Medial-*wh* Phenomena, Parallel Movement, and Parameters

TERJE LOHNDAL

University of Maryland

1. Introduction·

This paper will analyze the fact that some children produce intermediate copies when they form long-distance questions (hereafter 'medial-*wh*s'), as seen in (1).

(1) a. Who do you think **who** is in the box?

b. What do you think **what** Cookie Monster likes?

These data have been confirmed in several and different studies: Thornton [74, 75], McDaniel, Chiu and Maxfield [56], Crain and Thornton [22].[1] They also seem to be cross-linguistically robust; they are attested in Dutch (van Kampen [41, 42]), French

· Parts of this paper have been presented at the Workshop on Universal Grammar, Language Acquisition and Change (University of Oslo, August 2008), Syntax lunch (University of Maryland, October 2008), the Workshop on Typology and Parameters (University of Arizona, February 2009) and at Harvard University (February 2009). I'm grateful to the audiences for their helpful comments and to Jason Kandybowicz, Jeff Lidz, Jairo Nunes and Bridget Samuels. For their encouragement along the way and insightful comments, I'm indebted to Cedric Boeckx, Noam Chomsky, Elly van Gelderen, Rozz Thornton and Juan Uriagereka.

[1] See also de Villiers, Roeper and Vainikka [23] for a comprehension study that looks at somewhat similar data.

(Strik [71]), Basque and Spanish (Gutiérres Mangado [34]). In none of these languages are the sentences in line with the 'target' grammar and they are also not part of the input to the children who produce them.

Interestingly, the data in (1) is very similar to data from adults speaking German dialects and Romani, as shown in (2).

(2) a. Wen glaubt Hans **wen** Jakob gesehen hat? *German dialects*
whom thinks Hans whom Jakob seen has
'Who does Hans think Jakob saw?' (McDaniel [55:183])

b. Kas misline **kas** o Demìri dikhlâ? *Romani*
whom you.think whom Demir saw
'Who do you think Demir saw?' (McDaniel [55:182])

On the surface, English-speaking children and the adult German dialects and Romani appear to be very similar. In this paper I will argue that this is not the case and that the derivations underlying (1) and (2) are different in important ways. In particular, I will argue that (1) follows from children's analysis of null complementizers in English (cf. Jeong [39]), an analysis that plausibly can be extended to similar data from children speaking other languages. The data in (2), on the other hand, will be argued to follow from parallel chain formation, which is an extension of the notion of parallel movement in Chomsky [21]. Lastly I will argue that the analysis aligns well with recent developments on how to think about parametric variation within the Minimalist Program.

2. Multiple Pronunciations in English-Speaking Children

In this section I will discuss medial-*wh*s in English-speaking children. Before I go into the details of the phenomenon, I first present some arguments as to why medial-*wh*s are traits of competence rather than performance. I also discuss arguments showing that medial-*wh*s are really derived through successive cyclic movement. Then I go on in the following sections to discuss ways to account for multiple Spell-Out of copies, in particular focusing on Nunes [58] and Jeong [39].

It was first discovered by Thornton [74] that children are able to pronounce medial-*wh*s. Some representative data are given in (3).

(3) a. Who do you think **who**'s in the box?

b. What do you think **who**'s in that can?

c. Which animal do you think **what** really says 'woof woof'?

d. Why do you think **why** Cookie Monster likes cookies?

In Thornton's [74] study, nine out of twenty children (aged 2;1 – 5;5) frequently produced medial-*wh*s during an elicited production task. McDaniel, Chiu and Maxfield [56] also found that several children reported that these sentences are grammatical. Notice that the latter researchers asked for acceptability judgments from the children whereas Thornton used an elicitation technique. I will return to this difference below.

Before we look closer at the restrictions children seem to obey, let us first ask the question whether medial-*wh*s are a reflection of competence or a trait of performance. Thornton [74:331-33] discusses this question, and here we will just mention some of the

issues that she brings up, which favor the view that medial-*wh*s are competence phenomena.

We know that performance effects often result from memory overload and that this subsequently leads to deletion of material. This is as expected if memory limitations prevent us from keeping a certain amount of items in working memory. Deleting items would take away some of the burden imposed on working memory. In the cases we are discussing here, material is inserted instead of being deleted. Thus it is less plausible that we are dealing with performance effects in cases involving medial-*wh* structures.

It has also been shown that performance errors typically are found with object extraction. However, children that produce medial-*wh*s do so more often in subject extraction than in object extraction. In fact, there is an important developmental trajectory here. Children often start out producing medial-*wh*s in both subject and object extraction, but at a later stage they only produce medial-*wh*s in subject extraction, until they converge on the target grammar. If performance errors are mostly found with object extraction, it is clear that these data cannot be analyzed as such errors.

A final argument against medial-*wh* structures being performance effects comes from parsing considerations. We know that resumptive pronouns are more likely to occur with depth of embedding. This is due to memory reasons: resumptive pronouns make it easier to recover the dependencies and we do not need to store a gap in working memory. The prediction emerging from this is that the lower clause is more likely to be filled as it comes late in the parse. Consider (4).

(4) Who do you really think who Grover wants to hug?

If parsing considerations determine the spell-out of copies, we would expect there to be a copy in the infinitival clause, as in (5).

(5) *Who do you really think who Grover wants **who** to hug?

Such data are not attested. As I will discuss below, Thornton [74, 75] has argued that children never produce a medial-*wh* in infinitival clauses. Even if that were true, it would not affect the argument we are making here. Parsing considerations are generally thought to work outside of grammar proper[2], thus the parser would presumably not know that medial-*wh*s cannot appear in infinitivals – if the goal of the parser is to aid memory retrieval. Taken together, these arguments make it hard to justify the claim that medial-*wh*s are performance effects.

There is another issue that has to be discussed before turning to the similarities between English child language and other languages that display medial-*wh*s. The issue is whether the medial-*wh*s really are reflexes of long-distance movement.[3] Thornton [74] already asked that question, though she only discussed one piece of evidence. She asked whether (6a) is better represented as (6b).[4]

[2] For an important exception, see Phillips [60].

[3] I am assuming that successive cyclic movement exists in the traditional sense. See den Dikken [24] for arguments to the contrary. See also Koster [49] for critical discussion.

[4] One question not raised by Thornton is whether the first part of (6b) really is grammatical: *Who do you think?* At least in adults' English grammar, this structure is not well-formed. This is another reason to be suspicious towards an analysis of (6a) along the lines of (6b).

(6) a. Who do you think **who** is in the box?

b. Who do you think? Who is in the box?

However, as will become evident momentarily, this piece of data is not well suited to settle this question. Consider instead (7).

(7) What do you think **what** pigs eat?

This sentence would have the bi-clausal structure in (8) if (6b) is the correct representation for (6a).

(8) What do you think? What do pigs eat?

As we see, *do*-support is an ideal case here. Whereas we obligatorily get *do*-support in the main clause, we are not supposed to get *do*-support in the embedded clause in (8). This is also the case: *do*-support in each clause as in (8) does not occur in children's speech. To my mind, this seems like a good argument in favor of not analyzing these structures as two-clausal structures.

However, there are still some other alternatives to consider before we can move on. One is whether (9b) is an adequate representation for (9a).

(9) a. Who do you think **who** left?

b. Who do you think is the one who left?

That is, is (9a) really a reduced relative of some kind where *who* is a relative pronoun? There are several issues that render an analysis such as (9b) untenable. First, we have to bear in mind that the context where these medial-*wh*s were produced, namely a context highly favorable to elicitation of questions. Since children are otherwise very good at producing questions in this context, it seems odd that they suddenly should start producing relative clauses. Second, one wonders what would trigger the change from the relative pronoun structure to a long-distance question structure. The latter is arguably the correct structure for the target grammar that the child eventually converges on. It seems hard to come up with anything reasonable to say about this question. However, there is an empirical prediction that a reduced relative clause approach makes. In (9) we have an animate *wh*-element. If we instead have an inanimate *wh*-element, we would predict *that* to occur instead of a *wh*-element. That is, (10a) would be something like (10b) instead of (10c), contrary to fact.[5]

(10) a. What do you think is in the box?

b. What do you think is the thing **that** is in the box?

c. What do you think is the thing **what** is in the box?

[5] This is true for subject and object reduced relatives. Thornton [74] shows that some English-speaking children show that-trace violations, which means that there are some cases of *that* appearing, but only for subject extraction.

This is a serious problem for any analysis claiming that medial-*wh*s are reduced relatives. Of course, additional questions would emerge as to how one goes about to analyze a reduced relative of this kind, but I take it that the analysis is already too dubious to merit further discussion.

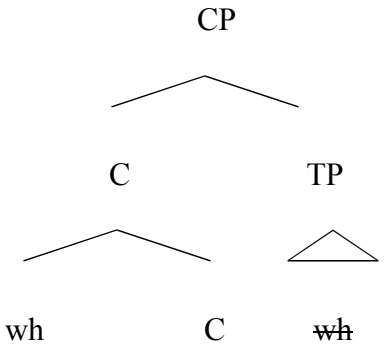
So far I have presented a few arguments that that medial-*whs* are competence effects and I have argued that the best analysis is one that views these medial-*wh*s as reflexes of successive-cyclic movement. This is entirely uncontroversial within the generative literature. However, so far nothing has been said about how exactly medial-*wh*s should be analyzed. I now turn to that question by reviewing two proposals that have been put forward in the context of the Minimalist Program.

3. Nunes on Spell-Out of Multiple Copies

Nunes [57] attempts to give a comprehensive theory of Spell-Out at the PF interface. He argues in favor of a Chain Reduction operation that ensures that only one copy is pronounced, based on Kayne's [46] Linear Correspondence Axiom (LCA). Nunes derives the ban on chains where multiple members of a chain are phonetically realized from linearization requirements. Syntactic items intervening between two phonetically realized chain links must both precede and follow the same element, thus resulting in a contradiction. As Nunes also points out, structures with two phonetically realized chain members violate the irreflexivity condition on linear order, i.e., if A precedes B, then A ≠ B. Furthermore, the fact that usually the highest copy is pronounced is derived through the claim that lower copies usually have checked fewer features compared to higher copies. Thus is it more economical to delete these copies. Notice that one cannot delete

all copies because, according to Nunes, that would involve too many applications of the operation Chain Reduction. However, structures like those mentioned above are attested, and we need to account for them. The way Nunes does it is by saying that in cases where medial-*wh*s are allowed, movement proceeds through adjunction to an intermediate C head. A simplified representation is given in (11).

(11)



In this case of adjunction to another head, Nunes [57:40] follows Chomsky [19:337] who says that '[the structure's] internal structure is irrelevant; perhaps [the structure] is converted by Morphology to a 'phonological word' not subject internally to the LCA, assuming that the LCA is an operation that applies after Morphology'. Put differently, Morphology converts [$_C$ wh [$_C$ C]] into a single terminal element through a process of morphological reanalysis akin to the fusion operation of Distributed Morphology (Halle and Marantz [34]). This morphological reanalysis makes the big structure invisible to the LCA.

This story seems to work quite well, but there are problems. One problem relates to children who produce medial-*wh*s. For these children, successive cyclic movement has to be able to proceed along the lines above, namely involving adjunction to the

intermediate head. There is no evidence in the input that this is the case, and furthermore, the kid will have to learn one way or other that successive cyclic movement in English happens by way of phrasal movement through specifiers. It is not clear what the relevant cue would look like here.

However, there are more serious problems, namely empirical problems. According to Nunes, adjunction can only happen if the *wh*-phrase is a head. He used data such as the following from German dialects to support this view:

(12) *Wessen Buch glaubst du **wessen Buch** Hans liest?
whose book think you whose book Hans reads
'Whose book do you think Hans is reading?' (McDaniel [55:183])

Whereas this seems to be true for German, Felser [27] observes that data such as the following from Afrikaans are problematic for Nunes [27].

(13) met wie het jy nou weer gesê **met wie** het Sarie gedog **met wie**
with who did you now again said with who did Sarie thought with who
gaan Jan trou?
go Jan marry (du Plessis [25:725])
'Whom did you say (again) did Sarie think Jan is going to marry?'

In this case we have PPs that are doubled in intermediate positions. If these were the only kind of complex constituents that were allowed in intermediate position, one could

perhaps say that these PPs are reanalyzed somehow. The data given in (14)-(16), however, suggest that this is implausible.[6]

(14) Van watter vrou het jy gedink **van watter vrou** het hulle
of which woman have you though of which woman have they
gister gepraat?
yesterday talked
'Which woman do you think they talked about yesterday?'

(15) met watter meisie het jy gese **met watter meisie** wil Jan trou?
with which girl have you said with which girl wants John marry
'Which girl do you say John wants to marry?'

(16) a. Watter meisie sê hy **watter meisie** kom vanaand kuier?
which girl say he which girl come tonight visit
'Which girl did he say is coming to visit tonight?'

b. Watter mooi meisie sê hy **watter mooi meisie** kom vanaand
which beautiful girl say he which beautiful girl come tonight
kuier?
visit
'Which beautiful girls did he say is coming to visit tonight?'

[6] I am grateful to Theresa Biberauer (p.c.) for the Afrikaans data. Note that these structures are best with emphatic intonation on all the *wh*-words. However, given that the structures are still possible without such an intonation, it does not seem plausible to argue that emphasis somehow creates a unit of sorts, which could be sufficient for Nunes's purposes to count as a 'word'.

(14) and (15) are cases where a complex PP occurs in a medial position whereas (16) shows two examples of complex DPs. All of these are fine in (colloquial) Afrikaans. Together, these data clearly show that narrow syntax needs to allow for even complex DPs and PPs to be generated in such a way that we can get multiple pronounced copies. Of course, this also raises the question how we deal with the difference between German and Afrikaans, a topic that I will return to below. But as a general analysis of medial-*wh* phenomena, it should be clear that Nunes' account is inadequate. Let me now turn to a different account that only tries to derive the child data, namely that of Jeong [39].

4. Jeong and Null Complementizers

Jeong [39] suggests a very interesting minimalist analysis of the medial-*wh* data produced by English children. In this section, I am going to adopt Jeong's proposal and I will also show how this can derive the absence of medial-*wh*s in infinitives. Lastly I discuss an asymmetry between production and comprehension regarding the appearance of medial-*wh*s in infinitives.

Jeong's point of departure is that phonetically null complementizers in English (and possibly other languages, cf. Richards [65]) are affixes. In particular, these null complementizers are affixes that need to attach to the immediately dominating verb. This assumption is able to account for the contrast in (17), as shown by Pesetsky [60] and Bošković and Lasnik [14] (see also Stowell [73] and Kayne [45]).

(17) a. John expected that/Ø Mary would come.

b. That/*Ø Mary would come was expected.

In (17a), the null complementizer affix can attach to the verb whereas in (17b), the affix is not close enough to the verb. Jeong argues that children don't know the exact specification of null complementizers. Whereas they know that null complementizers are affixes, they don't know which elements these affixes can be attached to. In English, null complementizer affixes cannot attach to nouns, as shown in (18).

(18) a. the claim that Mary was ill upset me.

b. *the claim Ø Mary was ill upset me.

However, for children that produce medial-*wh*s, they need to allow affixes to combine with *wh*-phrases.[7] Jeong assumes that children know that affixes cannot be stranded (Lasnik's [51] Stranded Affix Filter) and that null complementizers attach to Vs. What they also do, then, is to entertain the possibility that *wh*-phrases can attach to the null C. As Jeong argues, the fact that the null C can attach to the *wh*-phrase plausibly forces pronunciation of the medial-*wh* because an affix cannot attach to something that is not pronounced. Interestingly, this also accounts for why children only pronounce copies of *wh*-phrases in SpecCP and not in the original landing site or any other intermediate positions, as in (19a). A simplified representation of the derivation is given in (19b)-(19c) where COMP illustrates the affixal null complementizer.

[7] Jeong [39:14] actually says 'nouns', but this cannot be right since we have seen cases where adjuncts are pronounced as medial-*wh*s (cf. (3d)).

(19) a. **Who** do you think [$_{CP}$ **who** [$_{IP}$ the cat [$_{vP}$ <u>who</u> chased <u>who</u>]]

b. [who do you think [$_{CP}$ **who** [$_{C}$ COMP [$_{IP}$ the cat [$_{vP}$ <u>who</u> chased <u>who</u>]]]]]

c. [who do you think [$_{CP}$ **who** + COMP [$_{IP}$ the cat [$_{vP}$ <u>who</u> chased <u>who</u>]]]]

It is clear why the underlined copies cannot be pronounced: the affix that forces pronunciation of a *wh*-copy is located in C and not in *v* or V.

Notice that Jeong's approach makes two predictions. One is that medial-*wh* elements are not expected to show up in children's production in languages with embedded inversion, assuming C is filled, as in Belfast English. Or more precisely: medial-*wh* elements should not show up at the time when embedded inversion is acquired. The second prediction is that medial-*wh* should not appear in languages with overt complementizers co-existing with *wh*-elements, as in Basque and other languages. As far as I know, the first prediction is borne out. The second one is a bit trickier. Giving data from a Spanish child, Gutiérres Mangado [34] shows that this child produces medial-*wh*s with co-occuring complementizers. An example is given in (20).

(20) Dónde crees **que dónde** ha ido el señor?

where think.2SG that where has gone the man

Target: 'Where do you think the man went?' (Gutiérres Mangado [34:269])

However, it's not clear that Jeong's analysis cannot be applied to such data. It may be that the *wh*-element is in a different functional projection, say SpecFocP (Rizzi [68])

from the complementizer *que*, which may be lexicalizing the Force head.[8] If that is the case, there could be a silent complementizer that would force pronunciation of the medial-*wh*. Such an analysis does not seem to be far-fetched, and insofar as it can be maintained, the second prediction goes through.

In what follows, I basically adopt and extend Jeong's proposal. In particular I am adopting the idea that what triggers the pronunciation of medial-*wh*s in English child language is the null complementizer affix. I also suggest that the kids assume that only simplex *wh*-elements can be pronounced as medial-*wh*s, maybe for phonological reasons along the lines of the adult grammars; see below. This also accounts for the fact that children never produce structures like (21).

(21) *Which boy do you think **which mouse** the cat chased?

Instead, a few children would produce structures like (22).

(22) **Which boy** do you think **who** the cat chased?[9]

---

[8] This raises the question of how we know where affixal null complementizers are located in a split-CP. I assume that languages can differ, just as they may differ concerning where overt complementizers are merged. Independent evidence that may bear on this issue could be *that*-trace effects, cf. Lohndal [53].

[9] Similar cases can be found in German dialects:

(i) **Welchen Mann** denkst du **wen** er kennt?
which man think you who he knows
'Which man do you think he knows?' (Fanselow and Ćavar [26:18])

The problem, though, as Jeong [39] points out, is how *which boy* can leave a copy like *who*, given that copies are identical. Jeong suggests a Distributed Morphology solution to this problem. Distributed Morphology assumes that there is a distinction between syntactic features, which are the only features that the syntax operates on, and morphophonological features that are added to syntactic feature bundles upon Transfer. Late insertion of this kind makes a discrepancy between syntactic and morphophonological features possible. Jeong suggests that in *which boy*, the relevant syntactic features are [+wh, +singular, +human, +masculine], which are also features shared by *who*. *Who* can thereby function as an exponent of the syntactic features that characterize *which boy*.

Structures like (22) are less frequent than medial-*wh*s with simplex *wh*-phrases (Thornton [74, 75]). Jeong suggests that this is because D-linked *wh*-phrases involve an extra morphological operation, namely that of turning a complex *wh*-phrase into a head-like element in the morphophonology.

The flipside of (22), which also occurs in children's production, is what looks like partial movement (23).

(23) What do you think **who** jumped over the fence? (Thornton [74:213])

In partial movement constructions, the topmost *wh*-phrase (*what* in (23)) acts like a scope marker for the downstairs *wh*-phrase (*who* in (23)). Similar constructions are found in German dialects, Romani and other languages, and I will briefly return to them in section 3. Felser [27] argues convincingly that these structures should be analyzed in a different

way than medial-*wh* structures (see also Bruening [15] for a similar and strong argument from Passamaquoddy). Therefore I will not have anything to say about them in this paper. Whatever one's favorite analysis of partial movement may be will be compatible with the data in (23).

There is one fact that Jeong [39] does not discuss. Thornton [74:213] says that no child ever produced a medial-*wh* in infinitival clauses (24). Instead, children produced adult forms as in (25).

(24) a. *What do you want what to eat?

b. *Who do you want who to eat the pizza?

(25) a. What do you wanna eat?

b. Who do you want to eat the pizza?

Thornton marks (24) as ungrammatical. It's straightforward to account for this by extending Jeong's proposal. Infinitivals don't have a null complementizer in English (though see Kayne [45] and Pesetsky [60] for complications), so there is nothing for the medial-*wh* to attach to. Consequently the medial-*wh* cannot be pronounced.

It should be added that other experiments have shown that some children judge these sentences grammatical. McDaniel, Chiu and Maxfield [56] tested children's (aged 2;11 – 5;7) judgments on the two sentences in (26).

(26) a. Who do you want **who** to cook dinner?

b. Who do you want **who** to kiss?

I have excerpted the relevant results from McDaniel, Chiu and Maxfield [56:724] in Table 1 below.[10]

Infinitive type	Session 1 (N = 32)	Session 2 (N = 32)	Session 3 (N = 24)	Session 4 (N = 15)
Subject	28%	22%	21%	13%
Object	22%	13%	13%	20%

*Table 1: Percentage of acceptance for medial-*wh*s in infinitives.*

As McDaniel, Chiu and Maxfield [56:732, fn. 25] point out, there is no direct contradiction between their data and Thornton's [74]. It is perfectly possible that subjects find a particular sentence acceptable but they are not producing it in a task. It is not clear *why* there is such an asymmetry, though production-comprehension asymmetries are quite frequently attested in child language. It is likely that they have a variety of reasons, ranging from performance to competence reasons. I do not have anything new to add here. What is important for present purposes is that Jeong's account can be easily adapted to account for the lack of medial-*wh*s in infinitives.

We have now seen how we can analyze medial-*wh*s in English child language. In the next section I will discuss how medial-*wh*s in adult grammars can be analyzed.

[10] The study consisted of four sessions, which were separated from each other by a period of three to four months. Not all children took part in all sessions, which can be seen in Table 1. There was some attrition, and any child who manifested adult knowledge of the constructions investigated three sessions in a row and was at least five years old as of the third session was not seen for a fourth session.

5. Medial-*wh*s in adult grammars

The aim of this section is to account for the data in (2), repeated here as (27).

(27) a. Wen glaubt Hans **wen** Jakob gesehen hat? *German dialects*
whom thinks Hans whom Jakob seen has
'Who does Hans think Jakob saw?' (McDaniel [55:183])

b. Kas misline **kas** o Demìri dikhlâ? *Romani*
whom you.think whom Demir saw
'Who do you think Demir saw?' (McDaniel [55:182])

The first task is to explore whether Jeong's [39] account of the child English data can be extended to the German dialects and Romani. Recall that her hypothesis was that the child overgeneralizes the licensing requirements on null complementizers so that they could combine with *wh*-phrases. This even goes beyond just nouns since we have seen that adjunct *wh*-phrases can be pronounced medially. If we were to use this analysis for adults who have medial-*wh*s as part of their I-language, that would mean that somehow these adults figured out that only verbs (cf. the data in (17)) and *wh*-phrases can attach to a null complementizer affix. This is clearly not a natural class, i.e., it is not clear why this should be the relevant class as opposed to some other class.

There is a more serious problem, though. First, some German dialects allow medial-*wh*s with an overt complementizer, as shown in (28).

(28) Wen denkst du **wen dass** du eingeladen hast?

who think you who that you invite have

'Who do you think that you have invited?'

(modeled on Fanselow and Ćavar [26:127])

This is similar to the Spanish case above in (20). However, those data come from one single child, and it is always hard to build too much of an analysis on one case study. More importantly, though, the medial-*wh* structures for adults are typically optional. This sets the structures apart from the ones that children produce. For children, they often produce the medial-*wh*s for a very limited time, and when they do so, they often do it frequently (Thornton [74]). To the extent that this is a general property across languages, it seems like medial-*wh*s in child language is less optional than for adults.

Another potential problem faces us if we want to use Jeong's theory for the adult structures. We would have to say that there are two silent complementizers in these grammars. One is able to attach to *wh*-elements and the other is not. That would give us the facts, but at the cost of giving an account that basically just re-describes the facts. In that sense, it would not be a very deep account.

Although these problems are not in and of themselves lethal for Jeong's analysis, they seem to suggest that the pronunciation of medial-*wh*s is not due to properties of the complementizers in the grammar of the adults that produce these structures. Below I will explore a different analysis, which says that a quite different derivation takes place in grammars that allow medial-*wh* structures. Specifically, I will suggest an extension of Chomsky's [21] parallel movement (cf. independent work by Kandybowicz [43] and

Aboh and Dyakonova [1]), and I will argue that this accounts for the adult medial-*wh* data. Let me first present Chomsky's concept of parallel movement and its motivation, before I go on to extend it.

Chomsky [21] is concerned with how structures like (29) should be analyzed.

(29) Who saw John?

He suggests the following representations:

(30) a. C [T [who [*v** [see John]]]]

b. Who$_i$ [C [who$_j$ [T [who$_k$ *v** [see John]]]]] (Chomsky [21:149])

Here there is *parallel movement* from Spec*v**P to SpecTP and SpecCP. This means that *who* moves simultaneously to SpecTP and SpecCP, thus we have two chains, namely (*who*$_i$, *who*$_k$) and (*who*$_j$, *who*$_k$). The motivation for introducing parallel chains is that it gives us the distinction between A-chains and A-bar chains, and it is triggered by two kinds of features. The movement to SpecTP is related to phi-features that are inherited from C whereas the movement to SpecCP is driven by an Edge Feature on C. However, if there are two chains, and if, as is a common assumption, only one element of a chain is pronounced, how come (29) isn't pronounced as (31)?[11]

[11] Cf. Chomsky [20:13] on why only one element is pronounced: 'If language is optimized for satisfaction of interface conditions, with minimal computation, then only one will be spelled out, sharply reducing phonological computation'.

(31) *Who who saw John?

Chomsky [21:150] says that: 'By the usual demand of minimal computation, the A-chains contain no pronounced copy'. Instead, I will assume that (31) is ruled out by a syntactic version of the Obligatory Contour Principle (*XX), which is independently motivated, cf., Grimshaw [32], Ackema [2], van Riemsdijk [66] and Ott [59].

In what follows, I will suggest an extension of this notion of parallel movement. Chomsky uses it to derive the distinction between A- and A-bar movement, and I will suggest that we extend this analysis to also involve parallel movement of two A-bar chains. There does not seem to be anything that would bar this extension, rather, preventing it would require a motivation.[12] I will also follow Nunes [58] in assuming that the highest member of a chain is pronounced, presumably because this member has checked more features than the lower chain members. It is not clear how to reconcile this with a view where all copies are identical (Chomsky [21]), but I set that issue aside for present purposes.

It is easiest to see how this would work by considering a specific derivation. I will first work through two concrete examples, and then I will consider the difference between German dialects on the one hand and Afrikaans on the other hand can be derived, namely why only Afrikaans allows D-linked *wh*-phrases to occur in medial-*wh*s.

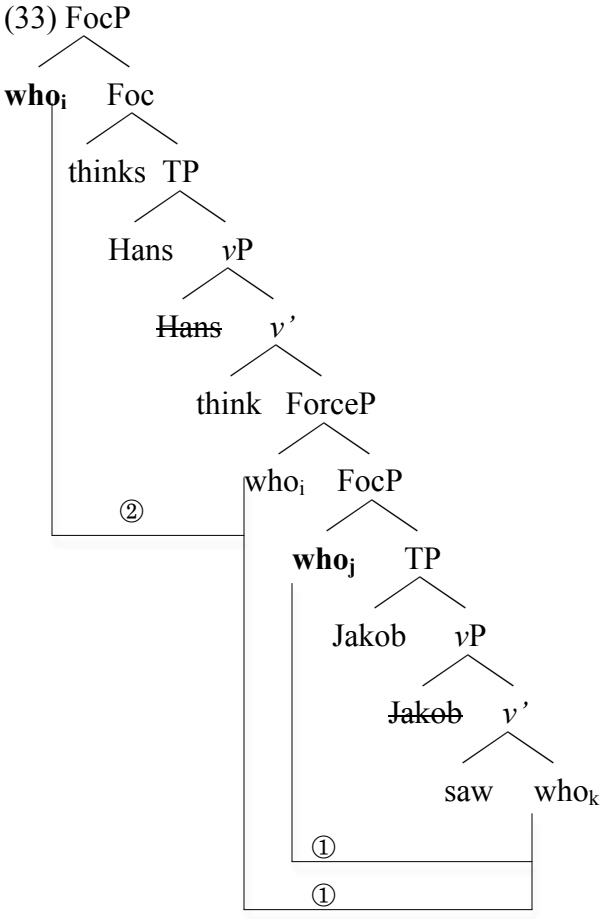
Let us look at a case where there is one medial-*wh*. In the tree in (33), I use bold to indicate which *wh*-phrase is spelled-out, and I have numbered the movement

[12] This is in particular true if one believes that the A/A-bar distinction should be eliminated.

operations so that it is easier to keep track of them. Indices are also used to aid exposition, but they have no theoretical importance. Lastly, I do not show possible intermediate landing sites (e.g. the left edge of *v*P) in order to simplify the structures.

(32) Wen glaubt Hans **wen** Jakob gesehen hat? *German dialects*
whom thinks Hans whom Jakob seen has
'Who does Hans think Jakob saw?' (McDaniel [55:183])

(33)



Let us go through what happens in (33) step by step.[13] Parallel movement applies to who_k, creating who_i and who_j, which each moves to the left periphery of the embedded clause (movement ①). The two *wh*-phrases target different projections in the left periphery, namely SpecFocP and SpecForceP. Rizzi [68] argues convincingly that interrogatives and focus compete for the same position, namely SpecFocP (see also Stoyanova [72] for further and interesting developments of this idea in a cross-linguistic perspective). I follow Nunes [58] who argued that whenever an element moves to SpecFoc(us)P, it has to be pronounced because of the focus properties associated with it.[14] Finally, the topmost *wh*-phrase in SpecForceP moves to the left periphery in the main clause (movement ②).[15] The fact that the highest moves is derived by a principle such as Attract Closest, which says that the closest target always is attracted. Assuming that the Force head is a phase head (cf. Rizzi [69], Julien [40]), one also derives this result because only the highest *wh*-phrase is available for further operations since the lower *wh*-phrase is within the Spell-

[13] Boeckx [9] argues persuasively that movement to intermediate landing sites is not triggered by a feature (see also Bošković [13]), which I will adopt here.

[14] An interesting question arises concerning cases discussed by Chomsky [17] where focus can be assigned to various constituents in a structure. An example is shown in (i), where capital letters indicate focus.

(i) a. JOHN eats a banana.

b. John eats a BANANA.

As Chomsky discusses, this is related to the intonation contour. See Frascarelli [28] for an analysis of cases like (ib). Thanks to Mark Baker (p.c.) for raising this point.

[15] Grewendorf and Kremers [31] argue that minimality does not hold on Chomsky's [21] assumptions. Since this potential problem is irrelevant for present purposes given that all the *wh*-phrases are identical, I set it aside.

Out domain of the Force head. As a result, we see that the heads of both chains are in SpecFocP, which is a position that requires pronunciation.

There are some potential questions that we should consider before moving on to a more complex case. One issue is whether anything would change if parallel movement took place from an adjoined *v*P position. The answer is no (though see Grewendorf and Kremers [31] for a different answer). It makes no difference whether parallel movement happens from the base position or from an intermediate landing site. I have only shown one derivation, but parallel movement from an intermediate site yields another convergent derivation. Another question is what happens if there is parallel movement of who_i in (33). Wouldn't that predict two occurrences of *wh*-elements in the main clause left periphery? Such a derivation would be ruled out by a syntactic version of the Obligatory Contour Principle, on a par with **Who who saw John*. Another issue concerns the semantics. Usually it is assumed that LF cares about chain membership, but in this case we only want to say that the *wh*-phrase is interpreted once, despite there being two chains in the syntax. There are various ways one can get this result, either by stipulating that the semantics only sees one copy (Hornstein [37]), or by invoking some notion of chain collapse (Martin and Uriagereka [54]). The exact details do not matter much for present purposes; it is only important that there is a way to ensure that the semantics recognizes that there is one *wh*-phrase despite there being multiple chains in the syntax.

Let us now move on to consider a case with two medial-*wh*s, that is, a case like (34).

(34) met wie het jy nou weer gesê **met wie** het Sarie gedog **met wie**

with who did you now again said with who did Sarie thought with who

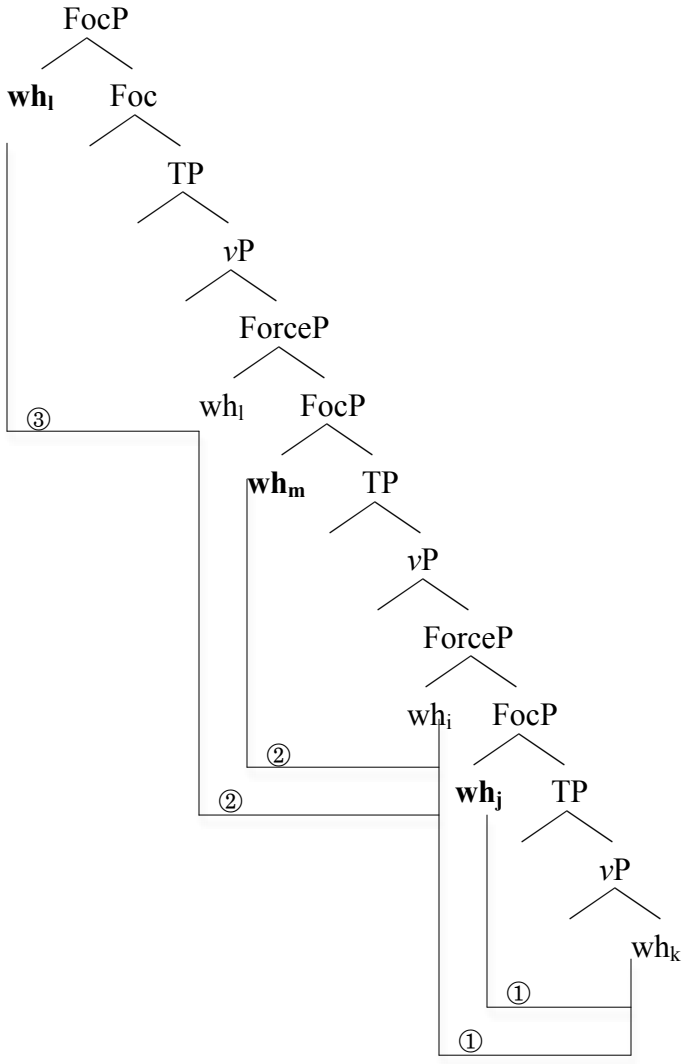
gaan Jan trou?

go Jan marry (du Plessis [25:725])

'Whom did you say (again) did Sarie think Jan is going to marry?'

In order to simplify the tree, I have only put in the *wh*-phrases. This makes it easier to see exactly how the derivation works. The conventions for illustrating the derivation are the same as for (33). Boldface is used for spelled-out *wh*-phrases, the movement steps are numbered, and possible intermediate landing sites are not shown.

(35)



The difference between (33) and (35) is that in (35), there is parallel movement from SpecForceP in the most embedded clause. There are no restrictions on where parallel movement can occur, so it can just as easily occur from SpecCP, Spec*v**P and from the base position. Otherwise the derivation is very similar to (33) and no new technology is introduced.

A question that we need to deal with is how we can account for the difference between German dialects and Romani on the one hand and Afrikaans on the other (below I will deal with the difference between medial-*wh* languages and English). Recall that only the latter allows complex *wh*-phrases to be pronounced in medial positions. It is likely that language-particular rules determine the size of the medial-*wh* element (cf. Benincà and Poletto [6] and Poletto and Pollock [62] for Romance, see also, more generally, Bošković [12], Landau [50] and Grohmann [33]). It is also plausible that these are phonological rules, i.e. that there are restrictions on how many syllables the *wh*-word may have. These rules will then apply after the syntactic rules.[16] I leave the specific implementation of this for future work.

We saw above that English child language has structures that look very much like partial *wh*-movement of the kind one finds in German (36).

(36) Was glaubt Hans **mit wem** Jakob jetzt spricht? *German*
what believes Hans with whom Jakob now talks
'What does Hans believe with whom Jakob is now talking?'
(McDaniel [55:111])

[16] An alternative is to say that D-linked *wh*-phrases have a different syntax, i.e., that they involve a referentiality phrase that is located at the top of the tree. (cf. Thornton [75] and references therein). In the present context, one would have to say something like the following: whether an embedded referential phrase can be pronounced or not is a matter of variation. This seems to be nothing but a re-description of the facts.

Again, I will not present any analysis of these for the same reason as mentioned above, namely that partial-*wh* movement probably should be analyzed differently than medial-*wh*s (Felser [27]). There is, however, a fact that is common to both partial *wh*-movement and medial-*wh*s, namely that negation cannot intervene. Long extraction when only one copy is pronounced is acceptable.

(37) a. Wie denk je niet dat zij uitgenodigd heeft? *Dutch*
who think you not that she invited has
'Who don't you think she has invited?'

b. *Wat denk je niet wie zij uitgenodigd heeft?
what think you not who she invited has

c. *Wie denk je niet wie zij uitgenodigd heeft?
who think you not who she invited has
(Barbiers, Koeneman and Lekakou [5])

As Barbiers, Koeneman and Lekakou [5] point out, it is not clear how these data should be analyzed, in particular because other operator structures (like pure quantifiers) do not show similar patterns. Rett [63] also argues that the restriction is not syntactic but rather semantic. One could also add that if analyses that do not analyze the *wat* in (37b) as originating together with *wie* are correct, then that adds another argument against (37b, c) being the result of a syntactic violation. My goal in this paper is not to provide an account of the restriction seen in (37), and given the lack of consensus in the literature as well, it

seems that a larger investigation into these restrictions and their nature is required. Such an investigation clearly goes beyond the scope of the present paper.

There are some remaining questions that we need to consider. Above I have presented some reasons for why the analysis of children's medial-*wh*s does not carry over to adult medial-*wh*s. Here I want to consider the reverse question, namely: If the parallel movement analysis is the correct one for adult medial-*wh* grammars, why isn't this the correct analysis of children's grammar? First it should be made explicit that the parallel movement analysis would give us the same result for the cases in English child language as for the adults. However, given that children are only producing medial-*wh*s for a limited amount of time (obviously setting aside children who acquire medial-*wh* languages), it is not at all clear what would make them entertain the parallel movement derivation and then later discard it when they discover that the derivation yields sentences that are not part of what we call English. Jeong's [39] analysis clearly states that children are misinterpreting properties of the null complementizer, which is a likely misinterpretation/overgeneralization given that children nevertheless need to figure out the restrictions on affixation to null complementizers. However, if children are at first entertaining a non-medial-*wh* derivation, what would make them entertain a medial-*wh* derivation if the latter is derived through parallel movement? There is no obvious trigger for them to change their hypothesis about the grammar. However, on the null complementizer analysis, there is an indirect trigger in that the children have to figure out what the licensing properties of null complementizers are. The latter can come closer to providing an explanation for why children do what they do, whereas the parallel movement analysis would not do that. Instead, the parallel movement hypothesis would

just state that children are exploring possibilities of which Universal Grammar provides. Although that is a perfectly valid hypothesis, we should prefer a theory that can come closer to saying something about why things happen the way they do. As far as I can see, only the null complementizer analysis does that.

What I have just argued, amounts to saying that the parallel movement derivation needs an overt trigger. When children are acquiring Romani or these German dialects, they presumably get medial-*wh*s as part of their input. Therefore it should come as no surprise that these children will grow up producing these structures using parallel movement. English children presumably never hear medial-*wh*s as part of their relevant input, and although they may produce them for a short amount of time, they never end up doing so permanently. Therefore *wh*-movement does not proceed in parallel in English, with the exception of the structures that Chomsky [21] claims are derived through parallel movement.

This is a natural place to pause and consider a bigger issue that relates to the topic of this paper, namely how we think about linguistic variation. In the next section I will discuss some implications of the analysis I have suggested in sections 4 and 5.

6. Consequences for parametric theory

I have argued that when English children produce medial-*wh*s, they do it in a different way than adult speakers of German dialects. Most of the previous literature has argued that the English children are doing roughly the same as German adults (Thornton [74], McDaniel, Chiu and Maxfield [56], Gutiérres Mangado [34]), though they disagree on the details. McDaniel, Chiu and Maxfield [56] argue that there is something like a *wh*-

parameter distinguishing medial-*wh* languages from non medial-*wh* languages, whereas Thornton [74] argued that there is no such parameter. Instead children are using the medial-*wh* as a complementizer, signalizing that Spec-head agreement has taken place, which is necessary to satisfy the Empty Category Principle in the framework she is assuming (roughly that of Rizzi [67]). Contrary to these proposals, I am arguing that the derivation underlying English children's medial-*wh*s and the derivation underlying German adults' medial-*wh* are different. The former involves overgeneralizing the licensing requirements for null complementizer affixes whereas the latter involves parallel movement. If this is on the right track, it has implications for how we conceive of parametric variation. In the remainder of this paper, I will discuss this issue.

In the late 1970s, the question of variation among languages became more and more pressing within Chomskyan generative grammar (see Lasnik and Lohndal [52] for some discussion of the history). A theory that distinguished between principles and parameters was developed. The principles were assumed to be universal and part of UG. Concerning parameters, Chomsky [18:4] said that: 'If […] parameters are embedded in a theory of UG that is sufficiently rich in structure, then the languages that are determined by fixing their values one way or another will appear to be quite diverse'. The idea was that the parametric space is finite and innately specified through UG. Put in the words of Baker [3:19], 'we may think of parameters as the atoms of linguistic diversity'. On this approach, a fundamental assumption was that UG should reflect typological generalizations. However, recent work has questioned this assumption. In particular, Newmeyer [57] has questioned the empirical foundation and argued that parameters cannot capture typology the way we used to think. I will not review all his evidence in

favor of this, but just add that several other researchers have come to the same conclusion (for useful discussion, see Baker [4], Gallego [29], Richards [64], Hornstein [38], Kandybowicz [44], Boeckx [10] and van Gelderen [30]). Besides the empirical issues, there are also theoretical reasons why one should be suspicious towards encoding typological variation in UG, especially if one is wearing one's minimalist hat. Since the first minimalist papers in Chomsky [19], Chomsky has argued that there should be no variation among languages at the level of Logical Form, known as the Uniformity Principle: 'In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety restricted to easily detectable properties of utterances'. Recently, Boeckx [10] has strengthened this to what he calls the Strong Uniformity Thesis:

(38) Principles of narrow syntax are not subject to parametrization; nor are they affected by lexical parameters

Put differently, narrow syntax itself is entirely universal, thus it does not correspond to any given language. The question then arises where we put the variation we know exists.

A prominent proposal is to locate the variation in the lexicon, as differences related to lexical elements (Borer [11:3]; cf. Chomsky [19], Kayne [47, 48]). In addition we have word order variation etc., which can be related to lexical elements (e.g., as strong and weak features as in Chomsky [19]). The strongest hypothesis would say that all variation is related to the externalization component, namely PF (Chomsky [20], Boeckx [10]). On this approach, narrow syntax is universal and not the locus of variation.

Going back to the medial-*wh* cases, it is clear that the analysis I have given above is only consistent with the latter view of parameters. The overgeneralization I have argued that children who are producing medial-*wh*s are making is a property of the null complementizer, that is, a lexical element. I have also argued that a similar analysis cannot be used for the adults. Instead I have suggested a specific implementation of parallel movement. This computation is 'input driven' in that it only seems to appear when there is positive evidence; it is not something that the child (or an adult English speaker) just decides to do at random, if I am right. Thus, English children have the same computational system as German children, but because of different input, they end up doing what look like similar things in different ways. There is no contradiction between this and the continuity hypothesis, which maintains that child language can differ from the language of the linguistic community only in ways that adult languages can differ from each other (Crain and Thornton [22]).

The empirical arguments against encoding typological variation in UG appear to be solid, as Newmeyer [57] convincingly argues. But notice that there is also a deeper motivation behind why it is better to put the variation outside of UG. To see this, let us look at what Chomsky [20] calls the three factors in the design of language. Chomsky says that: 'Assuming that the faculty of language has the general properties of other biological systems, we should, therefore, be seeking three factors that enter into the growth of language in the individual' (Chomsky [20:6]). These three factors are given in (39).

(39) a. Genetic endowment

b. Experience

c. Principles not specific to the faculty of language

It is possible to say that whereas the Government and Binding view was that parametric variation is located in (39a), the more minimalist view is that the variation is related to experience and possibly to third factors, as (39c) is commonly called (see in particular Richards [64] on the latter, and also Roberts and Holmberg [70]). If it is possible to show that all variation belongs to (39b) and (39c), then we would be close to a view of UG as underspecified, contra the view that UG is overspecified (Yang [76], Baker [3]). Such a view of UG as underspecified was actually explored in the 1980s by Richard Kayne and Juan Uriagereka, both inspired by Changeux [16], but has not been developed in any great detail since (though see Holmberg in press).

Summarizing, I have pointed out that the view of parameters within Chomskyan generative syntax is changing, and I have argued that the cases discussed in this paper lend further support to this change. On the new view, the range of variation among languages is no longer innately specified, though it is still innately given. Parameters are trivialized as being part of what a child needs to learn, which can be taken to imply that they are no longer necessarily binary. UG is the core universal part that is common to all languages, and as such it does not encode variation.

7. Conclusion

The goal of this paper has been to analyze medial-*wh* structures in both child grammars where medial-*wh* occurs despite not being present in the input and in adult grammars

where medial-*wh*s occur regularly. I have argued that children are overgeneralizing the licensing requirement on null complementizer affixes (following Jeong [39]), and that this is different from what adults are doing in languages that have medial-*wh*s. In the latter case, I have argued that medial-*wh*s are derived through parallel movements combined with the assumption that only one element per chain is spelled out. I have also claimed that the current analysis is compatible with a minimalist view of parametric variation.

Works Cited

1. Aboh, E. O. and M. Dyakonova. 2009. Predicate doubling and parallel chains. *Lingua* 119: 1035-1065.

2. Ackema, P. 2001. Colliding complementizers in Dutch: Another OCP effect. *Linguistic Inquiry* 32: 717-727.

3. Baker, M. C. 2001. *The Atoms of Language*. New York: Basic Books.

4. ----. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press.

5. Barbiers, S., O. Koeneman and M. Lekakou. In press. Syntactic doubling and the structure of *wh*-chains. *Journal of Linguistics*.

6. Benincà, P. and C. Poletto. 2005. On some descriptive generalizations in Romance. In *The Oxford Handbook of Comparative Syntax*, ed. by Guglielmo Cinque and R. S. Kayne. Oxford: Oxford University Press.

8. Berwick, R. C. and N. Chomsky. 2008. The Biolinguistic Program: The

Current State of its Evolution and Development. Ms., M.I.T. [Forthcoming in *Biolinguistic Investigations*, ed. by A-M. Di Sciullo and C. Aguero. Cambridge: M.I.T. Press.]

9. Boeckx, C. 2007. *Understanding Minimalist Syntax*. Malden: Blackwell.

10. ----. In press. Approaching Parameters from Below. *Biolinguistics: Language evolution and variation*, ed. by A-M. Di Sciullo and C. Boeckx. Oxford: Oxford University Press.

11. Borer, H. 1984. *Parametric Syntax*. Dordrecht: Foris.

12. Bošković, Ž. 2001. *On the nature of the syntax-phonology interface: cliticization and related phenomena*. London: Elsevier.

13. ----. 2007. On the Locality and Motivation of Move and Agree: An Even More Minimal Theory. *Linguistic Inquiry* 38: 589-644.

14. ---- and Howard Lasnik. 2003. On the distribution of null complementizers. *Linguistic Inquiry* 34: 527-546.

15. Bruening, B. 2006. Differences between the Wh-Scope-Marking and Wh Copy Constructions in Passamaquoddy. *Linguistic Inquiry* 37: 25-49.

16. Changeux, J-P. 1981. Genetic Determinism and Epigenesis of the Neuronal Network: Is There a Biological Compromise between Chomsky and Piaget? In *Language and Learning: The Debate between Jean Piaget and Noam Chomsky*, ed. by Massimo Piattelli-Palmarini. Cambridge: Harvard University Press.

17. Chomsky, N. 1971. Deep Structure, Surface Structure, and Semantic

Interpretation. In *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology*, ed. by D. Steinberg and L. Jakobovits. Cambridge: Cambridge University Press.

18. ----. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
19. ----. 1995. *The Minimalist Program*. Cambridge: M.I.T. Press
20. ----. 2005. Three factors in language design. *Linguistic Inquiry* 36: 1-22.
21. ----. 2008. On Phases. In *Foundational Issues in Linguistic Theory*, ed. by R. Freidin, C. P. Otero and M. L. Zubizarreta. Cambridge: M.I.T. Press.
22. Crain, S. and R. Thornton. 1998. *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. Cambridge: M.I.T. Press.
23. De Villiers, J., T. Roeper and A. Vainikka. 1990. The Acquisition of Long distance Rules. In *Language Processing and Language Acquisition*, ed. by L. Frazier and J. de Villiers. Dordrecht: Kluwer.
24. den Dikken, M. 2009. On the nature and distribution of successive cyclicity. Ms., The Graduate Center of the City University of New York.
25. du Plessis, H. 1977. *Wh* Movement in Afrikaans. *Linguistic Inquiry* 8: 723-726.
26. Fanselow, G. and D. Ćavar. 2001. Remarks on the Economy of Pronounciation. In *Competition in Syntax*, ed. by G. Müller and W. Sternefeld. Berlin: Mouton de Gruyter.
27. Felser, C. 2004. *Wh*-copying, phases, and successive cyclicity. *Lingua* 114: 543-574.
28. Frascarelli, M. 2000. *The Syntax-phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht: Kluwer.

29. Gallego, Á. J. 2008. The Second Factor and Phase Theory. Ms., Universitat Autònoma de Barcelona.

30. van Gelderen, E. In press. *The Linguistic Cycle: Language Change and the Language Faculty*. Oxford: Oxford University Press.

31. Grewendorf, G. and J. Kremers. In press. Phases and cyclicity: Some problems with phase theory. *The Linguistic Review* 26: 385-430.

32. Grimshaw, J. 1997. The best clitic: Constraint conflict in morphosyntax. In *Elements of Grammar*, ed. by L. Haegeman. Dordrecht: Kluwer.

33. Grohmann, K. K. 2008. Copy Modification and the Architecture of the Grammar. Paper presented at the LAGB, University of Essex, September 10-14.

34. Gutiérres Mangado, M. J. 2006. Acquiring long-distance wh-questions in LI Spanish. In *The Acquisition of Syntax in Romance Languages*, ed.by V. Torrens and L. Escobar. Amsterdam: John Benjamins.

35. Halle, M. and A. Marantz. 1993. Distributed Morphology and the pieces of inflection. In *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, ed. by K. Hale and S. J. Keyser. Cambridge: M.I.T. Press.

36. Holmberg, A. In press. Parameters in minimalist theory: The case of Scandinavian. *Theoretical Linguistics*.

37. Hornstein, N. 1995. *Logical Form*. Malden: Blackwell.

38. ----. 2009. *A Theory of Syntax*. Cambridge: Cambridge University Press.

39. Jeong, Y. 2004. Children's question formations from a Minimalist Perspective. Ms., University of Maryland.

40. Julien, M. 2007. Embedded V2 in Norwegian and Swedish. *Working Papers in*

Scandinavian Syntax 80: 103-161.

41. van Kampen, J. 1997. *First Steps in Wh-movement*. Delft: Eburon.

42. ----. 2010. The learnability of A-bar chains. In *The Linguistic Enterprise: From knowledge of language to knowledge in linguistics*, ed. by M. Everaert, T. Lentz, H. De Mulder, Ø. Nilsen and A. Zondervan. Amsterdam: John Benjamins.

43. Kandybowicz, J. 2008. *The Grammar of Repetition: Nupe grammar at the syntax-phonology interface*. Amsterdam: John Benjamins.

44. ----. 2009. Externalization and Emergence: On the Status of Parameters in the Minimalist Program. *Biolinguistics* 3: 93-88.

45. Kayne, R. S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.

46. Kayne, R. S. 1994. *The Antisymmetry of Syntax*. Cambridge: M.I.T. Press.

47. ----. 2000. *Parameters and Universals*. Oxford: Oxford University Press.

48. ----. 2005. *Movement and Silence*. Oxford: Oxford University Press.

49. Koster, J. 2009. IM not perfect: The case against copying. Ms., University of Groningen.

50. Landau, I. 2006. Chain resolution in Hebrew V(P)-fronting. *Syntax* 9: 32-66.

51. Lasnik, H. 1981. Restricting the theory of transformations. Explanations in linguistics, ed. by D. Lightfoot and N. Hornstein. London: Longmans.

52. ---- and T. Lohndal. 2010. Government-Binding/Principles and Parameters Theory. *Wiley Interdisciplinary Reviews: Cognitive Science*.

53. Lohndal, T. 2009. Comp-t effects: Variation in the position and features of C. *Studia Linguistica* 63: 204-232.

54. Martin, R and J. Uriagereka. 2008. Uniformity and Collapse. Paper presented at Ways of Structure Building, University of the Basque Country, November 13.

55. McDaniel, D. 1986. Conditions on *Wh*-chains. City University of New York doctoral dissertation.

56. McDaniel, D., B. Chiu and T. L. Maxfield. 1995. Parameters for Wh Movement Types: Evidence from Child English. *Natural Language and Linguistic Theory* 13: 709-753.

57. Newmeyer, F. J. 2005. *Possible and probable languages: A generative perspective on linguistic typology*. Oxford: Oxford University Press.

58. Nunes, J. 2004. *Linearization of chains and sideward movement*. Cambridge: M.I.T. Press.

59. Ott, D. 2009. Stylistic fronting as remnant movement. *Working Papers in Scandinavian Syntax* 83: 141-178.

60. Pesetsky, D. 1992. Zero Syntax, vol. 2. Ms., MIT.

61. Phillips, C. 1996. Order and Structure. Doctoral dissertation, MIT.

62. Poletto, C. and J-Y. Pollock. 2009. Another look at wh-questions in Romance: The case of Mendrisiotto and its consequences for the analysis of French wh-in situ and embedded interrogatives. In *Romance Languages and Linguistic Theory 2006*, ed. by D. Torch and W. L. Wetzels. Amsterdam: John Benjamins.

63. Rett, J. 2006. Pronominal vs. determiner *wh*-words: evidence from the copy construction. In *Empirical issues in Syntax and Semantics 6*, ed. by O. Bonami and P. Cabredo Hofherr. Colloque de Syntaxe et Sémantique à Paris.

64. Richards, M. 2008. Two Kinds of Variation in a Minimalist System. In *Varieties of Competition*, ed. by F. Heck, G. Müller and J. Trommer. Linguistische Arbeits Berichte 87.

65. Richards, N. 2001. *Movement in language*. Oxford: Oxford University Press.

66. van Riemsdijk, Henk. 2008. Identity Avoidance: OCP Effects in Swiss Relatives. In *Foundational Issues in Linguistic Theory*, ed. by R. Freidin, C. P. Otero and M. L. Zubizarreta. Cambridge: M.I.T. Press.

67. Rizzi, L. 1990. *Relativized Minimality*. Cambridge: M.I.T. Press.

68. ----. The fine structure of the left periphery. In *Elements of grammar: A handbook of generative syntax*, ed. by L. Haegeman. Dordrecht: Kluwer.

69. ----. 2005. Phase theory and the privilege of the root. In *Organizing Grammar. Studies in Honor of Henk van Riemsdijk*, ed. by H. Broekhuis, N. Corver, R. Huybregts, U. Kleinhenz and J. Koster. Berlin: Mouton de Gruyter.

70. Roberts, I. and A. Holmberg. 2010. Introduction: parameters in minimalist theory. In *Null Subjects: the structure of parametric variation*, ed. by T. Biberauer, A. Holmberg, I. Roberts and M. Sheehan. Cambridge: Cambridge University Press.

71. Strik, N. 2006. L'acquisition des phrases interrogatives chez les enfants francophones. *Psychologie Française* 52: 27-39.

72. Stoyanova, M. 2008. *Unique Focus: Languages without multiple wh-questions*. Amsterdam: John Benjamins.

73. Stowell, T. 1981. Origins of Phrase Structure. M.I.T. doctoral dissertation.

74. Thornton, R. 1990. Adventures in Long-distance Moving: The Acquisition of Complex Wh-questions. University of Connecticut doctoral dissertation.

75. ----. 1995. Referentiality and wh-movement in child English: Juvenile D-Linkuency. *Language Acquisition* 4: 139-175.

76. Yang, C. 2002. *Knowledge and Learning in Natural Languages*. New York: Oxford University Press.