ductions. If, as Roeper (1990) suggests, children base-generate the Wh-phrase, projecting pro at the 'extraction site' in their Crossover questions, the appearance of the medial-Wh cannot be explained. According to this non-movement analysis, there would be no reason to expect a Wh-form to appear in what would normally be the path of Wh-movement. Even assuming a movement analysis, it remains to explain why an overt Wh-phrase appears where we might expect a Wh-trace. This is the topic of Chapter III.

Chapter III
Children's Exceptional Questions

This chapter focuses on several non-adult question types that children produced in the experiment eliciting Wh-extraction from tense embedded clauses. These non-adult utterances all involve an "extra" element in the Intermediate CP domain. There were children who asked questions with two occurrences of the Wh-phrase — the expected one in the Spec of the matrix CP, but, also, a second unexpected one in the CP domain of the embedded clause. That is, in situations when long-distance questions like (1) and (2) were appropriate, some children produced questions like (3) and (4), with two matching Wh-phrases. These questions will be referred to as 'medial-Wh' questions.

1. Who do you think is in the box?
2. What do you think Cookie Monster likes?
3. Who do you think who is in the box?
4. What do you think what Cookie Monster likes?

The term "CP domain" is intended to be neutral with respect to the position of the extra Wh-phrase being in CPSpec or the head, C.
In addition to medial-Wh questions, two of the children tested sometimes produced questions with two Wh-phrases that did not match in features. The question word that appeared in the matrix CP was ‘what’. In these children’s productions, a question with the meaning of (5) was phrased as (6):

(5) Who do you think jumped over the can?
(6) What do you think who jumped over the can?

Finally several other children produced questions with an overt complementizer in the intermediate COMP, giving rise to a ‘that-trace’ configuration that appears to violate the Empty Category Principle (ECP):

(7) Who do you think that’s in the box?

All of the question types with an ‘extra’ element in the CP domain appeared in questions with extraction from a tensed embedded clause. No child ever put an extra element in the intermediate CP domain in the experiment eliciting questions involving extraction from infinitival clauses. Instead, children produced the appropriate adult forms as indicated in (8) and (9). Utterances such as (10) and (11) did not occur.

(8) What do you wanna eat?
(9) Who do you want to eat the pizza?
(10) *What do you want what to eat?
(11) *Who do you want who to eat the pizza?

Since questions with two Wh-phrases and ‘that-trace’ questions are not grammatical in adult English, it is safe to assume that they are not present in the positive evidence children have for grammar formation. Why, then, do these non-adult utterances appear in some children’s speech? The goal of this chapter is to investigate whether the principles and parameters of Universal Grammar (Chomsky, 1981; 1986a; 1986b) can accommodate the findings. In this context, the ‘that-trace’ questions represent a challenge, since they are apparent violations of a principle of Universal Grammar.

To begin these deliberations, we ask whether children’s medial-Wh questions can be explained as a consequence of parameter theory. Perhaps the simplest story is that the non-adult responses reflect a parameter setting that is inappropriate for English, but appropriate for other natural languages. If so, the appearance of these questions in the absence of positive evidence could be explained.

Parameter theory has met with considerable success in explaining young English-speaking children’s non-adult ‘null subject’ utterances (Hyams 1986). Pursuing a similar line, we will weigh the evidence for interpreting children’s medial-Wh questions as the product of parameters that relate to Wh-movement. As the initial setting of a Wh-movement parameter, medial-Wh questions should reflect the appropriate setting for other languages. If so, we should find questions with an “extra” Wh-phrase in the embedded CP domain in other languages of the world. Thus, our point of departure is a survey of cross-linguistic Wh-movement options. Examination of data from German and Romani (McDaniel 1966; 1969) reveal two derivations that are superficially similar to children’s medial-Wh questions: (i) a short-distance movement structure with movement only to the embedded CP domain, i.e. ‘partial movement’, and (ii) a long-distance movement structure with movement to the matrix CP, termed ‘Wh-copying’.
Following the introduction to different Wh-movement structures, the parameters related to Wh-movement will be discussed. One set of parameters discussed in detail is proposed by de Villiers, Roepke and van Itterbeek (1988). In their view, ordered parameters initially limit child grammars to no movement, and then to a stage where only short-distance movement is available. Subsequently, children switch the setting of the parameter on the basis of positive evidence to allow long-distance movement. A parameterized version of the Subjacency Principle is also considered in this section. This parameter suggests that initially only short-distance movement is available. Returning to children’s medial-Wh questions, we consider whether these questions are like short-distance ‘partial movement’ or long-distance movement ‘Wh-copying’ structures. While the evidence points to questions like (6) as partial movement, the status of children’s medial-Wh questions like (3) and (4) is not clear. In this case, it is not obvious which Wh-phrase is the ‘true’ Wh-phrase. Our exploration of this structure begins by comparing the properties of children’s medial-Wh questions with the properties of partial movement and Wh-copying questions in German and Romani. A test is devised based on extraction of complex Wh-phrases (i.e. “which boy”) that reveals medial-Wh questions pattern like long-distance Wh-copying questions. This conclusion is further supported by a pattern that emerges in the developmental data of one child. In conclusion, the restricted movement options predicted by Wh-parameters are not supported. To the contrary, the full range of options made available by “Move a” appear to be present in the early grammars of some children.

This leads us to reject medial-Wh questions as reflecting a parameter setting restricting Wh-movement. The remainder of the chapter aims to explain children’s non-adult questions as their attempts to satisfy the Empty Category Principle (ECP). Proposals for ECP-related parameters are first examined within the Lasnik and Saito (1984) framework. In this framework, the proper government requirement of the ECP is achieved by either lexical government or antecedent government. One possibility that is investigated is that the clause of the ECP related to antecedent government is parameterized, with children initially adopting a setting that prohibits proper government across an S’ boundary. Thus the appearance of the medial-Wh is to circumvent the ECP. Other proposals which are considered relate the medial-Wh to ECP requirements on the initial trace: Children may think proper government must be satisfied by (only) antecedent government, or by a conjunctive form of the ECP, that initially requires traces to be both lexically governed and antecedent governed. On this view, positive data would cause a switch to the disjunctive formulation of the adult grammar. Ultimately these proposals are rejected because they fail to explain the complete pattern in children’s Wh-questions.

Having failed to account for the unexpected medial-Wh questions in parameteric terms, an alternative formulation of the ECP is considered. This is the version of the ECP proposed by Rizzi (1990). The ECP and the associated explanation of ‘that-trace’ effects in this framework are discussed in depth. Rizzi’s basic treatise is that the ECP is a requirement only of head government. According to his proposal, in questions extracting the subject from a tensed clause, a null complementizer can serve as a head governor for the subject trace, once it is licensed by a rule of Spec-Head agreement. The subject trace does not need an antecedent governor. Data supporting the rule of Spec-Head agreement and the way it varies across languages is presented. These data will be drawn on to argue for a particular formulation of the Spec-Head agreement rule in some child grammars.
Finally, after presenting the theory, we return once again to children’s non-adult questions. It is argued there that a unified account can be given for the appearance of all children’s exceptional questions within Rizzi’s framework of the ECP. The proposal is that children’s questions conform to the structural requirements of the ECP — children have, however, incorrectly concluded that the overt complementizer is the appropriate head governor for the subject trace. Although this conclusion is mistaken, this means that children’s questions reflect an endeavor to ‘preserve’ the ECP, not violate it. Similarly, medial-Wh questions are seen as an attempt to provide the syntactic representation necessary to meet the ECP head-government requirement. This explanation provides insight into the developmental sequence observed in children’s questions: Initially the medial-Wh and ‘that’ are obligatory in subject and object questions, but later they occur only in cases of subject extraction. This change is presumed to occur when children realize that Spec-Head agreement is only required to satisfy the ECP for subject traces. The chapter concludes by noting that the data of the children who ask exceptional questions bear none of the earmarks of performance errors, and that an alternative performance-based explanation of the exceptional questions is unlikely.

**Wh-movement Across Languages**

This section introduces cross-linguistic data on Wh-movement, focussing on the syntactic structures different languages use to ask ‘long-distance’ questions. The term ‘long-distance question’ will be used throughout the chapter to refer to questions in which the Wh-phrase is in an embedded clause at D-structure. It should be kept in mind, however, that the term is neutral with respect to the derivation of the question. Thus a ‘long-distance question’ may be a structure with (a) no movement, (b) short-distance movement, or (c) long-distance movement of the Wh-phrase. According to Lasnik and Saito (1984), no matter which of these options a language selects, the following are universally true at LF:

(12) A [+Wh] COMP must have a [+Wh] head
A [-Wh] COMP must not have a [+Wh] head
All Whs must be in a [+Wh] COMP at LF

We will survey, briefly the Wh-movement options made available by Universal Grammar. This survey will anticipate the options that parameters on Wh-movement would need to account for.

**Movement Options**

In languages such as Chinese, Japanese and Korean, Wh-movement does not take place in the syntactic component. All Wh-questions, including long-distance questions, are framed with the Wh-phrase in-situ. In these languages, Wh-movement is argued to take place at the level of Logical Form (LF) (Huang, 1982). The Wh-phrase is then in a position to c-command, and thus take scope over the sentence, as required for semantic analysis. The following long-distance questions from Japanese illustrate the no-movement option. In example (13), the Wh-phrase dare (who) remains in subject position at S-structure. In (14), the Wh-phrase dono hon (which book) remains as object in-situ.
Other languages, such as American Sign Language (ASL) (Lillo-Martin 1988), Polish (Willem, 1989) and Russian (Radford, 1981; Freidin and Quicoli, 1989) allow syntactic Wh-movement, but it is tightly constrained. For the most part, syntactic movement in these languages is restricted to short-distance movement of the Wh-phrase. ASL is unusual in that short-distance movement is restricted to matrix clauses. Long-distance extraction of the Wh-phrase from an embedded clause is not possible; the Wh-phrase must be left in-situ. (This option is also available for matrix questions). Example (15) below, shows the ungrammaticality of long-distance extraction of the Wh-phrase.94

American Sign Language

(15) [[WHO] BILL FEEL JOHN 'LIKE' t]  
'Who does Bill think John has a crush on t?'

Long-distance questions may be asked using two different structures; (i) a statement in the form of a simple declarative sentence can be followed by a matrix question as glossed in (16), or (ii) the Wh-phrase can remain in-situ, as in (17). ASL, then, has the option of filling the [+Wh] COMP at S-structure or LF:

94 In the ASL representations, the solid line represents a facial expression that indicates the scope of the question. The subscripts a, b, c show agreement, not relevant for our purposes.

(16) BILL FEEL JOHN 'LIKE' SOMEONE. WHO  
'Bill thinks John likes someone. Who is it?'  
'Who does Bill think John has a crush on t?'

(17) BILL FEEL JOHN 'LIKE' WHO  
'Who does Bill think John has a crush on t?'

In Polish, too, movement is restricted for the most part to short-distance movement. Long-distance movement is ungrammatical with one exception. Wh-phrases may be extracted out of infinitival clauses, as the following contrast makes evident.95

Polish

(18) Full Movement: Tensed Clause  
*Czolg Jan byl przekonany, ze Maria kupila t?  
'What was John convinced that Mary had bought?'

(19) Full Movement: Tensed Clause  
*Ktoz Jan wie, ze t? odwiedza Marie?  
'Who does John know that t visits Mary?'

(20) Full Movement: Infinitival Clause  
Czol Jan chcial [[ PRO kupic t?  
'What does John want to buy?'

95 Willem (1989) reports that long-distance movement is possible out of infinitival complements when the complementizer position is null. The class of verbs that subcategorizes for an infinitival complement (e.g. 'want') also allows long-distance movement out of tensed subjunctive complements.
The most natural counterpart to an English long-distance question involves two different Wh-phrases; how in the matrix clause and what in the embedded CPSpec.

(21) Jak myślisz, co Jan kupi?  
    "How do you think, what John will buy?"  
    = "What do you think John will buy?"

William (1989) is unclear about the structure, but seems to suggest that such questions are two matrix questions. In cases where the structure in (21) cannot be used, Polish frequently uses a structure where the questioned Wh-phrase moves to the embedded CPSpec, and no Wh-phrase is present in the matrix CPSpec.  

(22) Jan myślał, że co Maria przeczytała?  
    "John thought that what Mary read?"  
    = "What did John think that Mary read?"

The Polish data bear some similarity to the child data. They are alike in that long-distance extraction is possible out of infinitival clauses, and Wh-phrases only appear in the Intermediate COMP when extraction is from a tensed clause. In addition, there is a question form with two Wh-phrases. However, this form differs from children's medial-Wh questions in that the Wh-phrases do not match

66 William (1989) considers the possibility that jak (how) is the lexical spell out of a [ + Wh ] matrix CPSpec, in which case it would marking the scope of the questioned Wh-phrase that has moved to the embedded CPSpec. The partial movement structure, in which the Wh-phrase in the matrix CPSpec serves to mark scope will be discussed in the German and Romani section. William rejects the partial movement analysis for Polish because the structure in (21) can only be used with verbs of 'thinking' when the addressee is 'you'.

67 This structure could be likened to partial movement if one thinks of the scope marker as being null.

...in features, and the Wh-phrase that appears in the matrix COMP appears limited to a fixed form.

So far, we have examined data from languages which ask long-distance questions using structures with either no movement or short-distance movement. English contrasts with the languages surveyed so far. In English, long-distance questions are posed by obligatory long-distance movement of the Wh-phrase. Other languages are much more flexible in the movement options allowed. This raises an interesting question for parameter theory. Would we expect all movement options to be available in early grammars, or would expect movement options to be tightly constrained until positive evidence informs otherwise? I will argue in the course of the chapter that Universal Grammar makes all options available to the child. An example of a language which appears to allow every “movement” possibility is the Austronesian language, Palauan (Georgopolous, 1985). Palauan allows the questioned Wh-phrase to be either in-situ as in (23), in the embedded CPSpec position, as in (24), or in the matrix CPSpec, as illustrated in (25). The questions are all identical in meaning.  

66 It is inaccurate to report Palauan as having many movement possibilities since Georgopolous argues that Palauan is a language without movement. Her conclusion follows from her observations that Palauan allows violations of the Wh-Island constraints. To preserve Subjacency as a universal principle, her solution is to propose instead, that Wh-phrases in Palauan are base-generated. Since there is no movement, Subjacency is not violated. Naturally, this argument assumes that Subjacency is a condition on movement and not on representations, as suggested by McDaniel (1986,1989). I simply assume, for the purposes of illustration, that Palauan does have syntactic Wh-movement.
Partial Movement

In partial movement structures, the questioned Wh-phrase moves short-distance to the embedded CPSpec position, and another Wh-phrase is base-generated in the matrix CPSpec. The Wh-phrase in the matrix marks the scope of the ‘true’ Wh-phrase that is positioned in the embedded CPSpec.\textsuperscript{100} Examples of partial movement structures in German and Romani follow. In each case, the long distance movement counterpart (with a Wh-phrase in the matrix CPSpec only) is given for comparison. In the German example below, the extracted Wh-phrase is a PP ‘with whom’. This is because there are independent restrictions on movement out of tense-independent (indicative) clauses limiting extraction to PPs and adverbs. In both German and Romani, the scope marker corresponds to the Wh-word “what”; in German it is was, and in Romani, the scope marker is so.\textsuperscript{101}

**German**

(26) \textbf{Partial movement} \\
Was glaubt [\_Hans [\_mit was], \_Jakob jetzt tₙ spricht? \\
‘What does Hans believe with whom Jakob is now talking?’

(27) \textbf{Long-Distance Movement} \\
[\_mit was], glaubt [\_Hans \_tₙ dass \_Jakob jetzt tₙ spricht? \\
‘With whom does Hans think that Jakob is now talking?’

---

\textsuperscript{100} In Romani, the scope marker is also homophonous with the complementizer “that”.

\textsuperscript{101} McDaniel claims that Wh-extraction is restricted to adverbs and PPs for reasons of Case-Inheritance (see McDaniel 1986;1989 for details).
In German and Romani, partial movement is more tightly constrained than long-distance movement. Partial movement is not possible when the Wh-phrase is embedded from an infinitival clause or a tense-dependent clause. \(^{102}\) \(^{103}\) (The parallel situation cannot be tested with Romani, since it does not have infinitival clauses).

**German**

(30) **Long-Distance Movement: infinitival**

Wen 12 versucht [1p Hans [cp t 4 PRO t 4 zu bestechen?
Who is Hans trying to bribe?

(31) **Partial Movement: infinitival**

*Wen 12 versucht [1p Hans [cp t 4 PRO t 4 zu bestechen?
Who is Hans trying whom to bribe?

(32) **Long-Distance Movement: tense dependent**

Wen 12 will [1p Hans [cp t 4 des [1p Jakob t 4 besticht?
Who does Hans want Jakob to bribe?

(33) **Partial Movement: tense dependent**

*Wen 12 will [1p Hans [cp t 4 des [1p Jakob t 4 besticht?
Who does Hans want whom Jakob bribe?

Since Wh-phrases never occur in the Spec of infinitival CPs in German, McDaniel assumes that infinitival CPSpec positions are not specified [ +Wh ] . This makes the prediction that indirect questions should be impossible with infinitival clauses.

---

102 A tense-dependent clause is used to mean a clause that would be subjunctive in Romance.

103 As McDaniel (1986) notes, the partial movement structure violates the putative universals of Wh-movement of Lasnik and Saito (1984) in (13).

and, as McDaniel notes, the prediction is borne out, as the following German example shows.

(34) *Ich weiss nicht [cp t 4 PRO t 4 zu machen I don't know what to do.

This line of argumentation leads McDaniel (1989) to propose that infinitival and tense-dependent clauses in German, have no CPSpec position; C projects just one bar level. \(^{104}\) So far, the characteristics of the German and Romani partial movement data look very like the child data. As in these languages, when extracting from infinitival clauses, children do not allow a Wh-phrase in the embedded CPSpec. In particular, the partial movement questions look very similar to the questions like (6) ("What do you think who jumped over the fence?") that some children produced. Now let us turn to examine the second structure with two Wh-phrases, Wh-copying.

**Wh-Copying**

In addition to long-distance movement and partial movement structures, certain dialects of German (the Cologne area) and some dialects of Yugoslav Romani are reported to have a third long-distance question structure, Wh-copying.

In Wh-copying questions, the questioned Wh-phrase appears twice; in the Spec

\(^{104}\) From this it would follow that there could be no intermediate trace in full-movement from infinitival clauses. One would think that this would cause a Subjacency violation. But this cannot be so because the sentences are perfectly grammatical. McDaniel (1989) accommodates this problem in the Barriers (Chomsky, 1986b) framework by adjusting the notion of what constitutes a barrier. The proposal that there is no intermediate trace in the CPSpec of questions extracting from infinitival clauses is also made in Rampton (1990).
of the matrix CP and the Spec of the embedded CP. That is, the same Wh-phrase appears twice. Wh-copying occurs not only in long-distance questions, but also in relative clauses. Some examples of Wh-copying in German and Romani are given in (35) - (38):

**German**

(35) Wh-Copying: Long-Distance Question

Mit wen*] mit den*] spricht
\[gr] Hans*] mit whom/that Hans talks?

(36) Wh-copying: Relative Clause

Das ist [gr der Junge [gr mit [gr mit dem*] spricht
\[gr ich glaube [gr Hans*] That's the boy with whom I believe (that) Hans talked'.

**Romani**

(37) Wh-Copying: Long-Distance Question

\[gr ka*] mit [gr ku*] mit [gr Desari dikhla t*]
\[gr whom do you think whom Desar saw';

(38) Wh-copying: Relative Clause

\[gr a*] chava[gr ka*] mit [gr ku*] mit [gr Desari dikhla t*]
\[gr Here's the boy whom I think whom Desar saw'.

The data from German and Romani raise the question of whether Wh-copying is a form of partial movement or long-distance movement. That is, in Wh-copying structures, is the Wh-phrase in the Spec of the matrix CP the 'true' Wh-phrase, or is it a scope marker that happens to bear the same features as the true Wh-phrase that is in the embedded CP Spec? If Wh-copying is a variant of partial movement, and involves only short-distance movement of the Wh-phrase, then it should pattern the same way. There are points of similarity, but McDaniel shows that there are differences also. Let us begin with a point of similarity.

When extraction is from an infinitival clause, neither partial movement nor Wh-copying is possible, as is illustrated in examples (39) and (40):

**German**

(39) Wh-copying: infinitival

\[gr wen*] mit [gr versucht du [gr wan/dem*] PRO t_1
\[gr anzurufen?]]

\[gr Whom are you trying whom to call?]

(40) Partial Movement: infinitival

\[gr was*] mit [gr versucht du [gr wan/dem*] PRO t_1
\[gr anzurufen WHAT are you trying whom to call?]

The two structures diverge, however, in whether they allow extraction from tense-dependent clauses. Partial movement is not possible out of tense-dependent clauses, but the Wh-copying structure is grammatical. This contrast is shown in (41) and (42):\(^\text{105}\)

**German**

(41) Wh-Copying: tense-dependent

\[gr wen*] mit [gr willst du [gr wan/dem*] PRO t_1
\[gr ansruf?)]]

\[gr Whom do you want whom to call?]

(42) Partial Movement: tense dependent

\[gr was*] mit [gr willst du [gr wan/dem*] PRO t_1
\[gr ansruf WHAT do you want whom to call?]

There is also another point of contrast in partial movement and Wh-copying questions. The two structures differ in the kind of Wh-phrase that can be extracted. In partial movement structures, complex Wh-phrases like 'which boy' or 'whose mother' may be extracted and appear in the embedded CP Spec. In Wh-copying structures, on the other hand, it is ungrammatical to have a complex Wh-phrase in the embedded CP Spec position. (Long-distance extraction of a complex Wh-phrase with no Wh-copy in the embedded CP Spec is, of course,

\(^{105}\) McDaniel (1986) argues that the Imposibility of partial movement out of tense-dependent clauses is a consequence of restrictions on Case Inheritance that do not apply to Wh-copying.
possible). This contrast between the partial movement and Wh-copying structures is illustrated below:

**German**

(43) Partial Movement: Complex Wh-Phrase
Was du [welchen Mantel], [Ir]
Jako in den beute angesogen hat?
"What do you think which coat Jakob put on today?"

(44) Wh-Copying: Complex (PP) Wh-Phrase
*mit welchem Junge? [Ir]glaubt du [welchen Junge, [Ir] Hans, t1, spricht?
"With which boy do you think which boy Hans talks?"

(45) Wh-Copying: Complex Wh-Phrase
*Wessen Buch? [Ir]glaubt du [Wessen Buch, [Ir] Hans, t1, liest?
"Which book do you think whose book Hans is reading?"

**Romani**

(46) Partial Movement: Complex Wh-Phrase
*Ke [mesline] [Ke kasqoro] [Ir] o Demiri ctitol t1?
"What do you think whose book Demir is reading?"

(47) Wh-Copying: Complex Wh-Phrase
*Save chaves [mesline] [Kaswe chaves, [Ir] o Demir dihla t1?
"Which boy do you think which boy Demir saw?"

(48) Wh-Copying: Complex Wh-Phrase
*Kasqo l111 [mesline] [Ke Kasqo] [Ir] o Demiri ctitol t1?
"Which book do you think whose book Demir is reading?"

The two points of contrast in partial movement and Wh-copying that we have observed lead McDaniel (1986) to conclude that Wh-copying is a form of long-distance movement, with the true Wh-phrase in the matrix CPSpec.

It is worth considering briefly why complex Wh-phrases may not be used in Wh-copying structures. McDaniel (1986) does not offer an explanation but the suggests that the Wh-copy is an A*-resumptive pronoun. This can be thought of as a "spell-out" of the features of the Wh-trace. Chomsky (1981) suggests that Wh-trace carries only pronominal features. Pronominal features are specified as including person, number, gender, Case and perhaps the feature [±Wh]. Thus it may be that Wh-copying is ungrammatical with complex Wh-phrases because the Wh-trace in the embedded CPSpec cannot carry sufficient feature information to realize morphologically a complex Wh-phrase like which boy. This contrast in extraction of complex Wh-phrases will be exploited when we investigate whether children's medial-Wh questions are partial movement structures or Wh-copying structures.

Assuming Wh-copying questions to be long-distance movement structures, we can conclude the section by summarizing the possible derivations for short and long-distance movement questions:

(49) Long-Distance movement:
(i) Full movement
Wh-movement to matrix CP
[wh-W-1 [Ir]... [Ir] [Ir]...t1...

(ii) Wh-copying:
Wh-movement to matrix CP and Wh-copy in embedded CP
[wh-W-1 [Ir]... [Ir] Wh-1 [Ir]...t1...
Short-Distance movement

(i) Two matrix questions

\[
\begin{align*}
&\text{cpWh-1][zp}\ldots \\
&\text{cpWh-2][zp}\ldots
\end{align*}
\]

(ii) Partial movement

Wh-movement to embedded CP and scope marker in matrix CP

\[
\begin{align*}
&\text{cpWHAT } [zp-1][cp \text{ Wh-1}[zp-2]\ldots
\end{align*}
\]

In the next section, we turn to the parameters that have been proposed for Wh-movement.

**Medial-Wh Questions and Parameter Theory**

As things stand now, it would appear that children's medial-Wh questions have the characteristics necessary for a parameter theory explanation: We have observed that they appear in the absence of experience, and that similar questions (with two Wh-phrases) appear in other languages. In looking to parameter theory for explanation, we will need to account for the difference between tensed and infinitival clauses that was observed in a range of languages and also in the grammar of children who ask medial-Wh questions. Now let us ask which parameters might be implicated in the appearance of the medial-Wh.

**Wh-Movement Parameters**

Experimental research on children’s comprehension of long-distance questions involving extraction of arguments and adjuncts led de Villiers, Roeppe and Vainikka (1988) to propose the following developmental sequence in the acquisition of Wh-movement:

\[
\begin{align*}
(50) & \quad \text{(i) no movement} \\
& \quad \downarrow \quad \quad \quad \quad \text{"MOVEMENT LEVEL" PARAMETER} \\
& \quad \text{(ii) short-distance movement} \\
& \quad \downarrow \quad \quad \quad \quad \text{"DISTANCE" PARAMETER} \\
& \quad \text{(iii) long-distance movement}
\end{align*}
\]

They propose that these stages in acquisition of Wh-movement can be distinguished by two parameters. One parameter, originally proposed by Huang (1982), allows languages to choose the level of grammar at which Wh-movement takes place. The ‘movement level’ parameter distinguishes between languages like Chinese, Japanese and Korean which allow only LF movement on the one hand, and languages like Polish, English, German and so on, which have syntactic movement, on the other. The second parameter, original to de Villiers, Roeppe and Vainikka (1990) allows a choice in ‘distance’ of movement. It distinguishes between languages that allow short-distance movement from those which have long-distance movement. Note that the second parameter is based on the assumption that when the first parameter is set for a syntactic movement option, only short-distance movement becomes available. I will refer to these two parameters as the ‘movement level’ parameter and the ‘distance’ parameter respectively.

---

100 The ‘movement level’ parameter does not refer to multiple questions; that is questions in which there are Wh-phrases in argument position at D-structure. In multiple questions, even in languages with syntactic movement, Wh-phrases can be in-situ at S-structure. For example, in a multiple Wh-question like “Who read what”, only who moves to the CPSpec at S-structure. The Wh-phrase what is in-situ at S-structure, but moves and adjoins to CP at LF.
Let us imagine that the first parameter to be considered by the child encountering data in the environment is the ‘movement level’ parameter. The parameter can be expressed as the choice: [+ syntactic movement]. Assuming, as suggested by de Villiers, Rooper and Vainikka (1968) that the unmarked setting is [-syntactic movement], children learning English will hear questions in their positive input and change the parameter setting to [+ syntactic movement]. Once a child has set the ‘movement level’ parameter to [+ syntactic movement], data for the ‘distance parameter’ become conspicuous, and the choice between short or long-distance movement can be made. This parameter system predicts that until the ‘distance parameter’ is set, only short-distance movement is available to children, at least for some period of time.

Positive data in the form of questions with long-distance movement would inform children to set the parameter to [+ long-distance movement].

---

107 The ‘movement level’ and ‘distance’ parameters are presumably independent and can be considered by the child in either order. In fact, though, children will not have relevant data for the ‘distance parameter’ until the ‘movement level’ parameter has been set to the [+ syntactic movement] option. If children happen to consider the ‘distance’ parameter first, they will not have the data relevant to make a decision. On this view, one would expect children learning English to start out with Wh-in-situ and later switch the setting to [+ syntactic movement]. Oddly, though, the literature does not report that children learning English go through a stage where they ask questions like “Daddy likes what?” for example, with Wh in-situ. According to research by Hamburger (1961), however, there is a stage with no Wh-movement, but the empirical evidence is to be found in a precursor to the relative clause, which precedes the appearance of Wh-questions developmentally. In a longitudinal investigation of a two-year-old’s acquisition of relative clauses, Hamburger first observed ‘protorelatives’ like “that my did it” (meaning “that’s the thing such that I did it”), about a week after the appearance of many expressions of this form, there was an apparent grammatical change, and the child produced to expressions like “that my did”. Apparently, the child had learned an empty set category was required in relative clauses as the result of Wh-movement.

Does this system of movement parameters allow us to explain children’s medial-Wh questions? If they reflect a parameter setting that is not appropriate for English, we are led to conclude that they may be the consequence of the unmarked setting on the ‘distance parameter’, that is, short-distance movement structures. If children’s medial-Wh questions are short-distance movement structures, then they must be partial movement structures, as in Romani and German. The evidence for and against this analysis of the questions will be investigated later in the chapter.

Thus far, the Wh-movement parameters look promising as an explanation of children’s medial-Wh questions, but a broad range of data must be considered.

An immediate problem with the parameter system outlined above is that it only accommodates languages that allow short distance movement or long-distance movement. But as we saw, the languages such as German and Romani that have short-distance partial movement make available both movement options: Partial movement is a short-distance movement option, and long-distance movement and Wh-copying are both alternative long-distance movement options. This set of options can be handled by the Wh-movement parameter if we assume that hearing any long-distance Wh-extraction (from tensed clauses) in the input is sufficient to set the ‘distance’ parameter to [+ long-distance movement]. Positive input informs children of the marked options; Wh-copying and partial movement. These would have to be incorporated construction by construction into the grammar. But if this is the way the parameter system works, then children learning English should not produce medial-Wh questions, since there is no evidence for this structure in the input data.

---

108 The same question concerning optionality of distance of movement can be raised for the ‘movement level’ parameter. The ‘movement level’ parameter
The contrast between long-distance extraction from tensed and infinitival clauses must also be addressed. As conceived above, the Wh-movement parameter system would not distinguish between extraction from the two complement types. Let us suppose that children who produce medial-Wh questions have selected the short-distance movement setting on the ‘distance parameter’.

This selection will prevent any long-distance extraction, and so the prediction is that even when extracting from infinitival clauses, children will only be able to move the Wh-phrase to the intermediate CPSpec. This prediction is contrary to the facts, since no child ever produced a question like “Who do you want who to help you brush your teeth?”. It may be, however, that the facts concerning infinitival clauses are handled by an independent principle, such as Subjacency (Chomsky, 1977; 1981; Rizzi 1982). Indeed de Villiers, Rooper and Vainikka (1996), citing Otsu (1981), make it clear that they consider the Subjacency principle to be an innate principle that is operative in early child grammar.\footnote{They do consider, though, that maturation could contribute to a delay in the availability of long-distance movement. In particular, they suggest that the operation of successive cyclic movement matures, perhaps because the COMP node is not available, some feature of the COMP node is missing, or some feature about chain formation has not matured.}

Our next task is to investigate whether the Subjacency Principle, either on its own, or in combination with the proposed Wh-movement parameters, can predict the observed child data.

The Subjacency Parameter

Subjacency is a principle of Universal Grammar that constrains Wh-movement. It ensures that extraction, although apparently unbounded, takes place in a stepwise fashion. The Subjacency principle states that a constituent may not move across more than one bounding node in a single rule application, where bounding nodes are defined as NP and S.\footnote{This discussion will be in the framework of Chomsky (1981).} This can be represented as follows:

\[(51) \text{No rule can involve } X, Y, \text{ in a configuration:} \]
\[* X\ldots[a\ldots[b\ldots Y\ldots]\ldots]\ldots X\]
\[\text{where } a, b \text{ are cyclic nodes (Chomsky, 1977).}\]

The constraint in (51) leaves room to maneuver. In particular, the symbols \(a\) and \(b\), which refer to the bounding nodes can differ across languages. In English, the bounding nodes are considered to be NP and S (Chomsky, 1981), but on the basis of data showing extraction out of Wh-islands is grammatical in Italian, Rizzi argued that the relevant bounding node for Italian is not S, but S'. Thus, he suggests that the choice of bounding nodes is a candidate for parametric variation. Languages could be parameterized as having either \(<\text{NP, S}>\) or \(<\text{NP, S'}>\) as
bounding nodes.\textsuperscript{112} Data from other languages in which movement is more restricted than either English or Italian has led other researchers to propose a further parametric choice: In languages such as Polish (Willing, 1989), Russian (Radford, 1981; Quicoli and Freidin, 1989) and ASL (Lillo-Martin, 1988) the relevant bounding nodes have been proposed to be both \( S \) and \( S' \), giving \( <NP, S, S'> \) as the relevant set of bounding nodes.\textsuperscript{114} This suggests that Subjacency is parameterized three ways:

\begin{align*}
(52) & \quad (i) \quad <NP, S \text{ and } S'> \quad \text{(Polish, ASL...)} \\
& \quad (ii) \quad <NP, S> \quad \text{(English...)} \\
& \quad (iii) \quad <NP, S'> \quad \text{(Italian...)}
\end{align*}

How could such a parameter account for the variation across languages? It appears that the child cannot be allowed to initially choose any option, since the Italian setting would raise learnability problems. The Italian \( <NP,S'> \) setting in

\textsuperscript{112} There is some question about whether languages can be differentiated by a parameter that chooses \( S \) or \( S' \) as a bounding node. Grimshaw (1986) points out that a paradigm similar to the one Rizzi constructed for Italian can be constructed for English, and suggests that the \( S/S' \) bounding node does not define the difference between the languages. What Grimshaw proposes is that languages allow the initial movement to skip a COMP, but any subsequent movement cannot. To illustrate the point, she gives the following examples from English.

(i) Which book did the TAs tell the students that they shouldn't forget who wrote?
(ii) Which book did the students forget who told them that Dorothy Sayers wrote?

\textsuperscript{113} Even in the Barriers framework (Chomsky, 1985b), the difference between the \( S \) and \( S' \) bounding nodes is stipulated by saying that in some languages the lowest tensed IP is an inherent barrier, and in other languages the lowest tensed CP is an inherent barrier.

\textsuperscript{114} Van Riemsdijk and Williams (1986) claim that in Dutch, PP is a bounding node. I ignore this additional possibility here.

(iii) cannot be the unmarked setting, since all of the movement options allowed by English are also allowed in Italian, but not conversely. If \( S' \) is the unmarked setting, as in Italian, children learning English would never encounter positive evidence that would trigger setting the parameter to \( <NP,S> \). That is, they would never hear evidence to inform them that extraction from Wh-islands is ungrammatical.\textsuperscript{115} If \( <NP,S> \) was the unmarked setting, then children should start out allowing long-distance movement of Wh-phrases. Children learning ASL and Polish would begin by using long-distance extraction and later change the parameter setting to \( <NP,S, \text{ and } S'> \). This scenario, although logically possible, does not seem likely. Further, it would predict that children learning English start out with the \( <NP,S> \) setting appropriate for English, and it would offer no explanation of children's medial-Wh responses. At least, it would predict that medial-Wh questions are long-distance movement structures and the extra element in the Intermediate COMP would be unrelated to Subjacency. On the other hand, if the unmarked setting of the parameter is \( <NP, S \text{ and } S'> \), a stage in which children produce only short-distance movement structures is predicted. If (but only if) children's medial-Wh responses are determined to be short-distance movement structures, then their appearance could be explained.

So far, the Subjacency Parameter has the same level of explanatory adequacy as the Wh-movement parameters discussed in the last section. Both parameters predict a stage that could be equated with children's medial-Wh questions if these are short-distance movement structures. Can the Subjacency Parameter surpass the Wh-movement parameter in explaining the data on extraction from infinitival clauses?

\textsuperscript{115} The observation that \( S \), not \( S' \) must be the unmarked bounding node was made by M. Anderson (Lasnik, p.c.), and Freidin and Quicoli (1969).
Addressing this problem with respect to Polish, Wilkin (1989) proposes that the bounding nodes must be relaxed for infinitival clauses to allow long-distance Wh-movement. She suggests that in Polish S' should not be considered to be a bounding node when the complementizer is null, as in infinitival clauses. This would allow long-distance extraction without a violation of Subjacency. This dispensation for extraction from infinitivals could be built into the Subjacency parameter. This would predict, however, that all languages with the <NP, S, S'> set of bounding nodes allow long-distance extraction out of infinitival clauses. This is the case for children with the medial-Wh grammar, but would have to be verified for ASL, Russian and so on. If cross-linguistic data revealed that there are languages in which long-distance extraction is not possible even out of infinitival clauses, then in those languages in which extraction out of infinitival clauses has more freedom, the <NP, S, S'> setting could be amended on the basis of positive data.118

To sum up, the Subjacency Parameter can be adjusted to predict the contrast between extraction from tensed and infinitival clauses observed in some languages. Thus, if combined with the Subjacency Parameter, the Wh-movement parameter could explain the medial-Wh responses. However, since the Subjacency parameter alone accounts for the full range of data, the Wh-movement parameter seems superfluous. Whether the Subjacency Parameter is adopted to explain children’s medial-Wh questions now depends on whether medial-Wh questions are short-distance movement structures. This is the topic of the next section.

---

118 Positive data will also be necessary to allow different movement options. Presumably a language like German or Romani will set the bounding nodes to <NP, S> to allow long-distance movement. Positive input will allow the grammar to incorporate the additional short-distance movement option.

---

The Derivation of Medial-Wh Questions

In the experiment eliciting Wh-extraction from tensed embedded clauses, recall that two kinds of non-adult questions with two Wh-phrases emerged in children’s data. The majority of questions were medial-Wh questions that exhibited matching Wh-phrases, as illustrated in (53):

(53) Who do you think who is in the box?
    What do you think what Cookie Monster likes?

Two children who asked medial-Wh questions also asked questions in which the Wh-phrases did not match, as shown in (54):

(54) What do you think who jumped over the fence?

Let us begin by examining the data from the children who produced questions with non-matching Wh-phrases. One child, Tiffany, produced mostly medial-Wh questions, and a few questions with non-matching Wh-phrases; the other child, Kelly, produced both question types freely, at least in her initial testing sessions.117 Examination of the Wh-phrases reveals that, for the most part, when the Wh-phrases do not match the Wh-phrase in the matrix CPSpec is “what”. In each of the examples illustrated from Kelly, below, the experimenter’s lead-in was devised to elicit long-distance extraction of the Wh-phrase that ap-

117 Tiffany produced 38 long-distance questions, 5 of which were partial movement. Four of the 5 questions occurred with movement of a complex Wh-phrase.

118 Kelly was referred to as K.S. in the studies of Strong Crossover in the previous chapter.
pears in the Spec of the embedded CP. This is illustrated in (55), along with examples of Kelly’s questions in (56):

(55)  Kelly 3:11
Experimenter: We know that the frog lives in the pond, right, but ask the
turtle where he thinks
Kelly:  What do you think where this froggy lives?
(56)  What do you think where the marble is?
What do you think who ate this?
What do you think who had the toothbrush?

The fact that on occasion Kelly and Tiffany ignored the prompt for which Wh-
phrase to extract, and insisted on beginning their question with “what” suggests
that these are partial movement structures. It is interesting that in the absence
of positive data, both children chose “what” as the scope marker. This choice of
“what” corresponds to the Wh-word used for scope marker in other languages.
This may be because it is the Wh-phrase that is most ‘neutral’, and carries the
fewest features.\textsuperscript{119}

The partial movement questions produced by Kelly and Tiffany support the
proposal that children initially have restricted Wh-movement. Before settling on
this conclusion, the derivation of children’s medial-Wh questions needs to be
determined, however. There are two possibilities to consider. As short-distance
movement structures, children’s medial-Wh questions could be: (i) two matrix

\textsuperscript{119} The only possible exception to use of ‘what’ as a scope marker came from
an example from Kelly. In the relevant example, which she repeated twice, it
was difficult to hear whether the Wh-phrase in the matrix CPSpec was “who”
or “how” (in which case this is a medial-Wh question). The example at issue
was an adjunct extraction question:
(i) Who/how? do you think how her dries her feet?

questions following in succession as in Polish, or (ii) partial movement structures
which use any Wh-phrase for a scope marker. In the remainder of this section,
reason will be given to reject both of these possibilities. Two arguments will be
given against a partial movement analysis of medial-Wh questions; one based
on Kelly’s longitudinal data; and a second syntactic argument based on data in-
volving extraction of complex Wh-phrases.

The first possibility is to consider whether medial-Wh questions are two matrix
questions, as in the Polish example observed in (21). If children’s medial-Wh
questions are two matrix questions in succession, a medial-Wh question like (57)
would be more accurately represented as (58):

(57)  Who do you think who is in the box?
(58)  Who do you think? Who is in the box?

The status of the “second” clause in medial-Wh questions as a main clause or a
subordinate clause can be tested by examining children’s use of ‘do-support’ in
object extraction questions. If both clauses have the status of main clauses,
Subject-Aux inversion should be triggered in them both. Consequently, children
favoring medial-Wh responses should produce utterances like:

(59)  What do you think? What does Cookie Monster like?

While children consistently used ‘do-support’ in the matrix clause, no child ever
produced an utterance like (59) with two occurrences of ‘do-support’. This sug-
gests that the medial-Wh questions are two-clause structures. It now remains to
investigate whether medial-Wh questions are partial movement structures, with
only short-distance movement of the true Wh-phrase, or long-distance movement structures.

As noted in the section on the partial movement structure, partial movement questions use a fixed lexical item in the CPspec of the matrix clause to indicate the wide scope of the question. Children learning languages in which partial movement is a grammatical option (e.g. Hungarian, German and Romanian), naturally receive positive evidence for the item that corresponds to scope marker. But what of English speaking children whose grammars allow partial movement? Since partial movement is not grammatical for adults, children in an English-speaking environment do not receive any input about what item to use as a scope marker. Given this, they have to venture a guess. A reasonable guess would be to use the same item as the true Wh-phrase, hence two appearances of the same Wh-phrase in medial-Wh questions. While initially appealing, there is reason to doubt that children repeat the true Wh-phrase to serve as a scope marker. This is because the same children who produce medial-Wh questions also produce questions with consistent use of the scope marker “what”. Given that Kelly and Tiffany chose “what” as the scope marker in some questions, it is unlikely that other times they would choose a scope marker that matched the true Wh-phrase in features.

An interesting argument for medial-Wh questions as long-distance movement structures comes from longitudinal study of Kelly’s Wh-questions. Her medial-Wh questions and partial movement questions did not pattern in tandem, as might be expected if they were both in fact, variants of partial movement. In Kelly’s first session, 5 of her 7 questions were medial-Wh questions and 2 were partial movement questions. None of her questions were the form of long-distance question used by adults, with no Wh-phrase in the intermediate CP domain. In the second session, most of Kelly’s questions still had a filled intermediate CP domain, but she also produced some correct adult questions. In this session, Kelly produced 6 medial-Wh questions, 2 that-trace questions, 6 partial movement questions and 6 adult questions with no overt element in the intermediate CP domain. In the third session, partial movement questions persisted; 4 of the 12 questions she produced were partial movement structures. There was a change evident in this session, however: adult questions (with no filled CP domain) seemed to have replaced the medial-Wh option. Kelly produced only 1 medial-Wh question, but 7 adult questions. This suggests that the medial-Wh was a long-distance question, and Kelly had finally become aware of the adult form. Further, if Kelly’s medial-Wh questions were partial movement structures, we would expect them to drop out of her grammar at the same time as the questions that were unequivocally partial movement. Yet her partial movement questions persisted while the medial-Wh questions disappeared.

Further evidence for rejecting a short-distance movement analysis of medial-Wh questions comes from data on extraction of complex Wh-phrases. Recall that in the German and Romanian data, we observed that partial movement and long-distance Wh-copying structures differed in whether they would allow a complex Wh-phrase to appear in the intermediate CP domain. The Wh-copying

---

120 Kelly and Tiffany asked both medial-Wh questions and partial movement questions, but it does not appear to be the case that all children who produce medial-Wh questions also allow the partial movement option.

121 I have excluded questions with extraction of complex Wh-phrases from this discussion, since they appear to be treated differently from questions with extraction of a bare Wh-phrase.
structure does not allow a complex Wh-phrase to appear in the embedded CP domain. In questions extracting complex Wh-phrases, only two movement options are available: (a) long-distance movement of the Wh-phrase to the matrix CPSpec, or (b) partial movement. This pattern of facts in the German and Romani dialects that allow Wh-copying enables us to make an important prediction about children's grammars. If medial-Wh questions are long-distance Wh-copying structures, children should not allow a complex Wh-phrase like "which boy" to appear in the embedded CP Spec. Instead, they should opt for partial movement or the adult question form. In order to test these predictions, situations as illustrated in the protocol below were devised to elicit questions with complex Wh-phrases:

Protocol for Questions with Complex Wh-Phrase

(Several Smurfs are hidden under a blanket with only their heads showing)
Experimenter: We know that Smurfette is wearing the roller skates, right, but ask the Rat which Smurf he thinks.
Child: Which Smurf do you think is wearing roller skates?

This follow-up experiment was conducted on both Tiffany and Kelly. Both children's productions in this situation were in line with the prediction. Interestingly, despite her tendency to fill the Intermediate CP domain, in this situation Kelly produced adult question forms, moving the complex Wh-phrase to the matrix CPSpec, and inserting no Wh-phrase in the Intermediate CP domain.

(60) Kelly 3;11
Which Smurf do you think is holding a toothbrush?
Which baby you think have the bottle?

In the situation demanding extraction of a complex Wh-phrase, in contrast to Kelly, Tiffany produced some partial movement questions. In fact, 4 of the 5 partial movement questions documented for Tiffany were produced in this situation, that is, when the question form with two identical (complex) Wh-phrases was not available.

(61) Tiffany 4;9
What do you think who's in that can?
What do you think which animal says "woof woof"?
What do you think which Smurf really has roller skates?

In the same situation calling for extraction of a complex Wh-phrase, however, Tiffany also produced long-distance movement questions, with the complex Wh-phrase in the matrix CPSpec. These questions were not grammatical adult question forms because, contrary to expectation, a Wh-phrase was also present in the embedded CPSpec. The Wh-phrase in the Intermediate CP domain was not a copy of the complex Wh-phrase, however, but a reduced bare Wh-phrase that agreed with the complex Wh-phrase in features. Examples from Tiffany follow:

(62) Tiffany 4;9
Which Smurf do you think who has roller skates on?
Which animal do you think what really says "woof woof"?

Similar questions were also observed in the productions of children who participated in the Crossover elicitation experiment reported in Chapter II:

(63) Ben 3;11
Which one do the bear and the squirrel think who has two bears?
Since these children did not produce questions with a copy of the complex Wh-phrase in the embedded CP domain, the evidence suggests that medial-Wh questions are, indeed, long-distance questions like the Wh-copying structures in German and Romani. The only time any child produced an exact copy of a ‘complex’ Wh-phrase was when the extracted Wh-phrase was a complex Wh-phrase like “which one” or “what one”, with little feature specification.

What these data suggest is that in children’s medial-Wh questions, Wh-phrases appearing in embedded CPSpecs cannot be complex Wh-phrases that carry full feature information. The Wh-phrases children used are exactly the “unspecified” Wh-phrases such as who and what that German and Romani usually allow in Wh-copying structures. Likewise, the Wh-phrases which one and what one carry only limited feature information. Thus, the fact that children used Wh-phrases like who, what and which one turns out to be support for analyzing children’s medial-Wh questions as long-distance movement structures involving successive cyclic movement. The complex Wh-phrase is moved through the embedded CPSpec, but because Wh-trace does not carry sufficient information to realize a complex Wh-phrase, children realize whatever information is carried by its trace.122

To summarize, our investigation has led us to the conclusion that children who produce medial-Wh questions are not restricted to short-distance movement structures. On the contrary, medial-Wh questions are like Wh-copying questions, with long-distance movement of the Wh-phrase. In some children’s grammars, medial-Wh questions are the only form of long-distance question that is sanctioned, at least for a time. Others children, like Kelly, produce partial movement questions and medial-Wh questions in situations where long-distance Wh-extraction is appropriate in the adult grammar. The conclusion we have reached is that, rather than being restricted in the forms of long-distance question available, children appear to initially have a range of options available. No child tested was restricted only to partial movement questions involving short-distance movement.

Returning to the proposals of parameters for Wh-movement (de Villiers, Roeper and Vainikka, 1988) and Subjacency, can these parameters explain the appearance of medial-Wh questions? The answer is not if the unmarked setting on these parameters allows only short-distance Wh-movement. Even supposing that any setting on these parameters was available initially, they would not be able to predict why children do not use the same form of long-distance movement question as adults; the appearance of the medial-Wh would still not be explained. This suggests that movement parameters cannot be used to explain children’s

122 The possibility that the medial-Wh is not a trace, but a complementizer (not in CPSpec) will be considered in the section on children’s long-distance questions.
Wh-questions. In the remainder of the chapter, I pursue a different avenue, that children's medial-Wh questions are related to their conception of the Empty Category Principle (ECP). The analysis in the rest of the chapter will build on the conclusion of this section, that medial-Wh questions are long-distance movement structures.

**ECP Explanations of Medial-Wh Questions**

In the previous section, children's medial-Wh questions were analyzed as long-distance movement. Examination of empirical data led us to the conclusion that the medial-Wh is a trace of movement left by the Wh-phrase as it moves successively cyclically en route to the matrix CPSpec. It was suggested that the medial-Wh is an overt "Wh-trace" instead of an empty category. Interestingly, the overt Wh-trace seems to be obligatory in the grammars of the children who produce this form of Wh-question, at least for subject extraction questions. This contrasts with its optionality in adult grammars such as German and Romani, in which the structure with a Wh-copy is optional. This difference leads us to ask whether children have an alternative conception of where empty categories may appear. Since the Empty Category Principle (henceforth the ECP) (Chomsky 1981, 1986a, 1986b; Aoun, Hornstein and Sportiche, 1981; Lasnik and Saito 1984, forthcoming; Rizzi 1989) determines where empty categories are licensed, the purpose of this section is to investigate further children's use of an overt category in a position where an empty category normally appears. Investigation proceeds by further pursuing the idea that children's non-adult utterances are the consequence of an incorrect setting on a parameter. ECP parameters are considered within the framework of Lasnik and Saito (1984). Let us turn to the ECP in this framework before detailing the proposed parameters.

**The ECP in the Lasnik and Saito Framework**

The Lasnik and Saito framework assumes the X'-framework in Chomsky (1981). The category S' is considered to be a projection of COMP, which is the head. COMP has just one position, which may be filled either by a complementizer of the trace of movement. The Empty Category Principle can be stated as in Chomsky (1981):

(67) A nonpronounal empty category must be properly governed.

(Chomsky, 1981)

Lasnik and Saito (1984, forthcoming) show that the ECP applies to all traces of movement, including intermediate traces. Proper government of an empty category is attained in one of two ways — either by lexical government or by antecedent government. Lasnik and Saito define the two forms of proper government as follows:

(68) a lexically governs b if
   a. a c-commands b, and
   b. a assigns Case or 0-role to b

   a antecedent governs b if
   a. a binds b, and
   b. there is no γ (γ an NP or S') such that
      a. a c-commands γ and γ dominates b,
      unless b is the head of γ
Let us see how the ECP constrains the long-distance questions that are the subject this chapter. Assume for the moment, that Wh-movement through COMP leaves a trace and that the ECP applies at S-structure. First consider a long-distance question in which the Wh-phrase is extracted from object position.

(69) What do you think [t₁ Cookie Monster likes t₁]

The trace in the extraction site is lexically governed by the verb, and the trace in the intermediate COMP is antecedent governed by the Wh-phrase what. In a long-distance question in which the extracted Wh-phrase is an adjunct, both traces are properly governed by antecedent government. The trace in the extraction site is antecedent governed by the trace in the intermediate COMP and the intermediate trace is antecedent governed by the Wh-phrase why.

(70) Why do you think [t₁ Cookie Monster likes cookies t₁]

Given what has been said so far, when there is a complementizer in the intermediate COMP, however, proper government of the trace in the extraction site would be blocked.

(71) Why do you think [that Cookie Monster likes cookies t₁]

The presence of the complementizer in the intermediate COMP position would mean that the trace in the extraction site cannot be properly governed by an intermediate trace. The Wh-phrase why in the matrix COMP cannot serve as a proper governor either, since there is an S’ intervening, and the trace is not in head position of the S’. Yet the trace must be properly governed because the example is grammatical. This leads Lasnik and Saito to propose that the level relevant for the ECP is LF. They propose that at the level of LF, items such as that, which are devoid of semantic content, can delete. Returning to the derivation for (71), the deletion of that allows the Wh-phrase why to lower, leave an intermediate trace, and raise again. The intermediate trace created by back and forth movement is an antecedent governor for the trace in the extraction site. Thus at LF, the adjunct extraction example obeys the ECP.

Having established the way in which a trace in the embedded clause can be antecedent governed even when there is an overt complementizer present, we can now turn to examine how subject extraction ‘that-trace’ violations are handled.

(72) Who do you think [that t₁ likes Cookie Monster]

The same mechanism of that deletion and the creation of an Intermediate trace that let in the example in (71) would presumably apply to (72) also, and allow the subject trace to be antecedent governed at LF. Hence the ungrammatical example is not ruled out, as it would have been had the ECP applied at S-structure. To remedy this situation, Lasnik and Saito propose a γ-marking mechanism that differentiates between arguments and adjuncts, since they behave differently with respect to allowing complementizers. The γ-marking mechanism is accom-

---

123 Verbs subcategorize for arguments but not adjuncts, which are optional. The verb does not assign a 'role or Case to the adjunct, and as it cannot be lexically governed.

124 The Principle of Recoverability would prevent elements with semantic content from deleting at LF.
panied by the principle in (74) requiring only arguments to be γ-marked at S-structure.

(73) a. Assign [+γ] to a trace that is lexically governed or antecedent governed
b. Assign [-γ] to all other traces

(74) At S-structure, γ-assignment applies only to arguments.

The ECP can now be formulated as a filter on γ-marking that applies at LF.

(75) The ECP Filter

*[−γ]

Let us return to the three long-distance extraction cases of interest to see how the γ-marking mechanism works. In object extraction examples like (69), the trace in the extraction site is lexically governed and marked [+γ] at S-structure, allowing the ECP filter to be satisfied at LF. Turning to adjunct extraction examples like (71), no γ-assignment takes place until LF. At LF, after that deletion, Wh-lowering and raising, and creation of the intermediate trace, both traces have an antecedent governor and are marked [+γ], giving the correct result. Now, consider how the subject extraction example which provided the motivation for the γ-marking mechanism can be handled. This example involves extraction of an argument, so γ assignment takes place at S-structure.

(76) S-structure:

Who do you think [that[t₃ likes Cookie Monster

[−γ]

LF:

*Who do you think [t₃[t₃ likes Cookie Monster

[+γ][−γ]

At S-structure, the complementizer that prevents an intermediate trace from properly governing the subject trace by antecedent government, and the trace is accordingly marked [−γ]. At LF, there is an intermediate trace, but the indelible [−γ] marking on the subject trace means that the ECP filter is violated.

ECP Parameter-Based Explanations

Now let us turn to children's exceptional questions, and consider how parameterized versions of the ECP might be used to explain the appearance of the medial-Wh. Our proposals will adopt the version of the ECP in Lasnik and Saito (1984). Consider first, why the overt Wh-phrase appears in the Intermediate COMP. Suppose that children think, for some reason, that a trace in the Intermediate COMP is not properly governed. Their response is to make the trace overt, as a way of circumventing the ECP proper government requirement since it applies only to empty categories. But why would children think that an intermediate trace fails to be properly governed? One place to start may be with the definition of antecedent government. Recall that according to the definition of antecedent government in (68), S' is a barrier to antecedent government with one exception — antecedent government is permitted when the trace is in the head
position (COMP) of the potential S'-barrier. Suppose children are not aware of the
exception. Then they would have a restricted version of antecedent government
that is without the "unless" clause below:

(77) a antecedent governs β if
   a. β binds β, and
   b. there is no γ (γ an NP or S') such that
      a. c-commands γ and γ dominates β
         UNLESS β IS THE HEAD OF γ

If there are languages in which antecedent government is blocked over S' (and
NP, though I will not consider NP here), then the difference in adults' and chil-
dren's conception of antecedent government can be expressed as a parameter.
The unmarked setting of the parameter would not include the "unless" clause in
(77). Positive input in the form of long-distance questions (with or without a
complementizer) would trigger the parameter to be set to include the exception.

The restricted version of antecedent government entails that intermediate
traces cannot be properly governed. The fact that intermediate traces cannot
be properly governed raises a dilemma for children since they know from the
ECP that traces, including intermediate traces, are required to be properly gov-
erned. One way around the problem would be to make the intermediate trace
overt at S-structure. Making the trace overt at S-structure, will mean that it won't
be subject to the ECP at LF, since, after all, The ECP applies only to empty cate-

125 Diane Lillo-Martin (p.c) points out that if children think Intermediate traces
cannot be properly governed, then one should consider why they choose to
generate a medial-Wh question rather than not attempt long-distance move-
ment at all. I will assume here that children have concluded from matrix
questions they have heard, that the [+WH] COMP feature must be satisfied
at S-structure in English. The only way to satisfy it is to move the Wh-phrase
long-distance.

gories. On this view children's medial-Wh questions are long-distance questions
with overt traces, as below:

(78) What, do you think [what] C N likes t₁
(79) Who, do you think [who] t₁ likes C N
(80) Why, do you think [why] C N likes cookies t₁?

The proposal that some children think S' is an absolute barrier to proper gov-
ernment explains why they initially use the medial-Wh in all of their questions
extracting from a tensed embedded clause. One problem with this proposal,
however, is that it makes the wrong predictions for questions with infinitival
clauses: Extraction from infinitival clauses also results in an intermediate trace
that must be properly governed.126,127 Consider object extraction questions with
the verb "want".

(81) What, do you want [x₁ t₁ [z₁ Pro to eat] t₁?

For children who cannot govern across the S', extraction from infinitival clauses
raises the same problem as extraction from tensed clauses; an intermediate

126 Any proposal that states CP (S') is an absolute barrier to government for
children fails to explain the lack of a medial-Wh with infinitival clauses, no
matter what framework the proposal is based in. This problem can be allevi-
ated, however, if only tensed CPs are designated as barriers.

127 Not all frameworks require an intermediate trace to be left in questions ex-
tracting from infinitival clauses. For example, in a development of the Barriers
framework (Chomsky, 1985b), Frampton (1990) proposes that Wh-phrases can
adjoin to IP as well as to VP, and move through CPSpec in their path of
movement to the matrix COMP. In questions with an infinitival clause, the
COMP is "not syntactically active" with the consequence that an IP adjoined
trace can be head governed by the matrix verb, hence movement through
Spec of CP is not required and no intermediate trace is necessary.
trace is generated that cannot be antecedent governed by the matrix Wh-phrase. Given that the problem is the same, we might expect children to exploit the same solution. That is, the expectation is that children would make the trace overt to ensure that it doesn’t violate the ECP. But as we noted, no child ever produced a medial-Wh in questions extracting from infinitival clauses.

Another problem with this account of the medial-Wh is that it does not predict any stages of development. Once children realize that S’ is not a barrier for antecedent government, the medial-Wh should disappear from subject, object and adjunct extraction questions simultaneously. However, the data suggest that the medial-Wh tends to be present in subject extraction questions when it has already dropped out from the other positions. In sum, the parameterized version of antecedent government in (77) is not supported by the data.

Our rejection of a parameterized version of antecedent government suggests that it is not failure of the intermediate trace to be properly governed that is giving rise to the medial-Wh. This leads us to consider a different viewpoint, that children use the medial-Wh to ensure that the initial trace in the embedded clause meets the requirements of proper government. On this account, whether the question involves extraction from subject, object or adjunct position, the medial-Wh would be used to satisfy the ECP. This assumption seems reasonable for subject and adjunct extraction questions, since in these cases, the initial trace satisfies the ECP by virtue of being antecedent governed by the intermediate trace. But in object extraction questions, the medial-Wh in the intermediate COMP is not easily explained, since object traces satisfy the ECP by being lexically governed by the verb in the embedded clause. As a consequence, the intermediate trace plays no part in satisfaction of the ECP in object extraction questions, and therefore the medial-Wh in children’s object extraction questions is unexplained.

Given these facts, an explanation of the medial-Wh based on the ECP would need to attribute to children different assumptions about how object traces are properly governed. Perhaps children begin by thinking that the only form of government is antecedent government.

(82) Traces must be antecedent governed

Further, suppose that children have concluded, on the basis of evidence from matrix questions, that antecedent governors must be overt:

(83) Antecedent governors must be overt at S-structure

Now, given these assumptions, the medial-Wh appears so that the initial trace is properly governed. Still questions with infinitival clauses present a problem. In object extraction questions with the ECM verb want, we saw that the initial trace is assumed to be properly governed by the verb (via lexical government). However, on this proposal the child assumes antecedent government is the only form of government. Given this, we might expect the child to use a medial-Wh as a proper governor for the object trace. However, the data show that at the time children produce a medial-Wh with questions extracting objects from tensed embedded clauses, no medial-Wh’s appear in their questions with infinitival clauses.
Another option to consider is that children start out with a conjunctive version of the ECP. On this account, the ECP would be parameterized so that the initial setting would require both antecedent government and lexical government of a trace. Later positive evidence would cause children to realize that the ECP is disjunctive, i.e. it requires either antecedent government OR lexical government of a trace.\textsuperscript{128}

\textbf{(84) CONJUNCTIVE ECP:}

\begin{itemize}
  \item Traces must be antecedent governed \textit{AND} lexically governed
\end{itemize}

\textbf{DISJUNCTIVE ECP:}

\begin{itemize}
  \item Traces must be antecedent governed \textit{OR} lexically governed
\end{itemize}

This parameterized form of the ECP, in which children start out with a setting entailing a conjunctive ECP is more attractive than the proposal that assumes children think the only requirement of the ECP is one of antecedent government (as in (82)). However, along with the conjunctive ECP, it is still necessary to make the additional assumption in (83), that children think antecedent governors must be overt at S-structure. A parameter with the conjunctive ECP as initial setting is superior because it accommodates the developmental trend of the medial-Wh dropping out of object extraction questions before subject extraction questions.

On this account, then, children initially have a conjunctive version of the ECP. At first, they use the medial-Wh in all questions, no matter what the extraction site, because all traces need to be antecedent governed. The positive evidence to change the conjunctive formulation to a disjunctive form of the ECP would come from adults' use of complementizers in long-distance questions. Roughly, children would hear adults sometimes using complementizers in their object extraction questions. This would make them realize that the COMP position does not always have to be filled by an overt Wh-word acting as an antecedent governor. At this point, they would realize that in English, at least, the ECP must be disjunctive: object traces only need to be lexically governed. As a result of setting the ECP to the disjunctive setting, children will no longer use a medial-Wh in object extraction questions. The use of optional complementizers in adjunct extraction questions could likewise trigger a the parameter to be changed from the conjunctive to the disjunctive setting.

This proposal seems promising as an explanation of the medial-Wh questions, but like the ECP parameter discussed previously, it fails to explain the lack of a medial-Wh in questions extracting from infinitival clauses. If children assume a conjunctive formulation of the ECP, then they will still use a medial-Wh to provide an antecedent governor for traces in infinitival clauses. Further, neither of the proposed ECP parameters can explain the 'that-trace' violations of some children. For example, children with the conjunctive form of the ECP should never ask questions with complementizers in COMP, since the COMP position is needed at S-structure for an overt antecedent governor. The data call this prediction into question, since one child, Jennifer, was observed to vacillate between medial-Wh questions and questions with that-trace violations. Furthermore, children who have changed the parameter setting to the disjunctive

\textsuperscript{128} See Tiedeman (1989) for a proposal for a parameterized ECP in which children start out with a conjunctive ECP. Tiedeman appeals to the conjunctive ECP to explain data on Subject-Aux Inversion reported in the literature.