Type of error	Prototypical questions	Unprototypical questions	Prototypical declaratives	Unprototypical declaratives
Simple substitution in main clause	4*	14	9**	17
Simple substitution in subordinate clause	2	6	7	5
Reversal of main and subordinate verb	0	3	0	12

*2 of these are *think* for *say* substitutions
**5 of these are *think* for *say* substitutions

replacing the main clause verb with *think* or *say*, and 9 in prototypical variants, 5 of which were *think* for *say* substitutions. Thus, only six main clause verb substitution errors (out of a total of 44) involved replacing the verb in the hypothesised template with another verb.[7] Although there were not enough errors for statistical analysis, the trend suggests that children had a tendency to choose prototypical verbs in the place of non-prototypical ones, which would be consistent with the lexically specific template hypothesis.

Note, however, that there is a confound: *think* and *say* are also the most frequent complement taking verbs, and have more general meanings than *hope* and *expect*, so the observed pattern of errors could be a result of the child simply substituting a more basic verb for a less basic one. If this were the case, we would expect to find the opposite pattern in the subordinate clause, where the verb is *hope* or *expect* in the prototypical variants and *think* or *say* in the unprototypical variants: that is to say, substitution errors in the subordinate clause should be more frequent in the prototypical variants. This is clearly not the case (see Table 7): verb substitutions in the subordinate clause are, if anything, more frequent in unprototypical sentences.

A third type of substitution error involved reversing the two verbs, i.e., putting the main clause verb in the subordinate clause and the

7. In three of these, the child replaced the main clause verb with the verb that was used in the subordinate clause, thus using the same verb in both clauses (e.g., *I hope the funny old man will hope so* for *I think the funny old man will really hope so*). In the remaining three, the child used the main clause verb from the immediately preceding sentence. Thus, such responses are best regarded as anticipation and perseverance errors respectively.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

subordinate verb in the main clause. (Note that reversal errors were excluded from the counts of simple substitution errors.) Reversal errors in prototypical sentences would make them less prototypical. Reversal errors in unprototypical sentences would make them more prototypical. As shown in Table 7, there were 14 reversal errors in our data all in non-prototypical utterances (3 in questions, 11 in declaratives). Again, the trend is to replace a non-prototypical verb with a prototypical one. Thus, the pattern of lexical substitution errors also supports the hypothesis that the prototypical variants of both constructions are more basic.

4. Study 3

Study 2 revealed prototypicality effects for both questions and declaratives. These effects cannot be attributed simply to the lexical properties of the stimuli, since the prototypical and unprototypical variants of the experimental sentences contained exactly the same lexical material. However, the predicted interaction with age did not occur: both five- and six-year-olds performed better on prototypical variants of both constructions, suggesting that both groups rely on lexically specific templates for declaratives as well as for interrogatives.

It is possible, of course, that complementation constructions continue to develop after age 6. To investigate this possibility, we administered a version of the repetition task to adults. If adults rely on lexically specific templates to produce and understand LDD questions but have a more general complementation pattern for declaratives, we would expect to find an interaction between construction type and prototypicality: specifically, adults should perform better on prototypical than unprototypical LDD questions, while there should be no corresponding difference, or a much smaller difference, in performance on declaratives. Study 3 was designed to test this prediction.

4.1. *Method*

4.1.1. *Participants* Nine adults (3 males and 6 females) aged between 30 and 50 participated in the experiment. All spoke English as their first language and were employed by a university in the north of England, either as lecturers or as administrative staff.

4.1.2. *Design.* The experiment employed a 2×2 within-subjects design with two within-subject independent variables: prototypicality (with two levels: prototypical, unprototypical) and construction (declaratives and questions). The dependent measures were, for each participant, the number of correctly repeated sentences.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Table 8. *Mean number (standard deviation) of correctly repeated sentences (Study 3)*

Condition	No. correct	(*SD*)
Prototypical question	3.22	(0.83)
Unprototypical question	2.22	(1.48)
Prototypical declarative	2.56	(0.73)
Unprototypical declarative	2.67	(1.12)

4.1.3. *Materials and procedure.* The materials and procedure were identical to those for study 2 with one difference. Since a pilot study showed that the task used in Study 2 was too easy for adults, the participants were asked to count backwards from 10 to 1 before attempting to repeat each sentence. This introduced a delay of about 10 seconds and prevented them from rehearsing the test sentence during the delay. Also, we did not use the toy interlocutor: the participants were simply asked to repeat each test sentence. The 'strict' scoring method was used when coding the results: that is to say, any omission or change to the stimulus sentence was coded as incorrect.

4.2. *Results and discussion*

Table 8 shows the number of correctly repeated sentences in each condition. These results were analysed using a 2 (construction) $\times$ 2 (prototypicality) ANOVA. The main effects of construction and prototypicality were not significant. However, as predicted, there was a significant interaction between construction type and prototypicality, $F(1,8) = 8.16$, $p = 0.021, \eta_p^2 = 0.51$: prototypical questions were repeated correctly significantly more often than unprototypical questions ($t(8) = 2.68$, $p = 0.028$), while there was no difference in performance on declaratives ($t(8) = 0.26$, *ns*). This suggests that even adults make use of lexically specific templates for LDD questions, but not declaratives, with finite complement clauses.

5. General discussion

The differences in performance on prototypical and unprototypical variants described above are fully compatible with a lexical template account. However, it should be noted that there is yet another explanation which may be able to account for our results. The prototypical and unprototypical sentences in our experiments were not in fact identical in form, in that the adverb modified the subordinate clause verb in the former and the main clause verb in the latter. Thus, the observed differences

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

could conceivably be attributed to form frequency, **if** the prototypical structures (i.e., WH Aux NP V [NP Adv V] and NP V [NP (Aux) Adv VP]) turned out to be more frequent than the 'unprototypical' variants (WH Aux NP Adv V [NP V] and NP Adv V [NP (Aux) Adv VP]).

One problem with this alternative explanation is that such structures are exceedingly rare. None of the 325 LDD questions in child-directed speech that we extracted from the Manchester corpus contained an adverb premodifying either the main or the subordinate verb; there are also no instances of an adverb premodifying either the main or the subordinate verb in a declarative sentence with a finite verb complement clause. A somewhat larger sample of 423 instances of LDD questions extracted from Spoken BNC contained one question in which the main clause verb was premodified by an adverb and four in which there was an adverb premodifying the subordinate verb. The sample is too small to determine if the difference is statistically significant; however, even if it was, it is highly unlikely that speakers are sensitive to such tiny differences in frequency. Spoken BNC contains about 10 million words, which means that the normalised frequencies of the sequences WH Aux NP Adv V [NP V] and WH Aux NP V [NP Adv V] are 0.1 and 0.4 per million words respectively.

A second problem with the account is that it cannot explain the interaction between prototypicality and construction type found in the adults. As we saw in the preceding section, adults were significantly better at repeating prototypical than unprototypical LDD questions, but showed no corresponding differences for declarative sentences. However, the BNC data indicate that declarative sentences with finite verb complement clauses are similar to LDD questions in that they are about four times more likely to contain an adverb premodifying the subordinate verb than the main verb. (A matched sample of 423 declaratives with finite verb complement clauses from Spoken BNC contained 22 instances of the former and only 5 of the latter.) We conclude that form frequency cannot account for the observed pattern of results.

6. Conclusion

Our results indicate that children continue to rely on lexically specific templates for both LDD questions and declaratives with finite verb complement clauses as late as age 6. Study 1 revealed that four-year-old children imitate prototypical variants of both constructions more accurately than unprototypical variants (which had the same grammatical structure but differed in lexical content) as well as deeply embedded sentences (which had a more complex syntactic structure). Study 2 replicated this re-

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

sult with older children in a design which, importantly, controlled for the possibility that the prototypical sentences are easier to repeat because they contain higher-frequency verbs. Study 3 demonstrated that even adults show the prototypicality effect for questions, but not for declaratives.

These results are consistent with the predictions of the usage-based approach, according to which children's knowledge about complementation constructions is best captured in terms of lexically specific templates acquired by generalizing over attested instances of the relevant constructions. They are difficult to accommodate in the generative framework, which assumes that children have abstract syntactic representations and general operations such as WH movement. Generative accounts predict that children should perform equally well on prototypical and unprototypical variants of both constructions, since these were matched for syntactic complexity. Syntactic complexity may play a role in performance: in Study 1, the children performed slightly more accurately on 'unprototypical' questions and declaratives than on 'deeply embedded' sentences. Such effects, however, are relatively small in comparison to the purely lexical effects, and they can also be accommodated in a usage based framework: as pointed out earlier, deeply embedded sentences are also less prototypical instances of the relevant constructions.

More strikingly, our results suggest that even adults rely on lexically specific templates for questions with long distance dependencies. This accords well with previous research on LDD questions in the usage-based framework, which has shown that (i) LDD questions in adult speech and writing are also very stereotypical (Verhagen 2005; Dąbrowska in prep.); (ii) prototypical LDD questions, i.e., questions of the form *WH do you think S-GAP?* and *WH did you say S-GAP?*, are produced more fluently than non-prototypical questions (Dąbrowska in prep.); and (iii) prototypical LDD questions are judged to be more acceptable than non-prototypical questions—and LDD questions which depart from the prototype in several respects are judged as bad as some clearly ungrammatical sentences (Dąbrowska 2008).

The fact that speakers use lexically specific templates does not of course preclude the possibility that they have more abstract constructions as well. However, our results, and the research cited earlier in this paper, show the lexically specific variants have a privileged status, in that they are ontogenetically earlier, apparently easier to access, and preferred by speakers.

Received 28 January 2008
Revision received 14 July 2008

University of Sheffield
University of Liverpool
University of Manchester

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Appendix 1: Sentences used in Study 1

Prototypical LDD questions	*Prototypical declaratives*
What do you think the boys will really like?	I think the boys will really like their shoes.
Where do you think the girls will actually go?	I think the girls will actually go to school.
Who do you think really likes these smelly socks?	I think my neighbour really likes these smelly socks.
Unprototypical LDD questions	*Unprototypical declaratives*
What does the man really hope they will like?	The man really hopes they will like their shoes.
Where does the girl actually expect they will go?	The girl actually expects they will go to school.
Who does the boy really believe likes smelly socks?	The boy really believes my neighbour likes smelly socks.
Deeply embedded LDD questions	*Deeply embedded declaratives*
What do you think he said they will like?	I think he said they will like their shoes.
Where do you think she said they will go?	I think she said they will go to school.
Who do you think he said likes smelly socks?	I think he said his neighbour likes smelly socks.

Appendix 2: Sentences used in Studies 2 and 3

Prototypical LDD questions	*Prototypical declaratives*
What do you think the funny old man really hopes?	I think the funny old man will really hope so.
What do you think the tall young woman probably expects?	I think the tall young woman will probably expect it.
What did you say the pretty little girl really expects?	I said the pretty little girl will really expect it.
What did you say the scared little boy probably hopes?	I said the scared little boy will probably hope so.
Unprototypical LDD questions	*Unprototypical declaratives*
What does the funny old man really hope you think?	The funny old man really hopes I will think so.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

What does the tall young woman probably expect you think?	The tall young woman probably expects I will think so.
What does the pretty little girl really expect you said?	The pretty little girl really expects I will say it.
What does the scared little boy probably hope you said?	The scared little boy probably hopes I will say it.

References

Ambridge, Ben and Adele E. Goldberg
2008 The island status of clausal complements: Evidence in favor of an information structure explanation. *Cognitive Linguistics* 19(3), 357–389.

Bloom, Lois, M. Rispoli, B. Gartner and J. Hafitz
1989 Acquisition of complementation. *Journal of Child Language* 16,101–120.

Bybee, Joan
2001 *Phonology and Language Use*. Cambridge: Cambridge University Press.

Cheng, Lisa Lai-Shen, and Norbert Corver (eds.)
2006 *Wh Movement: Moving On*. Cambridge, MA: MIT Press.

Chomsky, Noam
1977 On *wh*-movement. In Culicover, Peter W., Thomas Wasow, and Adrian Akmajian (eds.), *Formal Syntax*. Academic Press, New York, 71–132.

Culicover, Peter W.
1997 *Principles and Parameters: An Introduction to Syntactic Theory*. Oxford. Oxford University Press.

Dąbrowska, Ewa
2004 *Language, Mind and Brain. Some Psychological and Neurological Constraints on Theories of Grammar*. Edinburgh: Edinburgh University Press.

Dąbrowska, Ewa
in prep. Prototype effects in questions with unbounded dependencies.

Dąbrowska, Ewa
2008 Questions with unbounded dependencies: A usage-based perspective. *Cognitive Linguistics* 19, 391–425.

de Villiers, Jill
1995 Empty categories and complex sentences: The case of wh-questions. In Fletcher, Paul and Brian MacWhinney (eds.), *The Handbook of Child Language*. Oxford: Blackwell, 508–540.

de Villiers, Jill, Thomas Roeper, and Anne Vainikka
1990 The acquisition of long-distance rules. In Frazier, Lyn and Jill de Villiers (eds.), *Language Processing and Language Acquisition*. Dordrecht: Kluwer, 257–297.

Diessel, Holger
2004 *The Acquisition of Complex Sentences*. Cambridge: Cambridge University Press.

Frazier, Lyn and Charles Clifton, Jr.
1989 Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4, 93–126.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Goldberg, Adele E.
2006 *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Hawkins, John A.
1999 Processing complexity and filler-gap dependencies across grammars. *Language* 75, 244–285.

Kidd, Evan, Elena Lieven and Michael Tomasello
2006 Examining the role of lexical frequency in the acquisition and processing of sentential complements. *Cognitive Development* 21, 93–107.

Kluender, Robert and Marta Kutas
1993 Subjacency as a processing phenomenon. *Language and Cognitive Processes* 8, 573–633.

Kuczaj, Stan A.
1976 *-ing, -s* and *-ed:* A study of the acquisition of certain verb inflections. Doctoral dissertation, University of Minnesota.

Lieven, Elena V., Julian M. Pine and Gillian Baldwin
1997 Lexically-based learning and early grammatical development. *Journal of Child Language* 24, 187–219.

Lust, Barbara, Suzanne Flynn and Claire Foley
1998 What children know about what they say: Elicited imitation as a research method for assessing children's syntax. In McDaniel, Dana, Cecile McKee and Helen Smith Cairns (eds.), *Methods for Assessing Children's Syntax*. Cambridge, MA: MIT Press, 55–76.

MacWhinney, Brian
1995 *The CHILDES Project: Tools for Analyzing Talk*. Hillsdale, NJ: Lawrence Erlbaum.

Ouhalla, Jamal
1994 *Introducing Transformational Grammar: From Rules to Principles and Parameters*. London: Edward Arnold.

Philip, William and Jill de Villiers
1993 Monotonicity and the acquisition of weak Wh-islands. In Clark, Eve V. (ed.), *Proceedings of the Twenty-fourth Annual Child Language Research Forum*. Stanford: CSLI, 99–111.

Santelmann, Lynn, Stephanie Berk, Jennifer Austin, Shamitha Somashekar and Barbara Lust
2002 Continuity and development in the acquisition of inversion in yes/no questions: dissociating movement and inflection. *Journal of Child Language* 29, 813–840.

Slobin, Dan I. and Charles A. Welsh
1973 Elicited imitations as a research tool in developmental psycholinguistics. In Ferguson, C. and D. Slobin (eds.), *Studies of Child Language Development*. New York: Wiley.

Theakston, Anna, Elena Lieven, Julian Pine and Caroline Rowland
2001 The role of performance limitations in the acquisition of verb-argument structure. *Journal of Child Language* 28, 127–152.

Thompson, Sandra
2002 "Object complements" and conversation. Towards a realistic account. *Studies in Language* 26, 125–164.

Thornton, Rosalind
2008 Why continuity? *Natural Language and Linguistic Theory* 26, 107–146.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM

Thornton, Rosalind and Stephen Crain
1994 Successful cyclic movement. In Hoekstra, Teun and Bonnie D. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar*. Amsterdam: John Benjamins, 215–252.

Tomasello, Michael
2003 *Constructing a Language: A Usage-Based Theory of Child Language Acquisition*. Cambridge, MA: Harvard University Press.

Verhagen, Arie
2005 *Constructions of Intersubjectivity: Discourse, Syntax and Cognition*. Oxford: Oxford University Press.

Verhagen, Arie
2006 On subjectivity and 'long distance *Wh*-movement'. In Athanasiadou, Angeliki, Costas Canakis, and Bert Cornillie (eds.), *Various Paths to Subjectivity*, 323–346. Berlin: Mouton de Gruyter.

Brought to you by | University of Liverpool Sydney Jones Library
Authenticated | 138.253.100.121
Download Date | 8/31/12 11:31 AM