CROSS-LINGUISTIC INVESTIGATIONS OF SYNTACTIC CREATIVITY ERRORS IN CHILDREN'S WH-QUESTIONS

by

Carolyn Jane Lutken

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland

February 2021

Abstract

This dissertation investigates the relationship between competence and processing in children's first language acquisition, particularly of biclausal wh-questions. English-speaking children make consistent errors in production and comprehension of these questions. In production, these errors surface in the form of medial wh-phrases as in (1) when the child wishes to express (2) (Thornton, 1990). In comprehension, children respond to questions such as (3) as if the medial wh-phrase *what* were the question to answer (de Villiers and Roeper, 1995).

(1) **What** do you think **who** the cat chased?

(2) **Who** do you think the cat chased?

(3) Q: **How** did the boy say **what** he caught? Response: A fish!

These errors are particularly interesting because they resemble "Wh-Scope Marking" (WSM), which is attested in languages such as German (as seen in (4)), but not in English.

(4) Was hat Stefan Selina erzählt, **was** er stehlen wird?
What *did Steven tell Sherry* ***(what)*** *he would steal?*

Together, errors such as those in (1) and (3) suggest children may adopt multiple UG licensed grammars (Yang, 2002; Legendre, Vainikka, Hagstrom, & Todorova, 2002). This would be an example of *syntactic creativity*: the use of a UG-licensed grammar which is not the target grammar (Schulz, 2011). This dissertation investigates whether these errors are true examples of syntactic creativity. We begin with a thorough investigation of the cross-linguistic variation in WSM and examine the particular pragmatic contexts which license it. We then describe a series of experiments which examine children's ability to produce and comprehend biclausal wh-questions. Not only do we find no correlation between the two error types (which we would predict under a parameter-based view of the grammar), we find that these errors are correlated with children's working memory. Finally, we find that German-speaking children's performance on these tasks shows a striking resemblance to the English-speaking children. Our findings suggest that these errors are not the result of variation in the target grammar or grammar competence, but rather the result of something these children share: their immature processing mechanisms.

Readers: Géraldine Legendre (advisor), Barbara Landau, Steven Gross, Kyle Rawlins, Dana McDaniel (external reader)

Acknowledgments

The use of the royal "we" throughout this dissertation is neither a mistake nor reflective of hesitance to take full responsibility for this work, both its flaws and its contributions. It is an homage to the multitude of people who have helped me to produce it. While I proudly acknowledge it as the product of more than five years of my own work, it would not have been possible without the contributions of numerous people. These contributions took many forms and I will not be able to cover them all, but I will do my best to recognize a few essential people who have made this work, and the accomplishment it represents, possible for me.

In the first place, this work would not have been possible without the Doctoral Dissertation Research Improvement grant I received from the National Science Foundation. This funding allowed me to perform crucial aspects of my research at the University of Konstanz in Germany. I am grateful to live in a place where such support of the sciences is valued and available.

On a personal level, I must begin by thanking Géraldine for the countless meetings, discussions, emails, skype calls and every act of support. I have appreciated every one of them. Géraldine often says that academia is not for the faint of heart, as it comes with very few pats on the back. I am thankful for this opportunity to take this space to give her a pat on the back. As her student, I have learned more than I imagined possible. I have made more progress than I knew I could. I leave this department feeling capable of pursuing the career I wish and it is in large part because of her. Géraldine has encouraged me to apply for conferences and publications I would have thought "out of my reach" if left to my own devices. I certainly would not have (successfully) applied for an NSF grant without her encouragement. This was, perhaps, the most successful method of showing me my "reach" was greater than I thought, which is a lesson too few graduate students are taught. Géraldine has helped me gain experience teaching and supported my efforts in every way. I think the most crucial thing to say is that Géraldine has made it clear throughout my five (plus) years working with her that she is on my team. That is a quality that cannot be too highly praised.

I must also acknowledge the guidance of Akira Omaki whose presence is still palpable in this work. Without him, this work simply would not be what it is. No one could have predicted that my first years in academia would be his last or that I would be one of the last students he would advise, but I will be forever grateful to have learned from him as much as I did. While I am sure this project would have been different had he been here to supervise it, I hope that he would be proud of what it has become.

Next, I wish to gratefully acknowledge the members of my committee who took the time to thoughtfully engage with this work. It is truly appreciated. Special thanks to Dana McDaniel for serving on this committee from outside of the Johns Hopkins community. Her particular expertise is invaluable to the discussion of this work.

Within the department of Cognitive Science, I would like to acknowledge Kyle Rawlins, who took time out of his busy schedule to meet with me countless times over the past five years and who provided extremely helpful feedback on both of my qualifying papers. Special thanks to Barbara Landau, who welcomed me into her lab meetings in my second year and has given me valuable feedback on my experimental methodology and has helped me navigate my understanding of children's processing and child development. I am also grateful to Tal Linzen and Bonnie Nozari for joining the team on my grant and for numerous meetings discussing the best way to pursue the questions I wanted to answer. I also wish to thank Colin Wilson for numerous discussions about the statistical methods I employ and Paul Smolensky for his thoughtful comments on the Optimality Theory sections of this paper. The help and expertise provided to me from so many in the department has been invaluable.

I must also acknowledge the help of two recent graduates of the program: Emily Atkinson, for years of conversations and "R" tutorials and Eleanor Chodroff for a serious discussion about the puzzle of children's prosody.

Outside of the Johns Hopkins community, a particular acknowledgement must be made to Bettina Braun and the BabySprachLabor at the University of Konstanz, Germany. In spite of many barriers and difficulties, Bettina has been enthusiastic about my project from the first email I sent her asking if she would welcome a stranger into her lab space. She opened her lab to me, donated her time, recruited research assistants, and arranged for me to give a talk at the University of Konstanz which provided me with much helpful feedback. Her expertise and support have been invaluable to me. Again, this work would simply not have been without her.

I must thank the members of The Language Acquisition Lab, The BabySprachLabor, The Language Processing and Development Lab, and the Language and Cognition Labs for listening to multiple iterations of this work as it progressed. The feedback of so many from various backgrounds has given this work substance it would have otherwise lacked.

This work would not have been possible without the help of several research assistants who have played key roles in its development. At the Johns Hopkins Campus, these include, most prominently, Caroline West and Laura Nugent. Without Caroline's assistance investigating German corpora, I might never have gotten the question of how German-speaking children would perform on these tasks rolling around in my head. Laura Nugent's investigation into Hungarian corpora as well as her assistance recruiting child participants was simply invaluable. Both Caroline and Laura assisted in the creation of the tasks and the stimuli (including the creation of the 'alien language'!) used in Chapters 5 and 6. I will be forever grateful for their help and I hope they learned from me as much as I learned working with them. I must also thank Tegan White-Nesbit who joined the team as an official translator from the German department at JHU. Her expertise proved vital and her enthusiasm and interest were most welcome. In Konstanz, Germany Carina Haase and Naomi Reichmann adopted my project as their own and treated it with all the care and enthusiasm I could have asked for. They worked diligently and carefully. I cannot thank them enough.

I would never have been able to present such an in-depth, cross-linguistic study without the help of numerous native-speakers who chose to help me purely because they are my friends. Thus, my heart-felt gratitude goes out to Anna Luescher, Edit Takacs, Beatrice Gregg, Ayushi Pandey, Rashi Pant, and Sadhwi Srinivas. Ladies, without your help translating and recording stimuli, without your willingness to sit with me and discuss “what that question really *means*”, I do not know where I would be. Anna deserves a second thank you for helping me navigate the streets of Konstanz upon my arrival in the fall of 2019.

Over the years I have reached out to numerous experts in my fields of interest and must thank the following people for meetings, emails, and discussions at conferences. These include wonderful courses and conversations with David Pesetsky, Coppe van Urk, Brian Dillon, and Norvin Richards at the LSA Summer Institute in 2017. I am grateful to Benjamin Breuning, Veneeta Dayal, and Matt Goldrick for some very interesting one-on-one meetings and to Dana McDaniel and Jeff Lidz for engaging conversations at BUCLD over multiple years. I am grateful to Joel Wallenberg who has always been ready to discuss my work or the process of getting a PhD. I also have to thank him for convincing me that I ‘had what it takes’ to pursue a PhD. Finally, I thank Julia Horvath, Gisbert Fanselow, and Martin Hakl for very engaging feedback over email. All of these conversations helped challenge my thoughts and expand my thinking.

I must offer my sincere thanks to all of my fellow graduate students for every time they listened to a practice talk, helped me understand some new problem I was grappling with, or invited me to have a cup of tea. They were all appreciated. Special thanks go to the members of my cohort who braved that first year together with me. In particular, thank you to Celia, who encouraged me from the moment we met at our interview weekend, shared an office with me for 5 years, and who recorded the voice of the (now internationally famous) Hillary Hippo.

I also must acknowledge the work and effort of over 200 kindergarteners and their parents who came to ‘play science with me’. I appreciate the work and energy you brought to my research. I also appreciate you reminding me that this work is fun as well as fascinating.

To my friends around the world who have maintained a friendship with me even when I have been swamped with work and who reminded me of life outside, I thank you for keeping me grounded.

I think it cannot be overstated how important a strong support group is. I have always been incredibly grateful for my family, but their support during these past five years has been invaluable. They listened with enthusiasm to every ‘fun fact’ I shared (even when it was nonsense to them). They told me I belonged in academia when imposter syndrome made me doubt it. They kept me sane and reminded me of what is really important.

Finally, thank you to Jeff for making everything better.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The relationship between linguistic competence and performance is at the core of psycholinguistic research. Research on adults' production and comprehension errors has shed light on what type of errors are driven by constraints on the language processor, and whether adults' grammatical knowledge may in part be shaped to facilitate language production or comprehension (e.g. Bock & Miller, 1991; Dickey, 1996; Ferreira & Swets, 2005; Kroch, 1981). However, the relation between grammatical knowledge and language production mechanisms in children remains understudied (cf. McDaniel, McKee, & Garrett, 2010; McDaniel, McKee, Cowart, and Garrett, 2015)). The general question of how linguistic knowledge and cognitive processes interact is, therefore, largely unanswered and the more specific question of what leads to consistent errors in production and comprehension of child language, in need of further investigation.

This question is of particular interest for language acquisition researchers because children's production errors are often treated as indicative of their immature grammatical knowledge, or evidence for rule-guided learning, such as overregularization (e.g., *I goed* vs. *went to the zoo*), or parametric learning, as suggested by the Principles and Parameters theory (Chomsky, 1981; for discussion, see Snyder, 2007; Yang, 2004). For example, Hyams (1986) proposed that English speaking children produce null subject sentences (e.g., *Want more apples*) by setting the wrong Null Subject parameter value (e.g., Spanish setting). However, those errors may instead reflect constraints on children's immature sentence production mechanisms (e.g. Bloom, 1990; Valian, 1990), or how the immature production mechanisms interact with their developing grammar as the sentence planning and production processes unfold during language use. In sum, detailed investigations of children's sentence production mechanisms are needed a) to start shedding light on the understudied question of how the sentence production mechanisms develop during childhood, and b) to enhance our ability as acquisition researchers to identify the source of children's syntactic production errors in general.

Appropriate understanding of the production error data, and how those errors are mechanistically produced by children's grammar and sentence production mechanisms, has broader consequences for theory development in language acquisition research. As a first step towards this better understanding of grammar and sentence production in children, this dissertation focuses on errors in children's use of biclausal wh-questions (e.g., *Who do you believe we are going to see in the cave*?). Previous research and our own preliminary work (Lutken & Omaki, 2017; Lutken, Legendre, & Omaki, 2020) have identified curious production as well as comprehension data, which have suggested that children between 3;11 and 6;1 are creatively using their knowledge of Universal Grammar to create complex wh-questions in English.

Specifically, Thornton (1990) found that 4 to 6-year-old English-speaking children produce 'medial wh-questions' like (1) when their intended utterance is a long-distance (LD) wh-question spanning two clauses such as (2).

(1) **What** do you think **who** the cat chased?

(2) **Who** do you think the cat chased? (Thornton, 1990)

In a comprehension task with questions about stories, de Villiers & Roeper (1995) found that English-speaking children answer questions which contain multiple wh-phrases with an answer to the embedded wh-phrase which functions as a relativizer rather than as a question word. For example, in (3) the child responds to a *what* question rather than a *how* question.

(3) Q: **How** did the boy say **what** he caught? Response: A fish!
(de Villiers and Roeper, 1995)

These questions containing medial wh-phrases resemble question formation in languages which employ so-called wh-scope marking (WSM), such as many dialects of German, as seen in (4) (McDaniel, 1989):

(4)	**Was** glaubt	Hans,	**mit wem**$_i$	Jacob jetzt t_i spricht?
	what think	Hans,	with whom	Jacob now talks
	Who does Hans think Jacob is now talking to?			

(ex 7b McDaniel, 1989)

While there are a variety of syntactic analyses for (4), which will be discussed in length in Chapters 2 and 3, the adult answer to the question is a response to the medial wh-phrase (Lutz, Müller, & von Stechow, 2000). Such productions and responses by English-speaking children, therefore, could be taken as examples of WSM use in English and have been suggested as evidence that children have access to UG consisting of universal principles and language-particular parameters and may initially select the wrong syntactic parameter (which we for simplicity's sake can call [+WSM]) (de Villiers & Roeper, 1995). This is an example of what we refer to as a *syntactic creativity error* (Schulz, 2011).

In spite of the fact that syntactic creativity errors suggest the use of grammars from more than one language, there has been relatively little cross-linguistic investigation of this phenomenon. With the notable exception of Jakubowicz & Strik (2008), we are unaware of any effort to compare performance by children with two different target structures such as WSM (German) vs. long distance (LD) wh-movement (English). The comparison between two such groups would be instrumental in determining whether such an error is a true example of syntactic creativity, or if it is the result of some external factor.

In order to more fully investigate the roots of this particular example of syntactic creativity, this literature review will begin with a brief review of previous research on the WSM structure itself (Section 1.2) and several aspects of the semantic and pragmatic context which leads to the use of WSM in languages in which it is viable. This will be followed by background relevant to our two alternative hypotheses: previous research on Syntactic Creativity Errors (Section 1.3) and models of comprehension and production processing mechanisms (Section 1.4). Section 1.5 will present some background and motivation for conducting this study with German-speaking children.

1.2 Syntactic and Semantic theories of Wh-Scope Marking

In order to more fully test whether English-speaking children's erroneous productions include the use of a WSM grammar, a more thorough understanding of WSM is necessary. In particular, a review of the existing literature on the syntax and semantics of WSM will reveal that a better understanding of its pragmatics is necessary to design better stimuli, which in turn are more likely to provide stronger evidence bearing on the two main hypotheses considered in this dissertation- the Parametric Acquisition Hypothesis and the Limited Processing Hypothesis, which we discuss in detail below.

WSM structures are typified by the use of multiple wh-phrases: one at the top of the matrix clause and one appearing at each clausal boundary following, as seen in (5).

(5) German

> **Was** glaubst du, **mit wem** Maria gesprachen hat?
> **What** think you, **with whom** Maria spoken has?
> *With whom do you think Maria spoke?*

Ex. From Dayal (1994, 1b)

Previous research on WSM falls into two primary categories. First, the Direct Dependency Approach (DDA) is primarily a syntactic analysis. McDaniel (1989), for example, considers the scope marker (henceforth, SMer), as an expletive element base-generated in SpecCP of the matrix clause. This expletive is then replaced at LF by the true wh-phrase, whose scope it marks (hence the name, "scope marker"). Fanselow and Mahajan (2000) also consider the SMer an expletive, but one which has an associate dependency with the embedded CP. Under this analysis it is the entire CP which replaces the SMer at LF.

However, because DDA analyses assume that the SMer is an expletive and therefore devoid of semantic content, this predicts a semantic analysis which would be equivalent to a long distance (LD) construction, i.e., a single wh-question. This, however, is not true of all languages which employ WSM. Dayal (1994) argues that the semantic analysis of WSM constructions must be distinct from that of LD constructions. She argues that the first wh-phrase in WSM constructions is a contentful question over propositions (Dayal, 1994; 2000). The 'second' wh-phrase limits the set of possible answers to the 'first' question. In an example such as (5) *with whom Maria has spoken* limits the set of possible answers to *what do you think?*

For reasons which we will expand upon in Chapter 2, neither the IDA nor the DDA successfully accounts for all languages which have been claimed to make use of WSM. Furthermore, the exact pragmatic situations which render WSM optimal have not been examined. In dialects of German which allow both LD and WSM constructions, what types of scenarios make one construction better than the other? In Chapter 2, we will describe a survey which establishes what type of syntactic structure speakers of German (as well as Hindi, Hungarian, and English) prefer in certain pragmatic contexts. In Chapter 3, we will use Optimality Theory (Prince & Smolensky [1993]2004; Legendre, Grimshaw, and Vikner, 2001) to account for the cross-linguistic variation in WSM use as well as the variation of optimal structures within languages. The contribution of the corresponding chapters of this dissertation will be two-fold. In the first place, we will provide a theoretically grounded, cross-linguistically valid account of the target grammar for children acquiring English and German respectively, that is informed by a broader sample of languages with WSM, including German, Hungarian, and Hindi. In the second, we will establish the semantic and pragmatic criteria necessary for rendering LD and WSM felicitous and will use these insights in the design of later experiments. This will be crucial in eliciting these structures from children.

Summing up, there remain a number of theoretical questions to investigate, which include, but are not limited to, the following:

- What cross-linguistically valid analysis of WSM should we adopt that sheds light on the mapping between semantics/pragmatics and syntax in complex questions?
- What pragmatic and semantic features are relevant to stimuli used to demonstrate that children produce/comprehend particular types of complex questions in both languages with and without WSM?

1.3 Previous Research on Syntactic Creativity

Syntactic creativity has been attested in many different forms in child English productions. Examples include children's production of possessor extraction such as (6)

(Gavruseva & Thornton, 2001), subject extraction out of a *that* complement, commonly referred to as a *that* trace violation in the syntax literature (7) (Thornton, 1990), resumptive pronouns (8) (McKee & McDaniel, 2001; for related observations in French, see Labelle, 1990), and subject-auxiliary inversion in embedded interrogatives (9) (Ambridge, Rowland, Theakston, & Tomasello, 2006; Pozzan & Valian, 2017).

(6) Whose do you think ___ ball went in the cage?
(cf. Whose ball do you think went in the cage?)
(7) Who do you think that won the race?
(cf. Who do you think ___ won the race?)
(8) This is the woman that Grover talked to her.
(cf. This is the woman that Grover talked to __.)
(9) Katie wanted to know what was her brother building.
(cf. Katie wanted to know what her brother was building.)

These structures are ungrammatical or marginal at best in (adult) English, but they are grammatical in other languages. Possessor extraction constructions are attested in Hungarian and Chamorro (Chung, 1991; Szabolcsi, 1983). Subject extraction out of an embedded complement is grammatical in Arabic, Basque, and many Romance languages (Kandybowicz, 2006; Rizzi, 1982). Resumptive pronouns can be used in sentences like (8) in Irish, Hebrew, and Palestinian Arabic (Borer, 1984; McCloskey, 1990; Sells, 1984; Shlonsky, 1992). Finally, subject-auxiliary inversion in embedded interrogatives is grammatical in Spanish (Torrego, 1984) and Italian (Rizzi, 1996).

Syntactic creativity is not limited to sentence production. For example, it has been observed that English-speaking children interpret subject-less sentences (e.g., *Play with blocks*) as null subject sentences (meaning *They play with blocks*) rather than imperatives (Orfitelli & Hyams, 2012). For sentences that contain disjunction under negation (e.g., *John didn't drink coffee or tea*), Japanese and Turkish children interpret them to mean neither coffee nor tea, like English speakers would, even though that interpretation is not available for adult Japanese or Turkish speakers (Geçkin, Thornton, & Crain, 2017; Goro and Akiba, 2004). Russian-speaking children also behave like English speakers as they allow backward coreference between a pronoun in a fronted adjunct clause and the matrix clause subject (e.g., *While he$_1$ was reading, John$_1$ ate an apple*), even though this coreference is not grammatical in the (adult) Russian counterpart (Kazanina & Phillips, 2001).

These examples of syntactic creativity have typically been taken as evidence for a Parameter-based theory of language acquisition. A core assumption in the Principles and Parameters theory (Chomsky, 1981) is that the innate language faculty provides learners with a list of all parameters that define potential syntactic variation, and learners only need to select the right parameter setting based on the language input. This suggests that innate linguistic knowledge provides top-down guidance that constrains the hypothesis space, such that children are informed of what type of grammatical variation they should consider and evaluate against input data (For further discussions, see Baker 2005; Fodor, 1998; Newmeyer, 2005; Snyder, 2007; Yang, 2004). In other words, a child might “mis-set” a parameter or leave a parameter unset and show evidence of this through these examples of syntactic creativity. In this sense, syntactic creativity errors may provide important insights on the language learning mechanism. We will refer to this perspective as the *Parametric Acquisition Hypothesis*.

However, the Parametric Acquisition Hypothesis predicts that children should make these errors in both production and comprehension[1] and despite the wide-ranging evidence for syntactic creativity, few of these phenomena are found in both production and comprehension data. Of the cases reviewed above, the null subject phenomenon is the only one that has been attested in production as well as comprehension. This may simply be due to practical constraints on methodologies. In production studies, it may not be methodologically feasible to elicit certain structures in experimental or naturalistic contexts (e.g., backward anaphora). In comprehension studies, children (as well as adults) may be able to coerce a plausible interpretation for ungrammatical sentences (e.g. *that* trace violations), even though their target grammar does not allow such structures.

The presence of cross-modal evidence is one of the reasons children's use of medial wh-questions has received much attention in the literature. In sentence elicitation experiments Thornton (1990) observed that English-speaking children between 3;11 and 5;9 produce non-adult-like utterances such as (10), which, descriptively speaking, contain an extra wh-phrase in the middle of the sentence. In comprehension, de Villiers and Roeper (1995) conducted a longitudinal study using a question-after-story task with a question like (11), and found that English-speaking 3- to 4-year-old children answered the medial wh-phrase/relativizer *what*, instead of the sentence initial wh-phrase *how* (see also de Villiers, Roeper and Vainikka, 1990; Thornton, 1995; Thornton & Crain, 1994). McDaniel, Chiu and Maxfield (1995) used an acceptability judgment task with 3 to 5-year-old children, and found that a subset of the children accepted questions like (12). Finally, although the

[1] By comprehension we mean interpretation. This is in contrast to grammaticality judgments, which would require children to make a meta-linguistic judgment of an utterance.

present work will focus on English, production of medial wh-question structures like (10) has also been reported in French (Demirdache & Oiry, 2007; Jakubowicz & Strik, 2008), where such questions are also ungrammatical.

(10) What do you think what is under this box?
(cf. What do you think is under this box?)

(11) How did the boy say what he caught?
Response: A fish!

(12) What do you think who is gonna climb up the steps?
(cf. Who do you think is gonna climb up the steps?)

These behaviors were interpreted as evidence that English-speaking children adopt a WSM grammar[2] at least part of time. English-speaking children's production or comprehension behaviors may in fact indicate that WSM questions are allowed in their (non-target-like) grammatical knowledge, in line with the Parametric Acquisition Hypothesis. However, the previous studies that argued for this perspective have offered little explanation for why English-speaking preschoolers entertain this non-target-like parameter setting. One plausible explanation comes from competition-based models (Yang, 2002; see also Legendre, Vainikka, Hagstrom, and Todorova, 2002). For example, under a probabilistic model of parameter-based learning (Yang, 2002), different parameter settings are simultaneously evaluated and the competing parameter settings will be probabilistically used to parse the input. When this parse succeeds, that parameter is rewarded by a boost in probability; if the parse fails, the parameter is punished and its probability decreases. As a result, this process will continue over time and across many utterances, until the parameter setting that is most compatible with the input eventually

[2] Throughout the remainder of this dissertation, we will use *medial wh* as a theoretically neutral, descriptive term indicating the use of a wh-phrase at the clausal boundary in production. While WSM always includes medial wh-phrases, the use of a medial wh-phrase does not entail WSM.

wins the grammar competition (for critical discussions of this approach, see Snyder, 2007). This could be because children rarely hear biclausal questions (discussed in Chapter 4) and thus do not receive adequate evidence either for or against WSM or LD until they are older: they have yet to accurately establish a parameter setting. Under this account, English-speaking children are predicted to use WSM structures in both production and comprehension, and presumably at a roughly comparable rate, until the English LD movement parameter setting is selected as the winner.

In summary, there are multiple examples of Syntactic Creativity which appear cross-linguistically. The example of medial-wh productions and WSM-like interpretations appear to constitute an excellent example of Syntactic Creativity because the error appears in both production and comprehension. However, there are outstanding questions which this dissertation aims to answer which will help clarify whether these errors are a true example of Syntactic Creativity and thus evidence for a Parametric Acquisition Hypothesis, or something else. These questions include:

- Are medial wh-productions and WSM-like interpretations correlated? If there is no group relationship, are there exceptional children who show evidence of WSM in both production and comprehension?
- Will improving the experimental methodology of previous work (detailed in Chapters 4 and 5) reduce or eliminate these errors?

1.4 Processing Models of Production and Comprehension

The systematic presence of a syntactic creativity error in both production and comprehension, as discussed above, would indeed be strong evidence in favor of a Parametric Acquisition Hypothesis. However, there is also the alternative possibility that children's errors are not the result of using alternative grammar rules, but rather the result of immature processing mechanisms. This we term the *Limited Processing Hypothesis.*

Recent developmental psycholinguistics research has uncovered how children can sometimes fail to comprehend or produce sentences despite their sophisticated linguistic knowledge (for reviews, see Omaki & Lidz, 2015; Snedeker, 2013). These errors may reflect either immature procedural mechanisms for sentence comprehension or production, or immaturity in cognitive resources (e.g., working memory, cognitive control) that are critical for supporting real-time language use. For example, Bloom (1990) and Valian (1991) have suggested that production of null subject sentences in English-speaking children may reflect children's attempts to reduce sentence length due to their immature sentence planning abilities (for discussions, see Frazier, 2015; Hyams & Wexler, 1993). McDaniel et al. (2010) elicited production of relative clauses by adults and 3 to 8-year-old children and assessed the distribution of disfluency indicators such as pauses and restarts. When adults produced object relative clauses with long-distance dependencies (e.g., *Pick up the duck that Big Bird thinks the princess was kissing__,* where __ marks the gap in the relative clause corresponding to the antecedent 'the duck'), the observed signs of disfluency (mostly in the form of pauses at the beginning of the utterance) suggested that adults were able to plan a large structural unit such as an upcoming clause while holding the filler-gap dependency active in working memory (for related findings, see Ferreira & Swets, 2005). On the other hand, children demonstrated signs of disfluency in a variety of positions throughout the sentence, suggesting that children plan only small pieces of their utterance at a time. Furthermore, McDaniel et al. (2015) found that children produce more *that*-trace violations (e.g. *What do you think that's under the swimming pool?*) than adults and further suggesting smaller, more limited planning units in children. Both studies suggest that children plan more locally, and more frequently during sentence production than adults do. This more localized planning can lead to production errors which are unrelated to grammar competence.

In spite of these recent developments, the relation between grammatical knowledge and language production mechanisms in children remains understudied. Here we will provide a brief description of the models of processing involved in production and comprehension. We will then discuss the use of working memory (WM) as a proxy for a measure of general processing as well as several studies which used WM as a way to investigate the roll of processing in language production and/or comprehension.

1.4.1 Models of Comprehension and Production Processing

We assume a model of comprehension based on that described by Phillips and Erehnhofer (2015), which states that comprehension of a sentence proceeds incrementally. Each word is in turn processed and integrated into the construction as a whole. Each word has its own requirements, e.g. pertaining to thematic roles, scopal relations, agreement relations, etc. The requirements must be met by other words in the construction. This can be done either by retrieving a word that has already been processed from memory or by holding the word in memory until its requirements have been met (as is the case with long-distance dependencies, including subject-verb agreement, wh-questions, relative clauses, etc.), all while continuing to process the words that follow. Cunnings and Felser (2013) showed that adults with lower WM abilities were more likely to mis-parse constructions with multiple potential antecedents for pronouns. They suggested that this provides evidence that there can be interference from superficially similar items during the comprehension process. In other words, in addition to holding an item in memory, the comprehender must keep other, similar items from interfering with it. This requires resource management and provides plenty of ways for the processing to fail. Indeed, adults often mis-parse sentences and must reanalyze, making many errors when they do, e.g. adults frequently misinterpret syntactically ambiguous noun phrases even after they reanalyze the construction. (Ferreira & Henderson, 1991).

The model we assume for production involves three levels (Dell, 1986; Bock, 1987; Levelt, 1989): first, a semantic, "meaning" level where the speaker establishes the message they wish to communicate and the relevant semantic relationships involved; second, a syntactic, 'structural' level where multiple syntactic constructions are generated in parallel (an idea we will return to) and the speaker must choose one of these constructions for best expressing their intended message; and finally a third, morphophonological level where lexical items are chosen and the appropriate phonemes are mapped onto the chosen structure. The separation of these levels is based on the fact that errors of production often occur 'within' a level, saying another phonologically similar word for instance. However, there is also clear evidence for influence across levels. Phonological errors, for instance, do not typically result in non-words, which we might predict if there were no cross-over whatsoever between levels (Dell & Reich, 1981).

Because this dissertation focusses on syntactic errors, we will discuss the syntactic level in a bit more detail. Evidence for the syntactic level and for parallel processing of multiple constructions comes primarily from work on syntactic blends (Garrett, 1980). These blends are mixtures of multiple constructions, which have been argued to reflect the fact that the speaker must have processed these alternative constructions simultaneously and failed to choose between them, perhaps because of processing load, as suggested by

Coppock (2010). These alternative formulations have been "intertwined in speech" to produce an ungrammatical combination of grammatical constructions (Garrett, 1980). Take for instance (13) from Fay (1982). (13) is ungrammatical, but was produced in natural speech by an adult, native-speaker of English. Fay (1982) analyzes it as a combination of (13′) and (13″) which are both grammatical constructions and viable options for communicating the fact that the speaker is frustrated by the fact that there seems no good reason for their presence and they could be spending their time otherwise:

(13) *What the hell am I doing here for?
(13′) What the hell am I doing here?
(13″) What the hell am I here for?

Utterances such as (13) seem to contain a substring from one target and a final substring from another, suggesting the two were produced simultaneously.

Jaeger (2005) thoroughly investigates blends produced by children ranging in age from 1;11 to 5;0. She discusses various types of blends including those involving blended words, phrases, and entire constructions. One particularly fascinating pattern she discovered is that children produced the most blends around age 4;0 and the majority of those were complex blends involving multiple phrases or constructions. This age is likely to coincide with the emerging mastery of more complex constructions which task processing ability to the max. Jaeger (2005) suggests that these errors are indicative of children's attention in planning. It is reasonable to presume the rate of such errors might be tied to their processing abilities: more should occur in children with lower or less mature processing abilities and as children become more adult-like, they should produce fewer blends.

These models of comprehension and production predict that processing load is relevant to both, though different aspects of processing might be relevant to each. However, according to Baddeley (2017)'s model of the executive control system, all action is mediated by a WM buffer between general processing and action-taking. In other words, while different processes may be relevant to production and comprehension to varying degrees, a measure of WM will serve as an indicator for general processing. Indeed, previous work has done just that: Johnson, McMahon, Robinson, Harvey, Hahn, Leonard, Luck, & Gold (2010) used WM as a proxy for general processing in adults. Various studies of language acquisition have shown a relationship between children's production and comprehension and their WM. Adams and Gathercole (2000), for instance, measured children's mean length of utterance, the total number of free morphemes they used, and the syntactic complexity of their utterances (essentially, they asked if they used full sentences, questions, and negations). They found these language abilities were indeed correlated with WM ability, including on WM tasks which were not language-specific such as the Corsi block tapping task, which requires participants to touch a series of blocks in the same order as the experimenter, thus testing their memory without tapping into their language abilities (Corsi, 1972). Like Johnson et al. (2010) for adults, Adams and Gathercole (2000) conclude that WM performance can be used as a diagnostic of general executive function in children. Willis and Gathercole (2001) report slightly different findings. They divided children into groups based on their phonological WM abilities and found children with higher WM were more successful at repeating sentences. However, they did not find a difference in performance on the comprehension task. This could be related to the type of WM-measure they gave, which they do not report in detail. Regardless, their findings suggest WM is related to children's ability to repeat complex constructions. Finally, Delage and Frauenfelder (2019) conducted a study assessing WM a variety of ways including both forward and backward digit spans. They measured children's comprehension by means of a forced-choice task: the experimenter would read a description of an image and the child was asked to choose the image it matched. The descriptions varied in their complexity, the most complex of which included relative clauses with and without wh-relativizers[3]. In production they analyzed spontaneous speech for the rate of embedded clauses, the rate of multiple embeddings, and the rate of subordinate clauses. Their conclusions suggest that WM is relevant to both comprehension and production. In fact, they suggest that forward digit span is relevant for comprehension and backward digit span is relevant to production. We have seen, therefore that WM has been discussed as relevant to production and comprehension mechanisms theoretically and this relationship has been empirically shown.

In summary, this brief literature review has discussed models of processing in production and comprehension. This includes the fact that some errors in both production (syntactic blends) and comprehension (e.g., mis-analyzing anaphora) in adults have been explained as the result of an 'overload' on these processing systems. We have reviewed both theoretical and empirical motivation for using WM as a proxy measure for general processing ability and have seen, indeed, that others have shown a relationship between comprehension, production and WM. Questions that this dissertation will strive to answer include:

[3] It should be noted that a wh-relativizer acts as a free relative as in *Nobody knows what really happened*. This is in contrast to embedded questions such as *I wonder what really happened*, which occur when the matrix verb is a verb such as *ask* or *wonder*. The constructions we discuss throughout this dissertation are free relatives rather than embedded questions.

- Are children's performance on production and comprehension of WSM correlated with their WM?
- Do children's errors resemble the errors adults make when their processing systems are overloaded? In other words, in production could these errors be syntactic blends? In comprehension, could their non-adultlike responses be the result of interference from other lexical items?

1.5 Cross Linguistic Study

A final prediction of the Limited Processing Hypothesis which we will test is that if these errors resembling WSM in child English are the result of immature processing, then evidence of processing limitations should also surface in the production and comprehension of children of comparable age, whose target grammar includes WSM. We will, therefore, conclude this dissertation with a version of both our linguistic and WM experiments conducted with German-speaking children in Konstanz, Germany.

Doing this study with German-speaking children, in particular, will be useful for several reasons. In the first place, of course, we want to establish how children perform on these tasks when WSM is in the adult grammar of the language they are acquiring. Second, there has been relatively little investigation of complex question formation in German-speaking children and none that we are aware of on this particular construction, so we will be adding to that literature.

As we will see, our English data will document that children use various work-around strategies to get around the complexities of bi-clausal questions, but specific properties of English to be discussed in Chapters 4 and 5 limit our ability to interpret what these strategies are (e.g. are they using sequential questions or WSM?). By virtue of its special word order properties related to the Verb-second phenomenon (Chapter 6), German will allow us to be far more conclusive about such strategies. In particular, we will be able to identify a favorite strategy of using sequential questions or two consecutive questions, as in *What does he think? Who should we ask?*. That cross-linguistic evidence will then be used to inform our interpretation of a similar move in English.

We address the following questions about German-speaking children in the course of this study:

- How do German-speaking children perform on these tasks requiring production and comprehension of biclausal questions?
- What types of errors will German-speaking children make? Will they use alternative constructions (such as sequential questions) and interpretations?
- Crucially: Are German-speaking children's performance on production and comprehension tasks correlated with their processing abilities (measured via WM)?

1.6 Summary

We have seen two alternative possible explanations for medial wh-production and WSM-like comprehension: The Parametric Acquisition Hypothesis and the Limited Processing Hypothesis. While we have presented them in contrast to each other, neither is necessarily exclusive in the ability to explain syntactic creativity errors. It is certainly plausible that some examples of syntactic creativity are the result of parametric learning while others are the result of under developed processing mechanisms. However, it is important to understand the role each plays in language acquisition. With regard to the error in question, the Parametric Acquisition Hypothesis predicts that these errors should appear in both production and comprehension and furthermore, within a participant there should be a correlation between the rate of errors in production and comprehension[4]. The

[4] The expectation of a correlation does depend on the particular model of "parametric" acquisition we assume. A traditional Principles and Parameters model (Chomsky, 1981) or even a model such as is suggested by Yang (2002) would predict such a correlation, but an Optimality Theory model of acquisition does not because production and comprehension are fundamentally different types of optimizations. In production, the candidate set is made of alternative expressions sharing one and the same target interpretation. In comprehension, the candidate set is made of alternative interpretations sharing one and the same expression

Limited Processing Hypothesis does not necessarily predict this correlation because different processing factors are involved in comprehension and production, but instead predicts that WM should predict performance in either. This highlights the significance of cross-modal evidence for syntactic creativity errors in understanding the actual source of the errors. We must establish whether these errors occur in *both* production and comprehension consistently as well as whether WM predicts these errors.

This literature review has discussed previous work on acquisition of both the production and comprehension of questions with multiple clauses as well as the theoretical debates surrounding the WSM phenomenon itself. We have highlighted key remaining questions as well as ways in which we hope to build upon previous experimental work. We have also briefly discussed the two major analyses of the (adult) WSM grammar: the DDA and the IDA, as well as the questions that remain unanswered by that literature. Finally, we have discussed models of processing in production and comprehension and previous work that has used WM as a proxy measure for general processing.

To summarize, the main specific questions which remain and which this dissertation will strive to address include the following:

- What cross-linguistically valid analysis of (adult) WSM should we adopt that sheds light on the mapping between semantics/pragmatics and syntax in complex questions?
- What relevant pragmatic and semantic features are relevant to children learning languages with and without WSM?
- Are children's errors in production and comprehension tasks the result of incorrect parametric choices or are they due to grammar-external factors such as experimental artifact or processing difficulties?
- Does children's performance change when we make methodological improvements to the tasks? And when we control for relevant pragmatic and semantic features?
- Are medial wh-productions and WSM-like interpretations correlated? If there is no group relationship, are there exceptional children who show evidence of WSM in both production and comprehension?

(Smolensky, 1996; Legendre, Hagstrom, Chen-Main, Tao, and Smolensky, 2004). We will briefly discuss the possibility of an Optimality Theory based explanation for some of these errors, though not all.

- Are children's performance on production and comprehension tasks correlated with their WM?
- Do children's errors resemble the errors adults make when their processing systems are overloaded? In other words, in production could these errors be syntactic blends? In comprehension, could their non-adultlike responses be the result of interference from other lexical items?
- What types of errors, if any, do children make when their target grammar includes WSM (German)?
- Are the errors made by German-speaking and English-speaking children similar?

In order to start addressing these questions, we first establish the cross-linguistic variation in WSM-use via a survey study (Chapter 2). We then discuss a cross-linguistic analysis of the target grammars using Optimality Theory (Prince & Smolensky [1993]2004; Legendre et al. 2001), which offers a unified characterization of WSM in Chapter 3. We move on to discuss an initial set of three experiments on English-speaking children (Chapter 4), which investigated whether improving the experimental stimuli would either eliminate or reduce the number of WSM-like errors seen (a resulting journal article recently appeared in *Cognitive Science*). This chapter confirms that children will still make these errors when these aspects of the methodology have been addressed, but it also shows that there is no correlation between children's performance on the comprehension and production tasks. In other words, the child who produces questions with medial wh-phrases is not necessarily going to show evidence of WSM interpretations in their comprehension. This would be difficult to explain under a Parametric Acquisition Hypothesis, given that we expect children to make use of the same grammar in production and comprehension. In Chapter 5, we directly test the Limited Processing Hypothesis by investigating whether children's performance on production and comprehension of biclausal questions is tied to their WM abilities. This section also includes a novel experiment designed to elicit questions with extraction from an embedded clause without the use of a "lead-in," where the experimenter gives the child the matrix clause (e.g. "The questions starts '*What do you believe…*') as has been used in previous work, including Chapter 4. Chapter 6 presents the results of a study run with German-speaking adults and children at the University of Konstanz. We first establish that adults performed as expected (using WSM constructions in the production task and responding to them appropriately in the comprehension task) and then we examine the errors made by German-speaking children. Furthermore, as in Chapter 5, we examine whether the children's WM abilities predict their performance on either task. Each of these chapters is designed to stand alone and thus each includes a more detailed introduction to the major issues relevant to that chapter.

By investigating the cause of this particular example of Syntactic Creativity, this dissertation delves deeper into our understanding of the relationship between processing and acquisition. We hope to add to our understanding of the complex process of acquisition and the delicate balance between processing and competence in child language.

Chapter 2: Wh-Scope Marking: A question of Semantics

Wh-scope marking (WSM) has been described in a variety of languages including German (McDaniel, 1989), Hindi (Dayal, 1994), and Hungarian (Horvath, 1995; 1997; 2000) among others. [5]WSM is characterized by the use of a (typically) distinct interrogative pronoun in each clause of the long-distance extraction of a single wh-phrase[6], as seen in (1)-(3). As is clear from these examples, WSM can occur in languages that front wh-phrases (1) to the left periphery of the clause they are an element of, as well as those which require the wh-phrase remain in situ (2).

(1) German
Was glaubst du, **mit wem** Maria gesprochen hat?
What think you, **with whom** Maria spoken has?
With whom do you think Maria spoke?
Ex. From Dayal (1994, 1b)

(2) Hindi
Jaun **kyaa** soctaa hai Meri **kis-se** baat karegii?
John **what** thinks Mary **who-with** will talk
Who does John think Mary will talk to? Ex. From Dayal (1994, 18)

(3) Hungarian
Mit mondtál hogy **mire** számítanak a gyerekek?
What$_{acc}$ said-2sg-indef.$_{DO}$ that **what**$_{al}$ count-3pl the kids$_{nom}$?
What$_{acc}$ did you way what$_{al}$ the kids expected?

Ex. From Horvath (2000, 13c)

---

[5] WSM has been described in Russian and Polish and potentially other Slavic languages (Stepanov & Stateva, 2005) as well as Passamaquoddy, an Algonquin language spoken in the North Eastern United States and the East coast of Canada (Bruening, 2004). It should be noted that it was originally claimed that Passamaquoddy had two forms of WSM; one which corresponded to the Indirect Dependency Account and one that corresponded to the Direct Dependency Account (Bruening, 2004). However, Bruening has since revised his conclusion to claim that these two types of questions in Passamaquody, in fact, correspond to WSM and Wh-copying (Bruening, 2006, as well as personal communication).

[6] This is distinct from wh-copying which is typified by the use of the same wh-phrase appearing in the sentence-initial position as well as in each embedded clause. We will not discuss copy constructions further as there is considerable evidence that they are distinct from WSM (Murphy, 2016). In our studies with children, we will provisionally subsume wh-copying under WSM-like constructions. Note that in contrast to previous work we found that most of children's productions were WSM-like rather than copy constructions.

Early analyses of German WSM suggest that the first wh-phrase, *was,* acts as an expletive (hence contentless) element whose sole function is to mark the scope of the true wh-phrase, *mit wem,* surfacing in the intermediate specCP position (McDaniel, 1989; Riemsdijk, 1983). Simplifying a bit as to the origin of the expletive (which is taken up further below), the syntactic representation is as follows, omitting traces of subject and verb movement, for simplicity:

(4) [$_{CP}$**Was**$_{\mathbf{expl}}$ glaubst [$_{TP}$ Du [$_{VP}$ **[**$_{CP}$ **mit wem**$_{\mathbf{i}}$ [$_{TP}$ Maria t_i gesprochen hat]]]**]]**?

Since then, however, there have been myriad analyses investigating the exact nature of WSM, particularly of the first wh-phrase, the so-called "scope marker". These analyses fall into two major categories: The Direct Dependency Approach (DDA) and the Indirect Dependency Approach (IDA). Analyses which fall under the DDA follow McDaniel (1989) and Riemsdijk (1983) and typically analyze the scope marker as an expletive. While specific analyses differ on the exact nature of the scope marker, they agree that it is not a contentful wh-phrase. This is a key difference between analyses of the DDA and IDA. Under the IDA, the first wh-phrase is a contentful question over propositions (Dayal, 1994; 2000). The 'second' question limits the set of possible answers to the 'first' question. In an example such as (1), *with whom Maria has spoken* limits the set of possible answers to *What think you* to only include thoughts about who Maria's interlocutor was. Under the IDA the bracketing of the example displayed in (4) is syntactically the same but the initial wh-phrase is linked with the embedded clause rather than bearing the expletive subscript.

Both types of analysis have aspects which work well for some WSM languages. However, to date there has been no single analysis of WSM which satisfactorily accounts for the cross linguistic variation seen. While DDA analyses account for most data in languages like German, they fail to account for Hindi. In contrast, IDA analyses account for Hindi, but do not satisfactorily explain all the data from German. As discussed below, neither account can fully explain the pattern of Hungarian. **Fundamentally, DDA accounts are grounded in syntactic patterns while IDA accounts are rooted in semantics.** A main objective of this chapter is to determine whether either account can explain the existing cross-linguistic variation and to establish what semantic and pragmatic circumstances lead speakers to use WSM in a sample of languages that license it.

In what follows, we will give summaries of the advantages of the DDA (Section 2.2) as well as the IDA (Section 2.3). We will then explain the puzzle presented by WSM in Hungarian and the failure for either account to fully explain that phenomenon (Section 2.4). In Section 2.5, we make the claim that the DDA and IDA have truly been examining separate phenomena, one syntactic (DDA) and one semantic (IDA) which are superficially similar, but do not necessarily align. We then introduce other aspects of semantics which we hypothesize to be relevant to when WSM is used syntactically (Section 2.6). Finally, we describe two surveys conducted to inform our understanding of the cross-linguistic variation in the use of WSM. In Chapter 3, we will turn to creating a unified account of WSM which explains both the syntactic and semantic patterns attested.

2.2 The Direct Dependency Approach (DDA)

DDA analyses are based on the syntactic patterns found in WSM languages. They analyze the scope marker, not as a true wh-phrase, but as an expletive which is replaced at LF in accordance with the principle of Full Interpretation (Chomsky, 1991). However, within the DDA, analyses differ. McDaniel (1989), for example, considers the scope marker (henceforth, SMer), as an expletive, base generated in SpecCP of the matrix clause. This expletive is then replaced at LF by the true wh-phrase, whose scope it marks (hence the name, "scope marker"). Fanselow and Mahajan (2000) also consider the SMer an expletive, but one which has an associate dependency with the embedded CP. Under this analysis it is the entire CP which replaces the SMer at LF.

There is substantial evidence that the SMer does not behave like a contentful wh-phrase, at least in German. One of the primary arguments comes from what Müller (1996) terms the "Anti-locality effect". This refers to the fact that, while *was* as a contentful wh-phrase can appear in questions with multiple other wh-phrases (5a), as an SMer it can never be clausemates with the wh-phrase whose scope it marks (5b):

(5)a [$_{CP}$ Was$_1$ hat [$_{TP}$ sie warum$_2$ t_1 getan]]?
What$_1$ has she why t_1 done?
What has she done and why?

b *[$_{CP}$ Was$_1$ ist [$_{TP}$ sie warum$_1$ gekommen]]?
[SMer] is she why come?

Why has she come?

Examples from Müller (1996, 10a and 10b)

Furthermore, while WSM constructions are possible with multiple wh-phrases as seen in (6), the SMer must appear in the initial SpecCP position. In German, you can never have the SMer in situ, for instance, as in (7).

(6) [$_{CP}$Was meint wer [$_{CP}$ wann sie gekommen ist?]]
[SMer] thinks who when she come is
Who does she think is coming when?

(7) *[Wo meint was [wann sie gekommen ist?]]
Who thinks [SMer] when she come is
Who does she think is coming when?

Examples from Müller (1996)

While this evidence does not prove the SMer is not contentful, it does suggest, overall, that the scope marker does not behave in the same way as a typical wh-phrase and therefore should not be analyzed as such.

More direct evidence that the SMer should be analyzed as a contentless expletive comes from the fact that, in German, the use of the SMer, *was,* alternates with the expletive object pronoun *es* (Fanselow and Mahajan, 2000). Fanselow and Mahajan (2000) suggest that *was* is simply the [+wh] form of *es*[7]. Consider 8 (a-c) below: (8a) illustrates that *es* can be used with the verb *glauben* 'believe' and it can be used with LD movement in questions with multiple clauses. As shown by the contrast between (8b) and (8c), the expletive *es* and the scope marker *was,* cannot co-occur.

(8a) Wen glaubst Du es mir nicht, daß sie liebt ?
Who believe you it me not that she loves?
Who don't you believe me that she loves?

(8b) *Was glaubst Du es, wen Maria liebt ?
[SMer] believe you it, who Maria loves?
Who do you believe Maria loves?

Examples from Fanselow and Mahajan (2000, 8&9)

(8c) Was glaubst Du, wen Maria liebt ?
What believe you, who Maria loves?
Who do you believe Maria loves?

Example confirmed with native speaker (p.c.)

[7] Mahajan (2000) makes a similar argument for Hindi, suggesting that *kyaa* appears in alternation with the expletive *yah* and is, thus, the [+wh] form of *yah*. Horvath (2000) indicates this is also true of the Hungarian expletive *az*+case. This would be further evidence for the scope marker behaving as an expletive, which would support a DDA account of SM. We will see that this supports the specific account of "Syntactic Scope Marking" we adopt in Section 2.5.

Fanselow and Mahajan (2000) claim that for most speakers of German, the contrast between (8a) and (8b) is quite strong (but c.f. von Stechow, 1996; 2000). However, this piece of evidence changes the syntactic analysis of the SMer. If *was* is the [+wh] form of the expletive *es,* this further implies that *was* does not, in fact, originate in SpecCP. Rather, they argue, it originates in the same place as *es,* that is, as the object of the matrix verb, taking a CP associate.

We have, therefore, seen evidence from German syntactic patterns that the SMer carries no content and behaves similarly to expletive *es.* Semantically, this predicts that in a language like German which has both LD and WSM available, LD and WSM sentences should have the same meaning. Indeed, this is argued to be the case by (e.g.) von Stechow (2000) who claims the meaning of (9) and (10) are both expressed logically in (11).

(9) *Was* glaubst Du, *was* Peter meint, *was* Hans sagt, *was* Klaus behauptet, **mit wem**
[SMr] believe you, [SMr] P. thinks, [SMr] H. says, *what* K. claims, **with whom**

Maria gesprochen hat?
M. spoken has?

With whom do you believe that Peter thinks Hans said Klaus claims Maria spoke with?

(10) **Mit wem** glaubst Du, daß Peter meint, daß Hans sagt, daß Klaus behauptet, daß
With whom believe you, that P thinks, that H. says, that K. claims , that

Maria gesprochen hat?
M. spoken has?

With whom *do you believe Peter thinks Hans says Klaus claims Maria has spoken?*

(11) λp∃x[person (x) ^ p = you believe that Peter thinks Hans says Klaus claims Maria talked to x]

Informally, this means that there is a single entity, x, which can be described as the x you believe Peter thinks Hans says Klaus claims Maria talked to. If these two sentences have precisely the same meaning (a proposal we will return to), then a syntactic analysis such as is presented by the DDA, is a natural explanation. At LF their representation is actually the same.

The DDA, in short, provides a satisfactory account for German, assuming sentences like (9) and (10) truly have the same meaning. However, WSM has been described in many languages and the DDA does not explain all patterns found cross linguistically. Thus, we turn to the IDA.

2.3 The Indirect Dependency Approach

Hindi has also been described as having WSM (Dayal, 1994). Here, however, the SMer appears in the in-situ object position of the matrix verb rather than in SpecCP, as seen in (12) from Dayal (1994,18).

(12) [$_{CP}$ Jaun **kyaa** soctaa hai [$_{CP}$ Meri **kis-se**$_i$ baat karegii?]]
John **SMer** thinks Mary **who-with** will talk
Who does John think Mary will talk to?

(13) *[$_{CP}$ **Kyaa** Jaun soctaa hai [$_{CP}$ **kis-se** Meri baat karegii ?]]
SMer John thinks **who-with** Mary will talk?
Who does John think Mary will talk to?

The grammaticality of (12) compared to (13) suggests both the SMer and the true wh-phrase must remain in-situ. This in itself could be accounted for under the DDA. Fanselow and Mahajan (2000) in fact suggest that, in both German and Hindi, the expletive starts in SpecAgrOP and in German it undergoes overt wh-movement (as shown in 14) while it remains in situ in Hindi (as seen in (15). This analysis crucially assumes that the semantics of Hindi should resemble the semantics of German shown in (11), which many have argued is not the case (e.g., Dayal, 1994; Lahiri, 2002).

(14) [$_{CP}$**Was$_k$** glaubst$_j$ [$_{TP}$ Du [$_{VP}$ t$_j$ [$_{AgrOP}$ t$_k$**[$_{CP}$ mit wem$_i$** [$_{TP}$ Maria t$_i$ gesprochen hat]**]**]]]]?

(15) [$_{CP}$ [$_{VP}$Jaun [$_{AgrOP}$ **kyaa**] [$_{V}$ soctaa hai **[$_{CP}$** [$_{VP}$ Meri [$_{AgrOP}$ **kis-se**] [$_{V}$ baat karegii]]**]**]]]?

Dayal (1994) gives a different analysis of the semantics of WSM based on Hindi as well as the more general point that cross-linguistically, the SMer is consistently the wh-form which is used for abstract entities such as propositions. She suggests that WSM constructions consist of two *contentful* wh-phrases and, thus, two sets of propositions (following Hamblin, 1973). The first is a question over propositions. It denotes a set of possible answers which consist of propositions. The second wh-phrase restricts the set of possible answers. In other words, in a question like (16), the first question asks what John thinks. The set of possible answers includes all the things that John thinks. The 'second question' restricts John's thoughts to those about Mary's conversation partner.

(16) Jaun kyaa soctaa hai, Meri kis-se baat karegii?
John what thinks Mary with-who will-talk?
Who does John think Mary will talk to?

Dayal (1994) expresses this formally in (17):

(17) λp∃q[∃x[q= will-talk' (m,x)] & p=think'(j,q)]

Crucially, Dayal (1994) claims that (18), which is the analysis presented for German (von Stechow, 2000) does not correspond to the appropriate meaning in Hindi.

(18) λp∃x[person (x) ∧ p = John thinks that Mary will talk to x]

Thus, it seems that the German semantic analysis is not applicable to Hindi. If we assume these are examples of the same syntactic phenomenon, we must return to German and ask whether Dayal (1994)'s analysis satisfactorily explains German WSM. That would, crucially, mean WSM constructions are semantically different from LD constructions. A distinguishing factor between the DDA and IDA, therefore, is whether differences appear between LD and WSM constructions in languages like German that allow both. Indeed, such differences have been claimed to exist. Herburger (1994) claims that in a WSM construction, the embedded clause must be interpreted *de re*. This means that in a WSM question such as (19), it is presupposed that Maria spoke with someone. This is not the case for LD in (20), which allows for the possibility that Maria spoke with no one and the belief that she did exists only in the minds of Peter, Hans, and Klaus.

(19) *Was* glaubst Du, *was* Peter meint, *was* Hans sagt, *was* Klaus behauptet, **mit wem**
What believe you, *what* P. thinks, *what* H. says, *what* K. claims, **with whom**

Maria gesprochen hat?
M. spoken has?
With whom *do you believe Peter thinks Hans says Klaus claims Maria has spoken?*

(20) **Mit wem** glaubst Du, daß Peter meint, daß Hans sagt, daß Klaus behauptet, daß
With whom believe you, that P thinks, that H. says, that K. claims , that

Maria gesprochen hat?
M. spoken has?
With whom *do you believe Peter thinks Hans says Klaus claims Maria has spoken?*

Examples from von Stechow (2000, 1a & 1e)

While this might be a dialectal difference as not all speakers of German agree with this intuition (Gisbert Fanselow, personal correspondence), this is evidence that LD and WSM constructions in German are not identical (at least for some speakers), contra the DDA. This difference, however, is predicted by the IDA because it assumes different interpretations for LD (which has a single wh-phrase) and WSM (which has two).

Furthermore, Lahiri (2002) suggests that the presuppositions observed by Herburger (1994) for German WSM constructions hold for Hindi. A sentence such as (21) presupposes that Ramaa saw someone.

(21) Raam **kyaa** soctaa hai ki ramaa-ne kisko dekha.
Ram **what** thinks that Ramaa who saw
Who does Ram think that Ramaa saw?

Further differences between LD and WSM have been identified. Dayal (1994) indicates that negation of the matrix clause is allowed in German LD, but not in WSM, as illustrated by (22) and (23):

(22) **Mit wem** glaubst Du nicht, daß Maria gesprochen hat ?
With whom think you not that Maria spoken has?
Who don't you think Maria has spoken to?

(23) * **Was** glaubst Du nicht, **mit wem** Maria gesprochen hat?
What think you not, **with whom** Maria spoken has?
Who don't you think Maria has spoken to?
Example from Dayal (1994, 14a-b)

Quantifier scope readings are also different in LD and WSM constructions, as described by von Stechow (2000).

(24) a. **Wo**$_i$ glaubt jeder, daß sie gerne t_i leben wurde ?
Where believe everyone that she readily live would?
Where does everyone believe she should readily live?

b. ***Was***$_i$ glaubt jeder, **wo**$_i$ sie gerne t_i leben wurde ?
What believe everyone, **where** she readily live would
Where does everyone believe she would readily live?

For (24a) both readings (25a and 25b) are possible, however, for 43b, only reading 46b is accessible.

(25) a. For which place x, does everyone believe that she would like to live at x?
b. For everyone y: for which place x, does y believe she would like to live at x?
Examples from von Stechow (2000, 63a&b and 64 a&b)

It should be noted, that these differences do not necessarily mean we should discount the DDA completely. It could be the case that the DDA is correct and these differences in semantic meaning are the result of phenomena unrelated to WSM, such as restrictions on movement (e.g. von Stechow, 2000).[8] Under the IDA, however, subtle differences in the interpretation and use of LD constructions and WSM constructions are predicted, as they have different semantics. While there is much disagreement on just how robust these judgments are and how they should be analyzed (von Stechow, 2000) they all suggest a difference of interpretation between the LD and WSM constructions in German. However, we cannot ignore the substantial evidence presented in Section 2 which indicate the SMer in German is *not* a contentful wh-phrase.

Finally, Lahiri (2002) points out that, in Hindi, embedded yes/no questions are possible with WSM as shown in (26):

(26) Tum **kyaa** socte ho ki Ramma ghar gayii yaa nahiiN
You **what** think that Ramma home went or not
Do you think that Ramaa went home, or do you think she didn't?

[8] Von Stechow claims that the data from negation can be explained via Beck's filter (Beck 1996), which states that movement at LF is blocked by both quantifiers and negation, but movement at S-construction is not. Thus, these exceptions would be compatible with DDA.

This is predicted by the IDA, because under the assumption that the “second” question is simply limiting the set of possible answers to the first, any type of question should fill that purpose. However, in contrast, embedded yes/no questions do not allow WSM in German, as has been pointed out by von Stechow (2000). Thus, embedded yes/no questions pose a puzzle for each type of account, particularly for the IDA in accounting for German. In short, it is impossible to simply accept the IDA as providing a satisfactory analysis of German WSM. This will be further discussed in section 2.5.

2.4 The puzzle of Hungarian

Up to this point, one might conclude (along with von Stechow (2000)) that some form of the IDA could work for Hindi and the DDA could be made to work for German. However, neither analysis satisfactorily explains the patterns attested in Hungarian.

2.4.2 Evidence against DDA from Hungarian

Horvath (1995; 1997; 2000) suggests that Hungarian WSM cannot involve direct syntactic linking between an expletive scope marker and the contentful wh-phrase. The most convincing evidence is that, in Hungarian, the scope marker carries the case assigned by the matrix verb. Consider (27 a vs b and 27 c vs d).

(27a) **Mire** számítasz hogy mit fognak mondani a gyerekek ?
What$_{al}$ count-2sg that what$_{acc}$ will-3pl say-Inf the kids$_{nom}$?
What$_{al}$ do you expect what$_{acc}$ the kids will say?

(b) *$\textbf{Mi(t)}$ számítasz hogy **mit** fognak mondani a gyerekek ?
What$_{(acc)}$ count-2sg that **what$_{acc}$** will-3pl say-Inf the kids$_{nom}$?
What$_{(acc)}$ do you expect what$_{acc}$ the kids will say?

(c) **Mit** mondtál hogy **mire** számítanak a gyerekek?
What$_{acc}$ said-2sg-indef.$_{DO}$ that **what$_{al}$** count-3pl the kids$_{nom}$?
What$_{acc}$ did you say what$_{al}$ the kids expected?

(d) ***Mi(re)** mondt-ál/-ad, hogy **mire** számítanak a gyerekek?
What(al) said-2sg-indef/def.DO that **what$_{al}$** count-3pl the kids$_{nom}$?
What (al) did you say what$_{al}$ the kids expected?

Examples from Horvath (2000, 13 a-d)

It is clear from (27a) and (27c) that either accusative *mit* or allative[9] *mire* may act as the scope marker. Furthermore, the case does not necessarily match the case of the contentful medial wh-phrase, as is made apparent in both (27b) and (27d). Horvath (1997, 2000) suggests the SMer behaves more like a typical wh-phrase because it carries case. Specifically, she suggests the SMer is base generated in the matrix SpecAgrOP above VP independently of its morphological (accusative/allative) case, and as the complement of the matrix verb. In other words, the SMer is acting as a dummy case marker. Because she analyzes the SMer more like a typical wh-phrase than an expletive (it moves from a lower position within the VP shell to its surface position in SpecFocP), [10]Horvath suggests the syntax of Hungarian WSM is inconsistent with the traditional syntactic DDA analyses but seems compatible with the IDA. (28) below illustrates this construction using the sentence from (27c), omitting traces of subject and verb movement, for simplicity:

(28) [$_{CP}$[$_{FocP}$**Mit$_k$** [$_{TP}$ mondtál[$_{vP}$ [$_{AgrP}$ **t$_k$** [$_{VP}$ [$_{CP}$ hogy **[$_{FocP}$ mire$_i$** [$_{TP}$ számítanak [$_{vP}$ a gyerekek [$_{AgrP}$ **t$_i$**]]]]]]]]]]?

[9] The allative case indicates motion to or towards the object it marks.
[10] In Hungarian FocP (Focus Phrase) is the typical landing place for wh-phrases, which for embedded wh-phrases is clearly situated below CP whose head is filled with the complementizer *hogy* (Horvath, 1997). See (30) below.

2.4.3 Evidence Against the IDA from Hungarian

Though the syntax of Hungarian WSM constructions is more compatible with the IDA approach than the DDA approach, the IDA does not explain the semantic or syntactic patterns of Hungarian completely. Horvath (2000) discusses three predictions made by the IDA, which are not met. Recall that the IDA claims that the set of possible answers to the first question (clause) is limited by the set of possible answers to the second. Horvath points out that an analysis such as this, which necessarily considers the two clauses independent of each other, predicts that any well-formed embedded questions of the language should be possible subordinate CPs in the WSM construction. Second, the contrary case should also be true: an ill-formed embedded question should not be a possible subordinate CP. Third, if there are multiple wh-phrases in the embedded CP, they should all receive equal scope, that is, the scope marker should mark scope for all wh-phrases equally.

As evidence against the first claim, Horvath (2000) discusses *whether* clauses which can appear as embedded yes/no questions in Hungarian, as seen in example (29). While the main verb and context is appropriate for WSM in the presence of an embedded *who* (29), its *whether* counterpart is not, as shown in (30):

(29) Mari nem mondta hogy beszélt-e már Jánossal.
M_{nom} not said that talked-3sg-Q already J.with.
Mary didn't say whether she had already talked with John.

(30) **Mit** mondott Mari, hogy **kivel** beszélt már?
What$_{acc}$ said-3sg M_{nom} that **who-with** talked-3sg already?
Who did Mary say she had already talked with?

(31) ***Mit** mondott Mari hogy beszélt-e már Jánossal?
What$_{acc}$ said M_{nom} that talked-3st-Q already J.-with?
What did Mary say whether she had already talked with John?

Examples from Horvath (2000, 23, 21a-b)

Thus, not every type of question can appear as the embedded clause in Hungarian WSM constructions. The interpretation of (30), according to the IDA, should be: for which proposition, p, in the set of possible answers to {Mary had already talked to John; Mary had not already talked to John}, Mary said p? The IDA predicts this sentence to be grammatical but it is not. Recall from section 3.0 that Lahiri (2002) discusses the grammaticality of embedded yes/no questions in Hindi as evidence *for* the IDA. Here, because these constructions are ungrammatical in Hungarian and German (Beck & Berman, 1996; von Stechow, 2000) it is used as evidence *against* the IDA. This conflict in what scenarios allow WSM suggests that a unified account will be more complicated than simply using either the IDA or the DDA for different languages.

Furthermore, Horvath indicates that interrogatives which, for independent reasons, cannot otherwise be embedded, e.g. (31), can appear as embedded questions in conjunction with WSM constructions in Hungarian, as in (33):

(32) *Kérdezték hogy **kivel** találkoztam-e
Asked-3pl that **who-with** met-1sg-Q
They asked with whom whether I had met.

Example from Horvath (2000, 28)

(33) **Mit** kérdeztek hogy **kivel** találkoztam-e?
What$_{acc}$ asked-3-pl that **who-with** met-1sg-Q
With whom did they ask whether I had met?

Example from Horvath (2000, 27)

Taken together, examples (29-33) indicate that there are syntactic restrictions on what can act as the embedded clause interrogative in WSM constructions, which are not predicted by the IDA.

Finally, Horvath (2000) notes that the IDA predicts that a WSM construction which contains multiple questions in the embedded clause should give equal scope to all the embedded questions. If this were the case, we would expect *kit* 'who' and *mikor* 'when' to receive equal scope in (34).

(34) **Mit** kérdeztek hogy **KIT** mikor látott Mari?
What$_{\mathbf{acc}}$ asked-3pl that **WHO**$_{\mathbf{acc}}$ when saw M.nom?
Who did they ask when Mary had seen?

However, this not the case. Horvath (2000) crucially indicates that obligatory prosodic emphasis (indicated with capital letters in (34)) gives the emphasized wh-phrase matrix scope, which would not be predicted by the IDA. Thus, the prediction of the IDA that *kit* and *mikor* in (34) should receive equal scope is not born out.

Overall, these data (which we verified with a native speaker) suggest that an IDA analysis would not adequately explain Hungarian WSM. If WSM constructions were truly questions over propositions which were further restricted by their embedded clause interrogatives, then speakers of Hungarian should be able to interpret either *kit* or *mikor* as the main question being answered. However, *kit* remains the primary question and any answer to this question should be the answer to a *who* question. Furthermore, the IDA predicts that questions with multiple wh-phrases in the embedded clause should receive equal scope, which is not the case.

2.5 Summary: What is WSM?

The viewpoints embodied in DDA and IDA can be simplified to one question: Is WSM a syntactic or a semantic phenomenon? We propose that the syntactic pattern described by (e.g.) McDaniel (1989) and typified by the use of a contentless wh-phrase marking the scope of a contentful wh-phrase, need not necessarily be related to the semantics proposed by Dayal (1994). As seen in the case of Hungarian, it is entirely possible to have the syntactic pattern which would be expected by the IDA, but the semantics which would be expected of the DDA.

We propose that these are, in fact, two phenomena which are superficially similar, but which do not necessarily overlap. This is a similar conclusion to the one reached by Horvath (1997, 2000) who suggested that the IDA and DDA patterns might be the result of separate phenomena in Hungarian. Let us consider them as separate ideas. In the first place, there is the semantic analysis proposed by Dayal (1994), henceforth, SemSM. To reiterate, SemSM suggests that in a WSM construction, a *matrix* question over propositions is restricted in its possible answers by a second, embedded question, expressed formally in (35) (17 above):

(35) $\lambda p \exists q[\exists x[q= \text{will-talk'}(m,x)] \,\&\, p=\text{think'}(j,q)]$

Informally, this means roughly, there is something John thinks, there is also someone that Mary talked to: Which (of the things John thinks) is who Mary talked to? Dayal (2000) claims that (35) is universally available, appearing in English and other languages in the form of *sequential questions*[11] as in (36):

(36) A: What does John think? Who will Mary talk to?
B_1: John thinks Mary will talk to {Sue, Peter, Paul…}
B_2: * John thinks chocolate ice cream is superior to vanilla.

The argument is that a response about John's thoughts (an answer to the first question), is not a felicitous response to the two questions presented together. A felicitous overall answer must be from the set of people John thinks Mary might have spoken to. While a response such as B_2 might indeed answer the first question, it is infelicitous because it is not relevant to the second. Even though the second question is not syntactically embedded in the first, they are linked semantically. SemSM, therefore, can be realized by various syntactic patterns, cross-linguistically: a single question with two true wh-phrases in Hindi, or sequential questions connected by discourse in English and other languages (including as an option in German and Hungarian). This will indeed be born out in our survey results.

The use of an expletive SMer, henceforth SynSM, is one option cross-linguistically for how SemSM (the restriction of a matrix question over propositions by a second embedded question) might be realized. The literature reviewed above clearly suggests that SynSM does not correspond to SemSM in German. Similarly, Hungarian uses SynSM, but not to express SemSM. Rather, in German, the semantic input which results in the SynSM output is closer akin to what results in LD constructions in that language, with further specifications regarding Contrastive Topic, as we will discuss in Chapter 3. In sum, no

[11] Rawlins (2015) suggests *about* constructions as a potential output for this meaning, as in *What does John think about who Mary will talk to?*

single account of WSM thus far satisfactorily explains the cross-linguistic variation seen. We have suggested that this is in part because the explanations which have been considered in previous work were, in fact, describing two separate phenomena, SynSM and SemSM. In some languages (e.g., Hindi) SynSM is indeed the syntactic construction used to express SemSM, while in others SynSM is a grammatical construction but does not express SemSM (German, Hungarian). In still others, SynSM is not a viable construction and SemSM is expressed using other constructions (e.g., sequential questions in English).

We conclude, therefore, along with Dayal (2000), that SemSM is a universal appearing in all languages, but with distinct syntactic realizations. The task of coming to understand how SemSM and a syntactic realization are variably paired cross-linguistically is postponed until Chapter 3, where a specific OT analysis is proposed, that takes advantage of its input-output mapping view of the grammar. For now, the remainder of this chapter describes two surveys we conducted which investigate further the general hypothesis arising from the literature review that there is variation in syntactic constructions used to express SemSM cross-linguistically.

Section 2.6: Survey of syntactic variation in expressing SemSM

The primary goals of this experimental study will be to answer the following questions:

- Is there variation between languages in the constructions speakers choose to use in certain circumstances? If so, this suggests that neither the IDA nor the DDA fully account for the findings. DDA predicts synSM and LD should occur in similar circumstances, while the IDA predicts synSM should pair with sequential questions as both would have multiple, true wh-phrases.

- Does SemSM surface in different ways cross-linguistically? If Dayal (2000)'s assertion is correct, we should see variation in the constructions which occur in circumstances that involve SemSM.
- Are LD and SynSM interchangeably used in German? This is a direct test of the DDA which claims that LD and synSM should have the same meaning.
- When are sequential questions (SeqQs) used, including in languages that license WSM? Previous literature only discusses SeqQs in a cursory way. Dayal (2000) suggests that they are the way that English expresses WSM, but they are grammatical in all four languages of interest.
- What semantic/pragmatic factors is SynSM sensitive to?

To answer these questions, we will examine how various aspects of the semantics and pragmatics which have not previously been considered can generally affect what syntactic construction a speaker will prefer in a given scenario. To establish the variation that occurs cross-linguistically, we designed a survey (described in detail below) which controlled for three variables identified as relevant through preliminary interviews with native speakers. These variables include *Question Under Discussion* (QUD) (Roberts, 1996), Contrastive Topic (CT) (Büring, 2003), and Register (Ure & Ellis, 1977). From the Optimality Theory (OT) perspective adopted in Chapter 3 these variables become elements of the semantic input to syntactic optimization, which potentially activate different constraints. We begin this section by describing each of these variables in detail (sections 2.6.2- 2.6.4). We then describe the two surveys conducted and the results obtained. Finally, we introduce the very general idea behind the OT analysis developed in Chapter 3.

2.6.2 Aspects of the semantic/pragmatic input to syntactic variation: QUD

Dayal's proposal that SemSM involves multiple propositions is based on the generally accepted analysis of questions as denoting a set of possible answers (Hamblin, 1973). Her analysis then allows the set denoted by the matrix wh-phrase to be limited by the set denoted by the second wh-phrase. As Dayal (1994) suggests, this indicates the two wh-phrases are both contentful. This means they act somewhat independently in spite of being linked, with each question addressing a different, unresolved *Question Under Discussion* (QUD), as described by Roberts (1996, 2012). Thus, the first element we examine is QUD.

The QUD model resembles a stack. There is an underlying QUD to every exchange, with smaller QUDs addressing subtopics. For instance if a group is discussing what to have for dinner, the base QUD is *what should we eat*?, while smaller QUDs might include *Are there any relevant allergies? Do we need to go grocery shopping? Does everyone like fish?* etc. Different levels of QUD are more relevant depending on the point of time in the exchange. For example, consider the exchange below:

Q: Should we eat shrimp or beefsteak tonight?
A: Emily is allergic to shellfish.
A': Beefsteak is on sale today.

This exchange is felicitous because, even though neither response gives a direct answer to the question, they each address one of the QUDs. Furthermore, certain QUDs become irrelevant as the exchange continues. For instance, at this point in the exchange returning to what we should eat is unreasonable and a statement such as *We should eat shrimp* is no longer felicitous if the person heard the dialogue and is not being openly hostile. The question of whether anyone is allergic to shellfish is resolved and returning to it is not felicitous.

Thus, we see that multiple QUDs may be addressed in a given utterance. Furthermore, a QUD can be resolved or unresolved, and thus open for discussion. This "openness" affects the felicity of the following utterances. Let us return to questions with multiple clauses. Consider a question such as (37). Relevant QUDs include (38) and (39). The "embedded" question, *What should we eat,* (and the propositions it denotes) can be either resolved or unresolved at any point of the exchange.

(37) What does Mary think we should eat?

QUDs which must be present to make the use of (37) reasonable include, but are not limited to:

(38) What should we eat? (which we refer to as the "second question")

(39) What are Mary's thoughts about what we should eat, in particular?

While other QUDs will be present in the exchange, these two are necessary for this utterance to be felicitous. However, it is *not* necessary that the second QUD (38) be unresolved. For example, consider the following scenario (40), where the relevant QUDs are (41) and (42).

(40) Carol and Cathy are paleontologists looking for valuable fossils at a collectors' fair. They see one table with larger rocks. Carol immediately recognizes that one of the rocks is, in fact, a dinosaur egg. After examining the rock, Cathy agrees. This is an incredibly rare find. Wondering if the collector knows what a find this is, Carol turns to him and asks, "what do you think this is?"

(41) What is this? (second question)

(42) What are your thoughts on what this is?

In this scenario, for the speaker, (41) is resolved because Carol knows what the rock in fact is. However, (42) is still unresolved because Carol does not know what the rock collector's thoughts are on the finding. Thus, *what do you think this is* remains a felicitous question.

Thus, we see that in a question with multiple clauses, it is possible for the second question (Q2) to be either resolved or unresolved. Whether Q2 is resolved will be a crucial aspect of the input to syntactic optimization in our OT analysis of WSM.

2.6.3 Aspects of the semantic/pragmatic input to syntactic variation: Contrastive Topic (CT)

The second semantic/pragmatic aspect we hypothesize to be relevant to syntactic variation is whether the subject is a *contrastive topic* (CT) (Büring, 2003). We use the term *topic* to mean the context of the discourse ('what the dialogue is about'), following Reinhart (1981). Topics can be construed as responses to QUDs: they are brought up in the discourse because they play a role in responding to the QUD. *Contrast* is a type of emphasis which is placed on an entity, which suggests there is (at least) one other entity which could have been mentioned (Repp, 2009). The contrasted entity is mentioned in contrast to other relevant entities. This emphasis can be indicated either by prosody or by syntactic movement or both, depending on the language. Büring (2003) suggests that the manipulation of CT and focus can render an utterance felicitous or infelicitous, independent of QUD. Consider the following example summarized from Büring (2003).

(43) QUD: What did {Mary, Fred, George} eat?

A: Fred_{CT} ate the $\text{Beans}_{\text{focus}}$

A * $\text{Fred}_{\text{Focus}}$ ate the Beans_{CT}

In contrast:

(44) QUD: Who ate the {squash, beans, carrots}?

A: *Fred$_{CT}$ ate the Beans$_{focus}$

A: Fred$_{Focus}$ ate the Beans$_{CT}$

Büring (1997b) and Wagner (2012) argue that in German the entity which receives CT must have higher scope than any other focused entity. For instance, though (46) represents the canonical word order, (45) would be the grammatical construction in scenario (44), though not (43).

(45) Die Bohnen$_{CT}$ hat Fred$_{Focus}$ gegessen
The Beans$_{CT}$ has Fred$_{Focus}$ eaten

(46) Fred hat die Bohnen gegessen

Examples from Wagner (2012; 14)

Wagner (2012) considers CT to actually be a form of focus consisting of nested focus features. The focused element with widest scope must be the one intended as the CT. Crucially, according to Wagner, focused items can come with associates. Thus, it is possible for a focused item to appear high in a syntactic construction, but for its focus to be interpreted farther down the tree. This is relevant to WSM in several ways. In the first place, an expletive element cannot be a CT because it does not carry semantic content. There are no alternatives which it is in contrast to. However, as a wh-phrase, the SMer does have a focus feature under most analyses (e.g. Wagner, 2012; Sabel, 2000). Wagner (2012)'s analysis allows for the SMer to have this focus feature, but for the actual focus to be on the associate CP complement of the matrix V.

In the second place, native speaker consultants of German have indicated to us in a pilot study that the SynSM construction is not as felicitous as LD when the subject of the matrix clause is CT. This means that the focus is on whoever's thoughts are being elicited rather than the content of the wh-phrase. For example, if Karen's thoughts are being elicited in contrast to others', as seen below, Karen is the CT:

(47) Three friends are deciding where to go to a movie. Margaret and Sue have different opinions so they will let Karen make the decision.
Where does Karen$_{CT}$ (rather than Margaret or Sue) think we should go?

The question here is less about where the friends should go and more about Karen's thoughts on the subject, in contrast to Margaret and Sue's. In English, CT is generally indicated with intonation, but in German, as indicated by Büring (1997b) and Wagner (2012), it must also be indicated by word order. CT is crucial to our analysis of German because it is relevant to whether LD and SynSM constructions are truly synonymous. It is also relevant to exactly how we analyze the associate of the SMer. If the SMer takes the entire embedded CP as its associate, as suggested by Mahajan (2000) and Fanselow and Mahajan (2000), then the embedded clause subject will be the entity with the highest scope. This should mean that SynSM is not felicitous when the speaker wishes to emphasize the matrix clause subject in contrast to another entity (as in (47)). In an LD construction, on the other hand, the matrix subject will have highest scope and can therefore be CT. If this is the case, then SynSM and LD, therefore, map to different semantic meanings: LD is optimal when the subject is CT, SynSM is optimal when the subject is not CT. However, to confirm this hypothesis, the way native speakers interpret CT in WSM constructions must be investigated.

2.6.4 Aspects of the semantic/pragmatic input to syntactic variation: Register

Register refers to specific language patterns that are used in certain situations, according to Ure & Ellis (1977). Adult speakers command a variety of linguistic registers and adapt their style of utterance to who the interlocutor is. Though elaborate lexical or morphosyntactic systems identifying Register are not available in all languages, someone who has mastered a language understands that the way they address their 5-year-old child is different from the way they address their employer. Even young children have been shown to adjust their register depending on their interlocutor. Shatz and Gelman (1973) showed that 4-year-olds use shorter, less complex constructions when addressing younger children than they use with their peers or their parents. Wagner, Greene-Havas, and Gillespie (2010) showed that children's ability to identify the likely interlocutor of an adult based on register improves greatly between ages 4;0 and 5;0. They further showed that these young children are able to grasp multiple registers at once. Register is, therefore, something speakers are highly sensitive to, even at a very young age. Furthermore, one of the changes in register when addressing children, as noted by Shatz and Gelman (1973), is that people tend to use shorter and less complex constructions. Considering the constructions we are interested in, therefore, it is reasonable to assume that adults speaking any language might prefer SeqQs such as (48) over longer, more complex constructions such as either WSM or LD when addressing young children.

(48) What do you think? Where should we go?

Adult/child-directed register will therefore be a factor we control for as we establish which syntactic constructions are used in which semantic contexts.

2.7 Cross-Linguistic Surveys

We have discussed three relevant aspects of the semantics which preliminary discussions with native speakers have suggested might be relevant to when synSM is preferably used. We proceed now to discuss two studies we have run with multiple goals. The first is a pilot survey of English speakers whose goal was to establish whether SemSM might be expressed via sequential questions in English, as suggested by Dayal (2000). The second is a cross-linguistic survey whose goals were three-fold. First, we hoped to establish what semantic attributes lead to the use of SynSM in those languages where it is felicitous (German, Hungarian, and Hindi). Second, we hoped to establish whether SynSM and LD are truly interchangeable in German. Finally, we have included sequential questions (SeqQs) in our work, which no previous work has thoroughly examined, though all four languages we have targeted (English, German, Hindi, Hungarian) allow them. It is not enough to assume they are interchangeable with either SynSM or LD constructions. A thorough investigation of SeqQs has yet to be undergone, as far as we are aware. We therefore hope to identify which scenarios result in the use of SeqQs as well.

2.8 Experiment 1: Pilot study with English-speakers

2.8.1 Participants

22 native speaking undergraduate students at Johns Hopkins University were recruited, interviewed in person, and compensated with course credit.

2.8.2 Procedure

This is a within-subjects design. The task included a survey of ten target questions as well as five filler questions. Every participant heard and answered every question. No participant was excluded.

Participants were read a short scenario and then listened to a recording of four potential constructions: LD, SeqQs, and two monoclausal questions which corresponded to the relevant QUDs (see example below). They were asked to rate these alternative questions on a scale of 1-7, where (1) meant "this sounds terrible in this context" and (7) "this sounds perfect in this context". Participants heard five scenarios in each of the four conditions (Q2 resolved/unresolved; subject +CT/-CT) as well as 5 filler scenarios[12] which were used to determine whether they were engaged and actively participating and whether they were native speakers. Note that only two of the variables introduced in Section 2.6.2 and 2.6.3 were controlled for in this task. The recordings were made by a research assistant trained to use the prosody relevant to CT in questions. In addition to hearing the scenarios, participants were presented with a printed version, as shown in (49). (49) includes a

---

[12] Filler scenarios were set up in exactly the same manner with a scenario followed by four alternative productions. However, they investigated the use of double modals, such as "might can". For example: June is making a cake for the first time. She has some experience baking, but has never made a cake. She finds two pans that look like the right size. She says:

I can use these pans.

I might use these pans.

I can maybe use these pans.

I might can use these pans.

Double modals elicit strong judgments from those whose dialect does not include them thus we were able to use them to determine if participants were paying attention and whether they were native speakers. The use of fillers which include truly ungrammatical options were necessary to identify participants who were not truly native speakers.

scenario where Q2 is closed (Carol knows what the rock truly is) and where the subject is CT (we are interested in what the collector thinks, as opposed to Carol and Cathy).

(49) Carol and Cathy are paleontologists looking for valuable fossils at a collectors' fair. They see one table with larger rocks. Carol immediately recognizes that one of the rocks is, in fact, a dinosaur egg. After examining the rock, Cathy agrees. This is an incredibly rare find. Wondering if the collector knows what a find this is, Carol turns to him and asks,

What do **you** think? What is this?	1	2	3	4	5	6	7
What do **you** think this is?	1	2	3	4	5	6	7
What do **you** think?	1	2	3	4	5	6	7
What is this?	1	2	3	4	5	6	7

It should be noted that consultants frequently commented that the dis-preferred construction was a fine construction for the relevant language, but felt awkward or simply "not as good" as the construction they chose as the best expression.

2.8.3 Results

Scores were z-scored[13] by participant and analyzed using a two-way ANOVA. The results indicate that, in every condition, English speakers preferred LD constructions over sequential questions. However, sequential questions received varying degrees of acceptability in each condition, as seen in Table 1 and Fig. 1:

[13] A z-score is the number of standard deviations a data point is from the mean of the entire set. In this case, the mean and standard deviation is found for each participant and then each data point is converted into the average number of standard deviations the data point is from the individual's average. In this way, we account for individual variation in how the scales were interpreted. This method is standard in acceptability judgment tasks (Schütze & Sprouse, 2014).

Table 1. *Average z-score of Sequential Question by condition*

	Q2 unresolved (SD)	Q2 resolved (SD)
SUBJ CT	.72 (.45)	-.53 (.91)
SUBJ NOT CT	.57 (.55)	-.81 (.65)

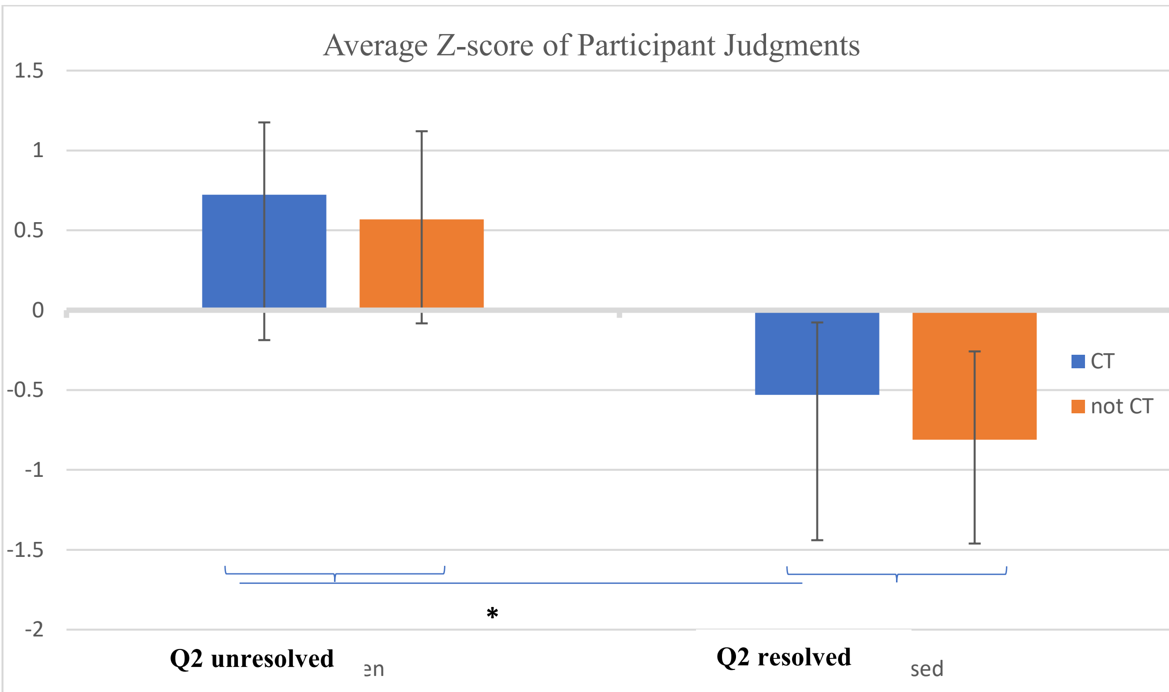


Fig. 1. Average Z-Score of participant judgments.

The use of sequential questions was judged significantly better when Q2 was unresolved than when it is resolved (p<.05). There was no significant difference between when the subject was CT and when it was not.

2.8.4 Discussion

Table 1 and Figure 1 provide evidence that sequential questions are possible in English when Q2 is unresolved, but never possible when Q2 is resolved. While LD is always the preferred construction in English, the possibility of using sequential questions is dependent on whether Q2 is resolved or not. Based on these results, we can conclude that QUD affects the use of SeqQs in English, but we cannot conclude when a participant would choose to use SeqQs instead of LD. However, this study did not control for Register. Multiple participants spontaneously suggested that SeqQs would be preferable when speaking to a child or in a teaching context. We also determined that a likert-scale task might not be the best strategy for our purposes as participants can judge constructions to be equally good. Experiment 2, therefore, will be a forced-choice task and will control for Register in addition to the resolved/unresolved status of Q2 and CT.

2.9 Experiment 2: Cross-Linguistic Study

2.9.1 Participants

Participants were recruited using Amazon Mechanical Turk (MTurk) and were paid for their time. We required the users to be native or near-native speakers of English, Hindi, Hungarian, or German. Filler questions were used to eliminate participants who did not speak the relevant language fluently. All participants were over the age of 18 and indicated that they considered themselves native speakers of the language of the survey they took. We saw varying levels of responses in the different language groups. Only data from participants who correctly responded to at least 85% of fillers (see footnote 8) were analyzed. 29 participants took the English survey, 12 met our criteria. Fifty participants took part in the Hindi survey, 10 met our criteria. Fifty-five took the Hungarian survey and 14 met our criteria. Finally, 47 took our German survey and data from 12 were analyzed. The small number of participants retained is not highly unusual for MTurk studies, as this type of study is susceptible to bots and participants who are not truly speakers of the relevant languages. The filler questions were not ambiguous: native speakers should have had no difficulties answering them. This allowed us to filter out participants who were not truly speakers of the relevant language.

2.9.2 Procedure

Participants read a short scenario and then were asked to "choose the best sentence to come next". They were then presented options which were all grammatical sentences in their language, but varied in syntactic construction depending on the language: LD, SeqQs, synSM with and without the complementizer (when available, e.g., in Hungarian), Q1 (*What do you think?*) and Q2 (the embedded question alone, as in *What is this?* from (70)). Items were a mixture of adjunct and argument questions.[14] However, unlike our pilot study, this forced-choice task required participants to choose only one of the options, rather than ranking them. This allows us to examine whether multiple constructions are chosen with equal frequency in our various conditions, or whether one is dominant. Furthermore,

[14] We recognize that some dialects of German do not allow the use of LD with argument questions (e.g., Lutz, Müller, & von Stechow, 2000). We do not consider this a problem because the variation in when participants chose LD vs. WSM was precisely what we were interested in. That being said, our primary focus is the use of WSM, thus the precise use of LD and other structures which reviewers have pointed out as grammatical (e.g., parenthetical constructions) we leave to future work.

forced-choice tasks are known for having high statistical power: judgments from a small number of participants, such as 10-14 per language, can be relied upon with relative certainty (Schutze & Sprouse, 2012). Because speakers of some of our languages of interest are not commonly found in the US (e.g. Hungarian), we decided this type of task would be optimal. The survey was created using Qualtrics online surveys and was made available via MTurk.

2.9.3 Stimuli

Recall that we have identified 3 variables of interest: the resolved/unresolved status of Q2; +/-CT of the subject of the question; and Register (adult/child-directed). This resulted in 8 combinations of the three variables. We created 4 scenarios for each of the 8 combinations, rendering 32 total items. A survey with all 32 items plus an equal number of filler items was deemed too lengthy. We therefore created two surveys (with items randomly assigned to each): Form A and Form B. Our filler to target item ratio was 1:1 in order to prevent participants from falling into a pattern of giving the same response type repeatedly. This meant that each form included 16 target items and 16 filler items. Filler items were the same in Form A and B and were identical to those used in the pilot study (see footnote 8).

While 10 of the target items were also identical to the pilot study, recall that that study did not include Register as a factor. We therefore added items which were child-directed. An example as shown below in (50), in which Q2 is resolved (because Jack knows the ride is off to the right), the subject is CT (because Christopher is being asked in contrast to his two brothers), and the Register is child-directed.

50) Jack has taken his three sons to a waterpark. They are looking for the way to their favorite ride. The two littlest boys are arguing at a fork in the path: should they go right or left? Stephen thinks he remembers the ride is to the left, while Andrew thinks it is to the right. Jack sees a sign that says the ride is off to the right. Christopher is just beginning to learn to read. Jack bends down and points at the sign, saying:

Christopher, what do **you** think? Where is the ride?

Christopher, where do **you** think the ride is?

Christopher, what do **you** think?

Christopher, where is the ride?

The stimuli and the Forms (A, B) were identical in all four languages. The German, Hindi, and Hungarian surveys were translated by native or near-native-speakers of the relevant languages and checked by a second native-speaker. Because CT is a prosodic as well as a semantic phenomenon, the options participants had to choose between were recorded as well, thus guaranteeing that participants made their choices based on the appropriate prosodic contour. The native-speakers were trained to read options with and without the prosodic contours required for CT subjects, depending on the conditions.

For the English survey, an additional pilot study was run to confirm the author's intuitions of the appropriate prosody for a question with a CT subject. Five adult, native speakers of English were asked to record questions which included CT subjects. They were shown the stimuli questions and the stories. They were also told to emphasize the subjects of the sentences, which were bolded to help them remember this. A PRAAT analysis of a characteristic sample of their recordings confirms that CT subjects in questions involve emphasis on the subject and a fall at the end of the question. The English options were then recorded by the author.

The number of options participants chose between varied based on the variety of viable constructions in the relevant language. English, for instance, included only LD, SeqQs and the two, relevant monoclausal questions (as shown in 71, above) while German included these as well as synSM.

To summarize, participants completed one of two forms (A or B) of a survey in their native language (English, German, Hindi, or Hungarian). The scenarios in the stimuli were designed to control for resolved/unresolved Q2[15], +/-CT subjects, and Register (adult/child-directed). The stimuli were identical in all four languages, barring small adjustments for things like the names of characters. Instructions and the consent form were all given in English; thus, all participants were bilingual, but the stimuli were all translated by native-speakers and the options participants chose between were pre-recorded by a native speaker of the relevant language. Participants listened to the options which included (maximally, when all were available in the language) LD, SeqQs, SynSM, SynSM with a complementizer, and the two monoclausal questions. They were asked to choose the sentence they thought sounded best for this circumstance. Surveys differed only in what language the stimuli were in. Any variation in the results should therefore be the result of language-specific variation in which syntactic construction is best in a given scenario.

---

[15] The resolved/unresolved nature of Q2 does not map directly onto the de re/de dicto distinction made in Herburger (1994). This is because the de re/de dicto distinction indicates a belief state, which is not necessarily related to whether the embedded question has been resolved or not. We did not pursue the de re/de dicto distinction in our stimuli because, as mentioned above, there is disagreement amongst native-speaker researchers on whether this distinction holds.

2.9.4 Coding and Analyses

We used a variety of statistical means to establish the effects of each condition on the use of each type of construction. First, we needed to establish whether these variables affected the use of each syntactic construction in the survey. To that end, the data were submitted to a mixed effects logistic regression model in the R statistical analysis environment (R Core Development Team, 2015). Each item was bimodally coded for whether it included a resolved (0) or unresolved (1) Q2 as well as whether the subject was (1) or was not (0) CT, and whether it was (1) or was not (0) Adult-directed. We then coded for each participant what syntactic construction they chose for each item. This coding was also bimodal as we marked whether each chosen response was (1) or was not (0) an example of LD, SeqQ, SynSM, SynSM with complementizer, Q1, or Q2. The models included Q2resolved/unresolved, CT, and Register as fixed effects and item number and participant as random effects. We used the maximally complex model that would converge which included Q2resolved/unresolved, CT, and Register as random slopes on participant and item number. Essentially, this model allowed us to determine to what extent each variable affected the choice of construction. Each construction was analyzed individually.

Second, we needed to establish whether the difference between the use of one construction was significantly more than other constructions within a given condition. For example, if 30% of responses indicate the use of SynSM in a particular context and 40% indicate the use of SeqQs, we need to establish whether SeqQs truly occur more frequently than SynSM or if that difference is noise and the two actually occur with equal frequency. To this end, we created a *construction* column which included SeqQs, SynSM, and LD as options and a *choice* column in which we indicated whether a participant did (1) or did not

(0) chose that construction in for that select that construction. We could then ask whether construction was predicted by choice. While this is somewhat counterintuitive, what the mode is really asking is whether a construction was more likely to be chosen than the other. Choice was, therefore, a fixed effect and we included participant and item number as random effects. We used the maximally complex model that would converge which included random slopes for item and participant. Effectively, this will tell us when one construction type occurs more frequently than another.

2.9.5 Results

English

Let us first turn to the results of the mixed effects logistic regression model examining the overall effect of the variables on each construction type. There was no effect of any predictor (or combination of predictors) on the use of either monoclausal construction alone. However, the use of SeqQs and LD both showed significant effects of the variables. Table 2 shows the results of the model on the use of LD.

Table 2. *Results of mixed-effects logistic regression examining LD.*

	Estimate Std.	Error	z-value	Pr(>|z|)
Register	-1.41	1.21	-1.16	.25
Q2	-3.14	1.12	-2.81	.004 **
CT	-2.51	1.09	-2.29	.02 *
Register:Q2	3.29	1.47	2.25	.03*
Register: CT	2.92	1.47	2.25	.05*
Q2:CT	4.42	1.39	3.19	.001 **
Register:Q2:CT	-4.45	1.85	-2.40	.02 *

(*) denotes significance

These results indicate that Register alone had no effect on the use of LD, the resolved/unresolved status of Q2 and CT did, as did the interaction of each combination of variables. However, when the same model investigated the effect of these variables on the use of SeqQs, the only effect found was that the interaction of Q2 and CT affected its use (p=.04). Fig. 2 illustrates this interaction, showing that for scenarios where the subject is contrastive, the use of SeqQ falls when Q2 is resolved compared to unresolved, while it rises when the subject is not contrastive.

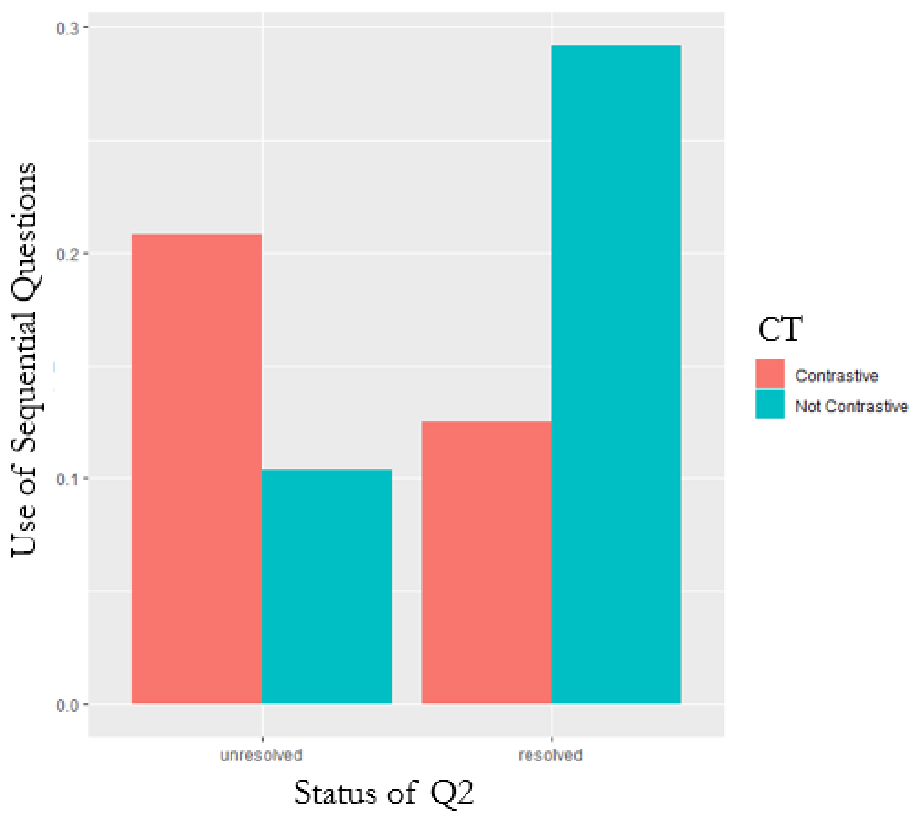


Fig. 2 The interaction between the resolved/unresolved status of Q2 and +/- CT significantly affects the use of SeqQs.

Though the total rate of SeqQ use is about 31% in scenarios where Q2 is unresolved and 41% in scenarios where Q2 is resolved, we can see that CT is also driving when SeqQs are chosen.

Fig 3, below, shows the variation in the choice of productions by condition. What is immediately apparent is that in all but 2 conditions, the grey bar illustrating the LD construction is dominant. The two conditions where LD is not dominant (First: Q2 unresolved, subject not CT, Child-directed and second: Q2 resolved, subject is CT, Child-directed) show no one construction occurring significantly more than the others.

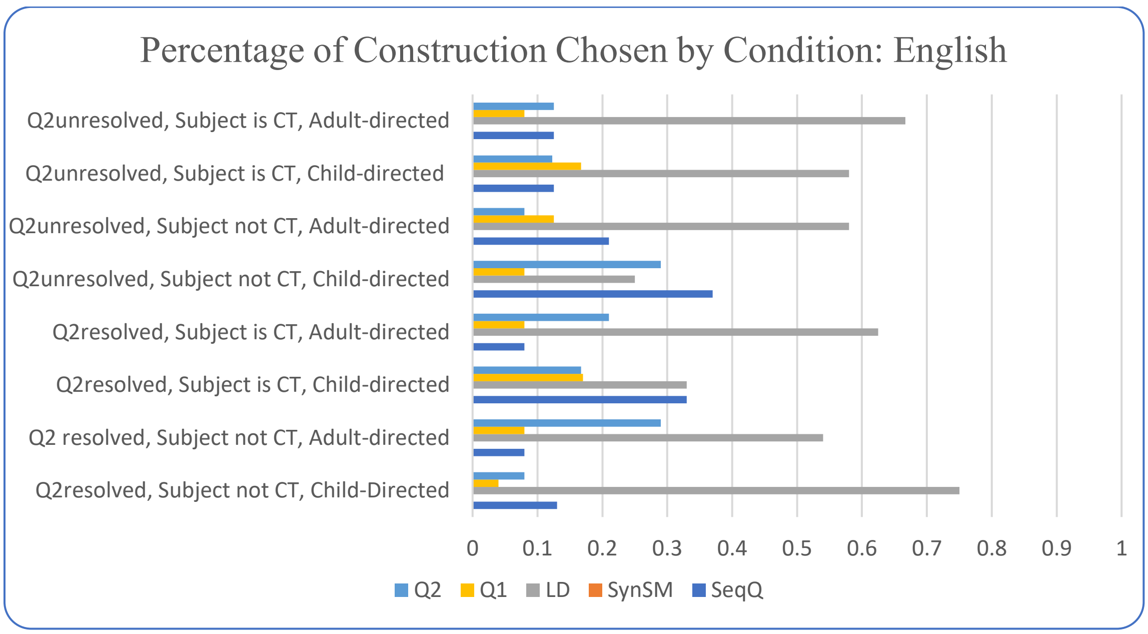


Fig. 3. Percentage of constructions chosen by condition. Color indicates construction type.

Overall, these data indicate that for English, LD is almost always the optimal choice, replicating the findings from our pilot study. There are two conditions where LD does not appear reliably more frequently than SeqQs: When Q2 is resolved, the Subject is CT, and the register is child-directed and when Q2 is unresolved, Subject is not CT and the register is child-directed. This is apparent when we consider Figure 3, as these are the two conditions where LD is roughly matched with SeqQ. In one condition, LD occurred significantly more frequently than SeqQs: the final condition, when Q2 was resolved, the subject was not CT and the utterance was child-directed. In all other conditions LD appears more than SeqQs to a degree approaching significance (p-values range from .06 to .09). There is no condition where SeqQ or any other construction appears significantly more frequently than the other constructions.

German

As with English, we will first turn to the overall pattern of constructions participants chose as optimal in German. The results of the mixed effects logistic regression model, presented in Table 3, illustrate the effects of each of our variables on the use of SynSM.

Table 3. *Results of Mixed-Effects logistic regression model investigating SynSM*

	Estimate Std.	Error	z-value	Pr(>|z|)
Register	0.23	0.94	0.24	0.81
Q2	-1.89	1.15	-1.65	0.099 .
CT	-3.06	1.52	-2.01	0.04 *
Register:Q2	2.48	1.45	1.71	.09
Register: CT	3.02	1.70	1.78	.075
Q2:CT	4.40	1.88	2.35	.019 *
Register:Q2:CT	-5.33	2.21	-2.41	.02 *

(*) indicates significance, (.) indicates approaching significance

Table 3 shows that CT, the interaction of CT and Q2, and the interaction of all three of the variables tested predict the use of SynSM constructions. On the other hand, none of these variables are reliable predictors for the use of either LD or SeqQs. Recall, that a primary argument of the DDA is that LD and SynSM constructions have the same interpretation. One prediction this makes is that they should not be distinguishable in any of the conditions we created. Based on the results above, this is clearly not the case.

Let us turn now to examining the dominance of each construction by condition. Again, the DDA predicts that the use of LD and SynSM should be roughly equivalent, so there should be no difference between these two constructions under that hypothesis. Furthermore, no literature reviewed above thoroughly investigated the use of SeqQs in German. With this in mind, consider Fig. 4. In the first place, there was clearly great variation in the construction chosen as no construction is ever chosen over 50% of the time in a given condition: this is in contrast to English, where LD was chosen over 60% of the time in most conditions. It is also immediately apparent that the construction chosen most frequently overall is SeqQs (in dark blue). [16]Finally, SynSM and LD (orange and grey, respectively) are not being chosen at an equal rate across conditions, counter to the prediction of the DDA.

[16] This unprecedented finding could be the result of multiple factors, and will certainly need to be replicated in order to give it proper credence. However, we provisionally believe this novel finding is the result of previous work failing to take SeqQs into account.

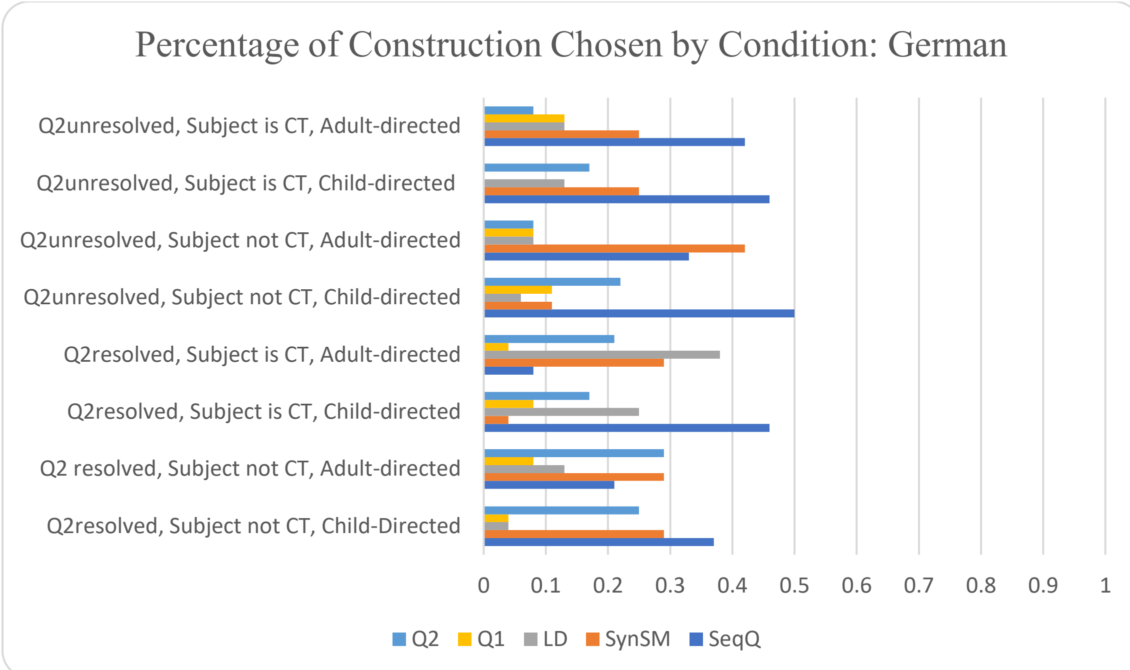


Fig. 4. Percentage of Construction type chosen by Condition. Color indicates construction.

To determine whether the differences which appear within conditions are significant, let us turn to the results of our mixed effects logistic regressions. In no condition was there one construction that appeared significantly more than all others. However, Table 4 presents those differences which were significant within conditions. For instance, in the third condition (where Q2 is unresolved, the subject is not CT, and the question is adult directed), SynSM and SeqQs both appear significantly more frequently than LD, but the difference between SynSM and SeqQs is not significant in this condition. In the last condition, where Q2 is resolved, the subject is not CT, and the question is child-directed, both SeqQs and SynSM occur significantly more frequently than LD, which was almost never chosen in this condition.

Table 4. *Significant differences in German constructions*

Condition	Significant differences
Q2 unresolved, Subject is CT, Adult-directed	No difference reaches significance
Q2 is unresolved, Subject is CT, Child-directed	No difference reaches significance
Q2 unresolved, Subject not CT, Adult-directed	SynSM occurs more than LD p=.02 SeqQs occur more than LD p=.05
Q2 is unresolved, Subject not CT, Child-directed	SeqQ occur more than LD p=.01
Q2 resolved, Subject is CT, Adult-directed	LD occurs more than SeqQs p=.03
Q2 resolved, Subject is CT, Child-directed	SeqQs occur more than SynSM p=.01
Q2 resolved, Subject not CT, Adult-directed	No difference reaches significance
Q2 resolved, Subject not CT, Child-directed	SynSM occurs more than LD p=.02 SeqQs occur more than LD p=.02

These tests suggest in the first place that SeqQs appear significantly more than either LD or SynSM in many conditions and in the second that there are multiple conditions where LD and WSM do not appear with equal frequency. But there are some conditions where SynSM and SeqQs are chosen with equal frequency. None of these outcomes are predicted by the DDA.

Hindi

The constructions in the Hindi survey included SynSM, SeqQs, and the relevant monoclausal questions, as with German, but we also included the use of SynSM with a complementizer, *ki*, (SynSM+ki) which is possible in Hindi. The two constructions are illustrated (51) and (52):

(51) Jahnavi, tumhein **kya** lagta hai, humein **kab** milna chahiye?
(52) Jahnavi, tumhein **kya** lagta hai **ki** humein **kab** milna chahiye?

While both constructions are generally accepted by speakers of Hindi, the difference between them (if any) is unclear from the literature reviewed (Dayal, 1994; Lahiri, 2002), but, as a possible construction of the language to be verified with native speakers, it was included as an option.

The mixed-effects logistic regression examining the data from Hindi show a very different pattern from either English or German. If we treated SynSM and SynSM+ki as distinct, then the use of no construction (SeqQs, SynSM, SynSM+ki, relevant monoclausal questions) was predicted by any of the variables. However, when we combined the SynSM constructions with and without the complementizer and considered them a single construction (which is what previous literature essentially does), then a pattern emerged, as shown in Table 5.

Table 5. *Results of the logistic regression model investigating SynSM with and without 'ki'*

	Estimate Std.	Error	z-value	Pr(>|z|)
Register	-.68	.82	-.83	.41
Q2	-1.39	.78	-1.78	.08 .
CT	-1.70	.81	-2.09	.04 *
Register:Q2	.98	1.03	.96	.34
Register: CT	2.56	1.04	2.46	.01 *
Q2:CT	1.83	1.12	1.64	.10
Register:Q2:CT	-1.63	1.47	-1.11	.27

(*)indicates significance, (.) indicates approaching significance

It is clear from these results that, as a whole, the CT and the interaction of CT and Register significantly affect the use of SynSM (with and without *ki*).

Turning to examining the constructions chosen by condition in Fig. 5, we see that SynSM (with and without *ki*, in orange) is dominant in multiple conditions, but it is not the primary type in every condition. However, the results of the mixed effects logistic regression model indicate that the difference between the constructions chosen within conditions reveal that the only condition where there is a significant difference between the use of SynSM and SeqQs is when Q2 is resolved, Subject is CT, and the register is Adult-directed (5th condition from the top). In this condition, SynSM(+ki) is used significantly more than SeqQs (p= .04).

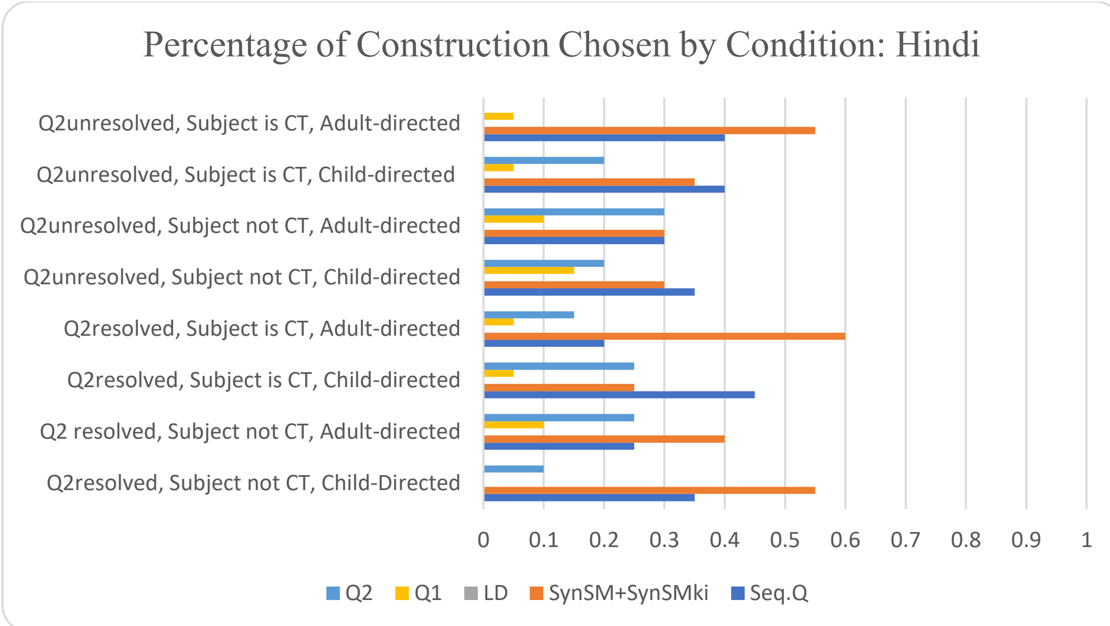


Fig.5. Percentage of Construction Chosen by condition. Construction indicated by color.

Because the IDA suggests that both SeqQs and SynSM are composed of two contentful wh-phrases, one might predict that SeqQs and SynSM could be used interchangeably. However, as indicated by the logistic regression model, we know that the use of SynSM is affected by CT and by Register, while SeqQs are not, indicating a difference between the two which is not predicted by the IDA. This is even apparent when we compare, for instance, the conditions where Q2 is resolved and the subject is CT (5th and 6th from the top). The use of SynSM drops dramatically when the Register is child-directed. However, it is not replaced purely by SeqQs, but by the monoclausal Q2 (in light blue) as well. The monoclausal Q2 is chosen with relatively high frequency in Hindi, compared with English or German.

Hungarian

As with Hindi, Hungarian allows for the use of SynSM with the complementizer *hogy*. We included this construction in the Hungarian survey in case there was a difference between when the two forms of SynSM (with/without *hogy*) were used. The results of the mixed effects logistic regression model indicate, however, that the only effect that approached significance for any construction was that the resolved/unresolved status of Q2 affected the use of SynSM (p=.07). This means, that overall, SynSM was used more when Q2 was resolved, but otherwise construction choice was unaffected by the conditions we created. We did not combine the use of SynSM with *hogy* and without, as we did in Hindi because the use of *hogy* was much more robust than *ki* in Hindi, as is apparent in Fig 6.

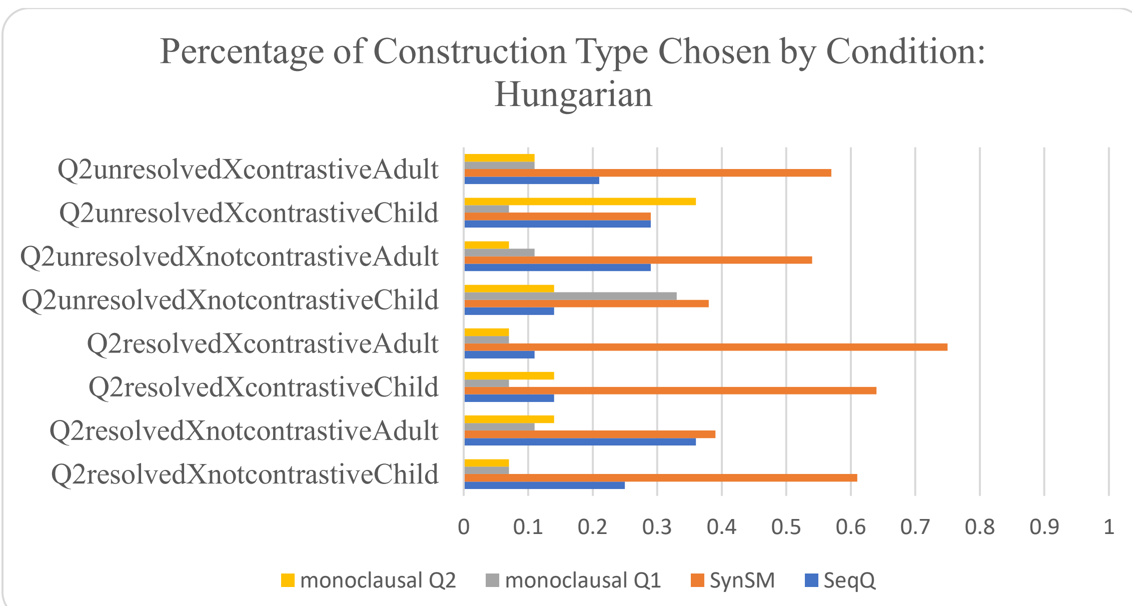


Fig. 6. Percentage of Construction Chosen in Hungarian.

In the Q2resolved, subject is CT, child-directed condition, SynSM was chosen significantly more than SeqQs (p=.02). In the condition where Q2 is resolved, subject is

CT and the register is adult-directed, SynSM was also chosen significantly more than SeqQs (p=.02). This suggests that in these two conditions we can conclude that SynSM constructions and SeqQ constructions are not considered equivalent in Hungarian. The exact role of constructions which include *hogy* is undetermined as it does not appear significantly more than other complex constructions in any condition. However, what can be determined is that, like Hindi and German, Hungarian allows SeqQs in most circumstances where it uses SynSM, but not always.

2.9.6 Results Summary

To summarize, native speakers of English, German, Hindi, and Hungarian judged from a given set of constructions which would be "best to follow" paragraphs describing scenarios which controlled for the resolved/unresolved status of Q2, the CT status of the subject, and the Register. The forced-choice judgements made by these participants resulted in distinct patterns for every language. In English, LD is almost always the optimal choice, but it is used with more or less frequency based on the conditions. In German, SeqQs are the most frequently chosen construction. The choice of SynSM constructions, however, was affected by conditions, appearing more frequently when the subject is not CT, when Q2 is unresolved, and when the register is adult-directed. In Hindi, SynSM(+*ki*) were used more when the subject is CT and the register is adult-directed. Finally, in Hungarian, we saw that SynSM(+*hogy*) were a frequently chosen construction. Only one variable approached a significant effect on any construction: Hungarian speakers chose SynSM more frequently when Q2 was resolved. We acknowledge that stronger effects would add credibility to our claims that these structures are used differently and hope to replicate our findings with further studies with more participants.

While no semantic variable affected SynSM in all four languages, we did find that at least one of the variables we controlled for was (minimally) approaching significance in all four languages of interest. In Hungarian, no variable reached significance, but the resolved/unresolved status of Q2 approached significance (p=.07). The status of Q2 significantly affects use of SynSM in German. The CT status of the subject and Register are both relevant in Hindi and German. We see, therefore, that our three variables were all relevant in multiple languages.

Our data also show that there are “default” constructions in each language. That is, there seems to one construction that is chosen most frequently and other constructions which are chosen when conditions are right. In English the default is LD. In Hungarian and Hindi it is SynSM. This is expected. However, in German, contrary to our expectation based on previous literature, we found that SeqQs were, in fact the most commonly chosen construction and that SynSM and LD were the most frequently chosen construction only in certain circumstances.

2.9.7 General discussion

The primary goals of this survey included the following: 1) Determine whether there was variation between languages in the constructions speakers chose to use in certain circumstances, 2) Determine whether SemSM surfaces as different syntactic constructions cross-linguistically, 3) Add to the findings from our pilot study by controlling for Register and determine whether Register affected the use of SeqQs in any language, 4) Directly test the DDA, by determining whether LD and SynSM are used with equal prominence in German, 5) Investigate the use of SeqQs, which had not been discussed to the same extent as the other constructions in previous literature. We discuss the findings relevant to these questions below.

1: Is there variation between languages in which construction is chosen under given semantics/pragmatics? There is variation in which construction is chosen as best for given semantics/pragmatics cross-linguistically. It is not surprising that languages like English and German should differ since one allows SynSM and one does not. However, even languages which allow SynSM vary in when it is the preferred construction. Let us take as an example the condition where Q2 is resolved, the subject is CT, and the Register is adult-directed. In English, LD is chosen most frequently in this condition. In German, participants chose SynSM and LD equally in this condition, while in Hindi and Hungarian, participants chose SynSM more frequently than the other constructions. However, in the contrasting condition (Q2 unresolved, subject is not CT, Register is child-directed), English-speakers chose SeqQs, LD, and the monoclausal Q2 with equal frequency. German-speakers chose SeqQs significantly more than any other construction. Hindi-speakers were at chance between SeqQs and SynSM and Hungarian speakers chose SynSM in every condition. In other words, the variation in conditions does not affect all languages in the same way. Unlike previous literature which has assumed that any language which includes SynSM as a possible construction should use it in the same way, this survey has shown that different contexts elicit SynSM to varying degrees cross-linguistically.

2: Does SemSM surface in different ways cross-linguistically? We chose to control for the resolved/unresolved status of Q2 based on the fact that SemSM should have two, contentful wh-phrases, meaning they should both be unresolved. Therefore, the answer to this question lies in whether different languages showed different constructions as optimal when Q2 was resolved vs. unresolved. While we have not seen a clear-cut distinction between when participants used constructions based on the resolved/unresolved status of Q2 alone, we have seen that it is a significant predictor of the use of SynSM in German, SeqQs in English, and SynSM (to a certain extent) in Hungarian. It does not affect SeqQs and SynSM differently in Hindi. We suggest that SemSM is tied to unresolved Q2 and our results show that unresolved Q2 indeed affects what type of construction is chosen by participants, and it does result in different constructions cross-linguistically.

3: Does Register affect the use of either SynSM or SeqQs? English-speaking participants from our pilot study suggested that SeqQs might be preferable to LD when they were addressing a child. We did not find that SeqQs were affected by Register in any language. However, the use of a complex construction (LD in English, WSM in German, Hindi, Hungarian) dropped when the Register was child-directed. However, it was unpredictable cross-linguistically what participants would choose instead: SeqQs or a monoclausal question (*What do you think?*) We suggest that, had participants not had the option to choose a monoclausal question, we might have seen a significant effect of Register on the use of SeqQs.

4: Are LD and SynSM used interchangeably in German? While SynSM and LD are used with similar frequency in some conditions, this is not always the case. There are conditions where LD is used more frequently and conditions where SynSM is used more frequently. Our data suggest that participants do not treat them as equivalent. It should also be noted that more than one participant said that LD was not grammatical to them[17]. We did have participants who never chose the LD construction, but all participants chose a SynSM construction at least once. In sum, even considering individual variation, LD and SynSM are not used interchangeably by German-speaking participants. Thus, analyses which treat them as such must be reexamined.

5: What is the role of SeqQs in expressing SemSM? SeqQs were a viable option in every language. One remarkable finding is that SeqQs were the option chosen most frequently overall by German participants. Based on previous literature, we had predicted that LD and SynSM would be the primary options chosen, but this was not the case. In Hindi, SynSM and SeqQs are close to equally dominant. In Hungarian, as well, SeqQs are the option preferred with second most frequency.

6: What semantic/pragmatic factors is SynSM sensitive to? The use of SynSM (in the languages which allow it) is significantly affected by CT in Hindi and German. The resolved/unresolved status of Q2 also significantly affects the use of SynSM in Hindi and German and approaches significance in Hungarian. Finally, Register affects the use of SynSM in Hindi and German. While no single variable affects the use of SynSM in all three languages who employ it, some combination of these variables is relevant to its use in all three.

[17] These participants indicated this in notes at the end of the survey where we asked for their own impressions of when they would use each type of construction. We did not eliminate these participants because we were primarily interested in WSM constructions. A future study should collect data from a larger pool of participants whose dialect includes both LD and WSM.

2.9.8 Conclusion

The two prominent analyses of WSM we started our discussion with (DDA, IDA) are grounded in intuitions surrounding both the grammaticality and limited aspects of the interpretation of WSM cross-linguistically. Overall, DDA accounts are grounded in syntactic patterns while IDA accounts are rooted in semantics. One main conclusion from the literature review in Sections 2.2 through 2.4 is that neither approach has cross-linguistic scope, and WSM largely remains a challenge, both semantically and syntactically. This led us to seek first a deeper understanding of the semantics of WSM by experimentally and cross-linguistically examining its use under a set of semantic/pragmatic factors which arguably affect the acceptability of complex questions, comparing WSM with other syntactic options licensed by the individual languages tested.

Our novel results end up posing a renewed challenge to both the DDA and the IDA, as briefly summarized here. On the one hand, the DDA makes the clear prediction that SynSM and LD constructions should have the same semantic interpretation. Our studies suggest not only that this does not hold cross-linguistically, but that within German, SynSM and LD are not always used with the same frequency. Furthermore, while SynSM and LD might be closer to interchangeable in German, in Hindi and Hungarian, it is SynSM and SeqQs which vie for the top position in each of our conditions, suggesting those constructions are more similar. On the other hand, the IDA would predict that SynSM and SeqQs should be the constructions which are interchangeable. Again, this is clearly not the case in German, but even in Hindi there is one condition where SynSM is chosen significantly more frequently than SeqQs. In addition, our results show that the use of SynSM is sensitive to the resolved/unresolved status of Q2 in Hungarian and German; CT status of the subject in Hindi and German; and Register in German and Hindi. Neither the DDA nor the IDA accounts for the variation observed.

In addition, our experimental investigation directly informs the next step of developing a unified account that is grounded in the idea that SemSM surfaces as different syntactic constructions cross-linguistically and that the semantics/pragmatics of a given situation can affect the syntactic expression of complex questions. A straightforward implementation is offered in Chapter 3 that exploits an input-output mapping view of the grammar and the claim that cross-linguistic variation in syntax is attributable to different rankings of a small set of universal but violable constraints regulating that mapping. These two properties lie at the core of Optimality Theory (OT) (Prince & Smolensky, [1993]2004; Legendre et al., 2001), to which we turn next.

Chapter 3: A unified account of Wh-Scope Marking, syntactic and semantic

Wh-scope marking (WSM) is a syntactic construction found in multiple languages including, though not limited to, German (McDaniel, 1989), Hindi (Dayal, 1994), and Hungarian (Horvath, 1995; 1997; 2000). WSM typically uses one interrogative pronoun in the matrix clause to mark the scope of the true wh-phrase found in an embedded clause (hence, "wh-scope marking") as seen below in (1) from German:

(1) Was glaubst Du, **mit wem** Maria gesprochen hat?
What think you, **with whom** Maria spoken has?
With whom do you think Maria spoke?

Ex. From Dayal (1994, 1b)

While *was* 'what' appears phrase initially, the question that needs an answer is *mit wem* 'with whom', which appears medially.

As discussed in detail in Chapter 2, there are two major accounts of WSM: the Direct Dependency Account (DDA) and the Indirect Dependency Account (IDA). The crucial difference between them is that the DDA suggests that the first wh-phrase (*was* in (49)) is an expletive element whose function is purely to mark the scope (the "Scope Marker", or SMer) for the true wh-phrase. Under this account, because the SMer has no meaning of its own, the semantic interpretation of a WSM sentence should be the same as if that sentence were a Long Distance (LD) construction. Formally, for a sentence such as (2), the proponents of the DDA (e.g. von Stechow, 2000) suggest its meaning corresponds to (3):

(2) Who does John think Mary will talk to ?

(3) $\lambda p \exists x[\text{person}(x) \wedge p = \text{John thinks that Mary will talk to } x]$

Informally, (3) has the meaning "who is the person (x) such that John thinks that Mary will talk to (x)". However, proponents of the IDA initially formulated in Dayal (1994) argue that this is not the case. Dayal proposes instead that both wh-phrases in a WSM construction are, in fact, genuinely contentful wh-phrases. She argues that the first wh-phrase is a question over propositions and the second limits the set of propositions to those relevant to the second wh-phrase. This is expressed formally in (4):

(4) $\lambda p \exists q[\exists x[q= \text{will-talk'} (m,x)] \;\&\; p=\text{think'}(j,q)]$

Informally, (4) has the meaning "there is a set of things that John thinks (p). There is a set of things within (p) relevant to whom Mary will speak with (q). In (q), x is the person Mary will talk to". Unlike German, English does not have WSM construction, and Dayal (2000) suggests that (4) would be expressed in English via Sequential Questions (SeqQs), as in (5):

(5) What does John think? Who will Mary talk to?

In Chapter 2, we concluded that neither IDA nor DDA accounted for the cross-linguistic variation found across English, German, Hindi, and Hungarian. We argued instead that these accounts were addressing two distinct phenomena: one semantic (SemSM), corresponding to the IDA as presented by Dayal (1994), and one syntactic (SynSM) in which the SMer is a true expletive, corresponding to the DDA as presented by McDaniel (1989). Following Dayal (2000), we concluded that SemSM is universally available and is realized in a variety of ways cross-linguistically: sometimes as SynSM as in Hindi, and sometimes as SeqQs, as in German. We discussed the results of a cross-linguistic survey showing that aspects of the semantics and pragmatics influence when native speakers choose SynSM and when they choose other structures. What remains to be investigated is how SemSM manifests in the languages of interest and what is formally responsible for the syntactic variation that has been empirically established in Chapter 2.

Our analysis will follow Optimality Theory (OT), which is a means of formalizing both a mapping between semantics and syntax (or input-output relation) and the resolution of conflicting constraints on linguistic systems (Prince & Smolensky, [1993]2004; Legendre et al., 2001). The input-output directionality of an OT analysis varies depending on what is being optimized -- the expression of a meaning or the meaning of an expression. In the former case, the input consists of a meaning and the candidate outputs are structural expressions (production-directed optimization) while in the latter, the input consists of a structural expression and the candidate outputs are interpretations (comprehension-directed optimization)[18]. Here, the focus is on syntax, hence our analysis will be a production-oriented one. The syntactic system is given an *input,* an intended semantic/pragmatic meaning, and evaluates multiple possible syntactic *outputs* or structural expressions of that meaning. Thus, the evaluation takes place at the semantics/syntax interface. The optimal output is determined by universal, violable constraints which are ranked by a language-specific hierarchy of importance. The optimal output varies cross-linguistically because of this variation in ranking. OT is an ideal means of formalizing the differences found in

[18] Either unidirectional optimization contrasts with alternative architectures under the umbrella of bidirectional OT, whereby candidate outputs consist of different expressions paired with distinct interpretations (Legendre, Putnam, de Swart, and Zaroukian, 2016.) Our interviews with native speakers have revealed multiple instances of neutralization (distinct semantic inputs resulting in a single syntactic output) which are incompatible with some bi-directional architectures, such as SuperOptimality (Blutner, 2000).

WSM as it allows for variation in what input yields a specific syntactic output: SemSM is universally available but it need not always result in SynSM[19].

It should be noted that this is not the first attempt to explain WSM via OT, Müller (1996) lays the groundwork we build upon below. However, his analysis of German WSM (incorrectly, based on our survey findings from Chapter 2) assumes that LD and SynSM should hold the exact same meaning and relies on a derivational version of OT, with optimization at every step in the derivation. Our analysis builds on subtle but real semantic and pragmatic differences between the two structures and is not derivational; optimization evaluates the global structure at once. We will show that the specific constraints operating on wh-scope marking and related syntactic structures relate to the vast generative literature on wh-questions and are for the most part independently motivated.

The analysis developed below crucially relies on the particular architecture of the syntax-semantics interface in OT motivated in Legendre, Wilson, Smolensky, Homer & Raymond (1995); Legendre, Smolensky, Wilson (1998); and Legendre (2010), which incorporates the following concepts: a "target" input containing scope information and discourse features as relevant, a candidate set that includes candidates that may not be faithful to input specifications, and input-output faithfulness constraints that, if violated, yield an output that is minimally unfaithful to a target input. This architecture contrasts, for example, with the one assumed in Grimshaw (1997), which restricts optimization to syntax, prior to standard semantic interpretation; all syntactic candidates share one and the same

[19] OT assumes that all related structures for a given target meaning allowed by UG are compared simultaneously and those which are not viable in a language are eliminated via constraint ranking. Children must, therefore, learn their language's relevant rankings. For a detailed discussion of the acquisition process in an OT framework, we turn the reader to Legendre et al. (2004).

interpretation; the candidate set includes only candidates which are faithful to the input specifications, there are no input-output faithfulness constraints in the evaluative component of the grammar. The main source of evidence in favor of the former architecture comes from ineffability (absolute ungrammaticality with no obvious 'repair' option) in syntax, one of several interpretation-form mismatches along with optionality and ambiguity (Beaver & Lee, 2004). For Grimshaw (1997), ineffability is the result of a construction that is simply uninterpretable.[20] Legendre (2010) counters this argument with the point that there is cross-linguistic variability in whether constructions are ineffable. For instance, *Who ate what?* is ungrammatical in Italian and Irish, though it is viable in English. To say that the Italian winner [Who ate what?] crashes at interpretation, while its English counterpart does not, entails that the competitors are not all interpretively equivalent after all. In this chapter, along, for example, with Bakovic & Keer (2001) and Wilson (2001), we follow Legendre et al. (1995, 1998) and show that our empirical findings cannot be modeled without the crucial use of semantic information in the input and input-output faithfulness constraints. Our findings are, thus, strong support for this specific theoretical approach to OT syntax and the syntax-semantics interface.

The chapter is structured as follows. First, we describe in further detail the input relevant to asking a question with multiple clauses (section 3.1) and then the syntactic structures manifested in English, German, Hindi, and Hungarian (section 3.2). In section 3.3, we launch our cross-linguistic analysis with a summary of the input-output pairings established in Chapter 2 (3.3.2) and the constraints relevant to our OT analysis (3.3.3). In

[20] Pesetsky (1997) takes ineffability to be evidence that OT is not suitable for syntax and should be restricted to the phonological spell out of syntactic structures. In other words, he assumes that input-output faithfulness constraints play no role in syntax.

(3.3.4), we discuss the universal candidate set and apply the evaluative procedure of Harmonic Bounding to reduce it to the active candidate set (containing only those candidates which can potentially be optimal cross-linguistically). Finally, in 3.4, we use the OT evaluation procedure to model the optimal input-output pairings in each target language. Overall, the variation in rankings of the constraints we have identified is able to account for the cross-linguistic variation seen in how SemSM manifests.

Because the chapter is designed as self-standing some repetition of assumptions and conclusions reached in Chapter 2 is necessary. The reader may want to skip sections 3.1-3.1.4 and turn to section 3.1.5, pg. 80.

3.1 Aspects of the Semantic/Pragmatic Input to Syntactic Optimization

Understanding the precise inputs relevant to syntactic optimization in multi-clausal information questions is crucial to our analysis and requires a detailed understanding of SemSM beyond previous semantic analyses. In Chapter 2, we reported on a survey conducted among native speakers of English, German, Hindi, and Hungarian in which we controlled for various aspects of the semantic input. These included speaker Register (adult-vs. child-directed), Question Under Discussion (QUD), and Contrastive Topic (CT). Below we will briefly discuss these aspects of the input again and refer the reader to Chapter 2 for in-depth discussions of each.

During the process of examining judgments from native speakers, it became clear that in addition to the aspects of the semantics we had controlled for in Chapter 2, there is also a need to differentiate where the matrix verb has scope over the embedded verb and where it does not. In some languages, like English, the relative scope of verbs involved results in different syntactic structures. For instance, if the matrix verb scopes over the embedded verb at LF, SeqQs are not representative of that relationship. We suggest that the scope represented in the syntactic output should match the scope of the verbs at LF. Because we are interested in biclausal questions, this analysis focuses on the scenario where V1 scopes over (>) V2. When V1>V2 we manipulate both QUD and CT, and map them to a variety of syntactic outputs.

As the goal is to identify multiple input to output pairings cross linguistically, the following sections will describe the different potential semantic/pragmatic inputs in terms of QUD and CT to the extent that they are relevant to the analysis. Section 3.1.2 will focus on QUD and section 3.1.3 will focus on CT. Section 3.1.4 describes speaker Register and section 3.1.5 will briefly address the role of relative Scope of the verbs. These aspects of the input are not exhaustive: rather they are the ones we have identified as relevant to WSM productions. Essentially, we suggest that the relevant semantic/pragmatic input specifies four crucial items: the set of unresolved QUDs, information about CT, Register, and structural information such as the relevant Scope of V1 and V2. The rest of the input to syntactic optimization consists of lexical info, predicate-argument structure, illocutionary features, and familiar functional features such as tense, per Legendre (2001).

3.1.2 Aspects of the input: QUD

The primary relevant aspect of the input is the QUD (for a review, see Roberts, 2012). The QUD model resembles a stack. There is an underlying QUD to every exchange, with smaller QUDs addressing subtopics. The following exchange is felicitous because both A and A' address one of the underlying QUDs.

Q: Should we eat shrimp or beefsteak tonight?

A: Beefsteak is generally too expensive

A': But it is on sale today.

While the base QUD asks *what we should eat,* underlying questions include things like *what do people prefer? What food do we have? What can we afford to buy?* Thus, a response to an underlying QUD can be a felicitous response to the base QUD, even when that response appears irrelevant, because it addresses an underlying QUD.

It is also the case that the relevant QUDs can be *resolved* or *unresolved* in the discussion. For instance, after the exchange above, if someone were to say *Beefsteak is too expensive*, it would be strange considering the question of cost has already been addressed. That particular sub-QUD was resolved. This resolved/unresolved status affects the felicity of the following utterances. Let us return to questions with multiple clauses and an example from Chapter 2. Consider a question such as (6). Relevant QUDs include, but are not limited to, (7) and (8). The "embedded" question, *What is this?,* (and the propositions it denotes) can be either resolved or unresolved at any point of the exchange.

(6) What do you think this is?

QUDs which must be present to make the use of (6) reasonable include, but are not limited to:

(7) What are your thoughts, in particular, on what this is?

(8) What is this?

No one would ask this question unless (7) were unresolved, but it *not* necessary that (8), the second QUD (Q2), be unresolved. For example, consider the following scenario (9), where the relevant QUDs are (7) and (8).

(9) Carol and Cathy are paleontologists looking for valuable fossils at a collectors' fair. They see one table with larger rocks. Carol immediately recognizes that one of the rocks is, in fact, a dinosaur egg. After examining the rock, Cathy agrees. This is an incredibly rare find. Wondering if the collector knows what a find this is, Carol turns to him and asks, "What do you think this is?"

In this scenario, for the speaker, (8) is resolved because Carol knows what the rock in fact is. However, (7) is still unresolved because Carol does not know what the rock collector's thoughts are on the finding. Thus, *what do you think this is* remains a felicitous question.

We have seen, therefore, that a question with multiple clauses can still be felicitous even when Q2 is resolved. However, as shown in Chapter 2, the resolved/unresolved nature of Q2 can affect which structures can be used. This is the crucial manipulation we argue is relevant to SemSM: when both QUDs are unresolved, they represent two, contentful questions over propositions (Dayal, 1994)). However, when only one QUD is unresolved, the question is closer akin to what has been proposed for the meaning of an LD structure (e.g. von Stechow, 2000).

3.1.3 Aspects of the input: Contrastive Topic (CT)

The second major aspect of the semantics which we propose to include in the input to syntactic optimization is what part of the sentence is a *Contrastive Topic* (CT) (Büring, 2003). We follow Wagner (2012) in considering CT to be a form of focus which must have the highest scope in the structure in order for the utterance to be felicitous. The expression of CT has been shown empirically to affect the grammaticality of German statements (Wagner, 2012; Büring, 1997b) and we suggest that this affects the grammaticality of synSM structures as well. We use the term *Topic* to mean the context of the discourse ('what the dialogue is about'), following Reinhart (1981). Topics can be construed as responses to QUDs: they are brought up in the discourse because they play a role in responding to the QUD. *Contrast* is a type of emphasis which is placed on an entity which suggests there is (at least) one other entity which could have been mentioned (Repp, 2009). The contrasted entity is one of several entities relevant to the topic and is mentioned in contrast to the others. This emphasis can be indicated either by prosody or by syntactic movement or both, depending on the language. Büring (2003) suggests that the manipulation of CT and focus can render an utterance felicitous or infelicitous, independently of QUD. Consider the following example, repeated from Chapter 2, summarized from Büring (2003).

(9) QUD: What did {Mary, Fred, George} eat?

A: Fred_{CT} ate the $\text{Beans}_{\text{focus}}$

A * $\text{Fred}_{\text{Focus}}$ ate the Beans_{CT}

In contrast:

(10) QUD: Who ate the {squash, beans, carrots}?

A: *Fred_{CT} ate the $\text{Beans}_{\text{focus}}$

A: $\text{Fred}_{\text{Focus}}$ ate the Beans_{CT}

Büring (1997b) and Wagner (2012) argue that in German the entity which receives CT must have higher scope than any other focused entity. For instance, though (12) represents the canonical word order, (11) would be the grammatical construction in scenario (10), though not (9).

(11) Die Bohnen$_{CT}$ hat Fred$_{Focus}$ gegessen
The Beans$_{CT}$ has Fred$_{Focus}$ eaten

(12) Fred hat die Bohnen gegessen Examples from Wagner (2012; 14)

Wagner (2012) considers CT to actually be a form of focus consisting of nested focus features. The focused element with widest scope must be the one intended as the CT. For example, if Karen's thoughts are being elicited in contrast to others', as seen below, Karen is the CT:

(13) Three friends are deciding where to go to a movie. Margaret and Sue have different opinions so they will let Karen make the decision.
Where does Karen$_{CT}$ (rather than Margaret or Sue) think we should go?

The question here is less about where the friends should go and more about Karen's thoughts on the subject, in contrast to Margaret and Sue's. In English, CT is generally indicated with intonation, but in German, as indicated by Büring (1997b) and Wagner (2012), it must also be indicated by word order, thus it is crucial to our analysis of German.

3.1.4 Aspects of the input: Register

Another aspect of the input we will consider is Register. *Register* refers to patterns of languages that are specific to certain situations (Ure & Ellis, 1977). Adult speakers adapt their register to their interlocutor and use different lexical items, constructions, and prosodic contours depending on the person with whom they are speaking. For our purposes, one linguistic aspect which changes when using a child register is the complexity of the syntactic construction, according to Shatz & Gelman (1973). Considering the fact that we are interested in complex constructions, it is reasonable to assume that an adult addressing a child might prefer sequential, monoclausal questions as opposed to structures involving long distance movement or an SMer.

3.1.5 Aspects of the Input: Verbal Scope

The final aspect of the input which must be introduced is the relevant scope of the verbs. In LD and WSM structures, which are single utterances with multiple clauses, the matrix verb has scope over the embedded clause verb. However, in sequential questions this is not the case. Thus in (14), *think* has scope over *win,* while in (15) this is not so:

(14) Who does he think will win the race?

(15) What does he think? Who will win the race?

This is perhaps most clear when we consider the types of responses to these questions which are felicitous, respectively (16) and (17):

(16) He thinks number 9 will win.

(17) Well, he thinks number 8 is the better horse, but doesn't like to run in the rain like 9, so given the weather, 9 will probably win.

In (14), *think* has scope over *win*, thus the response is limited specifically to who he thinks will win. (15), on the other hand, allows the response to include other thoughts related to who will win. Thus, when the semantic/pragmatic input to a syntactic optimization specifies that the matrix verb should scope over the embedded verb, SeqQs do not accurately reflect that in the output. The trade-off is typically with Register. SeqQs are preferred in the Child Register at the cost of violating Verbal Scope while LD or WSM satisfy Verbal Scope at the cost of violating Child Register.

3.1.6 Summary of the input to optimization

In summary, four input variables, QUD, CT, Register and the appropriate scope of the verb yield sixteen conditions relevant to the analysis. However, we did not control for verbal scope and only consider the situations where V1>V2. Thus, the eight relevant inputs are shown in Table 1.

Table 1. *Relevant Semantic Input Conditions*

	Register	Q2 resolved/Unresolved	Subject +/– CT	Verbal scope
1	Adult	Unresolved	+CT	V1 scopes over V2
2	Child	Unresolved	+CT	V1 scopes over V2
3	Adult	Unresolved	–CT	V1 scopes over V2
4	Child	Unresolved	–CT	V1 scopes over V2
5	Adult	Resolved	+CT	V1 scopes over V2
6	Child	Resolved	+CT	V1 scopes over V2
7	Adult	Resolved	–CT	V1 scopes over V2
8	Child	Resolved	–CT	V1 scopes over V2

These eight semantic/pragmatic conditions constitute different OT inputs to optimization and render different optimal syntactic outputs in different languages. The syntactic strategies employed by each language are described next.

3.2 Syntactic Strategies

We focus on three main syntactic strategies that are tied to the conditions/inputs to optimization discussed above. These include LD, SynSM, and sequential questions (SeqQs). It should be noted that SeqQs have only been seriously considered as instantiations of SemSM in semantic analyses (e.g. Dayal, 2000). Some languages, including some dialects of German, also allow Copy Constructions which resemble SynSM, except that they employ the same wh-phrase twice (e.g., Felser, 2004). In an OT analysis, all possible structures should be considered. However, an analysis of the relationship between the semantic input and syntactic output is not possible without a thorough specification of the semantic input. There is little consensus on either the semantic or syntactic analyses of Copy Constructions, which is likely due to a large amount of dialectal variation in German (Murphy, 2016). Thus, we leave Copy Constructions to further research and future discussion.

We assume a standard wh-movement analysis for SeqQs. They are simply two, monoclausal questions, linked by the discourse. A wh-phrase is a [+wh, +Focus] DP which moves to specCP (though not in all languages, see discussion of Hungarian, below) to check these features housed in C (assumed to encode interrogative illocutionary force). For LD structures, we assume a traditional cyclic movement analysis. This follows the same principle as wh-movement in monoclausal questions except that the wh-phrase stops at intermediate specCPs, leaving a coindexed trace.

The analysis of SynSM requires further specification. Building on Fanselow & Mahajan (2000), and Horvath (1997, 2000) we adopt a syntactic analysis in which the SMer is construed as an expletive element co-occurring with a CP associate. This expletive has the function of the direct object of the matrix verb and is base-generated in the complement of vP, in the specifier position of AgrOP[21]. The projections depicted in Fig. 1. are in keeping with current Minimalist Program hypotheses about syntactic structure and are crucial to our analysis. In German and other WSM languages that allow wh-fronting, the SMer moves from the specifier position of AgrOP to specCP. In Hungarian, the wh-phrase will instead move to the spec of FocP, located directly under the matrix CP (Bhatt & Yoon, 1992). In Hindi, the SMer remains in situ like all wh-phrases. Fig. 1 illustrates a generic, non-language specific, structure in which the contentful wh-phrase is the object of the embedded clause verb. Of course, this structure varies based on language specifications such as head-directionality and wh-strategy.

[21] Under a Larsonian VP shell analysis of double object constructions (Larson, 1988), the SMer object could alternatively be base-generated in the specifier position of a lower VP while the sister of V position is filled with the embedded CP. The SMer would then move to specAgrOP to check Accusative Case.

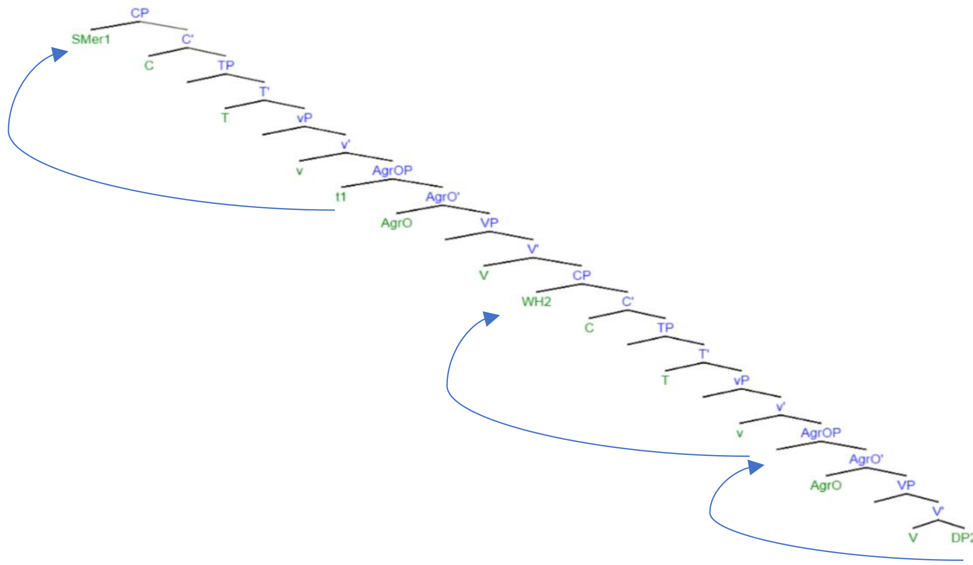


Fig. 1. Syntactic representation of SynSM equivalent to the following in bracket form:

$[_{CP\ +Q}\ \text{SMer}_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ t_1 [_{VP} \ldots [\![_{CP}\ \text{wh}_2\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_2\ ]]]]]\!]]]]]]?$

3.3 Cross Linguistic Analyses

Chapter 2 described the results of a survey conducted with speakers of English, German, Hindi, and Hungarian. This study added cross-linguistic scope to our understanding of the nature of WSM. We found in all languages except Hungarian that the pragmatic factors we identified (Register, the resolved/unresolved status of Q2, and whether the subject is CT) significantly affected which construction participants chose as optimal. We also found that in some cases, there was a significant difference between constructions within a condition. However, the difference did not always reach a level of statistical significance. This is probably because once responses had been divided by response types, there were relatively few items of each type. For this reason, the findings described here should be considered preliminary as a study with more participants will have to be conducted in order to determine whether the differences reported are, in fact, significant. For the purposes of the current analysis, therefore, the construction which appeared most frequently was considered optimal.

We also investigated the possibility of outliers in the participants. The survey was a forced-choice task and participants were asked to choose the best sentence. Options included SynSM structures, LD structures, SeqQs, as well as monoclausal questions. These monoclausal questions were solely included to verify whether the scenarios accurately prompted the use of biclausal structures, which they did. However, while most participants chose biclausal structures or SeqQs over 60% of the time, we found that there were some participants in every language studied that chose a monoclausal question over 50% of the time (n=5 across all languages). Because the OT analysis below focuses on the optimality of biclausal questions vs. sequential questions, these participants were eliminated from further investigation. It should be noted that, in spite of this, a monoclausal question does surface as optimal in one condition in Hungarian. This speaks to the fact that this construction is significantly preferable to the others in this condition even among speakers who otherwise always mostly choose biclausal constructions.

3.3.2 Summary of the input-output pairings

We have argued that four variables in the semantic/pragmatic input are relevant to questions with multiple clauses and that individual languages vary with respect to the syntactic strategies they allow. We have also discussed how each semantic input maps onto these various syntactic structures. Table 2, below, is a summary of native speaker judgments described in the previous sections when V1 > V2. In Hungarian there is one condition where the monoclausal embedded question (monoclausal Q2) is the winner, to be discussed when we discuss Hungarian. For all other scenarios either LD, SynSM or SeqQs were chosen as optimal. The reader will also note that in German, SeqQs surface as the optimal construction in most conditions. This was somewhat surprising given that the alternation we expected based on previous literature was between SynSM and LD.

Table 2. *Intra- and cross-linguistic variation in syntactic expression for a given input (R =resolved, -R = unresolved)*

Input	Q2 -R Sub CT; Adult	Q2 -R Sub CT; Child	Q2 -R Sub not CT; Adult	Q2 -R Sub not CT; Child	Q2 R Sub CT; Adult	Q2 R Sub CT; Child	Q2 R Sub not CT; Adult	Q2 R Sub not CT; Child
English	LD	LD	LD	LD	LD	LD	LD	LD
German	SeqQ	SeqQ	SynSM	SeqQ	LD	SeqQ	SynSM	SeqQ
Hindi	SynSM	SeqQ	SynSM	SynSM	SynSM	SeqQ	SynSM	SynSM
Hungarian	SynSM	Mono-clausal Q2	SynSM	SynSM	SynSM	SynSM	SynSM	SynSM

3.3.3 The relevant constraints

OT assumes universal constraints on the well-formedness of syntactic representations (or markedness constraints), as well as Input-Output faithfulness constraints requiring the content of the input to be accurately represented by the output. Constraints are frequently in direct conflict with one another. This is why variation in rankings results in variation in outputs. In what follows we describe the various constraints relevant to our OT analysis as well as the conflicts which arise between them. We start with Markedness constraints all of which were originally proposed in independently motivated analyses; for present purposes they may undergo some minimal adjustments to make them relevant to WSM structures.

Markedness constraints evaluate the inherent goodness of candidate outputs and include here:

a) An Economy of Structure constraint:

> **MINPROJ** (Grimshaw 1993, 1997, a type of *STRUCTURE of Prince & Smolensky, [1993] 2004): Syntactic projections should be minimal.

Any projection incurs a violation, meaning that structures with fewer projections are favored. This constraint eliminates adding syntactic structure for the sake of not violating a constraint (as is standard in the Minimalist Program, Chomsky 1981, 1995).

b) Three constraints pertaining to wh-movement (which, apart from the one involving Economy of Movement, have counterparts in Minimalist analyses):

> ***t** (Legendre et al, 1995, 1998; also called STAY, Grimshaw, 1997): No syntactic movement.
>
> **WH-CRIT(ERION)** (Müller, 1996; see also *Q in Legendre et al., 1995 and OpSpec in Grimshaw 1993): Wh-items surface in a specCP whose head is +Q. Its Minimalist counterpart is the wh-criterion (Rizzi, 1990).
>
> **MINLINK** (Legendre et al. 1995, 1998; Legendre, 2010): Links between traces should be minimal. Longer links are less harmonic than shorter links. Its Minimalist

counterpart is either the Minimal Link Condition (Chomsky, 1995) or Relativized Minimality (Rizzi, 1990).

One of the most widely attested and simplest markedness constraints is *t. This is a general Economy of Movement constraint. Any trace is a violation. Items with an overt copy of a trace incur an additional violation, see further discussion below. *t is in direct conflict, however, with WH-CRIT which requires wh-operators to surface in SpecCP. Any empty +Q specCP is a violation. Furthermore, any wh-phrase which is *not* in a +Q CP incurs a violation. This includes embedded CPs, which are typically –Q (Carnie, 2013). WSM structures as well as in-situ structures thus incur violations of WH-CRIT.

MINLINK is violated when a link of a movement chain crosses a barrier, which is defined in Legendre et al. (1995, 1998) as a phrasal category which is not theta-marked by a lexical head like V, based on Chomsky (1986). Under this definition, barriers include TP, vP, and VP, AgrOP, and FocP (where relevant) for the purpose of this analysis. MINLINK is also violated by 'covert movement', i.e., when the link between Q operators and wh-phrases which remain in-situ crosses a barrier. The role of MINLINK in distinguishing SynSM constructions from LD constructions should be highlighted. In LD constructions, movement of the wh-phrase from the embedded clause CP to the matrix CP will cross four barriers (violating BAR4). In SynSM, because we construe the SMer to originate in AgrOP, it crosses only three barriers (violating BAR3). This difference will be crucial.

MINLINK is generally presented as a subhierarchy of constraints: BAR1, BAR2, BAR3, such that BAR1 is violated when a single barrier is crossed, BAR2 when two barriers are crossed etc. These BAR constraints allow other constraints to be interspersed whose violation may be decisive, but those involving more barrier crossings are universally ranked higher than those involving fewer barrier crossings: BAR3≫ BAR2≫BAR1, no matter what other constraints intercede.[22] While some of our rankings require the splitting of the subconstraints, this is not always the case. Thus, in those situations where it is unnecessary, we subsume BAR1, BAR2, BAR3, etc. under MINLINK with a violation per barrier crossed.

To further understand the interaction of these three constraints, consider a simple example external to WSM, based on work from Legendre et al. (1995). Tableau 1 illustrates why English does not typically allow wh-in situ questions as information questions. Even though candidate (b) violates multiple constraints while candidate (a) only violates one, because WH-CRIT is highly ranked in English, it forces wh-movement. Note that in this simple example the three BAR constraints are displayed contiguously, effectively functioning as one monolithic MINLINK constraint. In this situation, a violation is incurred for every barrier crossed, thus in Tableau 1, candidate (b) incurs three violations for crossing 3 barriers. In all tableaux the optimal candidate is indicated with ☞; fatal violations are recorded as *!

[22] Wh-movement cases where Faithfulness constraints must be interspersed among BAR constraints are discussed in Legendre et al. (1995, 1998) and Legendre (2010).

.

Tableau 1. *Canonical English wh-questions*

Input: like (*he, what*), present tense	WH-CRIT	MINLINK	*t	FULL INT
(a) [$_{CP\ +Qi}$ [$_{TP}$ he [$_{vP}$ likes [$_{AgrOP}$ what$_i$]]]] (wh-in situ)	*!			
☞ (b) [$_{CP\ +Qi}$What$_i$ does [$_{TP}$ he [$_{vP}$ like [$_{AgrOP}$ t$_i$]]]] (wh-fronting)		***	**	*

In Tableau 1, we see that candidate (b) wins in spite of incurring many more violations overall than candidate (a). In particular, the use of *do*-support (an expletive form of *do* base-generated in T) incurs a further violation of *t upon moving to C as well as a violation of FULLINT (further discussed below) that (a) does not incur, but because WH-CRIT is highly ranked, (b) is still the optimal output. *t is low ranked in English, rendering syntactic movement ubiquitous, and WH-CRIT is of primary importance in the absence of other constraints at work.

The next constraint is a novel one we propose to further specify how wh-movement varies among languages. Hungarian has been claimed to have 'split' the function of C into pure subordination and a marker of clause type (Focus) (Bhatt & Yoon, 1992). Wh-phrases in Hungarian are typically analyzed as moving to FocP just below CP, as shown in indirect questions (see example 18 below). This focus type movement also incurs a violation of WH-CRIT as the WH-phrase is not in specCP. Thus, any analysis which includes Hungarian (and other languages that assimilate wh-movement to focus movement) requires the use of

c) A constraint on movement to FocP

MARKFOCUS: Syntactic phrases with a focus feature appear in FocP.

MARKFOCUS[23], which requires elements containing a focus feature be in the designated focus position, FocP, is in conflict with WH-CRIT, which requires Wh-phrases to be in SpecCP (the highest scopal position in a clause). MARKFOCUS is also in conflict with MINPROJ (Economy of Structure) which disfavors additional projections like FocP. All of the candidates relevant to this discussion actually have the same number of projections with the exception of those that include FocP. Thus, for the purposes of clarity we will only indicate this violation which is in excess of the other candidates. We present (18) as evidence in support of positing a separate FocP projection in Hungarian that houses wh-phrases (18a) (from Horvath, 1998) or focused DPs (18b), where bold indicates focus (Kiss, 1998). Note that the embedded wh-phrase or the focused DP follows the embedded complementizers *hogy* ('that', in 18a) and *ha* ('if', in 18b), supporting a sequence of CP-FocP-TP projections at the left clausal periphery.

(18) a. Kérdezték, hogy **kit** hivott fel Mari
Asked -3pl that **who-acc** called up Mary-nom
They asked who Mary had called up
ex. From Horvath (1998, 2)

b. Szeretnem ha **Peterrei** szavaznatok
would.like-1sg if **Peter.on** voted.you
I wish it was ***Peter*** *on whom you voted*
ex. From Kiss (1998, 28a)

The following tableau shows how the simple addition of MARKFOC allows for the Hungarian option to become optimal. For simplicity's sake, we illustrate only the

[23] MARKFOCUS is tenuously related to Choi (1996)'s prosodic [+N] constraint which requires new information be indicated with a pitch accent. Our constraint is a syntactic one which requires that the focused item be marked syntactically.

embedded clause ('who Mary called') and we only represent the path of the object wh-phrase which transits in specAgrOP to check Accusative Case.

Tableau **2. Illustration of the role of MARKFOC**

Input: called up (Mary, who)), that, past	MARKFOC	MINLINK	WH-CRIT	*t	FULL INT
(a) …⟦$_{CP}$ hogy[$_{TP}$ hivott fel [$_{vP}$ Mari [$_{AgrOP}$ kit …]]]]. (wh-in situ)	*!		*		
(b) …⟦$_{CP}$ Kit$_i$ [$_{TP}$ hivott fel [$_{vP}$ Mari [$_{AgrOP}$ t$_i$ …]]]] (wh-fronting to specCP)	*!	***		***	
☞ (c)…⟦$_{CP}$ hogy [$_{FocP}$ kit$_i$ [$_{TP}$ hivott fel [$_{vP}$ Mari [$_{AgrOP}$ t$_i$ …]]]]] (wh-fronting to specFocP)		***	*	***	

The five constraints discussed above constitute all the Markedness constraints required for our analysis. They solely evaluate the candidate outputs. For a complete OT analysis, we require Faithfulness constraints as well. Faithfulness constraints, in contrast to Markedness constraints, evaluate the relationship between the input and candidate outputs. They require that all input elements be realized in the output and that all elements present in the output be present in the input. These constraints interact with the Markedness constraints described above and determine the optimal output. They include:

d) A constraint disfavoring expletive elements (such as *do* support, non-referential *it*, etc.)

FULLINT(ERPRETATION) (Grimshaw, 1997): A lexical item has an interpretation at LF.

e) Four novel constraints evaluating the expression of relevant semantic/pragmatic properties (QUD, CT, and Verb Scope)

FAITH QUD: The number of true wh-phrases ([+wh] lexical items that carry content) reflects the number of unresolved QUDs, specified in the input.

FAITH CT: The item marked as CT in the input has highest focus scope in the output.

FAITH LF/SYN: The output reflects the relative scope of verbs in the input.

FAITH C(HILD) REG(ISTER): be simple

Expletives (by virtue of being contentless) are not present in the input. They only appear in the output to satisfy a high-ranked Markedness constraint. They themselves violate FULL INTERPRETATION, a principle originally proposed in Chomsky (1991). Any use of SMers incurs a violation.

Faith QUD is derived from PARSE (any element of the input has a counterpart in the output) and FILL (any element of the output has a counterpart in the input) from Prince and Smolensky ([1993] 2004). It requires that the specified number of relevant unresolved QUDs in the input be accurately represented via contentful wh-phrases in the output. A violation is incurred for the lack of [+wh] as well as for the lack of relevant content carried. For instance, in a question with a single contentful wh-phrase, but with two unresolved QUDs specified in the input, two violations are incurred: one for the lack of a wh-phrase and one for the lack of relevant content. Conversely, a violation is incurred for excessive contentful wh-phrases. If only one unresolved QUD is specified in the input, but two contentful wh-phrases are present in the output, two violations are incurred: one for the extra wh-phrase, and one for the extra content. As an expletive, an SMer is [+wh], but not contentful. A SynSM structure would incur a single violation in either of the cases above. Table 3 provides a summary. It should be reiterated here that there are many unresolved QUDs for every utterance. We are only concerned with those relevant to the matrix and embedded clauses, as described in section 3.1.2 and will therefore only consider either one or two unresolved QUDs.

Table 3. *Violations of FAITH QUD*

Input	Output	Violations incurred	Explanation
1 unresolved QUD	No wh-phrases	2	If no wh-phrases are used, there would be neither a [+wh] element nor a representation of the content carried. Thus, two violations.
1 unresolved QUD	LD	0	QUD and contentful wh-phrases match. No violations.
1 unresolved QUD	SynSM	1	Extra [+wh] item incurs one violation.
1 unresolved QUD	SeqQs	2	The extra wh-phrase incurs one violation for the extra [+wh] and one for the extra content.
2 unresolved QUDs	No wh-phrases	4	If no wh-phrases are used, there would be two [+wh] items missing and two instances of content missing. Thus, four violations.

2 unresolved QUDs	LD	2	Only one contentful wh-phrase. Violations incurred for both the extra [+wh] and the extra content.
2 unresolved QUDs	SynSM	1	The SMer satisfies the need for a second wh-phrase, but not the need for the content of the second QUD.
2 unresolved QUDs	SeqQs	0	Two QUDs, two contentful wh-phrases. No violations.

Faith QUD is also relevant to structures which are monoclausal. The content of the wh-phrase should reflect the relevant unresolved QUD. Consider (19). The relevant QUDs include Q1: what are John's thoughts on who won (regardless of who won) and Q2: who in fact won the race.

(19) Who does John think won the race?

If both QUDs are specified as unresolved in the input, then a monoclausal Q2 output (Who won the race?) reflects the content of the second QUD, but not the first and thus incurs 2 violations (one for lack of content, one for lack of [+wh]). However, if Q2 is resolved (we know who won the race) and there is only one unresolved QUD (namely the one inquiring after John's thoughts on who won) a monoclausal Q2 output not only fails to represent the relevant unresolved QUD, it presents an irrelevant question. Thus, a monoclausal Q2 output incurs four violations of Faith QUD (one for lack of relevant content, one for lack of relevant [+wh], one for irrelevant content, one for irrelevant [+wh]).

FAITH CT ensures that the output accurately represents whichever item is CT. Recall that Wagner (2012) argues that CT is the associate of a focus operator taking wider scope over a lower focus operator. Büring (2003) also indicates that CT is generally established first. Recall also that we are analyzing SynSM as an expletive with a CP associate (following Mahajan[24], 2000; see Fig.1): thus the SMer is a focused item (a wh-phrase) with an associate (the embedded clause) and the SMer takes scope over the matrix subject. This means that a SynSM construction should not allow for the matrix clause subject to be CT because it does not have the widest scope: the SMer does. Thus, when the subject of the sentence is CT, the use of an SMer in the matrix clause will incur a violation of FAITH CT because the SMer will then be the focused element with the highest scope.

FAITH LF/SYN concerns the interpretation of the relative scope of the two verbs involved. If the intended meaning is that the matrix verb has scope over the embedded clause verb at LF, then both verbs should be represented and the relative scope must be represented in the syntax. When the matrix verb (V1) should scope over the embedded verb (V2), this constraint is violated once by SeqQ and twice by monoclausal questions (once

[24] Mahajan (2000) suggests that the embedded CP is the associate of the SMer and is interpreted as having scope at LF. Among other things, Mahajan argues, this analysis explains more completely why WSM cannot be used with verbs such as *ask/wonder*. Under a standard DDA analysis, this is explained as the result of the wh-phrase itself being interpreted in the matrix CP position, thus leaving a verb which requires a [+wh] complement unsatisfied. However, in sentences such as (i) (13 in Mahajan, 2000) which include multiple wh-phrases, this requirement could be satisfied by the second wh-phrase.

(i) *Siitaa-ne mohan-se kyaa puuchaa ki kis-ne kis-ko dekhaa?
S.$_{erg}$ M.-from SMer ask that who$_{erg}$ who saw
Meaning either: *Who did Sita ask Mohan that who saw?*
Or: *Who did Sita ask Mohan that saw whom?*

Therefore, an analysis that assumes the embedded clause wh-phrase alone is the associate of the SMer incorrectly predicts this sentence should be grammatical. However, if the entire CP is interpreted in the matrix position, *puuchaa* will again be left with no [+wh] complement and the sentence will be rendered ungrammatical.

because both verbs are not present, once because the scope is not represented syntactically). This constraint is satisfied by both SynSM and LD.

Finally, FAITH C. REG simply states that, when the addressee is a child, the speaker should use the simplest construction. For our purposes, this has to do with the competition between complex constructions (LD, SynSM) and simpler constructions (SeqQs and monoclausal Qs since neither involves long distance movement). Use of LD or SynSM violate CHILD REGISTER.

3.3.4 Trimming the candidate set

Having defined the relevant constraints, let us now apply those constraints to all possible syntactic representations of relevance. From there, we will eliminate candidates which can never win due to *harmonic bounding* (Prince & Smolensky, [1993]2004; Samek-Lodovici & Prince, 2002), an evaluative procedure which eliminates candidates which can never be optimal under any ranking given these constraints. The set generated from relevant inputs consists of 82 candidates if all possible placements of FocP in one or both clauses are assumed. We eliminate all candidates which would always lose to their equivalent with no FocP because they incur additional violations of MINPROJ. This consists for example, of candidates which included FocP in the matrix clause, but not the embedded one. This reduces the candidate set to 25. We further remove any candidate which violates BAR5. As we will see below, an optimal output can violate up to BAR4 (e.g. Hungarian). The only circumstance we can think of where BAR5 (or more) would be violated would be if the output of a question specified as biclausal in the input included an additional clause and a wh-phrase did not move cyclically (moved from an embedded clause directly to the matrix SpecCP without stopping at the intermediate CP). We believe this is not possible. Hence, we treat BAR5+ as undominated and eliminate any candidate that violates it. This reduces our candidate set to 21.

These 21 candidates are then further reduced to 10 by considering further constraints and further applications of *harmonic bounding.* (Samek-Lodovici & Prince, 2002; Riggle, 2006,2009). This can be done one of two ways. In the first method, a candidate is eliminated if it has the same violations as another candidate plus at least one more. In the second method, given a set of 3 candidates {a,b,c} if *a* will always lose to either *b* or *c,* then *a* can also be eliminated. We begin by eliminating candidates using the first method and return to the second method later. Both methods amount to the same thing: limiting the candidates in the tableaux to those that can win under some condition and eliminating those which cannot.

As in any OT analysis, the exact definition of what constitutes a violation must be thoroughly understood in order to comprehend these tables. The violations are not always transparent. For instance, consider the Faithfulness constraints, the violations of which of course will change according to the relevant input. For the purposes of Harmonic Bounding, we consider the situations in which all constraints are active.

For the LD construction, illustrated in Table 4, we consider the input where V1> V2, so no LD structure incurs a violation of FAITHLF/SYN. There are two unresolved QUDs, so every candidate incurs at least two violations because FAITHQUD requires one contentful wh-phrase per unresolved QUD and these candidates only have one contentful wh-phrase. Candidates 4 and 5 incur additional violations for including the additional copies of the first wh-phrase. Even if we consider this copy to carry no content (which is debatable[25]), they will incur an extra violation in FAITHQUD. FAITHCT is satisfied in LD candidates and all candidates violate FAITHCREG equally. In our analysis, it is crucial that WH-CRIT be violated *both* when a wh-phrase appears outside of a CP with a +Q feature *and* when a CP+Q is empty. For example, Candidate 1 in Table 5 incurs two violations of WH-CRIT: one for the empty matrix CP, one for the wh-phrase in situ. The embedded CP is marked –Q and thus does *not* incur a violation for being empty. This is crucial for eliminating candidates such as 7 which contain a copy of the wh-phrase in their matrix CP, but also a wh-phrase in-situ. Similarly, MARKFOC incurs a violation any time an element which contains a focus feature, which we assume wh-phrases do (following e.g. Wagner, 2012; Sabel, 2000), is not in FocP. For instance, candidate (2) incurs two violations of MARKFOC: one for each wh-phrase not in FocP. Finally, *t violations are counted once for every trace and an *additional* violation is counted for each trace that is lexicalized. Candidate 2, for example has one violation for having movement and one violation for that trace being lexicalized.

Table 4 illustrates the LD constructions which remain before and after harmonically bounded candidates are eliminated (indicated by the grayed-out cells) by the first method of evaluating universally dispreferred candidates. Candidates 2,4, and 5 are all harmonically bounded by candidate 6 because they incur the same violations as 6 plus at least one more. For example, candidate 2 incurs an additional violation of MARKFOC plus three violations of WHCRIT. This means, no matter how we re-rank the constraints,

[25] The exact nature of the two wh-phrases in copy constructions is debatable and unclear, based on relevant literature. For our purposes we will consider the matrix wh-phrase as an overt trace which does not carry content, and therefore incurs single violations of FAITHQUD (similar to an SMer), but if they do carry content, they would incur two violations per copy. For an extensive review see Murphy (2016).

candidate 2 will lose to candidate 6 and we need no longer include it in our competition. Candidates 6 and 9 are attested constructions. Candidate 3 is not attested, but we will return to it later. The input is indicated in the top left cell. *-R QUD: x* indicates the number of unresolved QUDs. We indicate child vs. adult register with the words *child* or *adult.* We indicate whether the subject is or is not CT with *subj is/is not CT.*

Table 4. *Harmonic bounding in LD constructions*

–R QUD: 2 **Child Reg**	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C. REG	MARK FOC	MIN PROJ	WH-CRIT	*t	BAR 4	BAR 3	FULL INT
2) [$_{CP\ +Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$[$_{VP}$… ⟦$_{CP\ -Q}$ wh$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ …wh$_1$? **Harmonically Bounded (HB) by (6)**		**		*	**		***	**	**		
3) [$_{CP\ +Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$[$_{VP}$… ⟦$_{CP}$ wh$_{1-Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$?		**		*	*		**	*	**		
4) [$_{CP}$ wh$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … ⟦$_{CP}$ wh$_{1-\ Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ …wh$_1$? **HB by (6)**		** **		*	***		**	***	**		
5) [$_{CP}$ wh$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … ⟦$_{CP}$ wh$_{1\ -Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$? **HB by (6)**		** *		*	**		*	**	**		
6) [$_{CP}$ wh$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … ⟦$_{CP}$ t$_{1-Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$? **Attested : English**		**		*	*			**	**		
9) [$_{CP+Q}$ [$_{FocP}$ WH$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$… ⟦$_{CP}$ t$_2$ that $_{-Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ …t$_2$? **Attested : Hungarian**		**		*		**	**	**	**		

It is apparent from Table 4 that LD structures, with the exception of 9, have the advantage of no WH-CRIT violations and the disadvantage of not marking focus. These two constraints are at odds in LD constructions.

Table 5 illustrates the harmonic bounding of the SynSM constructions. Candidate 11 and 14 are eliminated by Candidate 10. The three remaining candidates are the three forms of SynSM attested in German, Hindi, and Hungarian.

Table 5. *Harmonic bounding in SynSM constructions*

–R QUD:2 Child sub is CT	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C. REG.	MARK FOC	MIN PROJ	WH- CRIT	*t	BAR 4	BAR 3	FULL INT
10) [$_{CP}$ SM$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ t$_1$[$_{VP}$ … ⟦$_{CP\ -Q}$ wh$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_2$ **?** **Attested: German**		*	*	*	**		*	**	*	*	*
11) [$_{CP}$ SM$_{1+Q}$[$_{TP}$ [$_{vP}$[$_{AgrOP}$ t$_1$[$_{VP}$…⟦$_{CP\ -Q}$ wh$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … wh$_2$ **?** **HB by 10**		* **	*	*	**		*	** *	*	*	*
13) [$_{CP+Q}$ [$_{FocP}$ SM$_1$[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ t$_1$[$_{VP}$... ⟦$_{CP-Q}$ that [$_{FocP}$ wh$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ …t$_2$ **?** **Attested: Hungarian**		*	*	*		**	***	**	*	*	*
14) [$_{CP+Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ SM [$_{VP}$ ⟦$_{CP-Q}$ wh$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$**?** **HB by 15**		*	*	*	**		***	*	*	*	*
15) [$_{CP+Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ SM [$_{VP}$ ⟦$_{CP}$__$_{-Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … wh-$_1$**?** **Attested: Hindi**		*	*	*	**		***		*	*	*

The SynSM candidates clearly have the advantage of having fewer, smaller links between coindexed elements. If we compare these structures to the LD constructions, we see that LD always incurs two violations of BAR 4 or more. The same will prove true for SeqQ constructions. SynSM structures, on the other hand typically have one BAR 4 violation and one BAR 3 violation. This is a critical difference between the two constructions. SynSM structures also have fewer violations of FAITHQUD when there are two unresolved QUDs specified in the input, which also gives them an advantage over LD constructions. However, SynSM constructions always have at least one violation of WH-CRIT, which LD almost never violates, and always violate FULLINT. Furthermore, when the input specifies that the subject is CT, SynSM incurs a violation of FAITHCT. Thus, the re-ranking of these particular constraints will determine which candidates are optimal in a given language.

Table 6 illustrates the harmonic bounding of SeqQs. Candidate 20 is harmonically bounded by 16. Candidates 16, 17, and 21 are the attested forms of SeqQs found in English/German, Hindi, and Hungarian. 18 and 19 are unattested, but we will return to them later.

Table 6. *Harmonic bounding in SeqQ Constructions*

–R QUD:2	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C. REG	MARK FOC	MIN PROJ	WH-CRIT	*t	BAR 4	BAR 3	FULL INT
16) $[_{CP}$ $WH_{1\ +Q}$ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ t_1? $[_{CP}$ $WH_{2\ +Q}$ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … t_2**?** **Attested : English**	*				**			**	**		
17) $[_{CP+Q}$ __ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ WH_1? $[_{CP\ +Q}$ __ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … WH_2**?** **Attested: Hindi**	*				**		****		**		
18) $[_{CP+Q}$ __ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ WH_1? $[[_{CP}$ $WH_{2\ +Q}$ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … t_2 **?**	*				**		**	*	**		
19) $[_{CP}$ $WH_{1\ +Q}$ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ t_1 ? $[_{CP\_+Q}$ __ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … WH_2 **?**	*				**		**	*	**		
20) $[_{CP}WH_{1+Q}$ $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ WH_1? $[_{CP}$ WH_{2+Q} $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … WH_2**?** **HB by 16**	*	****			**		**	****	**		
21) $[_{CP+Q}$ $[_{Foc}$ WH_1 $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ t_1 ? $[_{CP+Q}$ $[_{Foc}$ WH_2 $[_{TP}$ $[_{vP}$ $[_{AgrOP}$ $[_{VP}$ … t_2 **?** **Attested: Hungarian**	*					**	****	**	**		

While SeqQs have no clear structural advantage over LD or SynSM structures, they have clear advantages in the Faithfulness constraints. No SeqQ construction violates FAITHCREG and when there are two unresolved QUDs, SeqQs incur no violations (with the notable exception of structures like 20 which do not delete copies)[26].

Finally, Table 7 illustrates the harmonic bounding of monoclausal questions. Candidate 25 is harmonically bounded by 24. The three remaining candidates (22, 23, 24) are the attested monoclausal forms found in English/German, Hindi, and Hungarian.

[26] It should be noted that we did not consider a candidate which was SeqQ and included an SMer in each question such as (E). This is because it is not clear what the role of the SMer would be in such a structure. Furthermore, it would consistently lose to either candidate 15 or 16.

(E) $[_{CP+Q}\ SM_4\ [_{TP}\ [_{vP}\ [_{AgrOP}\ t_4\ [_{VP}\ WH_1\mathbf{?}\ [\![_{CP+Q}\ SM_3\ [_{TP}\ [_{vP}\ [_{AgrOP}\ t3\ [_{VP} \ldots WH_2\mathbf{?}]]]]]]]]]]$

Table 7. *Harmonic bounding in monoclausal constructions*

–R QUD:2	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C REG	MARK FOC	MIN PROJ	WH-CRIT	*t	BAR 4	BAR 3	FULL INT
22) [CP WH$_1$ +Q [TP [vP [AgrOP [VP t$_1$? **Attested :** **English**	**	**			*			*	*		
23) [CP +Q [TP [vP [AgrOP [VP WH$_1$? **Attested :** **Hindi**	**	**			*		**				
24) [CP+Q [Foc WH$_1$ [TP [vP [AgrOP [VP t$_1$? **Attested :** **Hungarian**	**	**					**	*	*		
25) [CP WH$_1$ +Q [TP [vP [AgrOP [VP WH$_1$? HB by 24	**	****			*		*	*	*		

Though monoclausal constructions are overall dispreferred by the Faithfulness constraints FAITHLF/SYN and FAITHQUD even when both QUDs remain unresolved, they are advantageous when the question is child-directed and they also include structural advantages: they incur no violations of MINPROJ and fewer violations of the MINLINK constraints.

So far, we have reduced the candidate set to the set of 13 presented in Table 8. We now examine whether the second method of evaluating harmonic bounding further limits the set. Recall that this is done by taking a set of three candidates and establishing whether there is one of the three which will always lose to one of the other two. For instance, consider 16, 17, and 18 in Table 8. If WH-CRIT outranks *t, 16 will always win among these three. If, however, *t outranks WH-CRIT, 17 will always win. These three candidates have the same violations in every other condition, so there is no way 18 will ever be optimal. Candidate 19 has identical violations to 18 and is thus also harmonically bounded by 16 and 17. These two candidates are therefore eliminated from further tableaux. Likewise, candidate 3 will always lose to either 6 or 15: if FAITH QUD or *t is high-ranked, 15 will win; if WH-CRIT is high-ranked 6 will win. The number of violations candidate 3 has of all other constraints are ties with at least one of the other candidates: there is no scenario where 3 is optimal. It is, therefore, also eliminated from further tableaux.

Table 8. *Remaining Candidates (after first round of harmonic bounding)*

–R QUD:2	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C. REG	MARK FOC	MIN PROJ	WH-CRIT	*t	BAR 4	BAR 3	FULL INT
3) [$_{CP\ +Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$[$_{VP}$... ⟦$_{CP}$ wh$_{1–Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$ **?** **HB by 6 plus 15**		**		*	*		**	*	**		
6) [$_{CP}$ wh$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … ⟦$_{CP}$ t$_{1–Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_1$ **?** **Attested : English**		**		*	*			**	**		
9) [$_{CP+Q}$ [$_{FocP}$ WH$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$... ⟦$_{CP}$ t$_2$ that $_{–Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ ...t$_2$**?** **Attested : Hungarian**		**		*		**	**	**	**		
10) [$_{CP}$ SM$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ t$_1$[$_{VP}$ … ⟦$_{CP\ –Q}$ wh$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_2$ **?** **Attested : German**		*	*	*	**		*	**	*	*	*
13) [$_{CP+Q}$ [$_{FocP}$ SM$_1$[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ t$_1$[$_{VP}$... ⟦$_{CP–Q}$ that [$_{FocP}$ wh$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ ...t$_2$ **?** **Attested: Hungarian**		*	*	*		**	***	**	*	*	*
15) [$_{CP+Q}$ __[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ SM [$_{VP}$ ⟦$_{CP}$__$_{–Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … wh$_1$**?** **Attested: Hindi**		*	*	*	**		***		*	*	*
16) [$_{CP}$ WH$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ t$_1$? ⟦$_{CP}$ WH$_2$ $_{+Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_2$**?** **Attested : English**	*				**			**	**		

Table 8. *Remaining Candidates (after first round of harmonic bounding)*

–R QUD:2	FAITH LF/SYN	FAITH QUD	FAITH CT	FAITH C. REG	MARK FOC	MIN PROJ	WH-CRIT	*t	BAR 4	BAR 3	FULL INT
17) [$_{CP+Q}$ __ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ WH$_1$? ⟦$_{CP+Q}$ __ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … WH$_2$**?** **Attested: Hindi**	*				**		****		**		
18) [$_{CP+Q}$ __ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ WH$_1$? ⟦$_{CP}$ WH$_{2\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_2$ **?** **HB by 16 plus 17**	*				**		**	*	**		
19) [$_{CP}$ WH$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ t$_1$? ⟦$_{CP\_+Q}$ -__[$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … WH$_2$ **?** **HB by 16 plus 17**	*				**		**	*	**		
21) [$_{CP+Q}$ [$_{Foc}$ WH$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ t$_1$? ⟦$_{CP+Q}$ [$_{Foc}$ WH$_2$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ … t$_2$ **?** **Attested: Hungarian**	*					**	****	**	**		
22) [$_{CP}$ WH$_{1\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ t$_1$? **Attested : English**	**	**			*			*	*		
23) [$_{CP\ +Q}$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ WH$_1$? **Attested : Hindi**	**	**			*		**				
24) [$_{CP+Q}$ [$_{Foc}$ WH$_1$ [$_{TP}$ [$_{vP}$ [$_{AgrOP}$ [$_{VP}$ t$_1$? **Attested : Hungarian**	**	**					**	*	*		

The candidates which remain are all candidates directly relevant to our cross-linguistic study[27]. See Table 9. Candidate 6 is LD wh-movement without a FocP position, as found in English. 10 is the typical German type SynSM. The SMer originates in specAgrOP, as per Horvath (2000) and Fanselow & Mahajan (2000), and moves to the matrix specCP. The true wh-phrase moves to the intermediate specCP. 13 is the Hungarian form of SynSM which includes FocP in both clauses, per (Bhatt & Yoon, 1992). The SMer moves to the matrix specCP and the true wh-phrase moves to the embedded specFocP. 15 represents SynSM as we find it in Hindi with both the SMer and true wh-phrase remaining in-situ. 16 represents SeqQs with wh-movement, as is found in English. 17 represents SeqQs as they appear in Hindi with both wh-phrases in-situ. 21 represents SeqQs in Hungarian with movement of both wh-phrases to FocP in each clause. 22 is a monoclausal question with wh-movement to CP, as see in English. 23 is a monoclausal question with the wh-phrase in-situ as seen in Hindi, and 24 is a monoclausal question with the wh-phrase moved to FocP as seen in Hungarian. Finally, 9 is LD as it would appear in Hungarian with a single wh-phrase moved to the specFocP of the matrix CP. However, we remove candidate 9 in further tableaux because this structure only appears with certain matrix verbs (Horvath, 2000). In OT terms, these constructions would be eliminated by a high-ranked lexical constraint.

[27] Note that candidate 5, a copy construction, is eliminated by harmonic bounding despite the fact that it is attested in some dialects of German and in Dutch. This accords with the observation in Murphy (2016) that the exact nature of the copy is far from established as is the semantics of the copy construction generally. It is therefore entirely possible that another constraint (likely an Input-Output Faithfulness constraint) would be relevant to the use of copy-constructions and would render them optimal in some condition. We leave this issue for future research.

Overall, harmonic bounding applied to WSM-related optimizations on the basis of the set of constraints listed above results in only 10 potential optimal candidates cross-linguistically, as shown in Table 9 (their numbering reflects the full set of outputs prior to trimming). Language-particular rankings of relevant active constraints will then determine which ones are optimal in the particular languages under consideration. For clarity, the relevant bracketing will be reduced to abbreviations in all following tableaux. The reader is referred back to Table 5 to verify details if necessary. Furthermore, of course, there are multiple monoclausal questions involved (Q1 and Q2), but we will only include one for illustration as they incur the same violations except regarding FAITHQUD. We will specify which monoclausal question when relevant.

Table 9. *The final set of cross-linguistic candidates attested in our survey*

Optimal in English and German 6) $[_{CP\ +Q}\ WH_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots \llbracket_{CP}\ t_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_1\ ]]]]\rrbracket]]]]]?$ LD movement
Optimal in German 10) $[_{CP\ +Q}\ SMer_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ t_1 [_{VP} \ldots \llbracket_{CP}\ WH_2\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_2\ ]]]]\rrbracket]]]]]?$ SynSM with movement
Optimal in Hungarian 13) $[_{CP+Q}\ [_{FocP}\ SMer_1 [_{TP}\ [_{vP}\ [_{AgrOP}\ t_1 [_{VP} \ldots \llbracket_{CP}\ \text{that}\ [_{FocP}\ WH_2\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_2\ ]]]]]\rrbracket]]]]]]?$ SynSM with movement to FocP
Optimal in Hindi 15) $[_{CP+Q}\ \_\_ [_{TP}\ [_{vP}\ [_{AgrOP}\ SMer\ [_{VP}\ \llbracket_{CP} \_\_\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots WH_1]]]]\rrbracket]]]]]?$ SynSM without movement
Optimal in German and English 16) $[_{CP+Q}\ WH_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ t_1\ ?\ \llbracket_{CP\ +Q}\ WH_2\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_2\ ]]]]\rrbracket]]]]]?$ SeqQs with movement
Optimal in Hindi 17) $[_{CP+Q}\ \_\_\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ WH_1?\ \llbracket_{CP\ +Q}\ \_\_\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots WH_2]]]]\rrbracket]]]]]?$ SeqQs without movement
Optimal in Hungarian 21) $[_{CP+Q}\ [_{Foc}\ WH_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ t_1\ ?\ \llbracket_{CP+Q}\ [_{Foc}\ WH_2\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP} \ldots t_2\ ]]]]]\rrbracket]]]]]]?$ SeqQs with movement to FocP
Optimal in English and German 22) $[_{CP}\ WH_{1\ +Q}\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ t_1]]]]]\ ?$ Monoclausal question with wh-movement to CP
Optimal in Hindi 23) $[_{CP\ +Q}\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ WH_1]]]]]?$ Monoclausal question with wh in-situ
Optimal in Hungarian 24) $[_{CP+Q}\ [_{Foc}\ WH_1\ [_{TP}\ [_{vP}\ [_{AgrOP}\ [_{VP}\ t_1\ ]]]]]]?$ Monoclausal question with wh-movement to FocP

3.4 OT Analysis by Language

Having reduced the universal candidate set to include only potential winners, we can now turn to optimizing within individual languages. Within each of the following sections we remind the reader of the optimal pairings collected from native speakers and then include tableaux illustrating the rankings which achieve the appropriate outputs. Candidates which are suboptimal are marked out with grayscale.

3.4.1 German

One participant was eliminated from the German-speakers based on their selection of monoclausal questions in over 60% of the items. This leaves 11 participants whose data we analyzed. The input/output pairings indicated by our survey are illustrated in Table 10.

Table 10. *German input/output pairings*

	Semantic input	Preferred syntactic output
German speakers	2 unresolved QUDs, Subj CT, Adult	SeqQs
	2 unresolved QUDs, Subj CT, Child	SeqQs
	2 unresolved QUDs, Subj not CT, Adult	SynSM
	2 unresolved QUDs, Subj not CT, Child	SeqQs
	1 unresolved QUD, Subj CT, Adult	LD
	1 unresolved QUD, Subj CT, Child	SeqQs
	1 unresolved QUD, Subj not CT, Adult	SynSM
	1 unresolved QUD, Subj not CT, Child	SeqQs

One novel finding from our survey we wish to highlight again is that German speakers overall used SeqQs most frequently, which is not what was suggested by previous literature on WSM. While both SynSM and LD also surface as optimal constructions, SeqQs are the overall preferred choice once particular semantic/pragmatic properties are taken into account. When Q2 is unresolved (resulting in 2 unresolved QUDs, including the matrix QUD) SeqQs is the optimal choice, except when the subject is not CT and the question is directed to an adult. This suggests an interaction of CT and Register. When Q2 is resolved (resulting in one unresolved QUD), then SeqQs are optimal so long as the question is in a child-directed register. When the question is directed to an adult, then CT will determine whether LD or SynSM is optimal. Table 10 overall suggests that FAITH C. REG is ranked high in German and that FAITH CT and FAITH QUD play crucial roles in obtaining the correct input-output mappings. With respect to WSM, our experimental results and analysis thereof below show that SynSM is more constrained in German than previously thought: it is optimal for only two (out of eight) semantic/pragmatic specifications.

When examining the tableaux in German, recall that MinLink is violated every time a link crosses a barrier, thus LD incurs two sets of four violations (because the wh-phrase moves twice across four barriers), SynSM incurs one set of four violations (for the movement of the contentful wh-phrase to the intermediate specCP) and one set of three violations (for the movement of the SMer from specAgrOP to specCP). SeqQs incur two sets of four violations and monoclausal questions incur a single set of four violations. We refer the reader to Table 5 for details on barriers crossed for each candidate. As previously mentioned, these sets of violations could be represented as Barrier constraints (BAR 4, BAR 3) and indeed they will appear as such in Hungarian, but German does not require that they split, they are therefore represented as their overarching constraint MINLINK.

Let us begin with the scenarios where Q2 is unresolved. Given the fact that Q1 is always unresolved in our inputs (otherwise, there would be no question), this means there are always two relevant unresolved QUDs which require two [+wh] lexical items that carry content. In these scenarios, we have seen that SeqQs are optimal in all but one condition: when the subject is not CT, but the register is adult-directed. This indicates that FAITH C. REG should be ranked higher than FAITHCT. Furthermore, the fact that monoclausal questions never surface as optimal indicates that FAITHLF/SYN must also be high-ranked. If we compare tableaux 3 and 4 we can see that when the subject is not CT, SynSM will be the optimal output if the question is directed to an adult, as in Tableau 3 where both FAITH C. REG and FAITHCT are vacuously satisfied, but when the question is directed to a child (as in Tableau 4), all constructions with multiple clauses incur a violation of FAITH C. REG and SeqQs are the optimal output.

T3. *German. Two unresolved QUDs, adult-directed, Subject not CT.*
Attested output: SynSM

Input: **–R QUD:2, Adult, Subj not CT**	FAITH C.REG	FAITH CT	FAITH LF/SYN	MIN-LINK	FAITH QUD	WH-CRIT	MIN PROJ	*t	MARK FOC	FULL INT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$				**** ****!	**			**	*	
☞SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$				**** ***	*	*		**	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t$_2$				**** ***	*	***!	**	**		*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$__… WH$_2$				**** ***	*	***!			**	*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?			*!	**** ****				**	**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$_… WH$_2$?			* !	**** ****		****			**	
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH2… t$_2$?			*!	**** ****		****	**	**		
Mono. Q [$_{CP}$ WH$_1$… t$_1$?			**!	****	**				*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!	****	**	**				
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?			**!	****	**	**	*		*	

T4. *German. Two unresolved QUDs, Child-directed, Subject not CT. Attested output: SeqQs*

Input: **–R QUDs: 2, Child, Subj not CT**	FAITH C.REG	FAITH CT	FAITH LF/ SYN	MIN - LINK	FAITH QUD	WH-CRIT	MIN PROJ	*t	MARK FOC	FULL INT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$	*!			*********	**			**	*	
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$	*!			*******	*	*		**	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ ... t$_2$	*!			*******	*	***	**	**		*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ __... WH$_2$	*!			*******	*	***			**	*
☞SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?			*	********				**	**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			*	********		****!			**	
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?			*	********		****!	**	**		
Mono. Q [$_{CP}$ WH$_1$… t$_1$?			**!	****	**				*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!	****	**	**				
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?			**!	****	**	**	*		*	

Tableaux 3 and 4 have illustrated the fact that FAITH C. REG must be ranked above FAITHLF/SYN because even though SeqQs incur a violation of FAITHLF/SYN, they remain optimal because they do not incur a violation of FAITH C. REG. Thus, we establish (20) in our constraint ranking.

(20) FaithCREG≫ FAITHLF/SYN

Let us now consider the situation where Q2 is resolved, meaning only one unresolved QUD is in the input (Tableau 5). Because FAITH C. REG is high ranked, we know SeqQs will always win when the question is directed to a child even though SeqQs now incur two violations of FAITH QUD. The interesting alternation happens when the question is directed to an adult. Tableau 5 illustrates that FAITH CT must be ranked above MINLINK in German because when the subject is CT, LD is the optimal candidate in spite of its violations of MINLINK because it has no violations of either FAITH CT or FAIT LF/SYN. All SynSM structures incur a fatal violation of FAITH CT and are thus eliminated, and SeqQs now incur two violations of FAITH QUD compared to no violations by LD, which emerges as the winner.

T5. *German. One unresolved QUD, Adult-directed, Subject is CT. Attested output: LD*

Input: **–R QUDs: 1, Adult,** **Subj CT**	FAITH C. REG.	FAITH CT	FAITH LF/SYN	MIN-LINK	FAITH QUD	WH-CRIT	MIN PROJ	*t	MARK FOC	FULL INT
☞LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t_1				**** ****				**	*	
SynSM with movement [$_{CP}$ SMer$_1$… t_1⟦$_{CP}$ WH$_2$… t_2		*!		**** ***	*	*		**	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t_1⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t_2		*!		**** ***	*	***	**	**		*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$__… WH$_2$		*!		**** ***	*	***			**	*
SeqQs with movement [$_{CP}$ WH$_1$… t_1? [$_{CP}$ WH$_2$… t_2?			*!	**** ****	**			**	**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			* !	**** ****	**	****			**	
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t_1? [$_{CP}$ [$_{FocP}$WH$_2$… t_2?			*!	**** ****	**	****	**	**		
Mono. Q [$_{CP}$ WH$_1$… t_1?			**!	****	****	**		*	*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!	****	****	**			*	
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t_1?			**!	****	****	**	*	*		

(21) {FAITHCT ≫ MINLINK}

Thus, when the subject is CT, LD is the optimal output, as attested in our survey, but when the subject is not CT, FAITH CT is no longer violated and SynSM emerges as the optimal output, as illustrated in Tableau 6.

T6. *German. One unresolved QUD, Adult-directed, Subject not CT. Attested output: SynSM*

Input: **–R QUDs: 1, Adult** **Subj not CT**	FAITH C. REG	FAITH CT	FAITH LF/SYN	MIN-LINK	FAITH QUD	WH-CRIT	MIN PROJ	*t	MARK FOC	FULL INT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$				**** ****!				**	*	
☞SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$				**** ***	*	*		**	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t$_2$				**** ***	*	***!	**	**		*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$__… WH$_2$				**** ***	*	***!			**	*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?			*!	**** ****	**			**	**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			* !	**** ****	**	****			**	
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?			*!	**** ****	**	****	**	**		
Mono. Q [$_{CP}$ WH$_1$… t$_1$?			**!	****	****	**		*	*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!	****	****	**			*	
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?			**!	****	****	**	*	*		

Tableau 6 provides crucial evidence for the relative ranking of MINLINK and FAITHQUD (shown in 22). If their ranking were reversed, LD would be incorrectly chosen as the optimal output for this input.

(22) MINLINK≫ FAITHQUD

Thus, we see the pattern indicated by the German native speakers is achieved by ranking FAITH C. REG, FAITH CT, FAITHLF/SYN, MINLINK, FAITH QUD and WH-CRIT at the top of the constraint hierarchy. Because FAITH CT and FAITHLF/SYN can never be violated by the same candidate, we cannot tell which is ranked above the other, but we know they both come between FAITH C. REG and MINLINK. The constraints that follow may be in any order.

(23) German ranking: FaithCREG≫ {FAITHCT, FAITHLF/SYN} ≫ MINLINK≫ FAITHQUD ≫ WH-CRIT≫ { MINPROJ, *t , MARKFOC, FULL INT }

3.4.2 HINDI

Let us now turn to Hindi. Hindi does not allow wh-movement, thus *t is the highest ranked constraint. However, there is alternation between the use of SeqQs (with wh-phrases in-situ) and SynSM (with wh-phrases in situ). Our logistic regression model discussed in Chapter 2 suggested that an interaction between CT and Register significantly affected the use of SynSM, but that the resolved/unresolved status of Q2 had no effect. We therefore do not consider QUD as a factor for Hindi. One participant was excluded for using monoclausal questions in over half of the stimuli, leaving 10 participants for consideration.

Hindi Speakers

Table 11 shows the input-output pairings for Hindi-speakers based on our survey results.

Table 11. *Hindi input/output pairings*

	Semantic input	Preferred syntactic output
Hindi Speakers	Subj CT, Adult	SynSM
	Subj CT, Child	SeqQs
	Subj not CT, Adult	SynSM
	Subj not CT, Child	SynSM

Of course, we know *t must be highly ranked such that all wh-phrases remain in situ. It is also clear that FAITH QUD must be a low ranked constraint because QUD does not affect the output. The fact that SynSM can win even when the structure is child-directed means that FAITH LF/SYN must be ranked relatively high. However, the alternation between SynSM and SeqQs when the subject is CT means that FAITH C. REG and FAITH CT must be high ranked as well. Tableau 7,8 and 9 illustrate that, in fact, these three constraints must be tied in ranking to obtain the attested output. Consider Tableau 7 where the question is adult-directed and the subject is CT. *t eliminates all but SeqQs and SynSM (both with in situ wh-phrases). In Tableau 7, because of the three-way tie of constraints, the optimization comes down to MINLINK for which SynSM incurs fewer violations than SeqQs. However, in Tableau 8 the construction is child-directed, thus SynSM incurs a violation of FAITH C. REG as well as its violation of FAITH CT. This eliminates it and SeqQs is the optimal candidate in spite of the violation of FAITH LF/SYN.

T7. *Hindi. One unresolved QUD, Adult-directed, Sub CT. Attested output: SynSM*

Input: **–R QUD: 1, Adult, subj CT**	*t	FAITH C. REG	FAITH LF/SYN	FAITH CT	MIN LINK	FAITH QUD	MIN PROJ	FULL INT	MARK FOC	WH-CRIT
LD [$_{CP}$ WH$_1$… ⟦$_{CP}$… t$_1$	**!				**** ****				*	
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$ ⟦$_{CP}$ WH$_2$… t$_2$	** !			*	**** ***	*		*	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$ ⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t$_2$	** !			*	**** ***	*	**	*		***
☞SynSM without movement [$_{CP}$_… SMer$_1$ ⟦$_{CP}$ _… WH$_2$				*	**** ***	*		*	**	***
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	** !		*		**** ****	**			**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			*		**** **** !	**			**	****
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?	**!		*		**** ****	**	**			****
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!		**		****					
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!		****					
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?	*!		**		****					

T8. *Hindi. One unresolved QUD, Child-directed, Sub CT. Attested output: SeqQs*

Input: **–R QUD: 1, Child subj CT,**	*t	FAITH C.REG	FAITH LF/SYN	FAITH CT	MINK LINK	FAITH QUD	MIN PROJ	FULL INT	MARK FOC	WH-CRIT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$	**!	*			**** ****				*	
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$... t$_2$	** !	*		*	**** ***	*		*	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ ... t$_2$	** !	*		*	**** ***	*	**	*		***
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ _... WH$_2$		*		*!	**** ***	*		*!	**	***
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	** !		*		**** ****	**			**	
☞SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			*		**** ****	**			**	****
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?	**!		*		**** ****	**	**			****
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!		**		****	****			*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!		****	****			*	
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?	*!		**		****	*****				

Tableau 7 and 8 illustrate that if FAITH CT and FAITH LF/SYN were not tied, we would not see the appropriate alternation between SynSM and SeqQs. If either were ranked over the other, we would see only SynSM or only SeqQs when the subject was CT. The ranking of FAITHLF/SYN and FAITHCT is established in (24).

(24) FAITHLF/SYN:FAITHCT

However, these tableaux do not explain why FAITH C. REG must also be tied. The reason for the three-way tie becomes clear when we turn to a condition where the subject is not CT. Without the three-way tie, we would again see alternation between SeqQs and SynSM, but this is not the attested pattern. Tableau 9 illustrates the scenario where the subject is not CT and the register is child-directed, which would result in SeqQs if these constraints were not tied. However, because FAITH C. REG is also tied with FAITH LF/SYN, we again have equal violations of the tie and MINLINK eliminates SeqQs, with SynSM emerging as the winner.

T9. *Hindi. One unresolved QUD, Child-directed, Sub not CT.*
Attested output: SynSM

Input: **–R QUD: 1, Child subj not CT**	*t	FAITH C.REG	FAITH LF/SYN	FAITH CT	MINK LINK	FAITH QUD	MIN PROJ	FULL INT	MARK FOC	WH-CRIT
LD [$_{CP}$ WH$_1$… ⟦$_{CP}$… t$_1$	**!	*			**** ****				*	
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$	** !	*			**** ***	*		*	**	*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t$_2$	** !	*			**** ***	*	**	*		***
☞SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ _... WH$_2$		*			**** ***	*		*	**	***
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	** !		*		**** ****	**			**	
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?			*		**** **** !	**			**	****
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?	**!		*		**** ****	**	**			****
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!		**		****	****			*	
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!		****	****			*	
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?	*!		**		****	*****				

Of course, if the production were not child-directed, there would be no violations of FAITH C. REG and SynSM would remain the winner, as attested in our survey. Tableau 9 establishes the ranking shown in (25).

(25) FAITH C. REG:FAITHLF/SYN:FAITHCT ≫ MINLINK

Thus, we see for Hindi-speakers, the ranking of *t and a three-way tie between FAITH C. REG, FAITH LF/SYN, and FAITH CT followed by MINLINK above all other constraints result in the pattern attested by native speakers. All other constraints may appear in any order.

(26) Hindi ranking: *t ≫ **FAITH C. REG:FAITHLF/SYN:FAITHCT** ≫ **MINLINK** ≫ **{FAITHQUD, MINPROJ , FULLINT, MARKFOC,WH-CRIT }**

3.4.5 Hungarian

Unlike the other languages discussed, Hungarian requires that wh-phrases be in FocP (Kiss, 1998). This means that MARKFOC is high-ranked, eliminating any structure which does not include FocP. The results of the linear regression model described in Chapter 2 indicated that none of the variables we controlled for had a significant effect on the output chose. However, this is unsurprising given that the results of the survey show that participants, as a whole, chose SynSM as the optimal construction in every condition except one. When Q2 is unresolved, the subj is CT, and the question is child-directed, the monoclausal embedded question (monoclausal Q2) is the optimal candidate. This suggests neutralization[28] in all cases except this one. In this scenario alone, SynSM violates too many constraints, critically FAITHCT and FAITH C REG, and the monoclausal Q2 surfaces as optimal. This can be achieved by two sets of equally ranked constraints. Crucially, recall that the use of monoclausal Q2 incurs two violations of FAITHQUD when there are two

[28] Neutralization (of input distinctions) obtains when distinct semantic/pragmatic inputs result in a single syntactic output (Legendre, 2010). This is the kind of pattern which cannot be modeled using SuperOptimality (Blutner, 2000), which crucially relies on the pairing of distinct inputs with distinct outputs, and is the most widely used Bidirectional OT architecture.

unresolved QUDs (because only the embedded clause QUD is represented) and it incurs four violations when the embedded clause QUD is resolved and the only unresolved QUD is the matrix clause QUD. This is because the content represented by monoclausal Q2 is not the content of the unresolved QUD. Therefore, it incurs two violations for the missing appropriate wh-phrase as well as two violations for the presence of the incorrect wh-phrase. Data from one participant were excluded as they chose a monoclausal Q in over 50% of cases. Data from 13 participants remained.

Tableau 10 illustrates the condition where Q2 is unresolved, the subject is CT, and the utterance is adult-directed. MARKFOC must be high-ranked to eliminate any competitor without wh-phrases in FocP. In Hungarian, the family of constraints MINLINK must be divided into BAR4 and BAR3. A violation of BAR4 indicates a chain link crossing 4 barriers (a set of four violations in MINLINK). BAR4 is high ranked, eliminating SeqQs. Both SynSM and monoclausal Q2 incur a single violation of BAR4, meaning neither is eliminated. FAITHCT and FAITHLF/SYN and FAITH C. REG are tied (which we will discuss below) and SynSM has only a single violation of the tied constraints where the monoclausal question has two, thus SynSM is the optimal candidate in this condition.

T10. *Hungarian. Two unresolved QUDs, Adult-directed, Subject is CT. Attested output: SynSM*

Input:–**R QUDs: 2,** **Adult,** **Subj is CT,**	MARK FOC	BAR 4	FAITH CT	FAITH LF/SYN	FAITH C. REG	FAITH QUD	BAR 3	*t	WH-CRIT	MINPROJ	FULL INT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$	* !	**				**		**			
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$... t$_2$	**!	*	*			*	*	**	*		*
☞SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ ... t$_2$		*	*			*	*	**	***	**	*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ _... WH$_2$	**!	*	*			*	*		***		*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	**!	**		*				**			
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?	**!	**		*					****		
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?		**!		*				**	****	**	
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!			**		**		*			
Mono. Q without movement [$_{CP}$_… WH$_1$?	*!			**		**			**		
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?		*		**!		**		*	**	*	

If FAITHLF/SYN ranked above FAITHCT, then SynSM (with FocP) would win in every scenario, but this is not the attested pattern. However, if either FAITHCT or FAITH C. REG were ranked above FAITHLF/SYN, then monoclausal Q2 would win in multiple scenarios, which is also not correct. The attested pattern shows the monoclausal Q2 is optimal in exactly one scenario which is illustrated in Tableau 11, below. The tie between FAITHCT, FAITH LF/SYN, and FAITH C. REG achieves the attested pattern because SynSM incurs one violation of FaithCT and one violation of Faith C. Reg (two violations total) and the monoclausal Q2 incurs two violations of FaithLF/Syn (two violations total). Because of the tie, neither candidate is eliminated. Thus (27) is established.

(27) {FAITHCT:FAITHLF/SYN: FAITH C. REG}

The same thing happens with the following set of tied constraints: FAITHQUD and BAR3 (tie justified below). SynSM incurs a violation of each (two violations total) where monoclausal Q2 incurs two violations of FAITHQUD but, again, because the constraints are tied, neither candidate is eliminated until SynSM is finally eliminated by its violation of *t and monoclausal Q2 is the optimal output.

T11. *Hungarian. Two unresolved QUDs, Child-directed, Subject is CT. Attested output: mono. Q2*

Input: –**R** **QUDs: 2,** **Adult,** **Subj is CT,**	M. FOC	BAR 4	FAITH CT	FAITH LF/SYN	FAITH C. REG	FAITH QUD	BAR 3	*t	WH-CRIT	MIN PROJ	FULL INT
LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$	* !	**			*	**		**			
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$	**!	*	*		*	*	*	**	*		*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ ... t$_2$		*	*		*	*	*	**	***!	**	*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ _... WH$_2$	**!	*	*		*	*	*		***		*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	**!	**		*				**			
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?	**!	**		*					****		
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?		**!		*				**	****	**	
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!			**		**		*			
Mono. Q without movement [$_{CP}$_… WH$_1$?	*!			**		**			**		
☞Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?		*		**		**		*	**	*	

Finally, we must justify the tie between FAITHQUD and BAR3. We saw above that in order for monoclausal Q2 not to be eliminated by its double violation of FAITHLF/SYN, we had to have a three-way tie between FAITHCT, FAITHLF/SYN, and FAITH C. REG. Simply ranking BAR3 as the next constraint would have eliminated SynSM and result in monoclausal Q2 as the optimal output. However, this cannot be correct because monoclausal Q2 is never optimal when Q2 is resolved and there is only one unresolved QUD in the input. The correct constraint ranking will allow the attested alternation between when SynSM and monoclausal Q2 when there is one relevant unresolved QUD. Tableau 12 illustrates the case with a single unresolved QUD, the subject is CT, and the utterance is child-directed. SynSM is the optimal candidate. Because monoclausal Q2 invokes four violations of FAITHQUD (rather than two as in Tableau 11), it is eliminated and SynSM is the optimal candidate, as attested in our survey. Thus the ranking of FAITHCT, FAITHLF/SYN,FAITCREG, FAITHQUD and BAR3 must be as shown in (28).

(28) FAITHCT:FAITHLF/SYN: FAITH C. REG ≫ FAITHQUD:BAR3

T12. *Hungarian. One unresolved QUD, Child-directed, Subject is CT. Attested output: SynSM*

Input: **–R QUDs: 1, Child, Subj is CT,**	M. FOC	BAR 4	FAITH CT	FAITH LF/SYN	FAITH C. REG	FAITH QUD	BAR 3	*t	WH-CRIT	MIN PROJ	F. INT
LD: [$_{CP}$ WH$_1$… ⟦$_{CP}$… t$_1$	* !	**			*			**			
SynSM with movement: [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$... t$_2$	**!	*	*		*	*	*	**	*		*
☞SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ ... t$_2$		*	*		*	*	*	**	***	**	*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ __... WH$_2$	**!	*	*		*	*	*		***		*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?	**!	**		*		**		**			
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$ _… WH$_2$?	**!	**		*		**			****		
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?		**!		*		**		**	****	**	
Mono. Q [$_{CP}$ WH$_1$… t$_1$?	*!			**		****		*			
Mono. Q without movement [$_{CP}$_… WH$_1$?	*!			**		****			**		
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?		*		**		****!		*	**	*	

Of course, we can also see from Tableau 15 that if the subject were not CT or if the utterance were not child-directed, and the violations of either FAITHCT or FAITH C. REG were no longer incurred, SynSM would remain the winner, as attested by our survey findings.

Tableaux 10 through 12 illustrate that the ranking required includes a highly ranked MARKFOC and BAR4, though either could outrank the other for the purposes of this analysis. This is followed by two sets of tied constraints FAITHCT, FAITHLF/SYN and FAITH C. REG followed by FAITHQUD and BAR3. Without either of these ties, the attested pattern cannot be attained. We have ranked *t next, though any other constraint could immediately follow with the same result. This interesting case of Hungarian has also illustrated how a highly dispreferred candidate, such as monoclausal Q2 in a scenario that really requires biclausal questions, can be the optimal output if the otherwise optimal candidate (SynSM) incurs enough violations of Faith constraints.

(29) Hungarian ranking: {MARKFOC, BAR4} ≫ FAITHCT:FAITHLF/SYN: FAITH C. REG ≫ FAITHQUD:BAR3 ≫ { *t , WH-CRIT, MINPROJ, FULL INT }

3.4.6 English

Finally, the use of LD and SeqQs in English can also be predicted from our analysis. Two participants were excluded from analysis as they selected monoclausal questions over 50% of the time, leaving 10 participants for analysis. According to the logistic regression described in Chapter 2, the output chosen by English speakers was only predicted by CT and the resolved/unresolved status of Q2. This means that Register actually played no role in English-speakers choice of output. We therefore do not consider Register in our conditions for English and we ignore its evaluation of the relevant constraint. In English, LD is the optimal output in every case. [29]

Several constraints must be highly ranked to eliminate the use of SynSM. MINPROJ excludes the use of constructions with a FocP (essential to Hungarian). FULLINT precludes the use of SMers and therefore eliminates all SynSM constructions. Finally, WH-CRIT excludes the use of wh in-situ constructions. All candidates are, therefore, eliminated except for LD and SeqQs. However, LD incurs a single violation of MARKFOC, where SeqQs incurs two, thus eliminating SeqQs with the result that LD will always be the optimal candidate. This is depicted in Tableau 13, where the input is an unresolved Q2, the subject is CT, and the utterance is adult-directed, but because all highly ranked constraints are structural, no change in condition will result in a different optimal output.

[29] The only change in output the model detects is just how great a difference there is between LD and SeqQs. In some conditions, we do see an increase in the use of SeqQs and a decrease in the use of LD (when the subject is not CT and Q2 is unresolved), but LD is still the most frequently chosen candidate. This suggests either that we have not identified the appropriate set of variables or the combination of factors that would result in SeqQs being optimal. However, for the conditions we tested, LD was always optimal. We will therefore proceed to model the output for these conditions.

T13. *English. One unresolved QUD, Subject is CT. Attested output: LD*

Input: **–R QUDs: 1, Subj CT**	MIN PROJ	FULL INT	WH-CRIT	FAITH LF/SYN	MARK FOC	FAITH QUD	MINK LINK	*t	FAITH C. REG.	F. CT
☞LD [$_{CP}$ WH$_1$…⟦$_{CP}$… t$_1$					*	**	**** ****	**		
SynSM with movement [$_{CP}$ SMer$_1$… t$_1$⟦$_{CP}$ WH$_2$… t$_2$		* !	*		**	*	**** ***	**		*
SynSM with movement to FocP [$_{CP}$ [$_{FocP}$ SMer$_1$ … t$_1$⟦$_{CP}$ [$_{FocP}$ WH$_2$ … t$_2$	**!	*	***			*	**** ***	**		*
SynSM without movement [$_{CP}$_… SMer$_1$⟦$_{CP}$ _… WH$_2$		*!	***		**	*	**** ***			*
SeqQs with movement [$_{CP}$ WH$_1$… t$_1$? [$_{CP}$ WH$_2$… t$_2$?				*	**!		**** ****	**		
SeqQs without movement [$_{CP}$_… WH$_1$? [$_{CP}$_… WH$_2$?			****!	*	**		**** ****			
SeqQs with movement to FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$? [$_{CP}$ [$_{FocP}$WH$_2$… t$_2$?	**!		****	*			**** ****	**		
Mono. Q [$_{CP}$ WH$_1$… t$_1$?				**!	*	**	****	*		
Mono. Q without movement [$_{CP}$_… WH$_1$?			**!	**	*	**	****			
Mono. Q with FocP [$_{CP}$ [$_{FocP}$WH$_1$… t$_1$?			**!	**		**	****	*		

Because LD is always the optimal output, we know that FAITHQUD must be ranked below constraints like FULLINT and MINPROJ because it is always the winner in spite of its two violations of FAITHQUD. We also know that FAITH LF/SYN must be ranked relatively high to eliminate SeqQs and monoclausal Qs even when they do not violate FAITHQUD and LD does.

We can conclude that for the English-speakers, the ranking of constraints includes **MINPROJ, FULL INT, WH-CRIT** in any order, followed by **MARKFOC** and the rest may follow in any order.

(30) English ranking: {MINPROJ, FULLINT, WH-CRIT} ≫ MARKFOC ≫ {FAITHQUD, MINLINK, FAITHLF/SYN, FAITH CT, *t, CHILD REGISTER}

3.5 Conclusion

In Chapter 2, we had argued that previous work had been examining two separate, superficially similar phenomena, semantic SM (SemSM, e.g. Dayal 1994) and syntactic SM (SynSM, e.g. McDaniel,1989). These phenomena are not wholly unrelated as in Hindi, for example, SemSM can map directly onto SynSM (and often does). However, this relationship does not hold for every language exhibiting either SM phenomenon. In German, for instance, SemSM frequently surfaces as SeqQs, as our survey in Chapter 2 revealed. This and other attested variation led us to build on the proposal from Horvath (1997, 2000) who had suggested that Hungarian had the syntax one would expect under the IDA SynSM, but the meanings associated with these constructions aligned with the DDA. We posited that SemSM is universal and addressed semantic/syntactic variation per se by taking two major steps. First, we broke down Dayal (1994)'s analysis of SemSM into two components we were able to manipulate separately: QUD and CT. We elicited judgments from native speakers which informed us as to how each aspect of the discourse background affects their syntactic expression in four languages. Second, we proceeded with specifying both QUD and CT properties in the input to syntactic optimization, using OT to formalize a unified and cross-linguistically valid analysis of SynSM and SemSM.

The focus of Chapter 3 has been to demonstrate that by reranking a relatively small set of universal constraints we arrive at the appropriate syntactic output for each input in English, German, Hindi, and Hungarian. Each constraint has proven critical for achieving the attested patterns in the languages examined. The I-O Faithfulness constraint, **FAITH C. REG,** is crucial to the analyses of German, resulting in the novel finding that SeqQs are the most frequently chosen construction for German speakers[30]. It is further impossible to obtain an adequate analysis of the German-speakers without the constraints **FAITH LF/SYN, FAITHCT, FAITHQUD**, or **WH-CRIT**. The ***t** constraint is crucial to the Hindi speakers, resulting in an option which leaves the SMer and wh-phrase in-situ. A three way tie between **FAITH C. REG, FAITH LF/SYN,** and **FAITH CT** followed by **MINLINK** are also critical to the analysis of Hindi, resulting in an alternation between SynSM and SeqQs. The analysis of Hungarian requires **MARKFOC,** forcing structures in Hungarian to include a FocP in their structure in spite of violations to other constraints such as MINPROJ and WH-CRIT, which are low ranked in this language. **BAR4** is also high ranked in Hungarian followed by two sets of tied constraints: **FAITHCT, FAITHLF/SYN** and **FAITH C. REG** followed by **FAITHQUD** and **BAR3** which together make it the only language analyzed where a monoclausal question is optimal in spite of **FAITHLF/SYN** violations. In English

[30] This novel finding will need to be replicated to eliminate the possibility that it is an artifact of the task. For now, we believe it is the result of including SeqQs which previous work has not seriously considered as an option for expressing questions with multiple propositions in German.

speakers, on the other hand, **MINPROJ** and **WH-CRIT** are critical for achieving the optimal outputs, as is **FULLINT,** which prevents the use of synSM in English altogether.[31] Finally, while **BAR 3** was not critical to any of the final analyses, it was critical in eliminating universally suboptimal outputs (resulting from harmonic bounding) and reducing the candidate set to a very manageable size. Thus, each constraint proved essential for the analysis and the analysis, in turn, illustrates the very basic claim that grammatical structures in languages of the world routinely violate both Markedness and Faithfulness constraints. Tableaux 3-13 are replete with such evidence.

This unified analysis of these four languages relies on a limited set of violable constraints whose ranking varies among the languages. This unified analysis is possible because of the specific framing of OT we adopt, which by its very architecture, exploits the interface of syntax and semantics/pragmatics. Furthermore, because OT syntax, under the architecture defined by Legendre et al. (1995, 1998) and Legendre (2010) allows for the semantic input to affect the syntactic output, this analysis illustrates how Faithfulness and Markedness constraints conspire to yield a variety of optimal strategies across languages. The success of this effort suggests not only the universality of SemSM, but the puissance of OT syntax in explaining cross-linguistic syntactic variation. Note that our optimization would not succeed under an alternative OT architecture which excludes Input-Output Faithfulness (e.g., Grimshaw 1997). It further allows us to understand that what has

[31] We acknowledge that **FULLINT** is violated in English (e.g. *do*-support, weather *it*). We further note that this constraint ranking does not align with Grimshaw (1997)'s analysis. Her OP-SPEC (related to our **WH-CRIT**) is ranked above **FULLINT**, thus future work is required to reconcile these two.

been called "Scope Marking" is not, in fact, a phenomenon limited either to syntax or semantics, but rather the result of a complex interaction between the two.

Standard OT (Prince & Smolensky [1993]2004) relies on strict ranking of violable constraints. However, the cross-linguistic analysis developed in Chapter 3 adopts constraint ties in both Hindi and Hungarian to explain the attested pattern of preferences in each language. Constraint ties are not uncommon in OT analyses of syntax (e.g., Grimshaw, 1997). They have most often been invoked to account for optionality, the situation where a single optimization yields two optimal competitors for a given input (e.g., embedded *that*-clauses vs. null complementizer clauses in Grimshaw 1997). The present use of constraint ties, however, serves a different purpose: it allows for the variation shown in syntactic expression when subtle aspects of the semantic input change. It should be noted that the ties we have invoked are generally ties of the various Faith Constraints (with one exception: **FAITHQUD** is tied with **BAR 3** in Hungarian). Furthermore, across languages, they appear in groups: In German **FaithCREG, FAITHCT, FAITHLF/SYN** appear contiguously in that order and it is these three constraints which are also tied in Hindi. In Hungarian, again these same three are tied. While further investigation would be required to make a solid claim, this pattern suggests that these semantic factors are a kind of meta-constraint requiring that the semantics and pragmatics be properly observed. It further suggests that the syntactic constructions used in different languages take advantage of these semantic factors to varying degrees. German exploits a strict ranking of these constraints and thus relies on the distinctness of the semantic factors involved to yield a variety of optimal syntactic outputs. Hindi and Hungarian, on the other hand, take these semantic factors to be interchangeable, with no one factor able to affect the optimization outcome;

instead, it is the overall number of violations that can result in the ultimate failure of a candidate. In other words, German opts for a qualitative evaluation while Hindi and Hungarian favor a quantitative one.

Relaxing OT's strict ranking by adopting constraint ties does not entail generating nonsensical patterns. For example, we examined whether random ties throughout Hindi and Hungarian would give us random results. We found that, generally speaking, no serious problems arise because the problematic candidates are often eliminated by more than one constraint. For instance, in Hungarian, a tie between anything lower ranked than **BAR3** makes no difference. A tie between **MARKFOC** and **BAR4** (above the tied Faith constraints) results in the same candidates being excluded, although for different reasons. In Hindi, If **MINLINK** and **FAITHQUD** were tied as they are in Hungarian, you would wind up with a monoclausal question being optimal in some scenarios, as we do in Hungarian (unsurprisingly). If any pair of constraints lower than **FAITHQUD** were tied, the results would not change. This exercise would need to be more thorough to be more conclusive, but it does confirm that using these ties in Faith constraints does not set a precedent for random ties that result in nonsensical outputs.

Importantly, these constraint ties do not cause problematic outputs when the input specifies a monoclausal question (instead of a biclausal one). The constraints **FAITHQUD** and **FAITHLF/SYN** essentially solve this potential problem for us by eliminating outputs which are biclausal or which use multiple wh-phrases when only one is necessary. These sanity checks give us further confidence that, in trying to solve the cross-linguistic problem of alternative optimal outputs (for different inputs) within a given language, we do not create problems for the analysis of simple wh-constructions.

Finally, it should be noted that it is possible that the need for tied constraints ultimately is the result of including monoclausal questions in the survey. Recall that this step was taken to establish whether we had appropriately created stimuli which resulted in QUDs being resolved or unresolved. However, it also created a situation where participants were sometimes split on whether they preferred one of the biclausal options or a monoclausal option. If the survey had not included this option then a different pattern would have likely emerged, which might not have required constraint ties. We plan to pursue this question in future work.

3.5.2 Back to language acquisition

This section has established input-output pairings relevant, among other things, to the adult WSM grammar for German. It has answered two of our major questions:

- What cross-linguistically valid analysis of (adult) WSM should we adopt that sheds light on the mapping between semantics/pragmatics and syntax in complex questions?
- What relevant pragmatic and semantic features are relevant to children learning languages with and without WSM?

Not only does this analysis establish the cross-linguistically valid target grammar (of complex wh-questions) German-speaking children are acquiring, it has informed our stimuli design for eliciting both WSM productions and responses to WSM vs. LD structures in German. The remaining chapters will return to experimental work on acquisition of these structures both in English and German. The first questions we address ask whether or not the previous evidence for these examples of syntactic creativity in production and comprehension could be related to the methodology of previous experiments which did not (as far as we know) take these relevant semantic and pragmatic aspects into account. We also address the question of whether performance in the production task is correlated with performance on the comprehension task. These questions are the impetus for the work presented in Chapter 4.

Chapter 4[32]: Syntactic Creativity Errors in Biclausal Questions

Decades of language acquisition research has aimed to uncover how children acquire language with such remarkable speed and accuracy. One valuable source of insight into children's powerful learning mechanism lies in their non-adult-like linguistic behaviors. A well-documented example is overregularization errors. For example, children some- times produce sentences such as *I goed to the zoo* rather than *I went to the zoo* (e.g., Brown, 1973; Marcus et al., 1992), suggesting that the past tense formation pattern has been overregularized to irregular verbs. Children also overregularize argument structure alternation, using intransitive verbs like giggle in transitive frames (e.g., *You giggled me*; Pinker, 1989). This deviation from the target grammar helps us understand how children are actively seeking rules and regularities in the input, which is a critical ability that allows them to flexibly acquire complex linguistic structures that vary across languages (Kam & Newport, 2009; Lidz & Gagliardi, 2015; Saffran, Aslin, & Newport, 1996; Yang, 2004).

While overregularization suggests children misapply the rules of their target language, some of the children's errors suggest they might also mistakenly apply a rule which is not applicable in their target language, but which could be used in another language. In particular, it has been reported that English-speaking children produce the so-called medial wh-questions like *What do you think who the cat chased?* when the intended utterance is a long-distance (LD) wh-question across two clauses like Who do you think

[32] This chapter is the recently published paper, *Syntactic Creativity Errors in Children's Wh-Questions,* which appeared in *Cognitive Science* in the spring of 2020. The reader will notice some repetition from the general introduction to this dissertation as well as some variation in labels. For example, we use WSM rather than SynSM to describe the construction. The reader may want to skip sections 1.0 and 2.0 and go directly to the experimental part of the paper (section 3.0) starting p. 149.

the cat chased? (Grolla & Lidz, 2018; Thornton, 1990). Here, the children's version resembles question formation in the so-called wh-scope marking languages, such as German or Hungarian, where the initial wh-phrase serves as a "dummy" wh-phrase, whose role is simply to indicate the scope position for the medial, "true" wh-phrase (Lutz, Meuller, & von Stechow, 2000). In this paper, we will refer to these unique non-target-like linguistic behaviors as *syntactic creativity errors* (Schulz, 2011).

At first glance, using a rule or structure that does not belong to the target language may appear to be a distraction for language learners rather than a demonstration of an effective learning mechanism. However, the syntactic creativity phenomenon may be interpreted as evidence that language learning is guided by innate knowledge and learning biases. For example, a core assumption in the Principles and Parameters theory (Chomsky, 1981) is that the innate language faculty provides learners with a list of all parameters that define potential syntactic variation, and learners need to select the right parameter setting based on the language input. This suggests that innate linguistic knowledge provides top-down guidance that constrains the hypothesis space such that children are informed of what type of grammatical variation they should consider and evaluate against input data (for further discussions, see Baker, 2005; Fodor, 1998; Newmeyer, 2005; Snyder, 2007; Yang, 2004). In this sense, syntactic creativity errors may provide important insights on the language learning mechanism in similar ways to the overregularization errors. We will refer to this perspective as the *parametric acquisition hypothesis*.

The present paper aims to evaluate this hypothesis with a focus on medial wh-questions in English-speaking children. As the review below illustrates (Section 2), this phenomenon is one of the most widely known examples of syntactic creativity in the language acquisition literature. This phenomenon is of particular interest because of reported evidence based on both production and comprehension behaviors. This would be a natural outcome if children had mis-set their wh-question parameter in such a way that medial wh-questions were part of their grammar. We report on a novel story-based sentence elicitation experiment (Section 3) and a comprehension experiment (Section 4) that strive to improve upon previously used methodology. We will show that children's overall performance is more adult-like than previous research has suggested but confirm that children produce a relatively small percentage of medial wh-questions. We also show that evidence for medial wh-questions in production and comprehension within subjects is not correlated (Section 5), and argue that this phenomenon receives a more straightforward explanation as a reflection of immaturity in sentence production mechanisms. Section 6 will present discussions of implications for theories of grammar development and language processing.

2.0 Syntactic creativity and medial wh-questions in child English

Syntactic creativity has been attested in many different forms in child English productions. Examples include children's production of possessor extraction such as (1) as described by Gavruseva and Thornton (2001); that-trace violations (2) as illustrated in Thornton (1990); resumptive pronouns (3) per McKee and McDaniel (2001); and subject auxiliary inversion in embedded interrogatives (4) described by Ambridge, Rowland, Theakston, and Tomasello (2006) and Pozzan and Valian (2017). For related observations in French, see Labelle (1990).

(1) Whose do you think ball went in the cage?

(cf. Whose ball do you think went in the cage?)

(2) What do you think that's under the swimming pool?

(cf. What do you think is under the swimming pool?)

(3) This is the woman that Grover talked to her.

(cf. This is the woman that Grover talked to.)

(4) Katie wanted to know what was her brother building.

(cf. Katie wanted to know what her brother was building.)

These structures are ungrammatical or marginal at best in English, but they are grammatical in other languages. Possessor extraction constructions are attested in Hungarian and Chamorro (Chung, 1991; Szabolcsi, 1983). That-trace sentences are grammatical in Arabic, Basque, and many Romance languages (Kandybowicz, 2006; Rizzi, 1982). Resumptive pronouns can be used in sentences like (3) in Irish, Hebrew, and Palestinian Arabic (Borer, 1984; McCloskey, 1990; Sells, 1984; Shlonsky, 1992). Finally, subject– auxiliary inversion in embedded interrogatives is grammatical in Spanish (Torrego, 1984) and Italian (Rizzi, 1996). Furthermore, syntactic creativity is not limited to sentence production. For example, it has been observed that English-speaking children interpret subject-less sentences (e.g., Play with blocks) as null subject sentences (meaning They play with blocks) rather than imperatives (Orfitelli & Hyams, 2012). For sentences that contain disjunction under negation (e.g., John didn't drink coffee or tea), Japanese and Turkish children interpret them to mean neither coffee nor tea, like adult English speakers would, even though that interpretation is not available for adult Japanese or Turkish speakers (Geçkin, Thornton, & Crain, 2017; Goro & Akiba, 2004). Russian-speaking children also behave like English speakers as they allow backward coreference between a pronoun in a fronted adjunct clause and the matrix clause subject (e.g., *While he$_1$ was reading, John$_1$ ate an apple*), even though this coreference is not grammatical in the adult Russian counterpart (Kazanina & Phillips, 2001).

Despite the wide-ranging evidence for syntactic creativity, few of these phenomena have gained support from both production and comprehension data. Of the cases reviewed above, the null subject phenomenon is the only one that has been attested in production as well as comprehension. This may simply be due to practical constraints on methodologies. In production studies, it may not be methodologically feasible to elicit certain structures in experimental or naturalistic contexts (e.g., backward anaphora). In comprehension studies, children (as well as adults) may be able to coerce a plausible interpretation for ungrammatical sentences (e.g., that-trace violations), even though their grammar does not allow such structures.

However, in the absence of evidence in both production and comprehension at the same stage of development, syntactic creativity errors could also result from children's immature language processing mechanisms. Recent developmental psycholinguistics research has uncovered how children can sometimes fail to comprehend or produce sentences despite their sophisticated linguistic knowledge (for reviews, see Omaki & Lidz, 2015; Snedeker, 2013). These errors may reflect either immature procedural mechanisms for sentence comprehension or production, or immaturity in cognitive resources (e.g., working memory, cognitive control) that are critical for supporting real-time language use. For example, Bloom (1990) and Valian (1991) have suggested that production of null subject sentences in English-speaking children may reflect children's attempts to reduce sentence length due to their immature sentence planning abilities (for discussions, see

Frazier, 2015; Hyams & Wexler, 1993). However, recent comprehension evidence for a null subject grammar has cast doubt on this performance explanation (Orfitelli & Hyams, 2012). This processing hypothesis, therefore, should be considered in absence of evidence supporting the parametric hypothesis. These alternative explanations highlight the significance of cross-modal evidence for syntactic creativity errors in understanding the actual source of the errors.

The presence of cross-modal evidence is one of the reasons children's use of medial wh-questions has received much attention in the literature. In sentence elicitation experiments, Thornton (1990) observed that English-speaking children between 3;11 and 5;9 produce non-adult-like utterances such as (5), which contain a medial wh-phrase. In comprehension, de Villiers and Roeper (1995) conducted a longitudinal study using a question-after-story task with a question like (6), and found that English-speaking 3 to 4-year-old children answered the medial wh-phrase what, instead of the sentence initial wh-phrase how (see also de Villiers, Kotfila, & Klein, 2019; de Villiers, Roeper, & Vainikka, 1990; Thornton, 1995; Thornton & Crain, 1994). McDaniel, Chiu, and Maxfield (1995) used an acceptability judgment task with 3 to 5-year-old children and found that a subset of the children accepted questions like (7). Finally, although the present paper will focus on English, child production of medial wh-question structures like (5) has also been reported in French (Demirdache & Oiry, 2007; Jakubowicz & Strik, 2008), where such questions are ungrammatical.

(5) What do you think what is under this box?
(cf. What do you think is under this box?)

(6) How did the boy say what he caught? Response: A fish!

(7) What do you think who is gonna climb up the steps?
(cf. Who do you think is gonna climb up the steps?)

These behaviors were interpreted as evidence that English-speaking children adopted the so-called wh-scope marking (WSM) grammar[33]. WSM questions are observed in languages including German, Romani (McDaniel, 1989), Hindi (Dayal, 1994), Hungarian (Horvath, 1997), Malay (Cole & Hermon, 1998), Passamaquoddy (Bruening, 2004, 2006), and Polish and Russian (Stepanov, 2000; Stepanov & Stateva, 2006), among others. These questions consist of two (or more) wh-phrases, and the structurally highest one is used to express the scope of the lower wh-phrase(s) over the entire sentence. In (8), we see the typical German WSM structure beginning with was ("what"). A related variation of WSM is wh-copying, which involves a repeated wh-phrase in the sentence initial and medial positions (for a German example, see (9)). These are found in dialects of Afrikaans, Frisian, German, and Romani, among others (e.g., Felser, 2004; McDaniel, 1989).

(8) **Was** glaubst Du, **wen** sie getroffen hat?
what think you **who** she met has?
Who do you think she has met?

(9) **Wen** glaubst Du, **wen** sie getroffen hat?
who think you **who** she met has
Who do you think she has met?

[33] Throughout the remainder of this paper, we use medial-wh as a theoretically neutral, descriptive term indicating the use of a wh-phrase at the left boundary of the embedded clause. We use WSM to identify the use of a Wh-Scope Marking grammar. While WSM always includes medial wh-phrases, the use of a medial wh- phrase does not entail WSM.

A variety of formal analyses have been proposed for the underlying syntactic and semantic structures of (8) and (9). Under the Direct Dependency Account, the medial wh-phrase *wem* ("whom") undergoes covert movement to replace the wh-expletive was (McDaniel, 1989). Alternatively, the Indirect Dependency Account posits that the sentence initial wh-phrase is a true wh-phrase that overtly moves from the complement position of the matrix clause verb to the sentential scope position. The indirect scope relation between the two wh-phrases is established through either clausal adjunction of the embedded clause (Dayal, 1994) or a doubling structure where the wh-scope marker is considered to form a complex phrase that takes the embedded clause as its complement (Stepanov & Stateva, 2006). In either of these analyses, the two (or more) wh-phrases interact with each other in the course of syntactic derivation to establish the sentential scope for the medial wh-phrase, via a grammatical process that does not exist in strict LD movement languages like English. Following Thornton (1990), we will generally group together WSM and wh-copying, as the use of either structure by children would constitute an instance of syntactic creativity. We will return to this issue in Sections 3.6 and 6.

English-speaking children's production or comprehension behaviors may in fact indicate that WSM questions are allowed in their (non-target-like) grammatical knowledge, in line with the parametric acquisition hypothesis. However, why English-speaking preschoolers entertain this non-target-like parameter setting remains unclear. One plausible explanation comes from competition-based models of syntactic development (Legendre, Vainikka, Hagstrom, & Todorova, 2002; Yang, 2002). For example, under a probabilistic model (Yang, 2002), different parameter settings are simultaneously evaluated and the competing parameter settings will be probabilistically used to parse the input. When this parse succeeds, that parameter is rewarded by a boost in probability; if the parse fails, the parameter is punished and its probability decreases. As a result, this process will continue over time and across many utterances, until the parameter setting that is most compatible with the input eventually wins the grammar competition (for critical discussions of this approach, see Snyder, 2007). This could be because children rarely hear biclausal questions and thus do not receive adequate evidence either for or against WSM or LD movement until they are older, they have yet to accurately establish a parameter setting. Under this account, English-speaking children are predicted to use WSM structures in both production and comprehension, and presumably at a roughly comparable rate, until the English long-distance movement parameter setting is selected as the winner. Alternatively, usage-based accounts of language acquisition suggest that children might not be integrating two clauses completely in production. Diessel and Tomasello (2000), for example, suggest that children's lack of experience with certain verbs (specifically those which take embedded clause complements such as think and say) leads them to treat these verbs more like epistemic markers than matrix verbs until they truly learn them. Under this account, children's production and comprehension performance need not be related: production of medial-wh phrases and responses to questions with medial wh-phrases would occur for different reasons.

A detailed comparison of WSM comprehension and production behaviors has rarely been attempted in previous work, partly because the production and comprehension studies were conducted by different labs who were not directly testing the same phenomenon and with different experimental settings (with the notable exception of Thornton & Crain, 1994). The present study aims to fill this empirical gap using a story-based experimental paradigm that taps production (Experiment 1) and comprehension (Experiment 2). In addition, Experiment 3 directly tests whether the production and comprehension of complex wh-questions are correlated either across participants or within individuals. As will be discussed below, our experiments aim to provide a more rigorous assessment of the reliability and validity of syntactic creativity errors in complex wh-questions. We will thoroughly test the parametric hypothesis and, in absence of evidence for it, we will consider a processing explanation.

3. Experiment 1: Story-based sentence elicitation task

Experiment 1 used a story-based sentence elicitation task, which was modeled after Thornton (1990) as well as research on adult second language learners by Schulz (2011). The parametric acquisition hypothesis predicts that some children will produce medial wh-questions due to incorrect or competing parameter settings, as has been reported in previous experimental studies (Thornton, 1990, 1995; Thornton & Crain, 1994). These studies on children's medial-wh production used a sentence elicitation task in which a child observed an event (e.g., the experimenter gives everyone a marble) while a puppet was hiding and unable to see what was happening. When the puppet came out, the child was prompted to ask the puppet a wh-question about what the puppet thought happened (e.g., Who do you think has the blue marble?). Because the puppet did not see the event and had to guess what happened, this procedure made it felicitous to ask biclausal wh-questions (for more discussions on the felicity of biclausal wh-question elicitation, see Demirdache & Oiry, 2007).

The present study hoped to build upon the methodology of previous work in a variety of ways. First, previous studies sometimes used a varying number of trials for each participant, which makes it difficult to assess how robust the WSM structure is within and across children. In Thornton (1990), 14 children participated in the relevant experiment and each produced several long-distance wh-questions, but most of the reported findings focused on two participants. Thornton and Crain (1994) report data from 15 children who participated in at least four trials each, but some of them participated in more trials, depending on how the child was performing.

Second, these studies did not systematically manipulate the gap position (e.g., subject vs. object) or the wh-phrase type (who vs. what), even though these structural details are important for understanding the generality of WSM strategies. For example, if the WSM structure is grammatical for children, it should be equally available for both subject and object extractions, but the lack of systematic manipulation for gap position makes it difficult to address this question. Relatedly, it is also plausible that the wh-phrase type (who/ what) affects children's question formation because in WSM languages, the scope marker is typically the counterpart of what (or how). For this reason, children may produce more medial wh-questions when they start the question with what, rather than who. To address these issues, the current experiment used the same number of trials with a greater number of children than in previous studies, and it systematically manipulated the gap position and wh-phrase type.

Finally, the medial wh-question responses in previous studies may have been inflated to some extent by the particular phrasing of the elicitation prompt. For example, when the elicitation target sentence was *Who do you think is in the box?*, children were told to ask Cookie Monster what/who he thinks. One issue with this phrasing is that this fragment sounds unnatural when the wh-phrase is *who*. In addition, the use of the matrix verb *think* may lead children to produce the frequently used combination of words *what do you think* as an idiomatic chunk without internal syntactic structure. As a result, even when children produce an apparent WSM structure (e.g., *What do you think what is in the box*?), this may in fact consist of two separate questions in sequence (e.g., *What do you think? What is in the box?*).

To address these issues, our sentence elicitation prompt was presented as a sentence preamble that children needed to repeat and complete. This technique is widely used in the adult psycholinguistics literature to investigate incremental sentence production mechanisms (e.g., Bock & Miller, 1991; Ferreira & Dell, 2000). In addition, this preamble prompt used the verb *believe*, which is less frequent in child-directed speech (see below) and therefore unlikely to be stored as a phrasal chunk. We note that the discouragement of chunking does not necessarily resolve the ambiguity between sequential questions and WSM questions, and we will return to this point in Section 3.6. Unlike the previous studies, the current study included an interactive story design where children were required to ask questions to the story characters to advance the narrative. This allowed the repetitive nature of the task to be a natural part of the story and provided entertaining motivation for asking the questions.

3.1. Participants

In all, 33 children between the ages of 4;0 and 6;1 (mean 5;0) were recruited from the Baltimore area. This age range was selected to approximately match the age range used in Thornton (1990). Data from three children were excluded from analyses because of the failure to complete the task (n = 1), or due to having been exposed to a language that productively uses WSM (n = 2). This left a remainder of 30 children. Twenty adult native English speakers served as a control group and were recruited from the Johns Hopkins community (*M*age = 19;6).

3.2. Procedure

Children were presented one continuous story through a Microsoft PowerPoint slide show on a Lenovo laptop computer. All sessions were filmed with an HDR-CX160 SONY Handycam. Additionally, audio recordings were made with Olympus Digital Voice Recorder WS-822.

Children were first introduced to three plush dolls. The experimenter explained to the child that they were going to help the three characters have an adventure in the fairy world inside the experimenter's computer. As puppets were placed behind a screen, the characters appeared on the PowerPoint slides. Once all three puppets were on screen, a fairy greeted them and explained the rules of the fairy world. The premise of the story was that one of the three characters (henceforth the target character) was under a witch's spell, and due to this spell, the target character was allowed to speak only when two conditions were met: (a) the child asked the target character a question and (b) the child and the fairy had to work together to formulate the question. This allowed the fairy to give the child the beginning of the question. Children spoke into a non-functional microphone, which served as a tool to make the child feel connected to the story.

Children were given two practice trials to ensure that they understood the task. Target trials were elicited in one of two forms in different orders. Version A elicited questions in the order who-object, what-object, who-subject, and what-subject (see Section 3.3. below). Form B elicited questions in the reverse order. The questions that we aimed to elicit were the same in each form, but the story changed slightly to accommodate the reversed order. The introduction, practice trials, and test trials were all incorporated into the continuous story. The experiment took 30 minutes on average. Children greatly enjoyed the task and were unperturbed by its complexity. No child was unable to complete the task because they did not understand it. Many asked if the story could keep going or if they could do it again to see if the characters ever changed their minds.

3.3. Materials

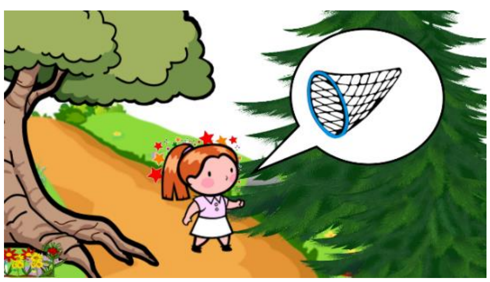
3.3.1. Materials: Story design

The story consisted of 12 problem-solving scenarios, in which two of the three characters expressed different opinions first. Children needed to ask the third character what she believed they should do to solve the problem. An example of a scenario including display images and event description is presented in Table 1, and the full story and question stimuli can be found in Appendix A. Each scenario was constructed using the following story design. Two of the three characters were always asked monoclausal versions of the question first (e.g., *What can stop the witch?*). The experimenter modeled the question for the child and had the child repeat it to the characters, in order to ensure that the child knew what information to elicit from the characters. The experimenter modeled the question explicitly for the first character, but encouraged the child to ask the second character on their own. This allowed the experimenter to help if necessary because it was not a target utterance, but it also gave the child the freedom to use what structure they liked. The two characters answered the question with an embedding sentence *I believe...*, which helped to establish that they held different beliefs. This conflict was instrumental in making the target question to the third character (e.g., *What do you believe...*) felicitous as a contrastive question about the character's belief. In order for a question about the character's belief to be felicitous, it had to be clear that characters were reporting their beliefs rather than their explicit knowledge. While the characters always did what the target character suggested, it was not the case that they always made the right decision.

3.3.2. Materials: Question design

The story was designed to elicit a total of 12 questions. The type of extraction (subject vs. object) and wh-phrase (what vs. who) were manipulated to create four different types of wh-extraction: who-object, what-object, who-subject, and what-subject. Examples of each type of question are illustrated in Table 2. This was a within-subjects design, and the same target questions were elicited for every participant in one of the two possible orders.

Table 1. *Sample trial sequence for 'What do you believe can stop the witch?'*

	The experimenter explains that the three friends are wandering down the path, when the witch appears. She is flying around and won't let them past. We need to find out what can stop the witch.
	The experimenter then prompts the child to ask characters 1 and 2 *What can stop the witch?* Each character responds with what they believe can stop the witch (Character 1: *I believe a net can stop the witch!* Character 2: *I believe a rope can stop the witch*). The experimenter points out that they need to know what the target character believes to break the tie.
	The experimenter reminds the child that to help the target character talk, they have to work with the fairy puppet. The fairy then gives the lead in. *Ok the question starts like this: What do you believe…* The child repeats the preamble *What do you believe* and then completes the question.
	In response to the target question, the target character gives her answer: *I believe a net can stop the witch, because a net covers lots of space*.

3.4. Data coding and analysis

Children's questions were transcribed by the first author, and syntactic details of the utterance were coded. There were a total of 354 child utterances and 240 adult control utterances. Two child utterances were excluded due to an obvious absence of clausal embedding (e.g., questions with the conjunction and, such as *Who do you believe and who should we trust?*), and two additional utterances were removed because the type of extraction could not be classified (e.g., *What do you believe that..umm... so we can break the spell?*). Responses with semantically appropriate lexical substitution were included in the analysis, as will be discussed in detail below. Our analyses will focus on the details of finite, biclausal wh-questions with a single wh-phrase (i.e., sentence-initial only) or sentence-initial and medial wh-phrases. In total, these utterance types accounted for 93.3% of the utterances. All other error types are reported in Appendix B (Other Production Errors, Experiment 1) along with the by-subject mean production rate. For apparent medial wh-questions, we also coded whether the two wh-phrases were the same or different, as well as whether the embedded question showed subject–auxiliary inversion (see below for the rationale).

Table 2. *Wh-extraction type examples*

Wh-phrase	Extraction Type	Example
What	Object	*What do you believe the bear wants to eat___?*
What	Subject	*What do you believe ____ can make the witch happy?*
Who	Object	*Who do you believe we should ask___?*
Who	Subject	*Who do you believe ____ can get the wand?*

There were 33 instances where children changed the type of extraction in the question. For example, in one trial, the characters must choose an instrument to play because music will make a snake fall asleep. For the target question in (10), some children produced (11), which was a contextually appropriate question, but the extraction changed from object to subject extraction.

(10) What do you believe we should play?

(11) What do you believe can make the snake fall asleep?

In these situations, which result from the open-ended nature of the production task, we included these utterances as belonging to the category which matched the child's production. An utterance such as (11) was considered in further analysis as "what-subject" rather than the target "what-object."

For the purposes of statistical analysis, Versions A and B were collapsed, as there was no difference in performance between the two. The adult data were at ceiling, with 99.6% using long-distance constructions without a complementizer (one adult included the complementizer *that* in an object extraction). Thus, we report the adult data to illustrate that the task successfully elicited the target structure in adults, without using population type as a factor. The utterances were coded in a binomial fashion, with medial wh-questions as (1), and other questions as (0). Data were submitted to a mixed effects logistic regression model (e.g., Kruschke, 2015), using the lme4 package version 3.2.2 (Bates, Maechler, Bolker, & Walker, 2015) in the R statistical analysis environment (R Core Development Team, 2015). The model used wh-phrase (*who/what*) and extraction type (subject/object) as fixed effects. The model also included a fixed effect of age, which was coded as a numerical variable and was centered as well as scaled. The random effects structure of the model included crossed participant and item effects. The model was run with the maximal random effects structure that would converge (Barr, Levy, Scheepers, & Tily, 2013), minimally including random intercepts for participant and item. Extraction type, wh-phrase, and age were used as random slopes for participants.

Finally, we coded for certain disfluencies in productions. These included the use of a Filler such as *uh/um*; Word Repetition or the repetition of a single word (*What do you believe can can help the fairy*); Restart or the repetition of more than one word (*What do you believe... What do you believe that can save the fairy*); and finally, Self-correction for cases where the child repeated a phrase and changed at least one of the words (*What should you... What do you believe that can save the fairy?*). Each utterance was coded as either including a disfluency (1) or not (0). We separately coded for when disfluencies appeared specifically at the embedded clause boundary and when they appeared elsewhere. Disfluencies have been used as indicators of processing load (e.g., McDaniel, McKee, & Garrett, 2010). Their frequency and locality could therefore be indicative of the way children are processing these structures. The data were submitted to a mixed effects logistic regression with disfluency as a fixed effect. This model also included *Age* as a fixed effect (again, centered and scaled) as well as *Participant* and *Item Number* as random effects.

3.5. Results

The elicitation task was an open-ended task affording subjects alternative acceptable options for answers. For example, while the target structures were questions like *What do you believe we should look for?,* children's only prompt was "*Can you ask* *her? The question starts, What do you believe…*" Children were, therefore, free to ask using whatever structure they chose, including the target but also including non-target structures which can be grammatical or ungrammatical. Furthermore, errors can be relevant or irrelevant to question formation. There are, therefore, four types of relevant utterances: grammatical, target-like (*What do you believe we should look for?*), grammatical, non-target-like (*What do you believe would be good to look for?*), ungrammatical for reasons unrelated to question formation (*What do you believe to look for?*), and medial-wh (*What do you believe what we should look for?*). Our results showed 78% of the total productions included only a single sentence-initial wh-phrase and were either grammatical or included errors unrelated to question formation. This suggests that the majority of children's productions were efforts at LD constructions. The errors that were not related to wh-questions will be described in detail later. We conclude from this general overview of the results that the task of eliciting LD structures was not overly difficult and that children are capable of producing these structures like adults. Children are able to do the task, but this also has the virtue of showing their limitations. This task, therefore, precisely taps into any relevant limitations children may have.

We now turn to the different types of syntactic creativity children showed in their wh- production. Table 3 shows the distribution of productions using a single, initial wh-phrase and productions which include a medial wh-questions in adults and children while collapsing the wh-phrase type (*who* vs. *what*) as well as the extraction type (*subject* vs. *object*). Additionally, when there was only one wh-phrase in the sentence initial position, the figure indicates whether the questions contained the complementizer that.

Table 3. *Distribution of initial and medial wh-phrases by children and adults*

		Adult	Child
Single, sentence-initial Wh-phrase	*By-subject mean productions with sentence-initial wh-phrase only*	.996	.45
	By-subject mean productions with sentence-initial wh-phrase + complementizer that	.004	.33
Sentence-initial + medial wh-phrases	*By-subject mean productions with both sentence initial and medial wh-phrases*	0	0.206
	By-subject mean productions with both sentence initial and medial wh-phrase + complementizer that	0	0.014

3.5.1. Complementizer production

While the use of complementizers was not our main question, the visual inspection of the data revealed an interesting adult–child contrast. Children produced a number of questions with overt complementizers (34.4%). One single adult participant also produced a question with an overt complementizer. Crucially, child participants were producing the complementizer *that* even though it was not included in the preamble given by the fairy[34]. This suggests that the children produced the complementizer independently, and that they actively integrated the preamble into their utterance generation process. We also note that in English the use of the complementizer is grammatical only with object extraction, but some of the children's (ungrammatical) utterances came from the subject extraction condition (21.4% of total, 75 of 120 productions with overt complementizers, *What do you believe that can stop the witch?). In other words, this is an instance of a that-trace violation (Chomsky & Lasnik, 1977), hence constituting another demonstration of syntactic creativity which results from the open-ended nature of the task. It should be

[34] Five of these utterances included a medial wh-phrase in addition to the complementizer: for example, *What do you believe that that that what what can make the bear full?* While these are not grammatical in either WSM languages or English, it may simply reflect children's disfluency in formulating complex wh-questions.

further noted that Table 3 illustrates the fact that while adults are consistent in their production types, children show more flexibility when faced with an open-ended task of this nature.

3.5.2. Medial wh-questions

Children as well as adults primarily asked questions with a single wh-phrase. However, children's questions included medial wh-phrases 22% of the time (Table 3). Examples of questions from each condition are shown in Table 2. The overall pattern is similar across conditions: Children produce target-like biclausal wh-questions with a single wh-phrase more often (78% overall) than questions with multiple wh-phrases (22% overall). Table 4 breaks down the distribution of medial wh-questions across wh-phrase and extraction types.

The mixed effects logistic regression analysis of utterance type is presented in Table 5. There was a significant effect of wh-phrase, indicating that significantly more medial wh- questions were produced when the initial wh-phrase was who. We return to this somewhat unexpected observation in Section 3.6. We also found a significant effect of age (see below), indicating that younger children were more likely to produce medial wh-questions than older children. There was no significant effect of extraction type, and no significant interaction among the fixed effects.

Table 4. *Distribution of medial wh-questions across wh-phrase and extraction types*

Condition	*By-Subject Mean Productions with Both Sentence Initial and Medial Wh-Phrases*
What-object	.22
What-subject	.11
Who-object	.29
Who-subject	.27

Table 5. *Results of mixed effects logistic regression model*

	B	Error	Z value	Pr(>|z|)
Wh-phrase	1.11	.50	2.2	.03*
Extraction	.56	.70	.79	.43
Age	-1.28	-.56	-2.29	.02*
wh-phrase × extraction	-1.36	.91	-1.50	.13
Wh-phrase× Age	-.37	.55	-.67	.50
Extraction× Age	-.05	.76	-.06	.95
Wh-phrase × Extraction × Age	.31	1.02	.30	.76

Note: † $p < 1$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

3.5.3. Effect of age and individual patterns

We found a main effect of age on the rate of production of medial wh-questions. To further investigate the nature of this effect, we examined the rates of medial wh-questions produced by individuals across children of different ages (Fig. 1). This individual analysis illustrates two important findings. First, production of medial wh-questions did occur, though inconsistently, both across and within individuals: Three participants produced medial wh-questions in more than 75% of the trials, while the majority of children produced this type of structure less than 25% of the time (n = 20). The remaining seven participants produced medial wh-questions between 25% and 75% of the time.

Second, those who often produced medial wh-questions were roughly at or below age 5 (60 months), suggesting that children learned to primarily produce target-like English wh-questions after age 5.

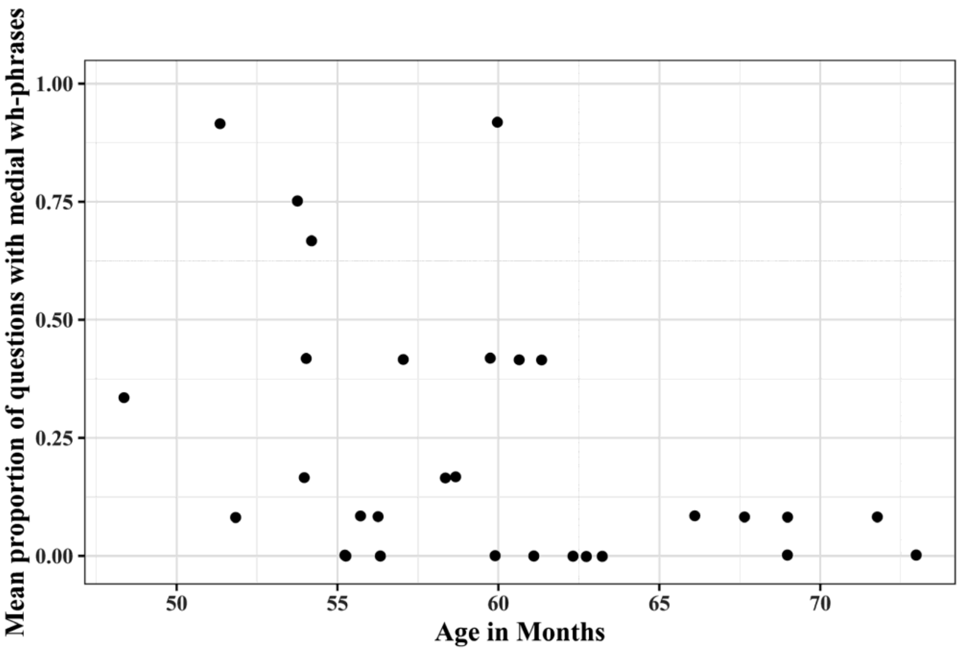


Fig. 1. Proportion of utterances with medial wh-phrases by participant.

3.5.4. Distribution of initial and medial wh-phrase type

In languages that use WSM structures, there is typically one wh-phrase that is used as a sentence initial, dummy scope marker, while any type of wh-phrase may be in the sentence medial position. On the other hand, in wh-copying languages, it is possible to produce a copy of the true wh-phrase in both initial and medial positions (e.g., Murphy, 2016). To investigate the distribution of wh-phrase types in our data, children's medial wh-questions were divided into questions that used the same phrase twice, and questions that employed two different wh-phrases (Table 6). Note that this is the distribution of only medial wh-questions, which account for 22.5% (78 out of 350) of total utterances.

Table 6 shows that, of the 350 utterances analyzed, there were only 14 instances of utterances which used two different wh-phrases. Thus, the majority of utterances which included medial wh-phrases used the same wh-phrase twice (82.1%). The 14 utterances using different wh-phrases varied in wh-phrase combination, and these combinations included what + where (1), what + how (2), who + what (5), who + how (1), and who + which (5).

Table 6. *Distribution of medial wh-questions*

	Utterances with medial wh-phrase	
Type of extraction	**Same wh-phrase (percentage)**	**Different wh-phrase (percentage)**
***What* object**	18 (94.7%)	1 (5.3%)
***What* subject**	9 (81.8%)	2 (18.2%)
***Who* object**	16 (84.2%)	3 (15.8%)
***Who* subject**	21 (72.4%)	8 (27.6%)
Total	64 (82.1%)	14 (17.9)

This is not a pattern typically found in WSM languages. Scope markers are typically a minimal wh-word, such as *what* (Horvath, 1997). Thus, while instances of *what* + *what* are common in WSM languages, *who* is not a typical scope marker. However, it is possible to begin a question with *who* in a copy construction (e.g., *Who do you believe who we'll meet here?*). Superficially, therefore, children's medial wh-questions appeared more similar to wh-copy constructions, which are attested in many of the same languages as WSM (Felser, 2004). It is important to note that *what* + *what* questions are ambiguous between WSM and copy constructions, as the initial what may be the scope marker. However, it is also possible that children produce questions resembling copy constructions due to reasons (such as immature processing mechanisms) that are unrelated to their grammatical knowledge per se (see Section 3.6.).

3.5.5. Subject–auxiliary inversion

Another possibility is that these apparent medial wh-questions consist of two individual questions, as discussed in Section 3 (e.g., What do you think? What is in the box?). One way to address this concern is to examine the rate of subject–auxiliary inversion. If the two clauses are intended as two separate questions, then the second clause is not embedded and as such should contain subject–auxiliary inversion. The rate of subject–auxiliary inversion in second clauses is shown in Table 7. Children were not consistent in whether or not they inverted the subject and auxiliary verbs in medial wh-questions. Of the 78 utterances with medial wh-phrases, 40 were subject extraction, which would not have elicited subject–auxiliary inversions. Of the remaining 38, there were 22 responses (57.9%) that included subject–auxiliary inversion.

Table 7. *Proportion of second clauses containing subject-auxiliary inversion*

Extraction Type	Total Utterances	Inversion	No inversion	Dropped Aux
Subject	40	*NA*	*NA*	3
Object	38	23	11	4

We further investigated the distribution of the cases of medial wh-phrase production that included subject–auxiliary inversion in the embedded clause, as shown in Table 8.

Table 8. *Distribution of inversion in the embedded clause*

Participant Age in months	Inversion/medial Wh-productions	Condition in which inversion occurred
54	6/8	*Who* Obj & *What* Obj
54	2/5	*Who* obj
54	2/9	*What* obj
57	3/5	*Who* obj & *what* obj
59	2/2	*What* obj
60	2/11	*Who* obj & *What* obj
61	5/5	*Who* obj & *what* obj
68	1/1	*What* obj

If we limit our analysis to these utterances which include no subject–auxiliary inversion in the second clause, the number of potential WSM structures is reduced dramatically. This will be further discussed in Section 3.6.

3.5.6. Disfluencies

The results of the mixed effects model testing whether disfluencies predict participants' final utterances did not suggest a relationship between speech disfluency and any particular error type, but did show a significant effect for target-like productions ($p < .05$). In other words, while a production could be target like even it included certain disfluencies (e.g., *What do you believe will help... will stop the witch?*), a participant with fewer disfluencies was more likely to produce adult-like structures. Furthermore, we looked specifically at disfluencies that occurred at the embedded clause boundary and made a similar finding. While these disfluencies at the clausal boundary did not predict any error type either, there was, again, a significant relationship between these errors and target-like productions ($p < .05$). The similarities in these results are not surprising given that 86% of disfluencies overall occurred at the clausal boundary. Table 9 shows the results of the mixed effects logistic regression model examining the relationship between target-like productions and disfluencies.

Table 9. *Results of logistic regression model examining if disfluencies predict medial-wh*

	B	Error	Z value	Pr(>|z|)
Disfluencies	-1.01	.51	-1.96	.05*
Scale(age)	1.67	0.50	3.32	.00***
Disfluency × Age	-.08	.39	-.22	.83

Note: † $p < 1$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$
Note. **$p \leq 0.01$; ***$p \leq 0.001$. -0.18 0.19 -0.96 0.34

3.6. Discussion

The results from Experiment 1 show that children primarily produced biclausal wh-questions with a single sentence-initial wh-phrase (78%). While these are not all grammatical as they include structures which are ungrammatical for reasons unrelated to question formation, they are all efforts at the adult, LD structure. However, the results also showed evidence for production of medial wh-questions (22%). Though medial wh-questions appeared much less frequently than questions with a single wh-phrase, a few children produced them systematically. In general, younger children produced medial wh-questions more frequently, though a greater number of children in the same age range, in fact, used very few medial wh-questions. Furthermore, the majority of medial wh-questions used the same wh-phrase twice, and many utterances included subject–auxiliary inversion in the second clause.

The results of our study were largely consistent with previous findings (e.g., Thornton, 1990; Thornton & Crain, 1994) that some children produced questions with medial wh- phrases. Furthermore, rather than being spread across all children, the medial wh-phrase pattern appeared consistently in a few individuals. While it is consistent with the parametric hypothesis that a few of the participants' grammars might include WSM, these observations are not conclusive evidence that WSM is allowed in Child English grammar. For example, the high rate of questions with subject–auxiliary inversion in the second clause suggests the possibility that at least some of these utterances contain two separate questions (*What do you think? What should we look for?*). The subject–auxiliary inversion rate per se may not, however, provide a reliable estimate of sequential questions, as children at this age range are known to make subject–auxiliary inversion errors in embedded clauses (see Section 1 above). Pozzan and Valian (2017) conducted a sentence elicitation experiment with 3 to 5-year-old children and found that children produced subject–auxiliary inversion in embedded wh-questions 27% of the time. Thus, we cannot assume that the presence of subject–auxiliary inversion indicates sequential questions.

Another way to determine whether these utterances are made of sequential questions is to examine their prosody. It would be reasonable to assume that the prosodic pattern of sequential questions would differ for questions with embedded clauses in child English. However, there are at least three problems with using children's prosody to determine whether these utterances are sequential questions. In the first place, comparing children's prosody of such questions to adults' is not viable as child prosody has been shown repeatedly not to follow typical adult patterns in syllable stress (e.g., Gerken, 1996). Thus, we cannot simply look for typical adult indicators of clausal boundaries and must instead establish what patterns occur at clausal boundaries in child English. This, however, leads to the second problem: Our results indicate that children's productions varied greatly in prosody between participants and thus comparing them to each other is problematic. Finally, the quality of the recordings was insufficient to conduct as thorough an investigation of prosodic contour as would be necessary. Nevertheless, an analysis of the prosody of a single participant whose recordings were of good enough quality to investigate and who also produced both questions with medial wh-phrases and questions with that at the embedded clause boundary (see Appendix C, Prosodic Analysis) suggests that *believe* + *that* and *believe* + *what* utterances are distinct from questions with embedded clauses and therefore likely not instances of WSM.

Experiment 1 also revealed that while the extraction type (subject vs. object gap) did not affect the production rate of medial wh-questions, significantly more medial wh-questions occurred when the initial wh-phrase was who rather than what, even though who is not reported as a possible scope marker in WSM languages. We offer two suggestions for what might lead to this difference. In the first place, the greater difficulty with who questions may reflect the fact that children generally use the wh-phrase who less often than what (Stromswold, 1995), and hence are less skilled in formulating wh-questions that start with who. In the second place, while our paradigm eliminated the possibility of asking *What do you think?* as a chunk, as previously described, we created a situation where asking

Who do you believe? is felicitous. The child has just seen character 1 state an opinion which character 2 contradicts. It is, therefore, possible that the child is posing the question *Who do you believe: character 1 or character 2?* and then adding a second, monoclausal, question. The result is the surface pattern *Who do you believe who will we meet here*, not because the child is using an alternative grammar, but because they have posed two questions in a row.

Another interesting finding, unrelated to WSM, was that some children produced (ungrammatical) extractions of an embedded subject wh-phrase as in *What do you think that can stop the witch?*; that-trace violations. This pattern was also observed in Thornton (1990) as well as in McDaniel et al. (2010). The replication of this finding in the present study suggests that this constitutes evidence for another instance of syntactic creativity. We will revisit this issue in Section 6.

Further results indicated that the more disfluencies present in a child's production, the less likely she was to produce a target-like question. This suggests that the processing load made apparent in the child's disfluencies also affects their ability to produce target-like questions, though further research would be necessary to solidify this claim. The further finding that disfluencies at the embedded clause boundary make up a significant proportion of these disfluencies (86%) suggests that the processing load is frequently heaviest at the embedded clause boundary, as suggested by McDaniel et al. (2010).

Finally, in our efforts to avoid *what do you think*, which children might have stored as a chunk phrase, we introduced a new verb (*believe*) with which they were almost wholly unfamiliar, as attested in our own corpus study (Section 3.3). However, some external evidence suggests children are treating *think* as a true matrix verb by this age (Diessel, 2004), while believe remains understudied. As suggested by a reviewer, this could have led children to simply repeat the unfamiliar *believe* phrase heard without necessarily integrating it with the overall question, resulting in a monoclausal question preceded by what do you believe as a kind of introductory tag rather than asking two questions in a row, as suggested above. This is in line with Diessel and Tomasello (2000)'s suggestion that prior to mastery as matrix verbs, phrases such as *I think* are akin to evidential markers which accompany otherwise independent clauses. Under this analysis, children in our study would be treating *Who/what do you believe* as such a marker and then adding their true, monoclausal question *what should we look for?* While we cannot eliminate this possibility for 22% of elicited questions resulting in questions with initial as well as medial wh-phrases, we again note that most questions (78%) produced by children contained only a single wh-phrase in the matrix clause (sometimes including an unprompted complementizer: e.g., *What do you believe that we should look for?*). This dominant, adult-like pattern in their production suggests that children were not typically repeating what they were told to say without integrating the two clauses.

In sum, our overall results support the conclusion that there is significant evidence for WSM structures in elicited production, given the open-ended nature of the task. While it is not as widely attested as previous literature suggested, the consistent production of medial wh-phrases by at least some children raises the possibility that these individuals may indeed entertain a (temporary) WSM grammar. Up to this point, therefore, our findings support a parametric explanation for these errors. However, under a parametric hypothesis, we predict evidence in comprehension as well as a correlation between production and comprehension performance. To this end, Experiment 2 will explore the possibility that children interpret grammatical English questions as if they were from a language with WSM.

4.Experiment 2: Question-after-story task

Prior evidence for WSM in children's comprehension comes primarily from their responses to question-after-story tasks. The prediction of the Parametric Acquisition Hypothesis is that some children will have WSM in their grammars and will respond to questions with multiple wh-phrases as if they were examples of WSM in English. De Villiers and Roeper (1995) report that an average of 56% of children's responses to questions such as (12) were answers to the medial wh-phrase; in other words, they answered *what was caught* rather than *how the boy said what he caught* (for related findings, see also de Villiers, Roeper, Bland-Stewart, & Pearson, 2008; de Villiers et al., 1990).

(12) Experimenter: How did the boy say what he caught? Child: A fish!

De Villiers and Roeper (1995) interpret this result as evidence that some children analyze the sentence initial how as a wh-scope marker, while the medial wh-phrase *what* constitutes the true target of the question. Taken together with evidence for production of medial wh-questions, responses to medial wh-phrases as if they were question words (instead of relativizers in the adult grammar) could be taken as evidence for robust presence of WSM in child English grammar and support for the Parametric Acquisition Hypothesis.

Experiment 2, therefore, will elicit responses to questions with multiple wh-phrases and determine whether these responses are consistent with WSM grammars. The task will be a question-after-story task following de Villiers et al. (1990) and de Villiers and Roeper (1995). Building directly on this previous work, the stories will be designed to render questions with embedded clauses felicitous and will be simple and engaging enough for young children to participate and understand. Our presentation highlights three specific desiderata for the task manipulations.

First, it is possible that children's responses to the embedded interrogative are driven by a number of contextual factors that make such responses more felicitous. For example, the manner adverbial how may be generally more compatible with caught the fish than say in the given context. De Villiers and Roeper (1995) addressed this possibility using multiple verbs including ask and learn in addition to say; we replicate this effort using *tell*. To make this *how-tell* association felicitous, it is critical that the context provides potential alternative means of telling. Following de Villiers and Roeper (1995) and de Villiers et al. (1990), our stimuli will also include plausible alternative answers to both an adult-like reading of the question and a WSM interpretation of the question.

Second, it is crucial that event prominence, or what is at issue in the story stimuli, be balanced. Many studies that use truth value judgment tasks have found that children are sensitive to such contextual information. For example, Conroy, Takahashi, Lidz, and Phillips (2009) found that children's pronoun resolution is affected by the discourse prominence of potential antecedents. Moreover, studies on scope ambiguity (e.g., *Every horse didn't jump over the fence*) have found that the availability of the inverse scope reading (e.g., *not every horse jumped over the fence*) varied depending on whether the number of horses that failed to jump over the fence was an important question to be addressed in the current discourse (question under discussion/QUD; for a review, see

Roberts, 2012; for empirical demonstrations of children's sensitivity to this, see Gualmini, Hulsey, Hacquard, & Fox, 2008; Lewis, Hacquard, & Lidz, 2017; Musolino & Lidz, 2006; Viau, Lidz, & Musolino, 2010). Thornton and Crain (1994) investigated this issue with respect to wh-questions like (12). Their results indicate that when stories emphasize the matrix clause event, children reliably give answers to the matrix clause wh-phrase. Together, the reviewed studies suggest that the prominence of events directly affects children's ability to respond to complex questions. Therefore, to provide a fair chance at attesting English-like responses or WSM-like responses, it is important to balance the prominence of matrix clause and embedded clause events.

Third, the medial question response in (12) may have resulted from children's selective attention to the embedded interrogative; children may have simply ignored or forgotten the matrix clause, and treated the embedded portion *what he caught* as a full-fledged question on its own. This strategy, which we dub a non-embedding analysis, will lead to the same response (e.g., *a fish*) as the WSM-like response in contexts where what the boy caught, as well as what the boy said he caught, are identical. On the other hand, if the boy caught one thing (e.g., *a butterfly*) while he said he caught something else (e.g., *a fish*), then the two analyses will yield different responses. Thus, by creating stories in which the boy catches one thing and says he caught another, we will be better able to understand how children are processing these stories, based on their responses.

A study that systematically controls these factors can prove useful to testing the robustness of WSM in child English. To this end, Experiment 2 used story stimuli that were designed not only to balance the prominence of matrix clause and embedded clause events but also to distinguish non-embedding analyses from WSM-like responses (see below for design details).

In addition to addressing these pragmatic details in the story stimuli, Experiment 2 introduced a question type manipulation as in (13) and (14) to measure the baseline bias to providing an answer that bears on the second clause event.

(13) How did Evil Steve tell Detective Sherry what he was gonna steal?
(*how* + *what*)

(14) What did Evil Steve tell detective Sherry that he was gonna steal?
(*what* + *that*)

The question in (13) resembles the type of questions used in previous work, and the question in (14) is an unambiguous long-distance wh-question that should elicit the embedded interrogative response (i.e., *what Steve said he was going to steal*). In other words, for children who are considering both English-like wh-question grammar and WSM grammar, (13) should be effectively synonymous to (14). While this question type manipulation is implemented as a between-participants manipulation, those children should provide the embedded clause interrogative response to (13) at a similar rate to those children who were given the baseline condition in (14).

4.1. Participants

Forty-two children between the ages of 3;11 and 6;1 (Mage = 4;9) were recruited from the Baltimore area. Ten children were excluded for failing to answer at least one filler question accurately. After this exclusion, participants in the medial wh-question group (henceforth, how + what group) ranged in age from 3;11 to 5;10 (mean 4;10). Participants in the single wh-phrase (*what* + *that*) group ranged from 3;11 to 6;1 (mean 5;0). There were 16 children in each group. This age range was selected to overlap with de Villiers and Roeper (1995)'s age range (3;4 to 5;1) and to match the age range of participants in Experiment 1. Twenty adult native speakers recruited from the Johns Hopkins community (Mage = 19;6) served as controls.

4.2. Procedure

The experimenter first introduced children to a puppet called Hillary Hippo, who then "went into the computer." Hillary provided the instructions and gave them a practice question. The practice story was designed to ensure that the children understood the procedure. The practice question, *Who brought Lilly the surprise?*, was a monoclausal question using the wh-phrase who, so that the structure did not overlap with the target questions. Regardless of the child's response, Hillary congratulated them on their performance and brought them a sticker. Each story was between 3 and 5 min long, and most children completed the task in under 30 min and asked to play again.

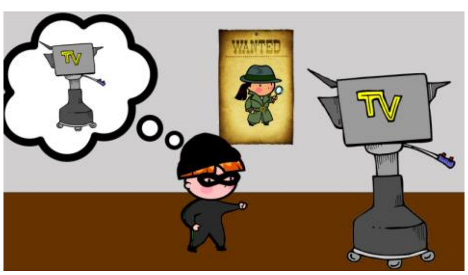
4.3. Materials

Children were presented with eight stories, including six targets and two fillers. To make the stories engaging and interactive, the experimenter told the story while navigating the PowerPoint slides. Each story was followed by a question from Hillary Hippo. Questions were presented in one of two orders: Versions A and B. The order of Version B was the reverse of Version A. Each version started with a filler and had another filler halfway through the experiment.

4.3.1. Materials: Story design

A summarized version of a story is provided in Table 10. Children saw images such as these while the experimenter narrated the story.

Table 10. *Sample story for Experiment 2*

	Detective Sherry wants to catch the famous thief, Evil Steve. She is pretty sure he's going to steal the queen's crown while it is on display at the museum.
	Meanwhile, Steve is planning how to steal the crown, but he knows Sherry is really smart. He decides he's going to trick her and tell her he's going to steal the queen's diamond ring at the queen's palace instead. Steve has invented a TV machine that can put him on people's televisions even when the TV isn't on! He tries using that to tell Sherry, but it breaks and doesn't work.

Instead, Steve writes Sherry a letter telling her he will steal the ring and there's nothing she can do to stop him.

Sherry goes to the Queen's palace, but the guard shows her that the ring is fine.

Sherry figures out that Steve must have been tricking her and goes to the museum where she catches him red-handed. Finally, Sherry has caught the famous thief, Evil Steve.

How+what question: *How did Evil Steve tell Detective Sherry what he was gonna steal?*

What+that question: *What did Evil Steve tell Detective Sherry that he was gonna steal?*

To review, the stories were designed with three goals in mind. First, in each story there were always two possible means of communication and two possible objects of the verb such that every question had a correct answer as well as an equally salient, incorrect option. The presence of an incorrect option renders asking a question about it felicitous

(Hamburger & Crain, 1982; Omaki, Davidson White, Goro, Lidz, & Phillips, 2014). This also allowed us to determine whether children were attending to the stories well. Table 11 indicates the two potential answers to the embedded what question and the two potential answers to the how question.

Table 11. *Contrasting possible answers for the story in Table 9*

Question	Alternative	Actual
How Steve told Sherry	The TV machine	A letter
What Steve stole	The ring	The crown

The second goal was to distinguish non-embedding question responses from WSM-like responses. A key difference, therefore, between our experiment and those of previous works, is that our characters always say something contrary to fact. This allows us to distinguish between a WSM-like interpretation of the sentence and answering the embedded question under a non-embedding question analysis. If children answer (13) with *the ring*, what Steve said he would steal, this indicates a WSM-like interpretation of the question. However, if they answer with what he actually stole, *the crown*, this would indicate that they ignored the first clause and answered the medial wh-phrase as if it were a mono- clausal question: the non-embedding response. Henceforth, we refer to responses indicating WSM interpretations as WSM-like responses. Stimuli can be found in Appendix D.

The third goal was to create stories in which the event pertaining to the *how* wh-phrase and the event pertaining to the *what* wh-phrase were given equal prominence in the story. To this end, we implemented two measures. First, to prevent a recency bias in children's answers, we ensured that the order of matrix clause and embedded clause events varied across different story stimuli. Thus, in four stories, the telling event happened before the embedded clause event, and the order was reversed in the other four stories. This meant half of the questions had the embedded clause in future tense and the other half had the embedded clause in past tense.

Second, we conducted a norming study with 17 adults to assess the relative prominence of the critical events in the stories. Participants were recruited from the Johns Hopkins undergraduate population and were given course credit. Stories were presented to these participants exactly as they were presented to children (see Section 4.2. above). After each story, participants were asked to judge the importance of the event pertaining to the what phrase and the event pertaining to the how phrase on a Likert scale from 1 to 7, where 1 indicated "*not an important theme of the story*" and 7 indicated "*most important theme of the story*." Scores were averaged across participants and then across items. For target items, the average rating of prominence of the how event was 5.2 (SD = 1.5), while the prominence of the what event was rated as 5.3 (SD = 1.5). A two-tailed t test revealed no significant difference between these two ratings (p = .85). Thus, there is no evidence to suggest that adult participants considered the events to be of differing prominence.

4.3.2. Materials: Question design

All questions used tell as the matrix clause verb. This verb can be used in WSM constructions in a variety of languages (Lutz & Meuller, 1996), and it can also be used more felicitously than say in English with a manner adverbial like how. *How* was selected as the initial wh-phrase for two reasons. First, *how* is the wh-phrase used in previous works by de Villiers and colleagues, in conjunction with a wh-argument like what in the embedded clause. Second, it is an attested scope marker in some WSM languages (Stepanov, 2000), yet can be used effectively and grammatically in biclausal English questions. While *what* is a more typical scope marker, English biclausal questions beginning with what do not allow for a second wh-phrase (e.g., **What did he say what he saw*?).

As discussed above, the experiment manipulated the question type as a between-participants factor (see (13) and (14) above), to assess the baseline frequency of embedded interrogative responses in an unambiguous question. Filler questions were matched to the target questions in length and complexity, and they did not involve extraction from the embedded clause (e.g., *Who told Elisha what the brontosaurus ate? Who told Alex that he would meet a mummy*?).

Instructions as well as questions were posed by the puppet, Hillary Hippo. Hillary's instructions and questions were pre-recorded by a female native speaker of American English. Recordings were made in a sound-attenuated booth on a Marantz professional solid-state recorder using a SHURE Beta 58A microphone. Recording the questions ensured that every child heard the exact same questions without prosodic variation. This also allowed the rate of speech to be controlled to be at an appropriate speed for children. To ensure that there is no prosodic variation at the onset of the embedded clause position (*what* or *that*), the words what and that were extracted from filler stimuli and spliced into recordings of the *how* + *what* questions.

4.4. Data coding and analysis

Children's responses to the target questions fell into six primary categories. A description and example of the children's errors based on the medial wh-question from Detective Sherry and Steve is given in Table 12. The key response type is False Object. If children have WSM in their grammars, they should give a comparable number of False Object responses in each condition. There was no difference in performance between Versions A and B, so the data from the two versions were collapsed for analysis. All data were binomially coded such that every utterance either was (1) or was not (0) an instance of a False Object response. These data were then submitted to a mixed effects logistic regression (Kruschke, 2015) using Question Type (*how* + *what* question or *what* + *that* question) and age as fixed effects. Age was coded as a numerical variable and was centered as well as scaled. Random intercepts included participant and item. The model was run with the maximal random effects structure that would converge, which minimally included random slopes for item and participant.

Table 12. *Example responses to:*
"How did Steve tell Detective Sherry what he was gonna steal?"

Response Type	Description	Example
Correct Manner	Adult like response	With a letter
Distractor Manner	The child responded with the alternative, incorrect answer to the first wh-phrase	The TV machine

Correct and Distractor Responses	The child provided both possible answers to the first wh-phrase	The TV and a letter
Irrelevant Answer	The child made up another answer to the first wh-phrase	On the telephone
False Object	WSM-like response. This response would indicate a WSM interpretation of the question (i.e. as if the question were *What did Evil Steve tell detective Sherry he was gonna steal?*)	**The ring**
True Object	Response indicating a non-embedding question analysis (i.e. *what did Steve Steal*?)	The crown

4.5. Results

Adult responses were 100% accurate, meaning all responses to the *how* + *what* condition were Correct Manner. Children's response types and percentage are provided in Fig. 2. In the *how* + *what* question group, 67.7% of responses were correct (SE = 5.9) compared to 80.2% correct in the what + that group (SE = 26.9). Thus, more errors did occur in the *how* + *what* condition, but performance overall was good, particularly in the single wh- phrase group. Considering the open-ended nature of the task which allowed children to give any answer they liked, this percentage of accurate responses is indeed high.

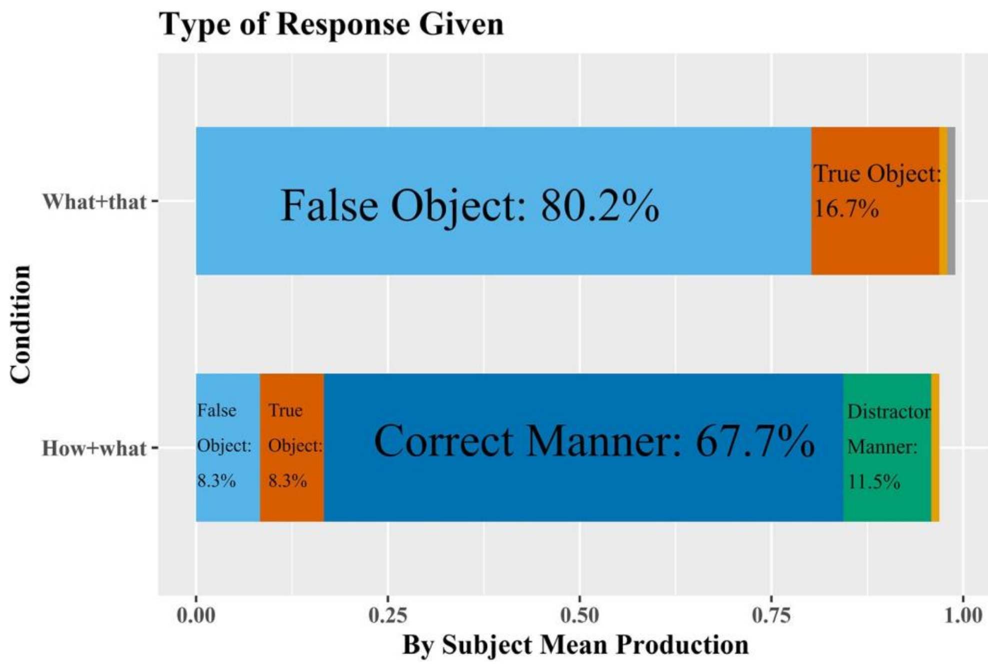


Fig. 2. By-participant mean production rate of responses observed in Experiment 2. Light orange, unlabeled responses indicate an irrelevant manner response (1% in each condition). Gray indicates the child gave both possible manner responses (1% in *what* + *that* condition). See Table 11 for the description of each response type.

There were a few responses to the medial wh-phrase what in the *how* + *what* condition. False Object responses (the signature responses indicative of a WSM-like interpretation in this condition) accounted for 8.3% of these responses (SE = 3.3) and True Object responses also accounted for 8.3% (SE = 3.5). There were a total of eight False Object responses given by five participants. Four of those participants only gave one False Object response, while one participant gave four False Object responses. False Object responses in the *how* + *what* condition occurred far less frequently than False Object responses in the *what* + *that* condition. Furthermore, in the how + what condition, False Object responses accounted for fewer responses than the Distractor Manner responses

(11.46%, SE = 2.5) that result from children misremembering the stories. This would not be expected if WSM were a part of children's grammars.

To establish whether there was a significant difference in the number of False Object responses between 40.*how* + *what* and *what* + *that* conditions, the data were submitted to a mixed effects logistic regression model. The results showed a significant effect of question type ($p < .01$) and no effect of Age.

The results of this mixed effects logistic regression model suggest that children were significantly more likely to correctly give the False Object answer in the control *what* + *that* condition than as a WSM-like response in the *how* + *what* condition. These results suggest that children do not consider these questions to be equivalent. Furthermore, to establish whether children were more likely to answer the question correctly in one condition or another, we used the same model to determine whether condition affected children's adult-like performance and the results showed no effect of question type ($p > .1$).

We further investigated the responses of the five individuals who did give WSM-like responses and include that data in Table 13. Under Other Responses Given a *why* response indicates the child gave a reason for the telling event rather than a response to either wh-phrase.

Table 13 also illustrates that WSM-like responses came from children ranging in age from 3;11 to 5;10. This is almost the full range of ages tested. Thus, we see that this error is not limited to the youngest children, though most errors were made by one of the youngest subjects.

Table 13. *Individual differences in children who gave WSM-like responses*

Participant Age in months	WSM total	Correct Manner	Other Responses Given (type)
47	1	1	4 (True Object)
48	4	0	2 (True Object)
55	1	3	2 (Distractor Manner)
55	1	2	3 (Irrelevant Manner; True Object; *why* response)
70	1	4	1 (Distractor Manner)

Finally, we examined whether the eight False Object responses in the *how* + *what* condition, which could potentially be evidence for WSM, may have occurred simply as random noise. There is not a clear threshold for chance in this experiment because a child could have given any response they liked. However, we assume that the Distractor Manner responses were the result of correctly interpreting the syntax and failing to accurately remember the story. Thus, we used the Distractor Manner responses as a comparison point for answers that indicate misremembering. We performed a paired samples t-test between Distractor Manner responses and False Object responses, the results of which indicated there is no significant difference between the two ($p > .4$). This suggests False Object responses occur with similar frequency to answers, indicating the child simply forgot the story.

4.6. Discussion

The main findings in Experiment 2 can be summarized as follows. First, we found that the majority of children's responses were adult-like in both *how* + *what* and *what* + *that* questions. The *what* + *that* condition, particularly, shows that the complexity of the stories did not cause the children difficulty. Performance in the *how* + *what* condition was not as high, but there was no statistically significant difference between the two groups. Thus, we can conclude that children performed well in both conditions.

Second, while children gave some non-adult-like responses in the *how* + *what* condition, there were more errors that resulted from misunderstanding of or inattention to the stories than those that could potentially serve as evidence for a WSM grammar. Third, our statistical analyses demonstrated that the critical response (False Object) was reliably greater in the *what* + *that* condition (where that is the only grammatical response in English) than in the *how* + *wha*t condition, which suggests that children did not treat these two questions types as semantically equivalent. Fourth, in the *how* + *what* condition, when children responded to the what question, they were evenly split between False Object (WSM-like responses: 8.3%) and True Object (non-embedding response: 8.3%). Finally, the occurrence of False Object responses was not significantly different from the occurrence of answers indicating memory lapses.

A possible objection to our third point, as suggested by a reviewer, is that children whose grammars allow both English-like LD and WSM would find the *how* + *what* condition to be ambiguous between the two readings. In contrast, they would find the *what* + *that* condition unambiguous. This would predict fewer False Object responses in the *how* + *what* condition than the *what* + *that* condition, which is what our results showed. However, this would also predict a roughly even split between False Object (WSM-like responses) and Correct Manner (Adult-like) responses in the *how* + *what* condition. Even if that pattern did not surface across children, it might be present within individual participants. We investigated the individual children who gave WSM-like responses. The results in Table 13 do not suggest an even split. Three of the five participants gave more Correct Manner responses than WSM-like responses. This suggests that children who were giving WSM-like responses were not treating them as ambiguous. However, it is possible that certain pragmatic factors within individual stories might have influenced questions which otherwise might have been ambiguous. Such a claim would require explicit testing, which we leave to future work.

A major result is that our findings (8.3% WSM-like responses) differ from those of de Villiers and Roeper (1995), as they found that the majority of children's responses indicated a WSM-like response (over 50%). One potential explanation for the discrepancy is the difference in the matrix clause verbs used (*say* in de Villiers & Roeper, 1995 vs. *tell* in the current study). The use of the verb *tell* in the current study may have strengthened the association of *how* with the matrix clause verb, and this processing bias masked their ability to entertain WSM structures. This is an important question, given that there are the so-called hybrid WSM languages (e.g., dialects of German) that are known to allow both

WSM and English-like LD wh-questions. If English-speaking children in fact had such hybrid grammars, the association of how with the verb *tell* might be heightened by their parsing bias toward a matrix clause interpretation (for a demonstration of this parsing bias in related wh-questions, see Omaki et al., 2014). At the suggestion of a reviewer, we conducted a follow-up study using the same story stimuli while changing the matrix verb to *say* to fully investigate the possibility that the verb change was the cause of the discrepancy between our work and previous work. Fourteen children ranging in age from 4;1 to 5;10 (mean: 4;10) took part in this follow-up study. Four were excluded for being unable to answer filler questions. We then compared children's performance on the *say* version of the task to the results from the *tell* version. The results showed no significant difference in performance with the verb say ($p > .9$). We conclude that the discrepancy between our findings and those of previous work cannot be ascribed to the choice of matrix verb and thus reject this alternative hypothesis.

One final finding concerns the prevalence of non-embedding responses. In both the *how* + *what* and *what* + *that* conditions, our study found that children sometimes gave a response describing the truth of the matter rather than what the question asked. This would include, for instance, answering the crown in response to either *How did Evil Steve tell detective Sherry what he was gonna steal?* (8.3%) Or *What did Evil Steve tell Detective Sherry that he was gonna steal?* (16.7%). We argue that this is not a WSM-like response in the *how* + *what* condition because children are only interpreting *what* as having scope over steal rather than say. However, the prevalence of this answer in the *what* + *that* should be explained. It is worth considering whether these children simply were responding to the truth values of the statement. In other words, rather than incorporating both clauses appropriately to answer What Steve said he was going to steal, children may have failed to integrate the two clauses and answered what in fact was stolen. A similar observation was made in de Villiers and Pyers (2002). This, too, we leave to future research.

In sum, the present study found that the use of carefully constructed story and question stimuli did not find evidence for WSM grammar in English-speaking children's language comprehension. The fact that we did not see a similar number of False Object responses in both conditions suggests that children did not consider LD and WSM as equally possible constructions in their grammar. Furthermore, the eight instances of False Object responses did not significantly differ from the responses which indicate forgetting the story. This finding poses an empirical challenge to the parametric acquisition hypothesis, which predicts that WSM structures would be active in both children's production and comprehension. Given the finding from Experiment 1 which indicates a large degree of individual variation, it is possible that the separate group of children tested in Experiment 2 simply did not have WSM in their grammars. In other words, those children who produced questions with medial wh-phrases in Experiment 1 might have shown evidence of WSM in their comprehension as well if they had participated in that study. We test this hypothesis in Experiment 3.

5. Experiment 3: Within-subjects experiment

The goal of Experiment 3 is to establish whether those children who produce medial wh-questions also show evidence of WSM in their comprehension, and vice versa. The Parametric Acquisition Hypothesis suggests that WSM should be a part of both production and comprehension and thus predicts a correlation between errors that indicate WSM in comprehension and production: Children who produce medial wh-questions should also respond to questions with multiple wh-phrases with WSM-like responses.

5.1. Participants

Twenty-nine children between the ages of 3;10 and 6;2 (*M* age = 5;2) were recruited from the Baltimore area. This age range was selected to overlap with previous work. All participants took part in both the production and comprehension tasks.

5.2. Procedure

As with Experiments 1 and 2, children were introduced to the puppets before performing both the production and comprehension tasks. The tasks were balanced for order (production then comprehension vs. comprehension then production) as well as version (A vs. B). Practice questions remained the same. Children generally took an hour to complete both tasks, with a break between the tasks. While both tasks are difficult, allowing us to investigate the boundary of children's capabilities, children enjoyed the tasks. They were unperturbed by the length and often had to be convinced to take a break rather than keep going.

5.3. Materials

The materials used were identical to those used in Experiments 1 and 2. For the comprehension task, we only used those questions which contained an initial and medial wh- phrase (*how* + *what*).

5.4. Data coding and analysis

Children were excluded from analysis if they failed to answer at least one of the filler questions correctly in the comprehension task or who simply would not produce questions (n = 9). This left 20 children (*M* age = 5;1). Productions were coded using the same system as in Experiment 1. Children's responses to the questions in the comprehension task were coded using the same system as in Experiment 2. The responses given were generally the same as those found in Experiment 2 with one additional response type to indicate that the main character was lying. For example, in response to *How did Evil Steve tell Detective Sherry what he was gonna steal*? Children would say *He didn't tell her! He told her he was going to steal the ring*. This was a response we did not find in Experiment 2, but which several children gave occasionally and one child consistently gave in Experiment 3. This is a reasonable response to the question, which indicates a nuanced understanding of the scenario and thus is included in the responses coded.

There was no difference in performance between Versions A and B, so the data from the two versions were collapsed for analysis. Comprehension data were binomially coded such that every response either was (1) or was not (0) a WSM-like response. These data were then submitted to a mixed effects logistic regression (Kruschke, 2015) using

Medial Wh-phrase Production (also binomially coded such that every response either did (1) or did not (0) include a medial wh-phrase) and age as fixed effects. Age was coded as a numerical variable and was centered and scaled. Random intercepts included participant and item. The model was run with the maximal random effects structure that would converge, which minimally included random slopes for item and participant.

5.5. Results

Of children's overall productions, 56% were grammatical, 46.7% were target-like, 11.7% included that-trace violations, and 15.4% included a medial wh-phrase (compared to 22% in Experiment 1). This is a similar finding to Experiment 1: Mostly children performed well (56%) in this open-ended task, but a small percentage of medial wh-phrase interpretations persisted. A detailed description of children's response types is given in Appendix E: *Other error types for Experiment 3*. By-subject means of the primary answer types given in the comprehension task are presented in Table 14. The six error types reported account for 96% of total error types. We saw two response types in Experiment 3 that we did not see in Experiment 2. The first was "Other item" responses, which indicate other items mentioned in the story, but not a felicitous response to any interpretation of the question. The second was an answer indicating that the character was not being truthful. These we label as "trick" responses.

Table 14. *By Subject means of response types for Experiment 3 comprehension task*

Answer Type	Example from the Evil Steve Story	% response type
Correct Manner (adult-like response)	The letter	40.8
False Object	**The ring**	**18.3**
Distractor	The TV	6.2
True Object	The crown	15
Trick	He didn't! He told her he'd steal the ring.	9.9
Other items	The museum	5.8

The proportions from Table 14 are represented in Fig. 3, below.

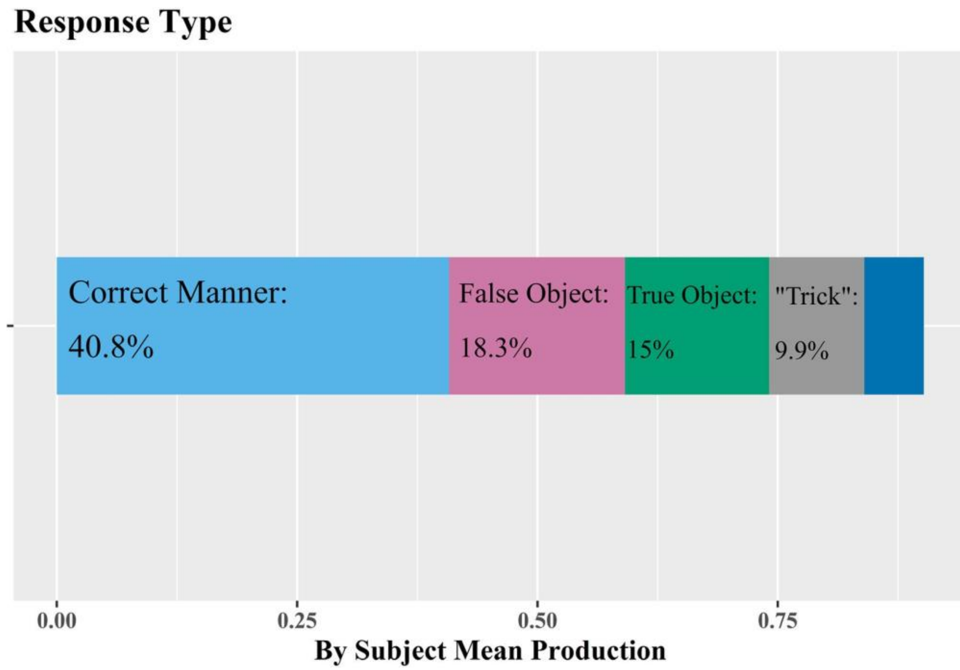


Fig. 3. By-subject mean of response types to How did Steve tell Sherry what he was gonna steal? Unlabeled blue bar represents the Distractor Manner responses, which accounted for 6.2% of response types.

We searched for a correlation between production and comprehension performance in two ways. First, we found the correlation coefficient between production and comprehension errors across the group. The results yielded an r^2 of .0016, suggesting no correlation between performance on the two tasks. Second, we submitted the data to a mixed-effects logistics regression model using performance on one task as a predictor of performance on the other. This allowed us to account for random effects such as participants and item number and therefore get a more complete picture. The results of the mixed effects logistic regression model investigating whether WSM-like comprehension is predicted by productions which include a medial wh-phrase found no relationship between production of medial wh-phrase and WSM-like comprehension ($p > .1$). This is illustrated in Table 15.

Table 15. *Logistic regression model results: medial-wh productions predictor of WSM responses*

	B	Error	Z value	*Pr(>|z|)*
Medial wh productions	-0.41	1.19	-.35	.73
Age	-.80	.60	-1.32	.18
Medial Wh ×Age	.55	1.16	.47	.64

Note: † $p < 1$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

Fig. 4 plots individual performance on the production task (x-axis), indicating the use of a medial wh-phrase, by performance on the comprehension task (y-axis), indicating a WSM interpretation ("false object") response.

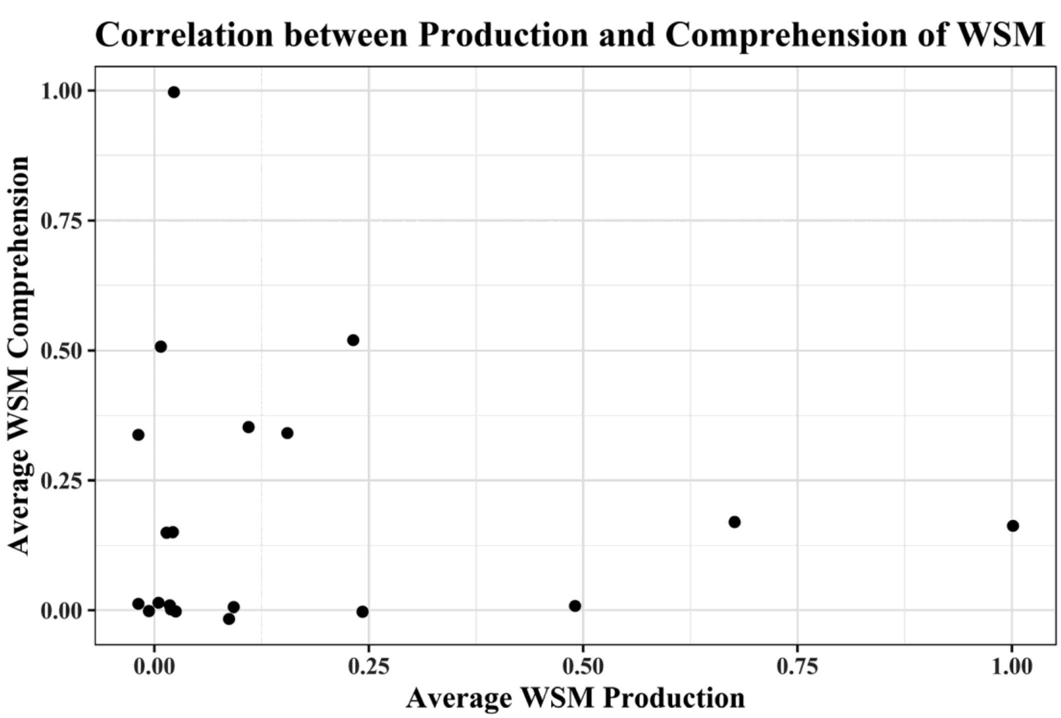


Fig. 4. Individual performance on the production and comprehension tasks.

A correlation between production errors and comprehension errors would result in a positive slope. Fig. 4 shows that while both error types occurred, an individual's performance on the production task did not predict how they performed on the comprehension task, and vice-versa.

Yet Fig. 4 is informative in several ways. One possible outcome of this study was that there might be a general trend of WSM in children. Another possibility was that only one or two children might show evidence of WSM. Not only does Fig. 4 show us that there is no evidence of a general trend, it shows that there are no individuals who seem to use a WSM grammar against the general trend of adult-like performance. There are two "exceptional children": one who gave 100% WSM interpretations but no medial-wh productions, and one child who used medial wh-phrases 100% of the time, but only gave one answer which indicated WSM interpretation in comprehension. This is not the pattern we would predict if either of these children used a WSM grammar. At the suggestion of a reviewer, we also examined the data after removing these two "exceptional children." After removing these participants, $r^2 = .0066$. The lack of correlation thus persists when the two potential outliers are removed.

Finally, as with Experiment 2, we wanted to verify that the number of WSM-like responses were not different from chance. Again, we performed a t-test between participants Distractor Manner responses, which result from misremembering the story, and WSM-like responses. The result of this t-test showed that there was no significant difference between Distractor Manner responses and WSM-like responses ($p > .05$).

5.6. Discussion

The production results of Experiment 3 replicated and enhanced those of Experiment 1. Most of children's productions included a single wh-phrase, while 15.4% included multiple wh-phrases. Again, most (57%) of children's productions were grammatical structures of English in the open-ended elicitation task, suggesting solid adult-like performance. When investigating individual performance, two exceptional children produced questions with multiple wh-phrases 50% of the time and one produced exclusively questions with multiple wh-phrases.

There were more than twice as many examples of WSM interpretations in Experiment 3 (18.3%) than in Experiment 2 (8.3%). The difference, however, is mostly driven by a single participant who gave a WSM-like response to every question. This indicates a large amount of individual variation in performance on this task as well. All other children gave this kind of response 50% of the time or less. In spite of the increase in answers of this type, there was still not a significant difference between these responses and responses indicating the child correctly analyzed the structure, but misremembered the story (as indicated by their Distractor Manner responses, 6.2% of total). This suggests they are not the result of children using a WSM grammar. While children gave adult-like answers only 40.8% of the time (compared to 67.7% in Experiment 2), this response type still represents the majority of responses. The next most common response type (False-object) accounted for only 18.3% of responses. Given the open-ended nature of the task and the variety of possible answers (see Table 14), we can consider 40% indicative of relatively successful performance.

The addition of trick-responses which did not occur in Experiment 2 also deserves discussion. These normally took the form of the spontaneous exclamation "but he didn't tell her! He was lying!" We interpret these to mean that (for instance) Steve never did tell Sherry what he was going to steal (the crown) and, therefore, answering how he told her cannot be done. This, we consider another form of adult-like response and direct evidence that children were following the stories. Indeed, while no adult gave such a response during the adult control in Experiment 2, many adults commented on this potential interpretation when hearing these stories presented. However, a reviewer suggested that children might have interpreted the question totally differently: They might have interpreted *how* as being extracted from the embedded clause. Therefore, they would interpret the question to mean *how did Steve say the stealing would happen*, which there is also no response for. We note that these responses almost always included the fact that Steve was lying, which indicates their focus on the fact that Steve lied about the object stolen rather than the fact that he never explained how the stealing would happen. It should further be noted that in the multiple iterations of this experiment performed previously (de Villiers & Roeper, 1995; de Villiers et al., 1990; de Villiers et al., 2019; Roeper & de Villiers, 2011), one consistent finding is that children never associate the matrix question word with the embedded clause verb. For this reason, we think it unlikely children were disturbed by the lack of possible response to this interpretation in our experiment.

A related reviewer comment suggests that children may have been giving accurate answers for a reason unrelated to syntax. The specific suggestion is that they hear and understand the word *how* and because these stories (intentionally) focus on how the action denoted by the matrix verb takes place (e.g., *how the telling happens*), but not how the embedded event takes place (*how the stealing happens*), children simply answer the only plausible *how* question. Children's responses never indicated such an interpretation in the previous literature listed above, but explicit testing of this suggestion would need to be done to entirely eliminate this possibility.

In summary, the results of Experiment 3 provided two crucial findings. First, there was no correlation between performance on production and comprehension. While both error types were found, there was no relationship between errors in the production task and the comprehension task. Second, there was no correlation within an individual. A Parametric Acquisition Hypothesis would predict that even if no general correlation was found across participants, an individual might have an incorrect parameter setting and therefore show evidence of a correlation in their behavior. That was not the case in our data. Those children who produced medial wh-phrases did not give WSM-like responses, and vice versa. The results of Experiment 3, therefore, do not support a parametric account of these error types.

6.0 General discussion

This study investigated the extent to which English-speaking children use medial wh-structures in language production and comprehension. Our experimental designs were intended to address a number of methodological concerns (partially) arising from previous studies. These resulted in new evidence first, that children showed more adult-like performance than has been reported in previous work, and second, that medial-wh question structures were observed 22% of the time in production (Experiment 1) and 8% of the time in comprehension (Experiment 2) in children between 3;11 and 6;1. Experiment 3, in particular, revealed that those children who showed evidence of WSM in production showed no evidence for it in comprehension, and vice versa. These findings are consistent with the independent observation that spontaneous productions of medial wh-questions are rare (Stromswold, 1995). Thus, any explanation for the use of medial wh-structures by children must account for (a) why WSM errors may appear either in a child's production or comprehension, (b) why these errors are not consistent across children, and (c) and why there is no correlation between these errors within participants. Below, we examine the theoretical implications of the present findings.

6.1. Implications for grammatical accounts of medial wh-questions in children

The lack of correlation between errors in comprehension and production within a participant provides substantial empirical challenges to the parametric acquisition account of medial wh-questions. If children have set an incorrect parameter setting and adopted a WSM grammar, this structure should be used in both production and comprehension (Yang, 2002), contrary to the present findings. While these results should not be construed as conclusive evidence against the parametric hypothesis in general, they do not provide evidence for it—contrary to what has been previously claimed. Within the parameter-based grammar acquisition approach, Snyder (2007) has proposed that children are conservative learners who never mis-set syntactic parameters, and observations of syntactic creativity behaviors result from performance errors or experimental artifacts. Our findings support the latter perspective, which we return to below when we present our performance-based explanation of the present findings.

An alternative implementation of the grammar hypothesis, suggested by Roeper and de Villiers (2011), is that children's WSM-like errors are the result of the way wh-phrases are integrated (realized and interpreted) at the left edge of individual clausal domains or Phases from the most embedded one up, according to more recent derivational versions of Minimalist Theory (Chomsky, 2001). If the child integrates the wh-phrase into the lower clause during the first Phase and there is a plausible answer to that question, it is possible the child simply answers that question before completing the higher Phase. This predicts, as Roeper and de Villiers (2011) point out, that children should give correct responses to this embedded clause question. In our study, this corresponds to a true object response. However, as we have seen in Experiments 1 and 3, true-object and false-object responses occur with equal frequency. This explanation, therefore, would only hold for a subset of the children tested.

There remains the possibility, suggested by a reviewer, that the lack of correlation found is because we considered only WSM constructions rather than copy constructions. If children do not in fact use WSM, but do use copy constructions, we would predict no WSM-like responses in the comprehension condition, but significant use of copy constructions in the production task. This is, in fact, what we found. In the production task, those productions which included multiple wh-phrases most often used the same wh-phrase twice. This resembles a copy construction. However, in the comprehension task, all questions had different wh-phrases, which resembles WSM, but not copy constructions. It is, therefore, possible that children have only the copy construction grammar and not WSM. This would explain why a child might show evidence of the error in production, but not in comprehension (as we saw in Experiments 2 and 3). Such an explanation, however, would fail to explain why some children made errors in comprehension, but did not make errors in production, as we also saw in Experiment 3.

We conclude, therefore, that the results we have obtained in these three experiments do not support a (parametric) grammar explanation of these particular syntactic errors, and turn to an alternative explanation for these errors. We proceed under the UG-based assumption that children have the appropriate English grammar settings by the age of 4;0 and propose that the specific errors which we observe in these tasks are best understood as the result of immature processing.

6.2. Syntactic creativity as a product of immature sentence production mechanisms

Our findings resulted in a range of performance patterns across children. In the following sections, we will provide explanations for each of these patterns. In the first place, we must account for children who produced questions with multiple wh-phrases, either consistently or intermittently. Relative immaturity in children's sentence production mechanisms (as reviewed below) leads us to hypothesize that children may use the medial wh-phrase to facilitate retrieval of the sentence initial wh-phrase at the clause boundary[35].

The first key ingredient of this proposal is the reactivation of the wh-phrase. The idea that the left-peripheral constituent in a filler-gap dependency is reactivated at the embedded clause boundary is well motivated by theories of syntax (e.g., successive

[35] Grolla and Lidz (2018) also propose a processing explanation for medial wh-phrase production. Their findings suggest that children with lower inhibition control tend to produce more questions with medial wh-phrases. This explanation is distinct from ours in that it explains only those productions which involve the same wh- phrase twice, which was the most common production type, but not the only one to be explained.

cyclicity assumed above; Chomsky, 1973), and it has also received ample support from adult psycholinguistics research. Gibson and Warren (2004) examined processing of relative clauses that span over multiple clauses, and they found that the subsequent integration of the filler at the ultimate gap position was facilitated when there was an earlier opportunity to reactivate this filler at an embedded clause boundary. (For additional evidence that production of long-distance dependencies involves memory retrieval and reactivation of early constituents, see Badecker & Kuminiak, 2007; Franck, Soare, Frauenfelder, & Rizzi, 2010.) Based on these findings, we assume that production of filler-gap dependencies in wh-questions proceeds in a similar fashion, and adults and children both reactivate the sentence-initial wh-phrase at the embedded clause boundary.

The crucial difference between adults and children is that adults can readily reactivate the wh-phrase covertly and move on to the articulation of the embedded clause, which contains the ultimate gap position where the wh-phrase is structurally integrated. We suggest that some children may struggle to do so and, as a result, produce another instance of the sentence-initial wh-phrase to strengthen their memory representation of the filler-gap dependency. The proposal that memory retrieval of the fronted constituent is facilitated by an overt articulation of the gap has been independently invoked to account for why sometimes adult English speakers produce resumptive pronouns in sentences with filler-gap dependencies (e.g., Frank had an operation which we just found out about it), despite their ungrammaticality (Dickey, 1996; Kroch, 1981). In fact, a child elicitation experiment by McKee and McDaniel (2001) has shown that children produce more resumptive pronouns than adults do. Furthermore, these resumptive elements frequently were repetitions of an entire Determiner Phrase (DP) rather than a pronoun (D. McDaniel, pers. comm.). This finding further corroborates the current proposal that children resort to overt articulation of an extra element to facilitate processing of long-distance dependencies.

Under the current proposal, the critical immaturity that distinguishes adults from children is in their ability to retrieve the target constituent at the clause boundary. This may simply indicate that the memory retrieval mechanism itself needs to undergo development. However, another important factor in determining the retrieval difficulty is the degree to which children's sentence production proceeds incrementally. If children's production is a slower, more effortful process with many pauses and interruptions, then the memory representation of the sentence-initial wh-phrase will have substantially decayed by the time the production mechanism arrives at the embedded clause boundary.

Independent evidence for children's immature incremental production ability comes from McDaniel et al. (2010), who elicited production of relative clauses by adults and 3 to 8-year-old children and assessed the distribution of disfluency indicators such as pauses and restarts. When adults produced object relative clauses with long-distance dependencies (e.g., *Pick up the duck that Big Bird thinks the princess was kissing* _), the observed signs of disfluency suggested that adults were able to plan a large structural unit such as an upcoming clause while holding the filler-gap dependency active in working memory (for related findings, see Ferreira & Swets, 2005). On the other hand, children demonstrated signs of disfluency in a variety of positions throughout the sentence, suggesting that children plan only the current phrase. In other words, children plan more locally, and more frequently during sentence production.

Let us now return to the current experimental findings. This new proposal can explain the observation that the production rate of medial wh-questions varies across individuals, and some children use these structures inconsistently. In particular, individual variance can be attributed to variation in the development of incrementality and memory retrieval for sentence production processes. The non-systematic use of medial wh-questions may result from fluctuations in the child's cognitive state and resources. In addition, this account also provides a more straightforward explanation of why most of the observed medial wh-questions (82% in Experiment 1) involved repetition of the same wh-phrase (e.g., *What do you think what should we look for*?). While this should not be seen as proof of our account, given that children have just been given the first wh-phrase and told to use it, a processing account gives a straightforward explanation for children's consistent production of the same wh-phrase twice. Under the current account, the reactivation of the sentence-initial wh-phrase is facilitated by overt articulation of the same wh- phrase, so the high proportion of wh-phrase repetition is actually expected.

For the few cases where the sentence-initial and medial wh-phrases were different (e.g., *What do you believe where do you want to go?*), we suggest that they may be instances of restart errors that also result from immature sentence planning. Children may start the matrix clause what do you believe without planning how the wh-question will continue, and when they realize that they would rather ask a question differently, they create a new question. This type of restart error may also account for the relatively high rate of subject–auxiliary inversion in the apparent embedded clause context. This is supported by our finding that children whose productions generally include disfluencies such as restarts and filler words are more likely to produce ungrammatical sentences than those whose speech is more fluid. Furthermore, this would account for the disfluency present in the prosody of the productions, though we leave a thorough investigation of prosody to future research. In sum, the observed variation in production of medial wh-questions is most likely to arise from children's immature sentence planning mechanisms. This explanation naturally extends to the empirical findings of Thornton (1990).

6.3. Syntactic creativity as a product of immature sentence comprehension mechanisms

The second pattern of relevant errors to be explained are those of the exceptional children who gave WSM-like responses on 50% or more of the trials in the comprehension task. Recall that the result of Experiment 2 showed 8.3% of responses were indicative of WSM interpretations, and in Experiment 3 one child gave WSM-like responses consistently, making the total WSM-like responses in that experiment 18.3%. We found that these errors do not occur significantly more frequently than distinct errors indicating children have forgotten the story in either Experiment 2 or Experiment 3. We have also argued (Section 5.6) that the lack of correlation between these children's performance on production tasks and comprehension tasks discounts the possibility that WSM-like responses are the result of a WSM grammar. However, in Experiment 2, one child gave WSM-like responses to four questions, and in Experiment 3, one child gave WSM-like responses to every question. This consistency should be explained. It is possible that a confounding factor, such as the lead-in provided by the fairy in the production task, simply did not allow them to perform as they otherwise would. Alternatively, children may struggle with integrating clauses together in comprehension as a result of limited memory capacity. We consider these two possibilities below.

Let us first consider the possibility that these few error responses actually are the result of a WSM grammar. Recall, that in Experiment 2, we found that children in both the *how* + *what* and *what* + *that* conditions gave adult-like responses over 60% of the time (67.7% in the *how* + *what* condition, 80.2% in *what* + *that* condition). These children, we can assume, are able to incorporate the embedded clause into the question in an adult-like way most of the time. These children also gave incorrect responses (Incorrect manner or True Object answers) to the appropriate wh-phrase less than 20% of the time (11.5% in *how* + *what*; 16.7% in *what* + *tha*t). Thus, we see that when children are capable of incorporating both clauses, they are also capable of answering the question correctly the majority of the time and only give responses that indicate misremembering the story occasionally.

As we consider the possibility that a False Object answer indicates the use of a WSM grammar, we would predict a similar pattern of mostly correct responses and a few which indicate misremembering the story. In other words, if a child has a WSM grammar and is successfully incorporating the embedded clause into the sentence, we expect that child to give mostly False Object responses with very few True Object responses. This, however, is not the case. In both Experiments 2 and 3, the rate of False Object and the rate of True Object responses were roughly equivalent (8.3% vs. 8.3% in Experiment 2; 18.3% vs. 15% in Experiment 3). This suggests that the children who gave these responses were not incorporating the embedded clause into the question even under a WSM grammar. Rather, it appears these children are simply those whose memories fail them reliably. We suggest this immature memory extends to the fact that they are unable to incorporate the embedded clause into the matrix and simply give an answer to the embedded clause question. This explanation is consistent with De Villiers and Pyers (2002)'s finding that children's memory of complements was a reliable predictor of their performance on comprehension tasks.

Having eliminated the possibility that these errors stem from WSM, we also need to explain why children who make errors in speech production often perform perfectly in comprehension. Here, we propose that children's preference for interpreting *how* as a matrix adjunct typically favors the correct interpretation of questions like *How did Evil Steve tell Detective Sherry that he was gonna steal the ring*? Cross-linguistic studies by Omaki et al. (2014) as well as Lassotta, Omaki, and Franck (2016) used a series of Question-after-Story experiments and found that in comprehension of biclausal wh-questions, 5-year-old children generally prefer to associate the sentence-initial wh-phrase with the first verb in the sentence (e.g., matrix clause verb for English, embedded clause verb for Japanese). This strong preference most likely reflects the so-called active gap filling bias (Crain & Fodor, 1985; Frazier, 1987; Stowe, 1986), which allows comprehenders to complete the dependency and assign an interpretation to the sentence-initial constituent as soon as possible (see Atkinson, Wagers, Lidz, Phillips, & Omaki, 2018 for critical discussion of whether children have this bias). Given this parsing bias, children in the current study should be biased to associate how with the matrix clause verb *tell* or *say*, as this is the first position in the sentence to grammatically complete the dependency. In other words, unlike the production process, there is no need to extend the dependency to the embedded clause boundary position.

In sum, the various syntactic creativity patterns found in the current study receive straightforward explanations as products of constraints on children's sentence processing mechanisms. In production, children must relate the sentence-initial wh-phrase to the embedded clause, but the memory representation of the wh-phrase may decay before the ultimate gap position due to their inefficient planning mechanism. To combat this difficulty, children resort to an overt articulation of the wh-phrase at the embedded clause boundary to strengthen the memory representation. In comprehension, even if the child struggled with the production task, the child parser incrementally associates the sentence-initial wh-phrase to the matrix clause verb to minimize the dependency and integrate the wh-phrase at the earliest possible position. It turns out that this heuristic leads to a response corresponding to the only adult-like analysis of the question stimuli in Experiment 2. For those children who struggled with the comprehension task in Experiment 3, we propose that limited memory capacity resulted in their failure to incorporate the matrix and embedded clause as suggested by de Villiers and Pyers (2002). They respond to the medial wh-question and are at chance as to whether they answer it accurately.

There remains an outstanding finding from McDaniel et al. (1995), which relies on a different experimental task. On the basis of an acceptability judgment task, McDaniel et al. (1995) found that some children consider WSM/Copy Constructions such as Who do you think who Grover kissed? acceptable sentences in English. This longitudinal study found that children are more likely to accept LD questions, but accept WSM/Copy Constructions significantly more than adults. Furthermore, rather than being evenly distributed, McDaniel et al. (1995) found that a subset of the children accepted these structures with consistency. Six children always accepted them, and seven accepted the structure at the beginning of the study and moved out of it. There were also several children who moved into and out of accepting the pattern. Two children accepted these structures on more than two observations. One possible explanation is that the divergent results are tied to different experimental tasks. Our story-based question elicitation and question-after-stories tasks, we suggested, tapped into children's processing limitations. In contrast, McDaniel et al.'s acceptability judgment task was a metalinguistic one, and errors may have had a different source. Given that such complex biclausal wh-questions incurred large processing demands, one possibility is that children's metalinguistic judgments were somewhat unreliable. We leave these questions open for future research.

6.4. Production of that-trace violation and other syntactic creativity phenomena

Experiment 1 demonstrated another instance of syntactic creativity, where some children produced that-trace violations (e.g., *What do you think that can stop the witch?*), which are ungrammatical for adult English speakers, but are grammatical in languages such as Italian (Rizzi, 1982). Similar observations were reported in Thornton (1990)'s production task as well as in McDaniel et al. (1995)'s acceptability judgment task. The current replication suggests that children's production of that-trace violations is a robust phenomenon.

In fact, these errors can straightforwardly be explained as a product of immature sentence planning mechanisms, as follows. McDaniel, McKee, Cowart, and Garrett (2015) argue that in adults the clause is a major planning unit, but that clauses can be planned separately or jointly. McDaniel et al. (2015) suggest that when adults plan the clauses separately, *that* is used at the embedded clause boundary. This extends to children, but because children have difficulty planning the clauses jointly, as suggested above, they end up producing *that* even when it is ungrammatical.

A related possibility is that because children's planning is much more local, they plan the embedded clause structure only after the matrix clause is completed. This is far later than when the filler-gap dependency is introduced. The lack of joint planning may prevent children from realizing that introducing the overt complementizer *that* would render the filler-gap dependency ungrammatical. In other words, children may have the knowledge of the relevant grammatical constraint on that-trace, but they may fail to access this knowledge due to the inefficient production mechanism.

Another possible explanation is that children may in fact not have adult-like knowledge of the that-trace effect. This lack of knowledge may be a product of overgeneralization, and there are two potential sources for this. First, when children observe LD extraction of an object wh-phrase out of an embedded clause headed by that (*What do you believe that we'll see in the cave?*), they may assume that the same extraction is possible when the subject wh-phrase is extracted. Second, either subject or object can be extracted from the embedded clause when the complementizer that is absent (*What do you believe can help the fairy?*). This can lead children to overgeneralize and assume that English has no constraint that blocks subject extraction from an embedded that-clause. In fact, this second source is more likely to influence children because object extraction out of embedded clauses with that is very rare (Pearl & Sprouse, 2013). It remains to be seen in future research which of these accounts provides the best explanation for children's production of that-trace violations.

These two lines of explanations—immature processing versus overgeneralization from input—may potentially explain other cases of syntactic creativity production. As we discussed above, medial wh-questions and resumptive pronouns may both reflect the same mechanisms that facilitate retrieval and reactivation of early constituents. With respect to subject–auxiliary inversion in embedded clauses, Pozzan and Valian (2017) suggested that children overgeneralize subject–auxiliary inversion from the matrix clause to the embedded clause. Valian (1990) suggested that null subjects may arise from overgeneralization of sentences that lack an overt expletive subject (e.g., *Looks like they are going to win*). A thorough examination of these possibilities goes beyond the scope of this paper. Nevertheless, this line of investigation raises the possibility that these two sources (i.e., immature production and overgeneralization) may explain a wide range of syntactic creativity phenomena in children's language production, and it provides open questions for future research.

7. Conclusion

The present study investigated the robustness of medial wh-questions in English-speaking children's sentence production and comprehension. The improved experiment designs led to a reduction in errors in production (Experiment 1) and essentially eliminated the errors in comprehension (Experiment 2). Experiment 3 indicated that some younger children do show evidence of WSM errors in comprehension, but the design of the stimuli allowed us to determine that their interpretations were at chance between a WSM-like response and an answer to the embedded clause question. Finally, Experiment 3 revealed that even when errors appear in both production and comprehension, there is no evidence for a correlation between production and comprehension of WSM in child grammars. We suggest instead that immature processing mechanisms (planning and memory) result in these errors allowing for asymmetry in performance on the two tasks within an individual. As such, this study illustrates that further integration of developmental psycholinguistics and syntactic development research can provide novel insights on child language acquisition.

Chapter 5: The relationship between Working Memory, Production, and Comprehension

Speech errors made by young children are often used as windows into the process of language acquisition (e.g. Kam & Newport, 2009; Lidz & Gagliardi, 2015; Saffran, Aslin, & Newport, 1996; Yang, 2004). For example, over-regularization errors in young people have been used as evidence that they know and use the rules of language, but have mis-applied a rule when faced with an exception (e.g. Brown, 1973). In other words, the use of *I goed to the zoo,* suggests that children know that adding the morpheme *-ed* denotes past tense, but they have yet to learn that the verb *to go* is an exception to this otherwise regular system. However, not all errors can be explained as examples of over-regularization. One classic, persistent error made by children is medial wh-productions in child-English, as seen in (1).

(1) ***What** do you think ***who*** the cat chased
c.f. Who do you think the cat chased? (Thornton, 1990)

In Chapter 4, we discussed two separate experiments in which we demonstrated that English-speaking kindergarteners systematically produce medial wh-phrases in biclausal questions about 20% of the time. This replicates findings from Thornton (1990) as well as others which suggests that this is a persistent error in children learning English. We considered two hypotheses which might account for this error. The first is the *Parametric Acquisition Hypothesis,* which suggests that children might apply an option available in Universal Grammar (UG), which does not apply in the language they are learning. In this instance, this error resembles Wh-Scope Marking structures (SynSM[36]) which are viable in languages like German and Hindi, but not in English. As discussed in detail in Chapters 2 and 3, SynSM languages typically include a true wh-phrase as well as an additional wh-phrase which marks the scope of the true wh-phrase (McDaniel, 1989). This is illustrated in (2), from German, in which the true wh-phrase is *mit wem* and the *Scope Marker (SMer)* is *was:*

(2) **Was** glaubst Du ___ mit **wem** Maria ____ gesprochen hat?
What think you ____with **whom** Maria _____ spoken has?
Whom do you think Maria spoke with?

Under a Parametric Acquisition Hypothesis explanation of this error, English-speaking children have either temporarily misjudged English as a SynSM language or perhaps SynSM is the initial 'setting' for a biclausal question. Regardless of the exact model, one clear prediction is that if a child truly had adopted SynSM as part of their grammar, one would expect children's comprehension to reflect this error as well: if the error is in children's grammar it should appear in both their production and comprehension. De Villiers and Roeper (1995) have suggested that this is, indeed, the case with medial-wh errors. Their work has repeatedly shown (e.g. de Villiers, Kotfila, & Klein, 2019; de Villiers & Pyers, 2002; de Villiers, Roeper, Bland-Stewart, Pearson (2008); de Villiers, Roeper, & Vainikka, 1990) that in questions-after-stories tasks, children will respond to the medial wh-phrase in questions such as (3):

[36] Throughout this chapter, we use SynSM to refer to the *syntactic* phenomenon of Wh-Scope marking (for reasons described in Chapter 3). However, when referring to productions made by English speaking children, we use *medial wh-productions*, in order to remain theoretically neutral. When referring to responses given by children which seem to indicate they have interpreted the question as if it were an instance of SynSM, we use *WSM-like interpretations* to distinguish them.

(3) **How** did the boy say **what** he caught?
A: a fish!

In other words, rather than responding to the true wh-phrase, *how,* they have mistakenly taken the wh-relativizer *what* as the true wh-phrase (the one to answer) and assumed *how* is acting as the SMer. They have interpreted (3) to mean (4):

(4) What did the boy say (that) he caught?

Taken together with the production of medial wh-phrases, this is apparent support for the Parametric Acquisition Hypothesis: the mistaken treatment of medial wh-phrases as the ones to respond to is appearing both in production and comprehension. This appears to be an example of *Syntactic Creativity* (Schulz, 2011) because children seem to be applying the grammar of another, UG-licensed grammar to English.

However, in Chapter 4, we showed that responses to the medial wh-phrase are greatly reduced when the questions address the relevant *Question Under Discussion (QUD)* (Roberts, 2012). QUD refers to the topic of discussion. In other words, when the story placed equal prominence on the part of the story relevant to the *how the boy did the saying* and the part of the story relevant to *what he caught,* children performed better. This further supports Thornton & Crain (1994)'s finding that placing all the emphasis on the matrix wh-phrase (making the story mostly about the *saying* event) swayed children to make errors in that direction. Furthermore, in Chapter 4, we saw that children's performance on our production task did not predict their performance on the comprehension task: there was no correlation between errors in production and errors in comprehension. This too is a similar finding to Thornton & Crain (1994). While these findings are not evidence *against* the Parametric Acquisition Hypothesis, they prompted us to seek another explanation: The *Limited Processing Hypothesis.*

The Limited Processing Hypothesis suggests that children's grammar is essentially adult-like at this stage, but their immature processing mechanisms limit their ability to produce adult-like constructions or give adult-like responses to questions. To understand precisely how these processing mechanisms might be limited, we first need to understand the differences between comprehension and production mechanisms, as they are described in the literature. The model of comprehension we assume is based on that described by Phillips & Ehrenhofer (2015), which states that comprehension is incremental and that as each new word is uttered, it must be integrated into the current interpretation of the utterance as a whole. Every word that is processed is related to the words preceding it and will sometimes satisfy requirements of those words and sometimes introduce new requirements. These requirements (thematic roles, scope relations, agreement relations, etc.) must be met either by accessing an item in memory or by waiting for a new item to be introduced. This requires resource management as well as many different processes operating in parallel. The comprehender must be able to manage these resources without allowing for interference by items that are superficially similar (for instance, reflexive attraction effects[37] (Cunnings & Felser, 2013)). This task requires non-negligeable use of

[37] Reflexive attraction effects happen when a reflexive pronoun is preceded by two noun phrases which could be antecedents. Though only one satisfies the appropriate c-command and locality constraints, comprehenders sometimes temporarily interpret the other noun phrase as the antecedent of the reflexive. For instance, in *Bill heard that the old man had injured himself in a cycling accident, the old man* is the appropriate antecedent for *himself*, and yet eye tracking data show that readers temporarily consider *Bill* to be the antecedent.

resources and adults often mis-parse sentences and must reanalyze, making many errors (e.g. Ferreira & Henderson, 1991; Sturt, 2007).

The processing model of production which we assume follows three levels (Dell, 1986; Bock, 1987; Levelt, 1989): first, a semantic level where the message is formulated; second, a syntactic level where possible syntactic structures are generated in parallel (Garrett, 1980; Bock & Ferreira, 2014; Coppock, 2010); and third, a morphophonological level where lexical items are selected and phonemes are mapped onto the chosen structure. These levels are established based on the fact that production errors in adults generally surface as either semantic errors, syntactic errors or morphophonological errors. Though they are clearly not completely separate since there are also examples of cross-level influence. For instance, phonological errors typically do not result in non-words, which one might predict if there were no connection between the semantic level and the morphophonological level (Dell & Reich, 1981).

Evidence that syntactic structures are generated in parallel come from yet another type of error: syntactic blends. Syntactic blends have been argued to reflect alternative formulations of the message the speaker wishes to communicate, which have been "intertwined in speech" (Garrett, 1980). For instance (5), an example from a corpus study by Coppock, (2010), is argued to be a single structure combination of (5′) and (5″).

(5) It's not so pretty bad.
(5′) It's not so bad.
(5″) It's pretty good.

The speaker wishes to communicate something along the lines of *This is alright*, and they come up with both (5′)and (5″) as alternatives for expressing this message. This cannot be argued to be a simple switch of words because simply swapping *good* for *bad*, for instance, will not repair the structure. It is, in fact, the *not so* and *pretty* both surfacing which suggest its blend status because they should not appear together. Syntactic blends have also been discussed in bilingual acquisition when speakers produce structures which include aspects of both of their languages, suggesting activation of both languages (e.g., Goldrick, Putnam & Schwarz, 2016). It has been argued that blends are the result of processing load (Coppock, 2010). The speaker has to select the structure to use which will optimally express their meaning and apply the relevant words. This minimally requires memory, inhibition, planning, and attention. When these systems trip, one outcome is a syntactic blend. Jaeger (2005) investigates blends of various types in child corpora and comes to a similar conclusion that these errors reflect the development of attention in planning. To reiterate Coppock (2010), given the number of cognitive processes involved, it is almost shocking that these errors are the exception rather than the rule.

To summarize, the models of processing involved in production and comprehension both require substantial work by various cognitive mechanisms. However, one aspect which both processes require is working memory (WM). Indeed, according to the Baddeley (2017) model of the executive control system, WM acts as a kind of buffer between the processing systems and action. In other words, a person's WM skills mediate between what they have processed and what they say. This suggests that WM should be able to be used as a proxy measure for processing mechanisms, generally and indeed this has been done (e.g. Johnson et al., 2010). Previous work such as Delage and Frauenfelder (2019) has even used WM as a measure of children's production and comprehension processing. This gives us a direct means of testing the Limited Processing Hypothesis: if these errors are the result of processing limitations, then children with lower WM should make more errors. Thus, we should expect a correlation between children's performance on the language tasks and their WM. However, we should not necessarily expect performance on the language tasks to be correlated with each other since it could be the case that low WM simply predicts non-adult like performance, and not necessarily medial-wh productions or WSM-like responses in a comprehension task. For example, a child with low WM might produce many questions with medial wh-phrases in the production task, but perhaps the error this leads to in this child's comprehension is not a WSM-like interpretation, as in (3), but instead simply responding with the incorrect response to the matrix (*how*) question.

We should also clarify that we are not the only researchers to consider the possibility that immature processing might be the underlying cause of medial wh-phrases in children's biclausal questions. Grolla and Lidz (2018), conducted a study very similar to our production tasks described in Chapter 4: a puppet provided a lead-in to elicit biclausal questions from children. They tested the specific hypothesis that medial-wh phrase production was the result of immature inhibitory control. Their suggestion is that children accurately move the wh-phrase from its base position in the embedded clause, pausing at the clausal boundary, but fail to inhibit pronouncing the wh-phrase that appears medially (supporting a claim made by Thornton & Crain, (1994) that these errors are evidence for cyclic movement). However, this is only a plausible explanation for those productions which include a copy construction (the initial and the medial wh-phrase are the same). They cannot account for constructions which employ two different wh-phrases.

In the following sections, we will describe a within-subjects experiment specifically designed to test the Limited Processing Hypothesis, using a measure of WM as a proxy for processing ability. In addition to our primary question (Is performance on comprehension and production tasks correlated with WM ability?), we will address several other questions about children's performance building on the work presented in Chapter 4 as well as the model of the adult grammar described in Chapter 3. These questions include:

- Will removing the lead-in used in the production task affect children's performance? Previous work, including our own, included the use of a prompt *The question starts 'Who do you believe…'* to elicit biclausal questions. However, it is possible this prompt affected children's productions.
- Will adjusting subtle aspects of the pragmatic background of the stories such that they reflect survey findings reported in Chapter 2 affect children's use of WSM-like responses in the comprehension task?
- Do we think children have the adult-grammar perfectly established (and all errors are the result of limited processing) or is there a possibility that children have yet to establish the adult-like grammar?

The results will show a lack of correlation between the production and comprehension task (replicating the findings of Chapter 4), but a strong correlation between each language task and WM skills. We conclude that children generally do have an adult-like English grammar with some caveats and that the errors they make are overwhelmingly the result of immature processing.

5.2 Production, Comprehension, and Working Memory

The primary purpose of this experiment is to establish whether children's WM abilities are correlated with their performance on production and comprehension tasks involving biclausal questions. We also want to establish whether removing the 'lead-in' used in the elicited production task in Chapter 4 (as well as previous studies such as

Thornton (1990); Thornton & Crain (1994); Grolla & Lidz (2018)) will affect children's production of medial wh-phrases. Finally, we want to establish whether we replicate the finding of Chapter 4 that production and comprehension errors are not correlated. To that end, we will include three tasks in our experiment: a production task; a comprehension task; and a WM task. While the comprehension task remains unchanged from Chapter 4 and the WM task is standardized, the production task is novel and hopes to build on that methodology by removing the lead-in.

In the novel version of the production task, we used a continuous, interactive story which encouraged the participant to ask questions to two characters (as in Chapter 4). To elicit biclausal questions without the lead-in, however, we created a child-friendly version of a "translation task" initially used by Schulz (2011) to elicit biclausal questions from L2 speakers of English. Because one character (Mindy) speaks English, but the other character (Mork) does not, it makes it impossible for the child to ask a simple monoclausal question to Mork. However, they can ask Mindy what Mork thinks about various things because she can speak to him. The success of this novel task is a significant addition to this area of research since the use of a lead-in is unnatural and limits children's productions in ways which will be illustrated below.

5.2.2 Participants

Fifty children participated in this study. As with our experiments in Chapter 4, data from children who could not perform filler tasks were excluded (n=18). This left 32 children remaining whose age ranged from 4;4 to 6;7 (average 5;5). This age range was selected based on the abilities of children to perform these tasks reported in Chapter 4, as well as several pilot tests (n=8). We extended the age range from the previous tasks in order to establish a maximum age at which children would make such errors.

It should also be noted that due to the COVID-19 pandemic, 12 participants participated in person, while the remaining 20 did so virtually in order to comply with social distancing restrictions. Necessary changes to the procedures will be described in the following section. The affect this had on children's performance will be described the results section.

5.2.3 Procedure

This experiment had a within-subjects design. Every participant completed three tasks: production, comprehension, and working memory (WM) tasks. The production and comprehension task were balanced for order: half of the participants did the production task first; half did the comprehension task first. All participants completed the WM task last because pilot trials revealed that when the WM task, which is a repetition task, preceded the language tasks, children tended to simply continue repeating what the experimenter said and had a more difficult time understanding the language tasks. All in-person sessions were filmed with an HDR-CX160 SONY Handycam. Sessions which were performed virtually were conducted via the Zoom interface and were recorded using that application. In the following section, we explain these three tasks in detail.

Production Task

Recall that the goal of the task was to elicit natural, biclausal questions from children. We used a continuous interactive story paradigm (similar to our production task reported in Chapter 4). The story involved an alien, Mork, and a friendly raccoon from earth, Mindy. The story was accompanied by PowerPoint illustrations and animation. Mork appeared as a cartoon, while Mindy appears first as a cartoon and then "comes out of the computer" as a puppet to communicate with the child. This establishes a tangible connection to the story. In the online version of the task, children were introduced to the puppet virtually and then she 'beamed up' into her spaceship on the computer, by simply adding an animated version of her to the image. The premise is that Mindy has been traveling around outer space and she wants to show her new friend, Mork, her home planet of earth. Their spaceship crashes and is spread all over the countryside. The child was asked if they would help Mork and Mindy find the pieces of their spaceship. The three need to decide where to look. In the in-person version of the task, the child was given a laminated outline of the ship and was told they will find pieces they will need to put on the spaceship to fix it. For the online version, a portion of the screen was dedicated as their "pocket" and pieces of the ship appeared there as they found them. This is an apparatus that appears frequently in children's television and online games, and children had no trouble understanding. After half of the trials, the child received a new piece of the spaceship. The piece would appear on the screen when the characters found it and then it was either taken "out of the screen" (from behind the computer) and given to the child to attach to their ship in the in-person task or it appeared in the game "pocket" in the virtual task. This was to maintain interest and connection with the story. Mindy explained that Mork does not understand English, but she can tell us what he was saying because she speaks his alien language. This means whenever we wanted to find out about Mork's ideas we needed to ask Mindy to help us. Based on Experiments 1 and 3, we had established that it was crucial that children know exactly what information they need to ask about. Therefore, we gave them multiple examples of the monoclausal question before they attempted the target. To this end, the experimenter always followed the pattern of asking the child's opinion about the problem first (e.g. *Who can fix the engine? Do you have an idea?*), followed by prompting the child to ask the monoclausal version of the question (*Can you ask Mindy "Who can fix the engine?"*), and then finally eliciting the target (*Who does Mork think can fix the engine?*). An example exchange is below. It should be noted that the experimenter never included the complementizer, *that,* in any practice or example.

> (6) Experimenter: Alright, so we know Farmer Bob was good at fixing things because he fixed Farmer John's fence. Maybe he could fix the engine? But we also know Jen the mechanic said she could fix anything. Hm. Farmer Bob or Jen the mechanic… Who can fix the engine? Do you have an idea?
>
> Child: Jen the mechanic!
>
> Experimenter: That's a great idea! Let's find out about Mindy. Can you ask her *Who can fix the engine?*
>
> Child: Who can fix the engine?
>
> Mindy: (always disagrees with the child to create a dichotomy) I think Farmer Bob can fix the engine.
>
> Experimenter: Good job, but it looks like you two have different ideas. Let's find out about Mork. Can you ask Mindy about Mork?
>
> **Child: Mindy, who does Mork think can fix the engine? (target)**
>
> Mindy: I don't know! I'll ask him! (speaks alien language)
>
> Mork: (speaks alien language)
>
> Mindy: He says he thinks Jen the mechanic can fix the engine!

Because the notion of translation was somewhat difficult for children, we incorporated four practice questions as well as three fillers into the story. The first four served the purpose of helping the child to understand the task of "translation" so they would know that they could not simply ask Mork the monoclausal version of the question. During these four practice trials, the experimenter modeled the structure for the children (e.g. Let's ask Mindy: *Where does Mork think we should look?*). Practice trials were questions children were familiar with when meeting a new friend: *What's his name? What's his favorite color?* Two further filler items were specific to the theme of the story, and which include more complex syntax, though not the target structure: *What does he want to see on earth? Where does he think we should look?*

Target structures were less rigidly defined than in the production task from Chapter 4. Children were allowed to use what verbs they preferred (*think, believe, want, etc.*). While we designed the task to include 6 object extraction targets and 6 subject extraction targets, we accepted any production that included more than one clause. Furthermore, all target constructions use the wh-phrase *who* because this allowed us to determine whether children were changing the wh-phrase to be an SMer (e.g. *what* in English; *was* in German) or producing a 'copy-construction', which is a variant on SynSM which uses the same wh-phrase twice rather than incorporating an SMer. Target structures are listed in Table 1, below. Numbers indicate the order of presentation for Form A. Half of participants will hear a story that elicits questions in this order, the other half will hear a story eliciting questions in the reverse order (Form B). These two orders (Forms A and B) will allow us to determine whether order of presentation affected participants' productions. A full transcript of the story that binds them together can be found in Appendix F.

Table 1. *Target Questions*

Subject Extraction	Object Extraction
(1) Who does he think has the wings?	*(4) Who does he think (that) we should ask?*
(2) Who does he think has a hammer?	*(5) Who does he think (that) we should help?*
(3) Who does he think can fix the fence?	*(6) Who does he think (that) we should give the tires to?*
(8) Who does he think knows the way?	*(7) Who does he think (that) we should invite?*
(11) Who does he think can show us the sharks?	*(9) Who does he think (that) we should listen to?*
(12) Who does he think has a better t-shirt	*(10) Who does he think (that) we should see?*

We did not give children a lead-in, so there was, as expected, a great deal of variation from these target structures. Because we were interested in errors, any thematically appropriate biclausal question was accepted, regardless of whether it was target-like. In other words, we were not looking for target-like or even adult-like constructions. Our focus instead was on the kinds of mistakes the children make. Children were given three chances to produce a biclausal question before the experimenter prompted them directly. There were several ways the child might respond, but which the experimenter could work with to direct them toward the target. For example, consider (7) which was the same scenario presented in (6), but with several typical, non-target responses:

(7)

Experimenter: Good job, but it looks like you two have different ideas. Let's find out about Mork. Can you ask Mindy about Mork?

Child (attempt 1): Mindy, what about Mork?

Experimenter (as Mindy): What should I find out?

Child (attempt 2): Who can fix the engine?

Experimenter: Remember, we know Mindy thinks Farmer Bob can fix the engine, but we need her to find out about Mork's ideas.

Child (attempt 3): Mindy, can you find out about Mork's ideas?

Experimenter: Remember earlier when we wanted to find out "where should we look?" you asked Mindy "Where does Mork think we should look?" Can you ask a question like that about who can fix the engine? It starts "Who does Mork think…"

Child (attempt 4): Who does Mork think can fix the engine? (target)

In this scenario, the fourth attempt would be coded as adult-like, but with the lead-in given. If the child did not use a biclausal question by the fourth attempt, the experimenter asked the child to repeat after them and gave the target. In that case, the response was coded as "excluded". This was to prevent frustration and to help elicit the target structure in the next trial. After completing all the questions, children would have received all the pieces necessary to complete their spaceship. At this point, the child was allowed to "put the spaceship in the computer" (place it behind the screen) and they see it appear on-screen. Mork and Mindy thank the child and return to their adventures.

The task was difficult, particularly for younger children participating virtually, but every participant completed at least half of the target questions. Crucially, the task proved enjoyable and entertaining for the children and many asked if they could continue playing with Mork and Mindy or go back and see what would happen if they decided to do something different at any point.

Comprehension Task

The comprehension task was the ‘questions after stories’ task seen in Chapter 4. Our survey of adult speakers, presented in Chapter 2, suggested that the Question Under Discussion (QUD) (Roberts, 2012) and contrastive topic (CT) (Büring, 2003) status of the subject plays a role in the use of SynSM. Recall that QUD essentially refers to the topic of conversation while CT refers to an entity presented in *contrast* with another (for details see Chapter 2). While we had already controlled for QUD in our stimuli (see discussion in Chapter 4), we had not accounted for CT previously. The stories were, therefore, changed such that the subject would be interpreted as CT. In this way, we could compare these results to the result from the comprehension task in Chapter 4 to determine whether the relative CT status of the subject affected children’s responses. According to our OT analysis, presented in Chapter 3, English-speaking children should not be affected by the CT status of the subject if they have an adult-like grammar of biclausal questions. Tableau 13 (in Chapter 3) suggests that FAITHCT is very low ranked in English and, thus, plays no role in determining the optimal outcome. In (8) below, the only difference in the story compared to its counterpart in Chapter 4, is therefore the addition of “Detective David” who tells Sherry something in contrast to what Steve tells her.

Children were presented with short stories followed by a pre-recorded question posed by a puppet, Hillary Hippo. The questions used prosody appropriate for child-directed questions that included a CT subject. All stimuli were the same as those used in Chapter 4 and were balanced for event prominence (the *saying* event and the *telling* event were equally important in the story). For reference, an abbreviated example story and question appear in (8):

> (8) Detective Sherry wants to catch the famous bad guy Evil Steve. Her boss, Detective David, comes to her office and says “Sherry, we have been told that Evil Steve is planning to steal the queen’s crown! It is on display at the museum, so be on your guard! Maybe you will catch him!” Meanwhile, Evil Steve really is planning to steal the queen’s crown, but he knows Detective Sherry will expect that, so he decides to trick her! He is going to *tell* Sherry that he will steal the queen’s diamond ring and that will distract her from the crown. Steve tries using his special TV machine to tell Sherry he’ll steal the ring, but it breaks. He can’t tell Sherry that way. Instead, Steve writes Sherry a letter. Sherry reads the letter and thinks the ring is in danger, so she goes to check on it, but when she gets there, the ring is fine. Sherry figures out that Steve was tricking her, so she goes to the museum where the crown is on display. She catches Evil Steve and the queen’s crown is safe.
>
> *Q: How did EVIL STEVE* tell detective Sherry what he was going to steal?*
>
> **Capital letters indicate prosody appropriate to CT*

This was an open-ended task and children could respond however they wanted. As a reminder, four responses which frequently appeared in the comprehension task in Chapter 4 were adult-like responses (the letter), alternative *how* responses (the TV machine), WSM-like responses (the ring), and embedded clause responses (the crown). We expected to see at least these four responses in the new version, but to varying degrees, given its open-endedness.

Working Memory Task

Finally, we included a task to assess children’s WM, which we will use as a proxy for general cognitive ability (Johnson et al. 2013). Following work such as Delage and Frauenfelder (2019), we will use two tasks: a Forward Digit Span (FDS) and a Backward Digit Span (BDS) (see description below), which have both been shown to correlate with performance on tasks with complex syntactic structures[38]. Delage and Frauenfelder (2019) suggests that FDS is relevant to comprehension while BDS is relevant to production because it requires manipulating the item held in memory. Our data analysis will primarily report children's *composite score*, which is simply their FDS and BDS combined. Wilde, Strauss, & Tulsey, (2004) suggests that composite scores give a more complete picture of the child's processing capabilities.

There are multiple ways of establishing digit span measures. We will use the Differential Ability Scales -II (DAS-II) (Elliot, 2007), which includes a standardized version of these tasks. Recall that digits may reflect the child's short-term auditory and sequential memory, and their oral recall of sequences, as well as their concentration and attention (Elliot, 2007). The BDS, while typically considered appropriate for children over 5;0 was included because it requires holding information in mind as well as manipulating it. In other words, to repeat 1-7-5 backward, you must remember the first digit was 1 and that it should be said last as you say *5-7*. This added layer of complexity could give us further insight into children's WM and planning abilities (Elliot, 2007).

During both tasks, the procedure was the same except that during the FDS participants were asked to respond with the digits in the same order as they were given them. In the BDS, participants must give the digits in the reverse order. Beginning with a two-digit string, the experimenter read a string of digits and the child's task was to say as many digits as they remembered in order. For the trial to be successful, the child had to

[38] The decision to use WM as a general measure was made after a discussion with Professor Bonnie Nozari (formerly affiliated with the department of Neurology, JHU; currently at the department of Psychology at Carnegie Mellon University), who is an expert in researching executive control and has experience using these tasks with both adults and children. Professor Nozari was also instrumental in deciding which standardization of the task to use.

repeat all the digits and in the correct order with no additional digits. If the trial was successful, the experimenter proceeded to the next block of digits until the child was unsuccessful. When the child was unsuccessful, the experimenter returned to the previous block and children were given four strings of digits for that block. Once the child reached a digit span where they successfully complete no more than one trial, the task ended. Children were given a point for every successfully completed trial. The raw score was then used to calculate an *ability score,* which was determined by total successful trials and the relative difficulty of those trials. While many studies use only the ceiling of the FDS, we consider the ability score established using the DAS-II to be a more complete picture of the child's performance.

The Limited Processing Hypothesis predicts that children's performance on production and/or comprehension tasks should be correlated with their performance on WM tasks. If immature processing mechanisms lead children to make errors resembling SynSM, then we should find a relationship between WM and their ability to produce our target structures accurately. In other words, a child with high WM should perform well on the production and/or comprehension tasks. A child who performs poorly on WM task should also perform poorly on our language tasks. However, if these errors were the result of a mis-set parameter or non-adultlike constraint ranking, then errors in production and comprehension should be correlated regardless of WM.

5.2.4 Data Coding and Analysis

WM

Individuals' BDS and FDS were combined to give them a composite score, which we will report as their WM ability.

Production

Children's questions were transcribed and coded by the author. There were a total of 384 attempted trials. Of these, only 26 were excluded because the experimenter had to prompt the child with the full structure. Any production that was pragmatically relevant was accepted whether it was target-like or not. There were 5 instances where the child never produced anything except a monoclausal version of the structure. Together, this means that only 31 were not attempts at biclausal questions. The primary production types were target-like LD constructions, constructions with medial wh-phrases, and those with overt complementizers, as we saw in our production tasks in Chapter 4. However, because of the open-ended nature of the task several other production types were also used and accepted. These included structures which were adult-like in terms of syntax, but could not be considered adult-like because of a lexical error. For instance, if the child used *what* rather than *who* in a question that was about a person as in (9). These were coded as 'lexical errors.'

(9) Target: Who does Mork think we should invite?
Production: What does Mork think we should invite?

Another common production involved swapping the matrix verb to *want* and using an infinitival clause. These were accepted since they involve more than just one clause and in many cases were adult-like, as in (10), below. These were coded as ‘infinitival responses’.

(10) Target: Who does Mork think we should invite?
Production: Who does Mork want to invite?

There were also utterances which contained errors unrelated to question formation (coded as “non-wh-errors”) as well as “think about” which included any structure with ‘think about’ or ‘think of’ as in (11).

(11) Target: Who does Mork think can fix the engine?
Production: Who does Mork think about to fix the engine?

In addition to these descriptive statistics we will present the results of several mixed effects logistic regression models. The 353 utterances which were not excluded and did contain more than one clause were coded for whether they did (1) or did not (0) include a medial wh-phrase. They were also coded for whether they were (1) or were not (0) *Target-like*, meaning they were an LD structure which used appropriate lexical items and verbs, and whether they were (1) or were not (0) *adult-like*. The distinction between *target-like* and *adult-like* is simply that an *adult-like* structure could include any grammatical structure of English (e.g. infinitival responses could be *adult-like*, but are not *target-like*), *target-like* is restricted to LD constructions.

Data were submitted to a mixed effects logistic regression model (e.g. Krushchke, 2015) in the R statistical analysis environment (R Core Development Team, 2015). The models included fixed effects of Composite WM score as well as age, which was coded as a numerical value and centered as well as scaled. The random effects structure of the model included crossed participant and item effects and was run with the maximal random effects structure that would converge (Barr et al., 2013) minimally including random intercepts for participant and item[39].

Finally, one of the goals of this experiment was to determine whether the lead-in affected children's performance. We therefore combined the data from this task with the data from our within-subjects experiment from Chapter 4 and ran another model to determine whether medial-wh production was predicted by the experimental paradigm. The model included Experiment (3 vs. 4) and Age (centered and scaled) as fixed effects and participant and item numbers as random effects and run with the maximal random effects structure that would converge, which included random slopes of Experiment and Age for participant, when possible. Furthermore, since the hypothesis here was that it was possible the mere presence of the lead-in might cause children to make more errors and there were some instances even in this version of the experiment where the lead-in was given if the child was really struggling, we also submitted this data to the same model with lead in (present, absent) as a fixed effect instead of Experiment.

Comprehension

As in Chapter 4, children's responses to the target questions fell into six primary types: correct manner (adult-like), false object (WSM-like response), distractor manner (alternative how response), true object (response to the embedded question), "trick" response (pointing out that the character lied), and 'other' responses. Comprehension data were binomially coded as being each of these response types (1) or not (0). The data were then submitted to several mixed effects logistic regression models in the R environment,

[39] The model would not converge with a fixed effect of order, however, a separate model confirmed that order was not a significant factor.

as with the production data. The model always included a fixed effect of composite WM as well as age (centered and scaled) when the model would converge with this effect. Again, the model was run with the maximal random effects structure that would converge including at least random intercepts for participant and item. Age and WM were used as random slopes for participant when the model would converge.

Comprehension and Production

Another goal of this experiment was to replicate the lack of correlation between production and comprehension found in Chapter 4. We therefore submitted the production data (coded binomially for whether there was (1) or was not (0) a medial wh-phrase) to another model with fixed effects of WSM-like comprehension (coded binomially for whether responses were (1) or were not (0) a WSM-like response). The model included the maximal random effects structure that could converge, which included random intercepts for participant and production task item and comprehension task item as well as random slopes of WSM-like comprehension and age for participant.

5.2.5 Results

WM

Children who participated in this experiment showed an average composite score of 228.5. Considering the range of ages (4;5 to 6;7) this is comparable to the standard reported in Elliot (2007).

Production task

We first present a descriptive illustration of the results of the production task. As described above, there were more production types than in our previous production tasks, which can be explained simply as a result of the lack of lead-in. Table 2 presents the primary types of constructions children produced during this task.

Table 2. *Production Types*

Production Type	**Example Production**	**Percentage of Total Productions**
Target-like	*Who does Mork think we should help?*	33.9% (all grammatical)
Medial Wh	*What does Mork think who has the wing?*	19.5% (all ungrammatical)
Infinitival Responses	*Who does Mork want to help?*	19% (11% ungrammatical; 8% grammatical)
Lexical Errors	*What does Mork think we should invite?*	12.5% (all ungrammatical)
Non-wh Errors	*Who does Mork thinks we should invite?*	10.2% (all ungrammatical)
Overt Complementizer	*Who does Mork think that knows the way?*	3.1% (2.8% ungrammatical; .3% grammatical)
***Think about* structures**	*What does Mork think about which one has the wing?*	2.5% (1.1% are grammatical; 1.4% ungrammatical)
***Can you ask Mork* "…"**	*Can you ask Mork who he thinks can fix the engine?* *OR* *Can you ask "Who do you think can fix the engine?"*	1.6% (.6% ungrammatical; 1% grammatical)
Q1 and Q2	*What does Mork think and who can fix the engine?*	1.1% (all grammatical)
Gave a choice	*What does Mork think? Farmer Hans or Dr. Mary?*	.8% (all grammatical)

Note that these percentages will not sum to 100, as construction types are not mutually exclusive and there was some overlap. For example, *What does Mork think about to fix the engine?* would be an infinitival response as well as a 'think about' structure. It would not count as a 'lexical error' because those are exclusively constructions which would be target-like if not for a word choice.

While the average production of medial wh-phrases (19.5%) is similar to Experiments 1 (22%) and 3 (15.4%) (the two production tasks from Chapter 4), there are a few differences we wish to highlight. While in both previous experiments, the majority of medial wh-phrases were the same as the initial wh-phrase, resembling copy-constructions (82% in Experiment 1), in this experiment, only 37% of medial wh-phrase productions resembled copy constructions. Most used the wh-phrase *what* initially and either *who* or another wh-phrase medially. There were also only 11 instances of productions with overt complementizers (*that*) (3.3%) compared with 33% from Experiment 1.

Does the lead in affect productions?

One of the goals of this experiment was to establish whether the lead-in affected children's productions. To that end we compared the findings of this experiment to the production findings from Chapter 4 (which did include a lead-in). We ran two separate models to determine whether one experiment elicited more SynSM structures than the other. The first examined whether which experiment the child participated in predicted whether they would use SynSM structures (This might have been the case if, for example, something about the pragmatics of one of the stories elicited SynSM more than the other).

The second model asked whether the presence of a lead-in predicted the use of SynSM. Neither experiment nor ‘lead-in’ was a significant predictor of SynSM productions (p>.1 for both). This result suggests not only that neither task was more likely to result in constructions with medial wh-phrases, but also that the presence or absence of a ‘lead-in’ did not affect the production of medial wh-phrases. As we will discuss below, however, this task has led to fewer productions which resembled copy-constructions.

Comprehension

The results of the comprehension task largely replicated the findings from the comprehension tasks in Chapter 4. A model testing whether experiment number (this version, or the experiments in Chapter 4) elicited more WSM-like interpretations found no significant difference (p>.1 for all). This suggests no difference between performance on the two tasks, in spite of the change to make the subjects CT. No additional answer types were found, and children generally performed well on this open-ended task. The percentage of the response types are displayed in Table 3.

Table 3. *Response Types*

Production Type	**Example from the “Evil Steve” story**	**Percentage of Total Responses**
Correct Manner (adult-like)	*The letter*	60.3%
Incorrect object (WSM-like response)	*The ring*	13.1%
Distractor Manner (distractor item)	*The TV machine*	4.6%
Correct object (response to “second question”)	*The crown*	5.5%
Trick response	*He didn’t! He said he stole the ring!*	8.9%
Other responses	*The museum!*	7.7%

Notable patterns apparent in Table 23 include the fact that, yet again, the adult-like *Correct Manner* response accounts for the largest group of response types. However, one difference is that while in Chapter 4, we found no significant difference between *Incorrect Object* (WSM-like) responses and the *Distractor Manner* (distractor) responses or the *Correct Object* (second question) responses, this time a paired samples t-test reveals that there are significantly more *Incorrect Object* responses than the other two ($p<.01$, for both comparisons).

Independently, the three tasks WM, production, and comprehension showed basically standard results: the WM scores are within range for the standard scores reported by Elliot (2007); the production and comprehension tasks showed very similar results to what we saw in Chapter 4, regarding SynSM production and WSM-like interpretations. There were a few minor differences in each. In production, we had fewer productions with overt complementizers and fewer productions which used the same wh-phrase twice (resembling copy-constructions). In comprehension, in contrast to our previous work, we did find significantly more WSM-like interpretations than other responses. However, this is not because there were more WSM-like interpretations, but rather that there were fewer of the other responses.

Correlations: WM and Language Tasks

We now turn to the comparisons we made between languages tasks and WM. We ran several models to establish the relationships between the language tasks as well as between each language task and the WM task. In what follows we present the results of these investigations.

Is performance on the language tasks correlated?

As in Chapter 4, we found no correlation between performance on the comprehension task and performance on the production task. This is made apparent in Fig. 1. As in Chapter 4, we find that there is neither an overall pattern suggesting children use SynSM in their production and comprehension, nor are there individuals for whom this seems to be the case. We see there are both children who use medial wh-phrases in their productions, but never give WSM-like answers to the comprehension task as well as children who give WSM-like responses in their comprehension, but never produce medial wh-phrases.

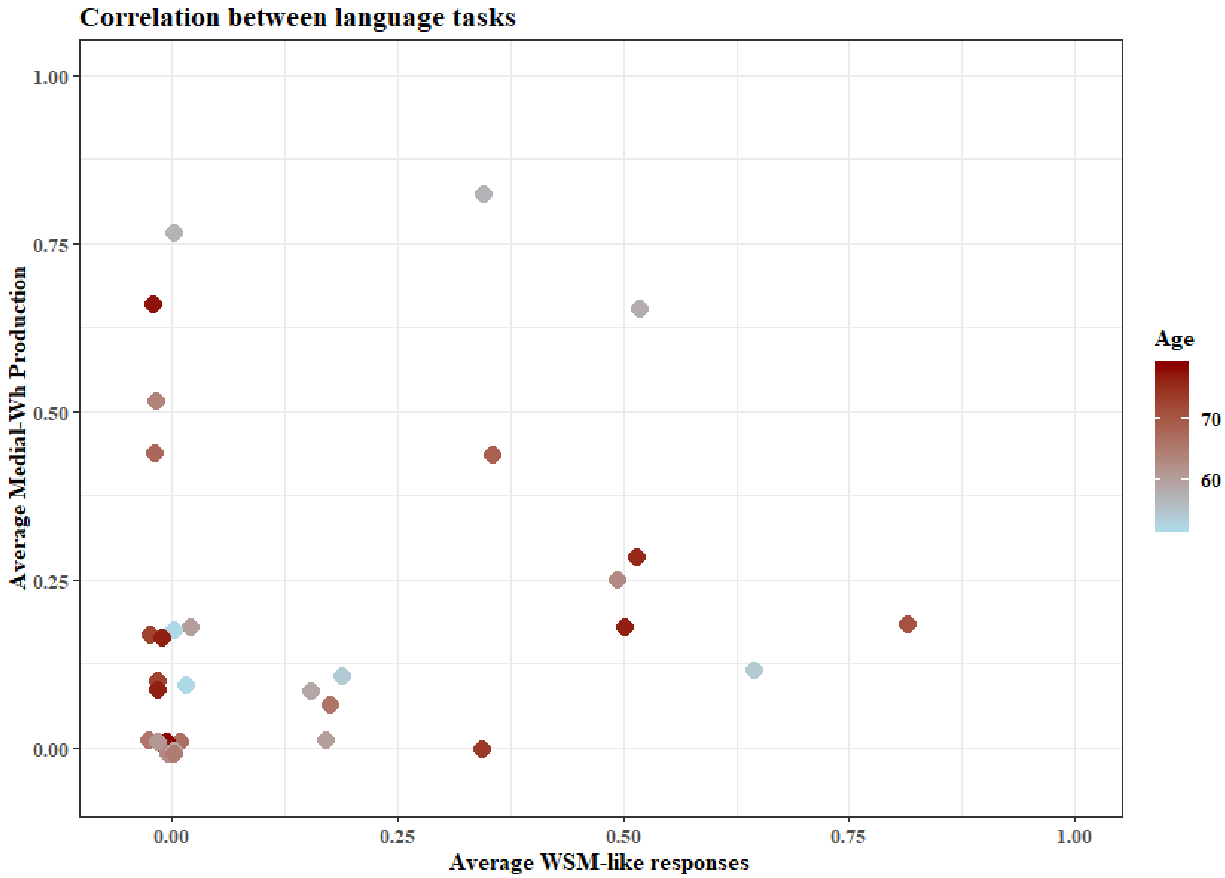


Fig 1. Average production and comprehension performance by individual. Age (in months) is illustrated by color.

The results of the mixed effects logistic regression model investigating whether performance on the production task was predicted by performance on the comprehension task further confirm this, finding no significant relationship (p>.1)

Does WM affect Production?

We now turn to the results of our investigations into a relationship between children's WM abilities and their performance on the open-ended production task (without a lead-in). Let us begin with target-like productions and adult like productions. By 'target-like', we mean specifically biclausal questions with LD movement with no other error types, while ‘adult-like’ productions refers to any grammatical construction of English (that was semantically relevant). We ran two models using WM as a predictor of target-like and adult-like productions. Both models found WM to be a significant predictor ($p<.05$, in both cases). Fig. 2 plots children’s working memory against the average of their target-like productions.

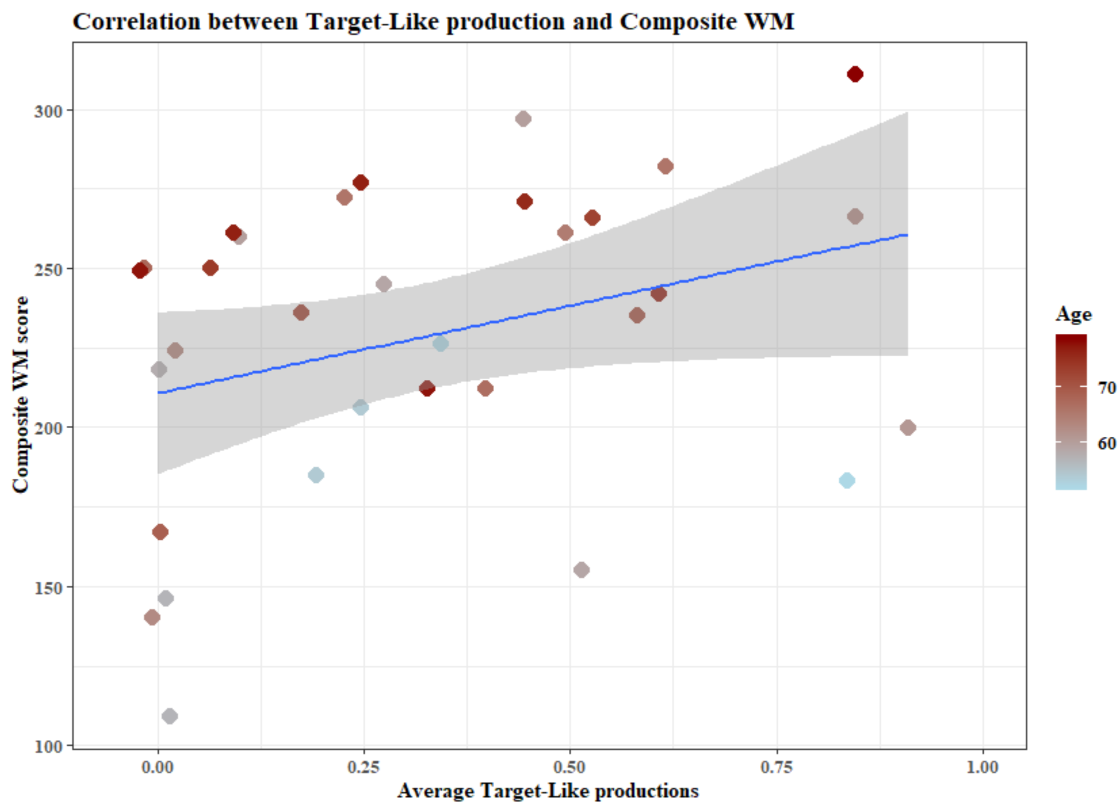


Fig 2. Average target like productions by composite WM score. Children’s age (in months) is illustrated by color. *p=.03, r= .2*

Fig. 2 illustrates that there is a relationship between composite WM and Target-Like productions: children with higher WM are more likely to produce target-like productions. In Fig. 2, we can see this general trend. We can also see that there are both young children (indicated in blue) who make mostly target-like productions as well as older children (indicated in dark red) who produce almost no target-like constructions. The results of the logistic regression model investigating this relationship are shown in Table 4.

Table 4. *Logistic regression results: WM as a predictor of Target-like structures*

	B	Error	z-value	Pr(>|z|)
Composite WM	.02	.01	2.13	0.03*

* denotes significance

We have seen, therefore, that there are positive correlations between both target-like productions and adult-like productions and composite WM scores. This suggests, unsurprisingly, that children with higher WM are more likely to produce target-like and adult-like questions. However, our primary interest lies in whether children with lower WM make more productions with medial wh-phrases. The next question, therefore, is whether there is a correlation between WM and the use of medial-wh productions. Fig. 3 plots children's average medial-wh productions against their composite WM score.

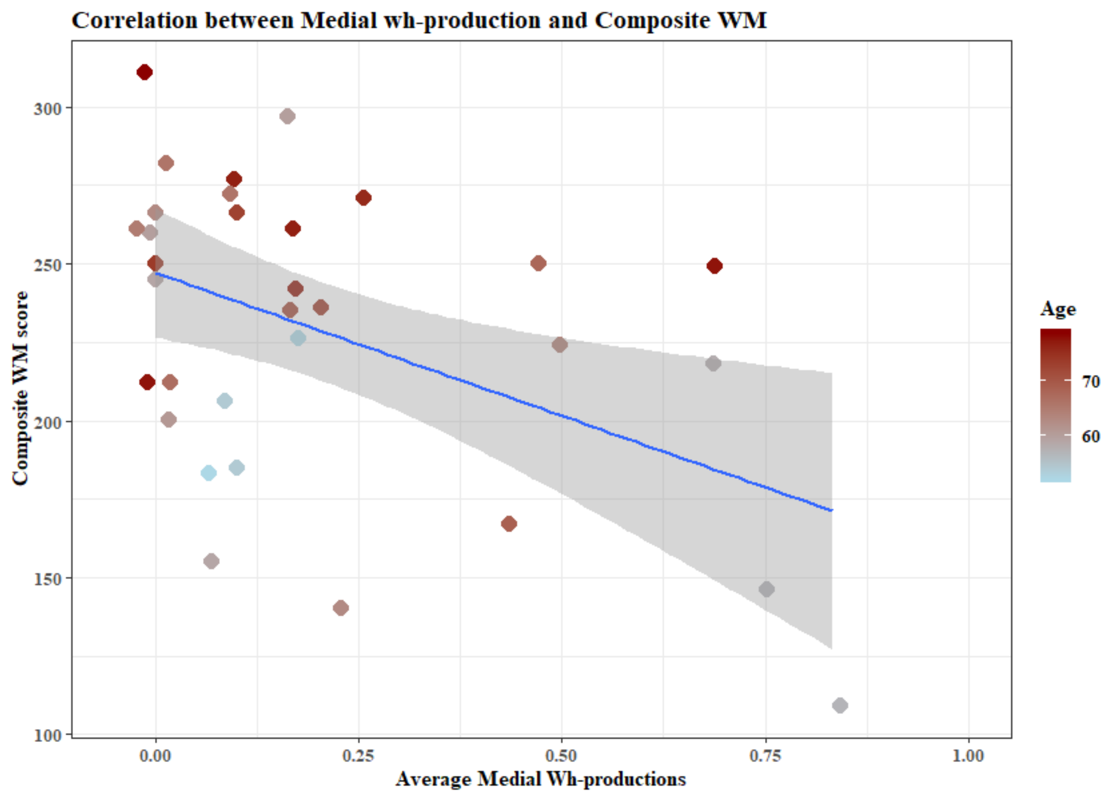


Fig. 3. Composite WM score and children's average number of Medial wh-phrase productions. Age (in months) is indicated by color. *p=.01, r=-.2*

Fig. 3 illustrates a negative correlation between medial wh-production and composite WM score, indicating that the higher a child's WM, the fewer productions with medial-wh phrases they made. Table 5 shows the results of the logistic regression model investigating this relationship. This model asked whether WM and age were predictors of medial-wh production and it found that while age was not a significant predictor (p=.33), WM was (p= .01). While the finding that age is not a significant predictor might, at first, be surprising, it is apparent in Fig. 3 that the youngest participants actually made relatively few productions of this type (all under 25%) and one of the oldest children (6;6) used medial wh-phrases in almost 67% of their productions. While most of these errors are made by children under 5;0 (as we also saw in Chapter 4), medial wh-productions are not restricted to the youngest children. However, Fig. 3 also makes it clear that those younger children have relatively high WM, while those children who produced many medial wh-phrases have lower WM.

Table 5. *Logistic regression results: WM as a predictor of medial wh-production*

	B	Error	z-value	Pr(>|z|)
Composite WM	-0.02	0.01	-2.75	0.01*
Scale (Age)	0.344	0.36	0.96	.33

**denotes significance*

In summary, we have seen that there is a positive correlation between composite WM scores and adult-like and target-like productions. We have also seen a strong negative correlation between composite WM scores and productions which include medial wh-phrases. In other words, higher WM means more adult-like and target-like productions. Lower WM means more productions with medial wh-phrases. WM significantly affects production. In the next section we will determine whether WM also affects Comprehension, as would be predicted by the Limited Processing Hypothesis.

Does WM affect comprehension?

We now turn to examining the relationship between WM and comprehension. As with the production task, we will begin with adult-like responses. Fig 4. Illustrates the relationship between composite WM and adult-like responses.

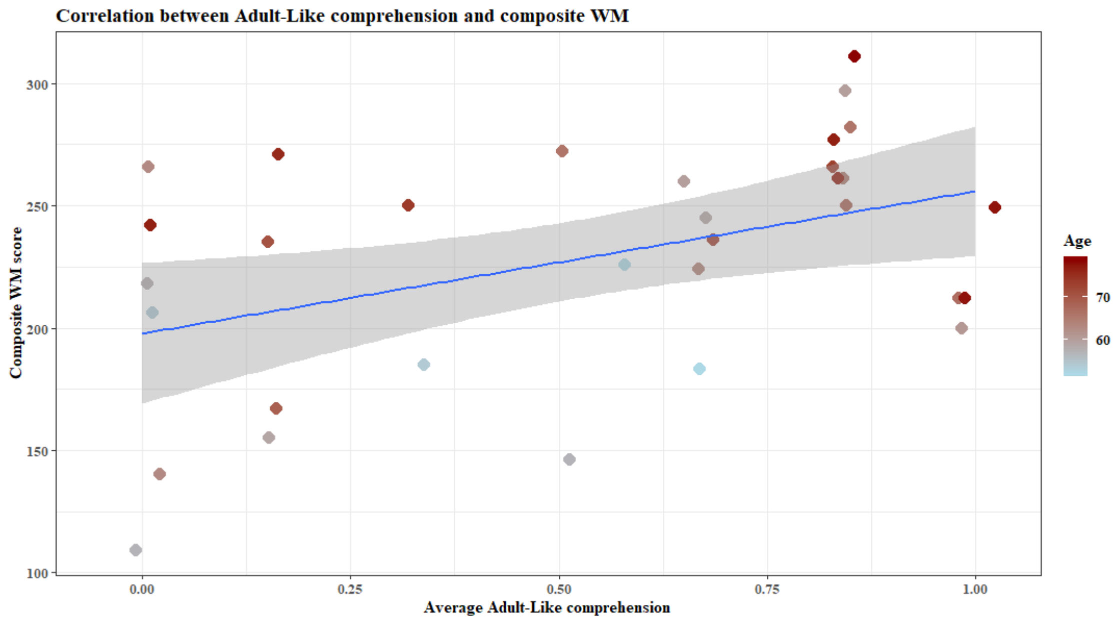


Fig. 4 Average Adult-like responses to the comprehension task by composite WM score. Each point represents an individual and their age (in months) is indicated by color. *p=.01, r=.3*

We see from Fig. 4 that there is a positive correlation between composite WM and adult-like responses to the comprehension task: the higher a child's WM, the more likely they are to give an adult-like response. Table 6 shows the results of the mixed effects logistic regression model used to investigate whether WM was a predictor of adult-like responses in the comprehension task. We see that indeed WM is a significant predictor of correct responses: a higher WM leads to more correct responses.

Table 6. *Logistic regression results: WM as a predictor of adult-like responses*

	B	Error	z-value	Pr(>|z|)
Composite WM	0.02	0.01	2.53	0.01*

However, the critical question is whether WM also predicts WSM-like interpretations. Fig. 5 illustrates children's average WSM-like responses against their composite WM scores. It is apparent in Fig. 5, that many children (regardless of WM) actually give no WSM-like responses. Those who do give WSM-like responses do show a weak relationship between their WM and the rate of WSM-like responses.

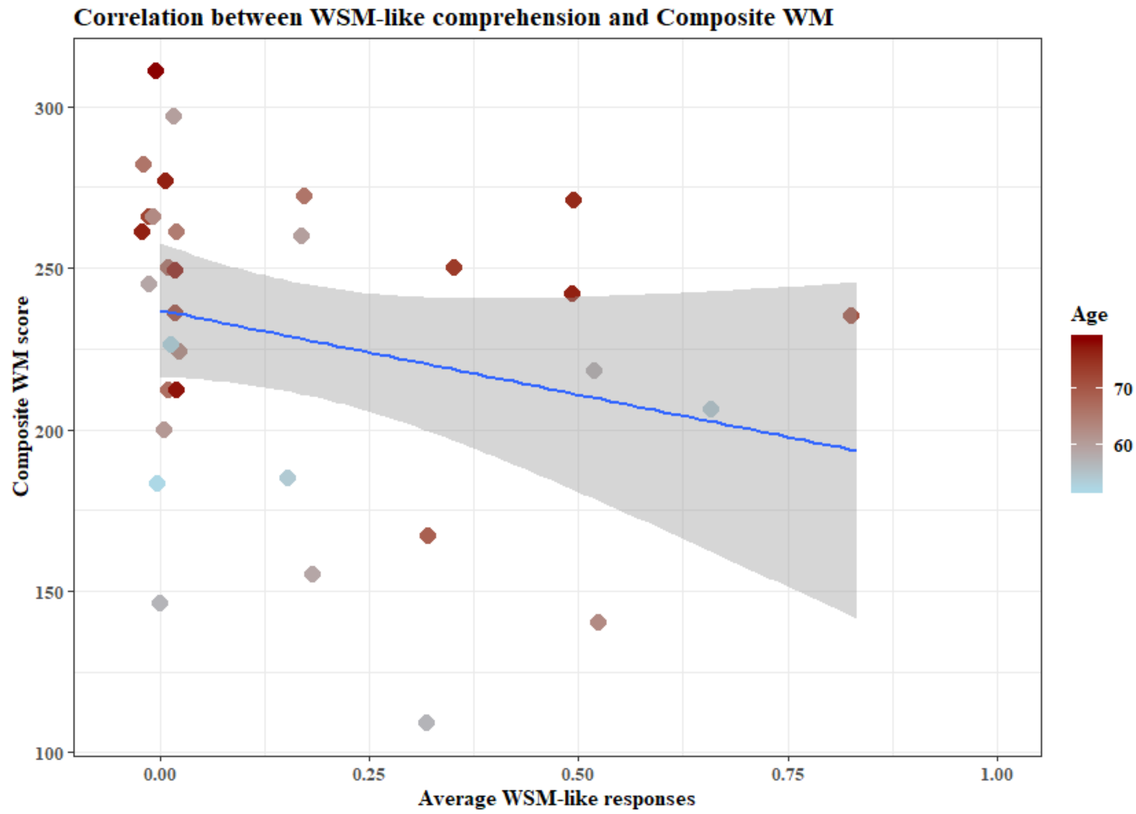


Fig. 5 illustrates the relationship between Composite WM score and average WSM-like responses. Children's age (in months) is indicated by color.

We also submitted the data to a mixed effects logistic regression model, which found that WM predicted WSM-like interpretations at a rate approaching the standard measure of significance (p=.08). In order to fully understand this finding, we decided to look into whether BDS and FDS gave similar results. It is possible that only one of these abilities was driving this task. Given that FDS might be more relevant to comprehension, while BDS might be more relevant to production (Delange and Frauenfelder, 2019), it is worth investigating whether these two measures make different predictions. We therefore submitted the data to two new models using BDS and FDS as fixed effects (predictors).

While BDS was not a significant predictor of either adult-like responses (p=0.9) or WSM-like responses (p=0.2), FDS was a significant predictor for both (p=.002 for adult-like responses and p=.01 for WSM-like responses[40]). We therefore present Fig. 6 as an illustration of FDS and WSM-like comprehension as well as Table 7 which shows the output of the logistic regression model investigating this relationship.

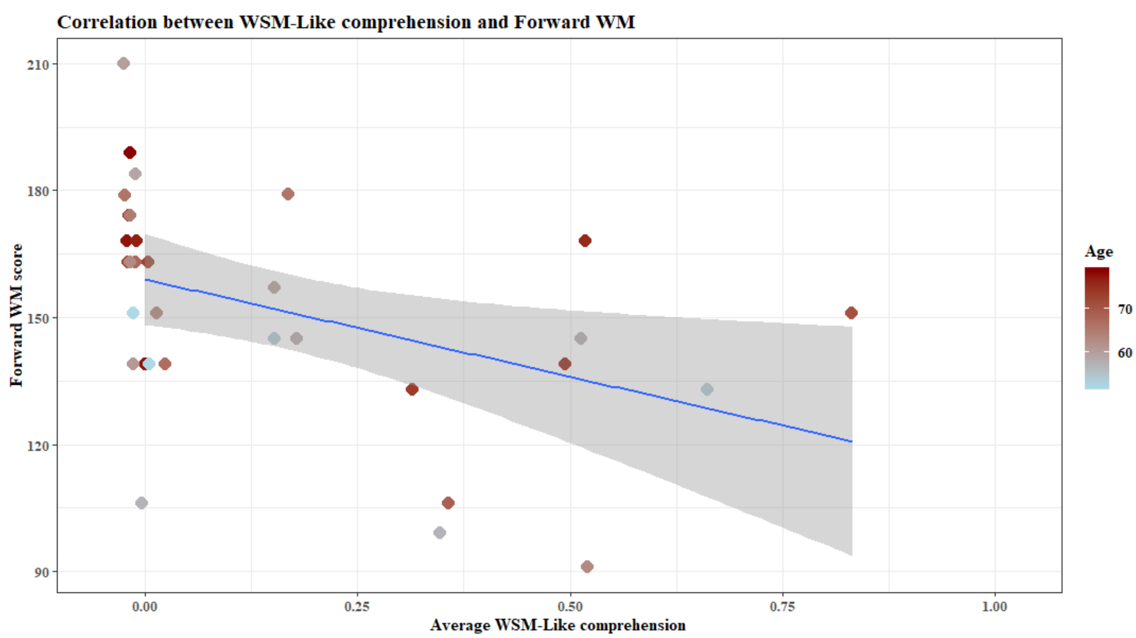


Fig 6. The relationship between FDS and average WSM-like responses. Children's age (in months) is indicated by color. *p=.01, r=-.3*

Table 7: *Logistic regression results: FDS as a predictor of WSM-like comprehension*

	B	Error	z-value	Pr(>|z|)
FDS	-0.05	0.02	-2.5	0.013*
Scale(Age)	2.05	3.24	0.64	0.53
FDS: scale(Age)	-0.01	0.02	-0.57	0.57

[40] We note that if children who made no WSM-like errors are removed from the analysis, this relationship is no longer significant, but this is likely due to the fact that only 13 participants remain after eliminating those children who made no WSM-like comprehension errors.

Both Fig. 6 and Table 7 illustrate the fact that FDS is a significant predictor of WSM-like responses: the higher a child's FDS, the fewer WSM-like responses the child will give.

We have seen that the composite WM score predicts adult-like and WSM-like responses, but upon further investigation, we saw that the relationship was being driven by the FDS score. FDS is a strong predictor of both adult-like and WSM-like responses. Again, as with the production task, WM is a stronger predictor than age is. For instance, in Fig. 6, we see multiple children under 5;0 who give no WSM-like responses, but the child who gave the most WSM-like responses (83%) was 5;11.

In summary, the higher a child's WM (specifically, their FDS), the more likely they are to give adult-like responses to complex questions and the less likely they are to give WSM-like responses. Furthermore, given the fact that we found that BDS was a significant predictor of production performance and FDS was a predictor of comprehension performance, we provide confirmation for Delange and Frauenfelder (2019)'s claim.

Did children who did the task virtually perform differently?

Because we had to do much of the study virtually, some aspects of the results are not ideally balanced. Children at the lower end of our age range had a greater amount of difficulty completing the task virtually. This means that the average age of those who completed the task virtually is higher (5;8) than those who completed the task in person (5;1). Because WM is affected by age, this also means there was a difference in WM for those who completed the task online (composite WM average of 241) and those who completed it in person (composite WM average of 207. Consequently, it would be reasonable to anticipate a difference in performance between those who completed the task online versus those who completed the task in person, crucially and perhaps counterintuitively, this predicts children who completed the task online (arguably with more distractions and uncontrollable factors), should do better than the younger children who participated in person. This prediction is borne out both in production and in comprehension. We ran two more mixed effects logistic regression models with fixed effects of modality (online or not) and age (centered as well as scaled) and random effects which minimally included participant and item number. The model indicates that participants who participated in person were more likely to use medial wh-phrases (*p=.06*) and less likely to use Target-like structures (*p=0.01*). In person participants were also more likely to give WSM-like responses (*p=0.01*) and less likely to give adult like responses (*p=0.01*).

In other words, yes, doing the task virtually did affect children, but not in such a way that invalidates any results. The difference boils down to an effect of age and WM: younger children had more difficulty with the online task and were more often excluded. This results in a higher average age and WM score for those who completed the task virtually and in turn those children performed better than those who completed the task in the lab. In fact, this finding is precisely in line with what we would predict, namely higher WM leads to better performance.

5.3 Discussion

The goal of this experiment was three-fold: First, to establish whether removing the lead-in from the paradigm would affect the results; Second, to determine whether WM was a predictor of how participants would perform on either or both of the language tasks. We also sought to replicate our finding from Chapter 4, which found that there was no correlation between production and comprehension errors.

Are children's productions affected by the presence of a lead-in? The results of our mixed effects logistic regression model show no significant effect either of experiment (whether the child participated in our experiments in Chapter 4 or those in Chapter 5 or lead-in (the presence or absence of a lead-in) on the presence of a medial wh-phrase. This suggests no role of the lead-in in the production of medial wh-phrases. Therefore, if these errors are caused by limited processing mechanisms, it is not because of the effort necessary to integrate the lead-in into the structure. In other words, this is not an error brought on by the task.

However, while the results from the production task essentially replicate our findings from Chapter 4, they also bring to light several other factors that did change when we removed the lead-in. First, though the overall rate of medial-wh phrase productions did not change, productions resembling copy constructions (two identical wh-phrases) were reduced from 87% to 32% and second, the open-ended nature of this task showed us many "work-arounds" that children use in order to communicate the necessary message, using a simpler structure than the target. These came primarily in the form of *Infinitival Responses* (19% of total productions), but also included the *Think about* construction (2.8%) and the use of the monoclausal question alone (1.4%), monoclausal questions joined by *and* (1%), and *Can you ask Mork "Who should we help"* (1.6%). Together these "work arounds" accounted for 25.6% of productions. It is possible that these work-arounds reflect an attempt to plan the material all in a single chunk forcing a simpler structure. We leave this suggestion to future investigation. In sum, removing the lead-in allowed children more freedom to ask these questions in the way they preferred, revealing a preference for simpler, work-around structures.

In comprehension, are children's responses sensitive to the pragmatic constraints discussed in Chapters 2 and 3? Children's performance was not different when we adjusted the stories to include CT subjects. This is, however, just what we would expect for English-speaking children because the relevant constraint (FAITHCT, see Chapter 3 for details) is too low-ranked to make a difference either way. This in turn tells us that children have mastered that aspect of the adult grammar.

In spite of changes we made to the two tasks, do we replicate the lack of correlation between language tasks found in Chapter 4? Yes, we replicated our finding from Chapter 4 that there is no relationship between medial wh-production and WSM-like responses in comprehension. That is to say that the child who produces many medial wh-phrases in their questions is not necessarily a child who gives WSM-like responses (and vice versa). This finding, again, is evidence that does not support a Parametric Acquisition Hypothesis. In Chapter 4, we suggested that if these errors were not the result of parameter-guided learning, they must instead be the result of immature processing. However, we had not directly tested that proposal. In what follows we discuss the findings which we argue directly support the Limited Processing Hypothesis.

Is there a correlation between WM and production? Comprehension? Yes, we found a significant correlation between children's performance on both language tasks and their WM scores. This finding is actually two-fold: first, there is a significant relationship between WM and target-like productions/responses; second, there is significant negative relationship between WM and medial wh-productions and WSM-like responses in comprehension. While it might seem that the one would necessitate the other, that is not the case: the open-ended nature of both tasks meant there were many ways to give a non-adult-like response. In other words, not only does WM affect a child's ability to use the grammar in an adult-like way, but a low WM also more frequently leads to a WSM-like error in both production and comprehension. Previous work such as Delage and Frauenfelder (2019) had also found a relationship between WM abilities and adult-like production as well as comprehension, though other work (such as Willis and Gathercole, 2001) found a relationship between WM and production, but not comprehension. Our findings therefore add to the body of literature which did find such a relationship. The different results across studies could come down to the exact type of WM measure used. While results vary, the relationship between adult like performance and WM is, perhaps, unsurprising. However, the negative correlation between this particular error type and WM is novel and suggests that lower WM (indicative of lower processing ability generally) is tied to this error type, directly supporting the Limited Processing Hypothesis.

We next examine the relationships between processing and production and processing and comprehension separately and in more detail. We have established that the relationship exists between WM and both production and comprehension. However, we also know that production and comprehension abilities are not correlated with each other.

This suggests that it is not simply the case that a child who performs well on WM performs well on both language tasks. This means the exact nature of the relationship between WM and each language task must be distinct. The goal of the next sections is to establish exactly what those relationships might look like.

Processing in Production

Let us first turn to the exact nature of the relationship between production and processing and what the Limited Processing Hypothesis means for medial wh-productions which, as a reminder, were originally construed as UG-based errors in the sense that Child English was erroneously admitting SynSM (as in e.g. deVilliers and Roeper, 1995). Recall that Grolla and Lidz (2018), conducted a study very similar to our production tasks described in Chapter 4, including the use of a lead-in. They also included a task measuring inhibitory control and established a correlation between inhibition ability and medial wh-phrase production. They suggested that these errors were the result of children accurately moving the wh-phrase from its base position in the embedded clause, pausing at the clausal boundary, but failing to inhibit pronouncing the wh-phrase that appears medially, supporting Thornton (1990)'s suggestion that these errors are the result of cyclic movement without deletion. However, we have seen in this chapter that, while previous work (including Experiments 1 and 3 from Chapter 4) found that most medial-wh productions were copies, when the lead-in is removed, the medial wh-phrases persist, but they no longer resemble copy constructions. Whereas in Experiment 1, 82% of medial wh-productions resembled copy constructions, in the production task without the lead-in, that percentage fell to 37%. We suggest that the prominence of copy constructions previously was the result of the lead-in simply because of the fact that the experimenter tells the child how to begin their question. In this version of the task, because we gave no lead-in, the child was allowed to begin their production however they chose, resulting in many more constructions which begin with *what* and included *who,* medially, as the contentful wh-phrase. The explanation given by Grolla and Lidz (2018) does not explain these constructions. We have said we do not believe these medial wh-productions to be examples of SynSM, but we need to establish what they are instead. We would like to suggest that these productions which include medial-wh phrases are actually the result of two different phenomena: (1) sequential question productions and (2) syntactic blends.

Sequential questions (SeqQs), we suggest, are another form of "work-around" which simply include two, monoclausal questions in a row (*What does he think? Who should we ask?*). We have seen that children express a preference for simpler constructions. Indeed, 25.6% of total productions represented work arounds. Furthermore, this is a completely acceptable utterance in English. While the adult data reported in Chapters 2 and 3 suggest LD constructions are always optimal in English, SeqQs were frequently reported as a second choice. It is, therefore entirely possible that these child utterances which were previously considered instances of SynSM (e.g. by deVilliers and Roeper, 1995) were, in fact, SeqQs. Because previous work almost exclusively used a lead-in to elicit these questions, they most frequently would have gotten copy constructions which cannot be interpreted as SeqQs (e.g. *Who do you think who we should ask* cannot be interpreted as **Who do you think? Who we should ask?*). However, we have seen that in our task without a lead-in, most of the productions with medial wh-phrases are not of this form, but could be two questions (e.g. *What does he think who can fix the fence* could be *What does he think? Who can fix the fence?*).

However, it appears that not every production including a distinct medial wh-phrase can be interpreted as an instance of SeqQs. In the first place, these structures do not always include the subject-auxiliary inversion necessary in SeqQs, which children are generally good at in monoclausal questions (Pozzan and Valian, 2017). If these are all SeqQs, then there are more examples of failed auxiliary inversion than would be expected. In the second place, this would not explain the structures which begin with *Who* and resemble copy constructions. These productions, we suggest, are examples of syntactic blends, such as those described in child productions by Jaeger (2005). Recall that syntactic blends are structures which essentially involve a mix of two possible structure types which would have been produced in parallel during the 'syntactic level' of sentence production. Crucially, the two structures should both be viable ways to express the semantic meaning the speaker wants to convey. The concept of syntactic blends has been established in the study of adult errors (e.g. Bock & Ferreira, 2014; Coppock, 2010) and is one of the primary pieces of evidence for multiple structures being produced in parallel. While blends can be as simple as a word-swap, they can also be *splices,* which consist of an initial substring from one target construction with a final substring from another target construction (Fay, 1982). In Jaeger (2005)'s account of blends in children's speech she characterizes *splices* as *cross over structures,* which indicate the use of two constructions blended together. While Jaeger (2005)'s data does not discuss any examples of questions appearing as blends, we suggest that productions in our data which include medial wh-phrases are an example of splices, specifically. For example, if in the process of production, a child's processing mechanism produces both SeqQs (*What does he think? Who should we ask?*) and an LD structure (*Who does he think we should ask?*), but they are unable to devote the necessary resources to choosing between the two (perhaps planning, attention, or inhibition), this might result in something like *What does he think who we should ask?*: the first substring coming from a SeqQ construction, the second coming from an LD construction.

Having established two possible alternative explanations for these constructions (medial-wh questions are either examples of SeqQs or syntactic blends), two tasks remain: we need to establish independent evidence for this proposal and we need to show that the use of these constructions might be related to WM. Independent evidence that there are two separate phenomena resulting in medial wh-phrase production might come from a number of sources. Initially, it might seem obvious to investigate the rate of subject-auxiliary inversion: structures with subject auxiliary inversion are SeqQs and those without are syntactic blends. However, as discussed in Chapter 4, based on the findings of Pozzan and Valian (2017), we cannot assume that any production without subject-auxiliary inversion is not biclausal because English-speaking children frequently produce biclausal questions with subject-auxiliary inversion. A second avenue might be to investigate prosodic cues: those productions which are SeqQs might have a pause or some other indication at the clausal boundary which would be absent in the blends. However, discussions with multiple experts[41] as well as preliminary prosodic analyses reported in Chapter 4 (particularly, Appendix C) have suggested that children's productions are too disjointed, in that their productions include long pauses throughout (not just at clausal boundaries, for instance), to provide clear-cut evidence that a given production is an example of sequential questions vs. a single structure.

[41] Eleanor Chodroff, Colin Wilson, Matt Goldrick, all in personal correspondence and discussion, with thanks.

As an additional source of evidence, therefore, we propose to analyze the rate of the constructions themselves, starting with the observation that these two potential phenomena are the results of totally different effects: SeqQs are a work-around strategy resulting in simpler syntactic structures, while blends are errors resulting from the mixing of two syntactic structures produced in parallel. These two phenomena make different predictions about behavior: work-arounds are deliberate and should happen with some consistency; blends are an unconscious error and should occur infrequently and potentially amidst other blends.

For example, consider the most prominent work-around construction: Infinitival Productions. We found that, typically, when a child produced one infinitival production, they would produce several. Indeed, on average, children that used the Infinitival Productions produced them for about one third of their utterances. This suggests that work-arounds come in groups. Children are internally consistent with the structures they use. If SeqQs are being used as a work-around, they too should be used consistently. In contrast, a child who is making a genuine processing error and produces a syntactic blend is not necessarily likely to do so continuously. The very notion of blends depends on the premise that the speaker has produced multiple constructions in parallel. This means a child who sometimes produces a blend should be just as likely to blend other constructions which they know are available and to use a variety of constructions.

Given this premise, constructions which resemble sequential questions (*What does he think who should we ask)* should be produced by children who produce many utterances with medial wh-phrases, while productions which do not resemble sequential questions (*Who does he think who we should ask*) should be produced by children who produce few medial wh-phrases and many other types of constructions. To determine whether this was the case, we needed a measure of "medial-wh consistency" as well as an indicator of whether the constructions could be SeqQs. To that end, each construction containing a medial wh-phrase was bimodally coded to indicate whether it could (1) or could not (0) be SeqQs. To count as potentially SeqQs, the word order had to be that of two monoclausal questions with no error. Structures which included a clear repair were considered to be of this category (e.g. *What does he think what – I mean - who should we invite?*), however, any other type of error would mean a production could not be considered SeqQs. Utterances which were *who*-initial for instance (*who does he think who can fix the engine?*) could not have been SeqQs because *Who does he think?* is not a good monoclausal question of English. For instance, consider (12) which was produced by participant 41, age (6;6):

(12) What does Mork think who has the hammer?

(12) was marked as a structure which could be an instance of SeqQs because *What does Mork think? Who has a hammer?* is an instance of SeqQs with no error: the initial wh-phrase is *what* but the medial wh-phrase is *who;* there is no error of subject-auxiliary inversion[42]; it is all pragmatically relevant. In contrast, consider (13) which also contains a medial wh-phrase but is not potentially SeqQs because of its initial *who:*

(13) Who does Mork think who can fix the fence?

For each utterance, we totaled the number of medial wh-productions made by the child who produced that utterance. This became their medial-wh number. Because participant

[42] In this example, because it is a subject extraction, no subject-auxiliary inversion is apparent, so this criterion does not appear relevant, but if the structure were object extraction and there was no subject auxiliary inversion, it could not count as potentially SeqQs.

41 also produced 7 other structures which included a medial wh-phrase, the medial wh-number associated with (12) is 8. (13) was produced by participant 13, age 4;6, who did not produce any other structures with medial wh-phrases. Thus, the medial wh-number for this structure is 1. In this way, we can measure whether a single instance of a medial wh-production is anomalous or the norm for that participant.

We submitted this data to a mixed effects logistic regression model with medial wh-number and Age (centered and scaled) as fixed effects and participant and item number as random effects. The maximally complex model which would converge was run, which included medial wh-number and age as random slopes for participant as well as item number. The results suggested that medial wh-number is a significant predictor of whether a structure could be two sequential questions (p=*.02*). In other words, if a child produces many questions with a medial wh-phrase, they are significantly more likely to production constructions that could be potential SeqQs. If a child only occasionally produces a medial wh-phrase, they are significantly more likely to produce a construction that could not be SeqQs. The prediction that a child who is producing SeqQs as a work-around should produce them consistently is, therefore, borne out. The prediction is that a child who is producing syntactic blends produces them in error and only occasionally is also borne out.

We also looked briefly into the prediction that a child who is producing medial wh-phrases as blends will also produce other types of constructions and blends of constructions. We took Participant 13, who produced (13), above, as an example. Participant 13 also produced (14- 16):

(14) What do you think has a hammer?

(15) Who does Mork think to fix the engine?

(16) Who does Mork think we should listen to?

As predicted from a child potentially producing blends, we see an example of a blend substitution in (14) (*what* for *who*) and a splice of LD and an infinitival construction in (15) (*Who does Mork think can fix the engine?* and *Who does Mork want to fix the engine?*) as well as an adult-like LD construction in (16).

While this was not one of the original goals of this experiment, this exercise has provided strong evidence that children who produce medial wh-phrases might do so for more than one reason. The first case is that of a child who prefers SeqQs as their work-around and produces them consistently. The second case is that of a child who occasionally produces a blend of LD and SeqQs.

The existence of a group of children who prefer SeqQs is, in fact, consistent with the constraint-based view of the grammar used to model SynSM production in Chapter 3. Those children who use work-around constructions, including SeqQs might be doing so because they have yet to establish the fully adult-like constraint ranking of English discussed in Chapter 3. It is certainly possible, that (a la Legendre et al. 2002, 2004) these children are availing themselves of several possible constraint rankings (resulting e.g., in some constraints 'floating' over the range of other constraints already fixed in ranking) and are sometimes operating under a constraint ranking closer akin to the target German constraint ranking than the target English one with a high ranked FAITHCREG and a lower ranked FAITHLF/SYN. This would result in SeqQs winning over LD constructions, periodically. Therefore, it is possible that these children who consistently use SeqQs and other work arounds are doing so because their constraint ranking is not yet fully adult-like for English. These results are also consistent with the possibility that children have a constraint on the complexity of their productions (which may or may not directly build on MINPROJ). An Economy constraint that favors SeqQs and other syntactically simpler work-arounds to LD constructions, for example, might be more highly ranked in those children who are consistently using SeqQs[43]. Both of these ideas are relevant to the question of just how "adult-like" we take children's grammars to be. Just why children prefer SeqQs and other work arounds is yet to be determined. The Limited Processing Hypothesis suggests that errors are the result of processing rather than grammar, but that does not mean the grammar is completely "adult-like". Our findings are certainly consistent with such an analysis, but we leave it to future work.

Finally, we wanted to establish whether WM was related to either of these findings. Of course, we have already established that WM significantly predicts medial-wh productions, but this time we want to know if WM predicts whether an individual has a high frequency of medial wh-productions. We ran a mixed effects linear regression which showed WM predicted medial-wh number to a degree approaching significance ($p=.07$). The negative correlation suggests that the higher a child's WM, the lower their medial-wh number. This suggests that children with lower WM produce consistent medial-wh phrase constructions (which we analyze as grammatical SeqQs) while children with higher WM are more likely to only produce medial wh-phrases occasionally (though these are more often blends and ungrammatical).

[43] Note that adding a new Economy constraint leads us to ask what happens to this constraint when the child reaches the adult ranking described in Chapter 3. From that perspective a better approach that does not assume child-specific constraints is that of a partial constraint ranking equivalent to multiple alternative rankings yielding multiple distinct UG-licensed productions a la Legendre et al. (2002, 2004).

In summary, we found a significant relationship both between WM and productions with medial wh-phrases and between WM and target-like productions. We suggested that productions with medial wh-phrases are not instances of SynSM but rather one of two possible things: first, SeqQs and second, examples of syntactic blends, specifically splices. We showed that children who produce utterances that *could* be adult-like SeqQs, produce them consistently while children who produce utterances that *could not* be adult-like SeqQs (for various reasons including not doing subject auxiliary inversion in what would be the second question) do so only rarely and also produce other constructions that could be blends. Our suggestion that at least some of these productions are blend constructions, as described in children this age by Jaeger (2005), is further evidence for parallel processing of syntactic structure and provides evidence for it in people as young as age 4;0. We would also like to highlight the fact that in this experiment with no lead-in, we had relatively few productions which included the overt complementizer, *that.* We suggested in Chapter 4 that the use of *that* might be a result of poor planning. Perhaps, those children who are not good at planning two clauses together perform differently with and without a lead in, producing a *that* complementizer when they are given a lead in and SeqQs when they are not given a lead in. We leave this question to future work.

Processing in Comprehension

A processing account of comprehension is comparatively simple. Following Phillips & Ehrenhofer (2015), we assume comprehension is incremental and that each new word integrated into the current interpretation of the utterance as it is uttered. This requires resource management as well as many different processes operating in parallel, plus WM. For the sake of example, recall the story about Evil Steve, who tells Sherry in a letter that he will steal the queen's ring, when in fact he steals the crown. After this story, children were asked a question such as (17) ((8), above).

(17) How did Evil Steve tell Detective Sherry what he was gonna steal?

Though the task was open-ended, primary answer types included the following:

Adult-like: In a letter
Alternative how responses: With his TV machine.
WSM-like: The ring.
Embedded clause responses: The crown

Our findings suggest a correlation between WM (specifically FDS scores) and adult-like responses as well as WSM-like responses. In Chapter 4, we found that WSM-like responses and responses to the second question were at chance. We suggested that children were, perhaps, simply responding to the most recent wh-phrase, but were at chance in their responses. However, we have found more WSM-like responses than second question responses. Furthermore, when we combined data from the within subjects experiments in Chapters 4 and 5, we found there were significantly more WSM-like responses than embedded clause responses across these two experiments ($p=.01$). Therefore, the explanation given for the results in Chapter 4 must be updated. The new results suggest that children were not simply answering the wh-phrase which appears most recently, but were actually giving more responses where the medial wh-phrase would have scope over the matrix clause (WSM-like). However, we also know that a child with lower WM is more likely to do this, which does not suggest an alternative grammar explanation. Let us, therefore, consider a processing explanation. According to the model presented by Phillips & Ehrenhofer (2015), interpreting a question requires the child to recognize a wh-phrase and keep it in mind as they search for the relevant corresponding gap in the embedded structure. They must also keep similar lexical items from interfering with their interpretation. For our purposes, this would mean they must maintain the matrix wh-phrase as the wh-phrase with scope rather than allowing the (wh)-relativizer (which is identical in form but not in meaning and function to a true wh-phrase) to interfere. This is similar to the claim made by Cunnings and Felser (2013) that adults with lower WM abilities are more likely to mis-parse constructions with multiple potential antecedents for pronouns. They claim that noun phrases with similar properties to the target antecedent interfere and lead to the mis-parse. [44] Our explanation is that children with lower WM scores simply are not as good at keeping the matrix wh-phrase in memory as the question to be answered. They integrate the two clauses and thus know that the question is not about the embedded clause, but they allow the relativizer to interfere with their memory of the wh-phrase to be answered and thus give a WSM-like response. For example, in processing (17), they know the question word has to do with Steve saying, but when they come to the relativizer *what* they allow it to replace the true wh-phrase, *how* and end up answering a question about *what* Steve said rather than *how* he said it.

5.4 Conclusion

Our primary, novel finding is that both production of medial wh-phrases and giving WSM-like responses to questions containing wh-relativizers are predicted by a child's

[44] Related ideas, pertaining to competence rather than performance, can be found in Featural Relativized Minimality. Friedmann, Belletti, and Rizzi (2009), for example, argue that children find some object relatives difficult to understand because of the structural similarity between the A'-moved element and the subject (which intervenes).

WM. We have suggested that in production these structures with medial wh-phrases are one of two things (1) SeqQs used as a work-around construction or (2) a syntactic blend. In comprehension, we have suggested that WSM-like responses (when they occur) are the result of children failing to keep the relativizer from interfering with the true wh-phrase and thus they answer a question which would give that wh-phrase scope over the whole structure. These findings are important for several reasons. First and foremost, they support the Limited Processing Hypothesis and suggest that these errors are the result of immature processing rather than immature grammar. Second, we have added to the body of literature suggesting WM is relevant to both production and comprehension. Third, our study provides (we believe) the first instance of evidence of syntactic blends in (monolingual) child productions.

Another question we have answered is whether the lead-in affected children's productions. The lead-in does affect production, but only in that it leads to more structures that resemble copy-constructions. Without the lead-in a greater proportion of structures with medial wh-phrases have an initial *what* and a medial, contentful wh-phrase.

Furthermore, we replicated the finding from Chapter 4 which showed no significant relationship between the use of medial wh-phrases in the production task and WSM-like interpretations in the comprehension task.

We also asked at the beginning of this chapter just how adult-like children's grammars are. Are all the errors the result of limited processing or is there a possibility that children have yet to establish the adult-like grammar? We have concluded that there are, in fact, two groups of English-speaking children who produce utterances with medial wh-phrases: those who are producing a syntactic blend and those who use SeqQs as their work-around. While the syntactic blends are unquestionably the result of processing, it is possible that those children who consistently produce SeqQs do so because they have yet to establish the full adult-like constraint ranking for English and simply use SeqQs more frequently than adults would as a result of a (potentially slightly) different constraint ranking. This would need to be directly investigated, but our findings are in line with such an argument.

Our overall claim that these errors are the result of processing mechanisms rather than grammar competence will be stronger once we have investigated the performance of children whose target grammar does include SynSM. If these errors are truly the result of processing, then children who are acquiring a language that includes SynSM should also make errors which we can identify as the result of immature processing. We would also expect to find a similar group of participants who use SeqQs consistently. To that end the following chapter will describe the results of this same series of experiments conducted in Germany with German-speaking children.

Chapter 6: Processing difficulties in complex question formation: evidence from German

English-speaking children have been shown to make consistent errors in the production and comprehension of biclausal questions up to the age of 6;6 (Thornton, 1990; de Villiers & Roeper, 1995; Lutken, Legendre, and Omaki, 2020). These errors involve the production of a medial wh-phrase as in (1) or the response to a medial wh-phrase as in (2).

(1) Who do you think **who** we're gonna see in the cave
c.f. *Who do you think (that) we're gonna see in the cave*

(2) Q: How did Steve tell Sherry what he was gonna steal?
A: The diamond ring! (as opposed to "in a letter)
Exs. From Lutken et al. 2020

These errors of production and comprehension have been interpreted as the result of children (incorrectly) assuming that the medial wh-phrase is the contentful one, the one to answer (e.g. deVilliers and Roeper, 1995). This behavior suggests that English-speaking children might be using a grammar that resembles the "Wh-Scope Marking" (WSM)[45] construction, viable in languages like German, as seen in (3). In (3), the correct response to give is an answer to the *with whom* question. This construction is viable in most varieties of German (Lutz, Muller, & von Stechow, 2000).

(3) **Was** glaubst Du ___ mit **wem** Maria ____ gesprochen hat?
What think you ____with **whom** Maria _____ spoken has?
Whom do you think Maria spoke with?

[45] We use WSM-language(s) to refer to languages which employ WSM. We use SynSM to refer to the syntactic co
nstruction described thoroughly in Chapters 2 and 3.

DeVilliers and Roeper (1995) argued that children's use of and response to medial wh-phrases is in fact direct evidence of UG-based and parameter-guided language learning, as would be suggested by Principles and Parameters theory (Chomsky, 1981). This could result either from a 'mis-set' parameter or it could be the result of competing grammar model as has been suggested by Legendre et al. (2002) or Yang (2002). These models vary in particulars, but each assumes that it is children's knowledge of the grammar which is imperfect. Each would predict that a child whose grammatical knowledge were imperfect in this way would make similar errors in production and comprehension (to varying degrees, depending on the model; see Smolensky 1996). We refer to any model which presumes the use of parameter guided learning in this broad sense as part of the *Parametric Acquisition Hypothesis.*

However, in Chapter 4 we saw that when the same children were tested on their abilities to ask and answer these types of questions, their performance on production and comprehension tasks was not correlated. In other words, a child who used medial wh-phrases in questions like (1) was not necessarily more likely to respond incorrectly to questions like (2). This finding does not support a Parametric Acquisition Hypothesis: if children made these mistakes because they hypothesized that English behaves as a WSM grammar, we should see a correlation between these production and comprehension errors. We suggested, instead, that these particular errors were the result of children's limited processing mechanisms, which we refer to as the *Limited Processing Hypothesis.*

In Chapter 5, we tested the Limited Processing Hypothesis directly by measuring English-speaking children's working memory (WM). Specific models of the processing involved in comprehension and production predict that a variety of processing mechanisms should be associated with performance in each task. The model of comprehension suggested by Phillips & Ehrenhofer (2015) suggests that when processing a question, a child has to hold in memory the relevant *wh*-phrase and keep this *wh*-phrase active, without allowing interference from other lexical items, until they find the corresponding syntactic gap in the embedded clause. A child with better WM would be better able to maintain this *wh*-phrase

The processing model of production is less clearly defined than that of comprehension, but we follow a 3-level model (Dell, 1986; Bock 1987; Levelt, 1989). This three-tiered process involves, first, a semantic level where the message is formulated; second a syntactic level where possible constructions are generated in parallel (e.g. Garrett, 1980) and finally a morphophonological level where lexical items are selected and the appropriate phonemes are mapped onto the structure. This model is discussed in greater detail in Chapter 5. Crucially, multiple facets of general processing have been argued to be relevant to production. In particular, McDaniel et al. (2010) suggest that children plan productions in smaller units than adult speakers do, leading to disfluencies, particularly at the clausal boundary. A child with better planning skills would be able to plan in more adult-sized pieces. Grolla & Lidz (2017) found that children with lower inhibitory control were more likely to produce constructions with medial wh-phrases and suggest that this might be because the child knows the wh-phrase moves, but fails to delete the intermediate trace. While this would explain "copy-constructions" which use the same wh-phrase twice (such as (1), this explanation would not cover those constructions which use different wh-phrases. In chapter 5, we have argued that WM itself is relevant to production because the speaker must maintain the chosen construction in memory without allowing it to blend with another construction, following Coppock (2010). Thus, various aspects of cognition have been claimed to be relevant to production and comprehension. However, under Baddeley (2017)'s model of the executive control system, which regulates cognitive processes, WM can serve as a general measure of broader cognitive function and it has been used as such in previous work (e.g. Johnson et al., 2013). The logic behind this is that WM acts as a kind of mediating step between cognitive processes and any action taken. In other words, regardless of how well a child has processed a sentence (either in preparation for producing it or in comprehending it), their performance will be mediated by their WM capabilities. Thus, while many separate cognitive processes might be involved in comprehension and production, WM will mediate both. This has been born out in previous research. WM has been shown to be related to children's production of complex sentences (Delage & Frauenfelder, 2019; Adams and Gathercole, 2000), repetition of complex sentences (Willis and Gathercole, 2001), and comprehension of complex structures (Delage & Frauenfelder, 2019; Cunnings & Felser, 2013 in adults). Furthermore, we saw that children's performances on both production and comprehension tasks were correlated with their WM in Chapter 5. We suggested that the correlation found between WM and the language tasks supports the Limited Processing Hypothesis.

However, there is one crucial prediction made by the Limited Processing Hypothesis which remains to be directly tested. If these errors made by English-speaking children are the result of immature processing mechanisms, then children in the same age range (and therefore with similar processing abilities) should show errors when trying to use these constructions, regardless of the target grammar. If integrating clauses puts a load on the processing system in English, it should do so in any language. Furthermore, while previous work has established the presence of the medial wh-production errors in other languages (e.g. French, Demirdache & Oiry, 2007), no one has done such a study with German-speaking children. While Jakubowicz & Strik (2008) do find Dutch-speaking children, whose target grammar has been argued to include SynSM, produce this structure, their work is on production alone. Overall, a study of the acquisition of production and comprehension of a WSM language has yet to be conducted. Investigating the production and comprehension of German-speaking children is a clear next step.

To that end, this Chapter examines the acquisition of biclausal questions by German-speaking children in Konstanz, Germany. This experiment was conducted as part of a Doctoral Dissertation Research Improvement Grant (BCS-1853297) from the NSF and with the help of the BabySprachLabor (Baby Speech Lab) at the University of Konstanz in Konstanz, Germany. This chapter addresses four specific questions:

1) Do German-speaking adults in this area use SynSM in their elicited productions? Will they use them in 'child-like' tasks?
2) Do German-speaking children use SynSM in their elicited productions? Do they respond to SynSM questions like adults?
3) Do German-speaking children, like English-speaking children, use simpler structures (work-arounds) to avoid complex SynSM constructions?
4) Are German-speaking children's performance on production and comprehension tasks correlated with their WM abilities?

The answers to these questions will further inform our understanding of the use of this complex structure by adults (discussed in detail in Chapters 2 and 3) as well as the acquisition of this complex structure by German-speaking children. Furthermore, our findings will add to the growing body of literature which investigates the relationship between cognitive mechanisms and language acquisition and will shed light on the relationship between grammar competence and processing abilities, from a cross-linguistic perspective.

We begin with a discussion of previous work on the acquisition of German questions (Section 6.2.) as well as a discussion of a small corpus study we performed investigating the presence of biclausal questions (or lack thereof) in children's spontaneous productions, as well as child-directed speech. We then report on a study of German-speaking adults in the Konstanz area (Section 6.3). This study establishes that the tasks elicit SynSM constructions appropriately and that adults of the area have SynSM in their dialect. Finally, Section 6.4 will report the results of a study using the elicited production and comprehension tasks used in Chapter 4 and 5 to examine production and comprehension of questions with complex clauses from German-speaking children. This work is, to our knowledge, the first attempt to examine the acquisition of these complex questions in German. As in Chapter 5, we also use a WM measure as a proxy for general cognitive processing, and test the WM of our participants. The conclusion we reach is that German-speaking children, like their English-speaking counterparts, find it difficult to ask and answer such questions, and their WM scores predict their performance on both language tasks. We conclude this is further evidence in favor of the Limited Processing Hypothesis.

6.2 Acquisition of wh-questions in German

Previous work on the acquisition of (monoclausal) questions by German-speaking children has suggested they are proficient at a young age. However, that work does not include investigations into biclausal questions. In what follows, we will briefly summarize those studies which have focused on the acquisition of German, particularly of *wh*-questions.

Like other language learners, German-speaking children begin asking questions before they have mastered either the appropriate lexicon or necessary word order and instead ask with simple, holophrastic questions (Wode, 1971). These questions are identifiable almost entirely from prosodic cues and context. However, children move on to using multiple words in yes/no questions by as young as 1;10 as seen in (4):

(4) Ist das Auto weg?
Is the car gone?
age 1;10 Ex. From Wode, 1971

Soon after, the first use of wh-phrases and verbs appear as in (5) and (6), though they also often leave out the wh-element completely as in (7) (Clahsen et al. 1993/1994; Wode, 1971).

(5) Wo (ist) Larsi?
Where (is) Larsi?
Age 2;4

(6) Wo ist (sie)?
Where is (it)?
Age 2;7

(7) Henning, (wo bist) Du gewesen?
Henning, (where have) you been? Exs. From Wode, 1971

Even from a very young age, German-speaking children's verb placement is mostly accurate despite its inherent complexity (Clahsen, 1982). Before delving into this more deeply, we wish to highlight that German children's accurate placement of verbs is crucial for our purposes because the placement of the verbs in verb second (V2) languages such as German indicates reliably whether a structure includes an embedded clause. In matrix clauses, the finite verb (or the auxiliary verb, as may be the case) goes into the C position (the 'V2' position) of the left peripheral CP while non-finite verbs are clause final. In embedded clauses, however, the C position is typically occupied (by a complementizer) and the finite verb or auxiliary must appear in the clause-final position. Since German-speaking children master this verb positioning early on (younger than 3;0), we will be able to distinguish SeqQs from SynSM constructions by using verb placement. This is in contrast to English-speaking children whose use of subject-auxiliary inversion in embedded clauses is roughly at chance (Pozzan & Valian, 2017). In Chapter 5, we claimed that at least some of the productions with medial wh-phrases produced by English-speaking children were SeqQs, but we could only make this claim with any confidence because children whose productions resembled SeqQs (i.e. had the appropriate subject auxiliary inversion) tended to use these productions frequently, whereas children whose productions could not be SeqQs (i.e. did not use the appropriate subject-auxiliary inversion) did so infrequently. Thus, we proposed that some of these productions were SeqQs, and some were processing errors (discussed in detail in Chapter 5, and to a lesser extent below). With German-speaking participants, however, we will be able to definitively determine when children are asking a question using a SeqQ construction vs. a SynSM construction. This will not only inform our understanding of how German-speaking children perform, but it will add credence to our argument that English-speaking children are sometimes producing SeqQs.

When German-speaking children begin asking questions with embedded clauses (between 3;0 and 4;0), according to Clahsen (1982), there are virtually no examples of verb placement errors in the embedded clause. In fact, there are more examples of errors occurring with verb placement in the matrix clause than in the embedded clause. Furthermore, when Poeppel & Wexler (1993) investigated the acquisition of finite vs. non-finite verb placement in a single child (age 2;6), they found that most verbs produced were finite except when one considers only those verbs which appear verb-finally, in which case they are mostly non-finite. This is, in fact, the adult-like pattern. In general, therefore, children as young as 3;0 are placing verbs appropriately, particularly in embedded clauses, and they are using finite and non-finite verbs when they should. Crucially, for our purposes, children younger than 3;0 know that in an embedded clause the verb should come in the verb-final position.

Because the construction we are interested in has been so little studied in German-speaking children, we first conducted a corpus study to further investigate whether SynSM constructions appear in either children's speech or in child-directed speech from parents and caregivers. We investigated the Leo (Behrens, 2006) and Wagner (Wagner, 1974) Corpora in CHILDES (MacWhinney, 2000). We examined a total of 7,300 utterances and interviews between parents and 4 children ranging from 4;0 to 12;4, and found no examples of either SynSM or long-distance questions (LD) in parents' child-directed speech, suggesting both structures are rare in spontaneous child-directed speech. Unsurprisingly, we have also found no examples of SynSM or long distance wh-questions (LD) structures produced by children. However, as with all corpus work, we cannot conclude that these structures are not used at all, as it could simply be the case that no discourse situation ever required a biclausal question. This further confirms the need for an elicited production task.

In summary, German-speaking children show an early aptitude for verb placement and we should expect the children who participate in our experiments, who are a full year older than the children discussed above, to place verbs like adults. We will use these children's adult-like placement of verbs to distinguish between questions with embedded clauses (SynSM) and SeqQs. We have also seen that there are a limited number of studies investigating German-speaking children's acquisition and that those studies are investigating acquisition broadly and thus do not investigate any particular complex structure such as SynSM. They are also all either corpus studies or longitudinal studies of a few children (sometimes only one). Our own corpus study indicates that the use of LD and SynSM is rare both in child-directed speech and in child productions. Thus, we are adding to this body of literature by conducting production and comprehension tasks which elicit a complex syntactic structure. We are also investigating an age range (3;0-6;6) that has not been thoroughly examined in German-speaking children.

The following section will describe a version of our production and comprehension tasks conducted with adult, native speakers of German. Given the lack of SynSM in our corpus study, this is an important step in establishing that adults will use this structure.

6.3 SynSM Production and Comprehension by German-Speaking adults

The goal of this study is to establish that adult German-speakers use SynSM in the contexts we will use in our experiments with children. We designed the stories bearing in mind the pragmatic factors that elicit SynSM constructions discussed in Chapter 2: Question Under Discussion (QUD) and Contrastive Topic (CT), we refer the reader to

Chapters 2 and 3 for details. We are also interested in whether they will use these complex constructions in tasks designed for children since the survey conducted online and reported in Chapter 2 indicated that adults prefer to use SeqQs with children. While none of the characters are explicitly children, the child-like nature of the task might influence adults to use a child-appropriate register.

We use the two story-based elicitation tasks based on the experiments used for English in Chapters 4 and 5 as well as the comprehension task used in both chapters. Not only will this experiment establish whether adult speakers of German will use SynSM structures in these conditions, we will use their responses as a base of comparison for the children's responses.

We conducted this experiment with the help of the BabySprachLabor (Baby Speech Lab) at the University of Konstanz in Germany. Two German-speaking RAs were trained in person to administer the tasks. They translated and conducted the experiment under the direct supervision of the author.

6.3.1 Participants

Thirty-eight students from the University of Konstanz took part in the experiment. They were rewarded either with SONA credit when their course allowed it or with 4 Euros (as is the policy at the University of Konstanz). All were native-speakers of German. 10/38 participants took part in the comprehension task. Recall that there are two tasks used to elicit production of SynSM. The production task described in Chapter 4 which employed a lead-in and the production task described in (Chapter 5) which did not. Both tasks will be briefly described again below. 10/38 participants took part in the task without the lead-in and 18/38 participated in the task with the lead-in (this imbalance in participants will be explained in the Results section).

6.3.2 Procedure

Participants were presented one of the tasks (described below) through a Microsoft Power Point slide show on a DELL XPS laptop computer. The sessions were recorded on an Olympus L100 audio recorder. All tasks began with practice questions to ensure the participants understood the task. The tasks generally took 20 minutes with adults.

6.3.3 Materials

The tasks summarized below were all direct German translations of those presented in Chapters 4 and 5, verified by native speaker Research Assistants (RAs). The complete German script of both the comprehension and production tasks is provided in Appendix G. Participants took part in only one of the following tasks.

Comprehension

The comprehension task consisted of nine short stories (1 practice, 2 fillers, 6 targets), accompanied by power point animation, followed by a pre-recorded question posed by “Nina Nilpferd”, a friendly hippo puppet. The stories were identical to those presented to the English-speaking children. Since our goal was to establish whether German-speaking adults would respond to WSM constructions in these scenarios, all questions were SynSM structures such as (8):

(8) **Was** hat der böse Stefan Selina erzählt, **was** er stehlen wird?
What has the evil Stefan Selina told, **what** he steal will?
What did Evil Stefan tell Selina (that) he will steal?

The crucial question is whether participants respond with the item Stephan *said* he would steal, indicating a SynSM interpretation, or with the item he actually stole, indicating they were answering the embedded clause question (what he will steal). Stories were balanced for past vs. present tense. This was crucial in order to eliminate the possibility of a recency effect. For instance, if the telling event was always first, participants might have a tendency to answer a question about the embedded clause event (here, the stealing) simply because they had just heard about it. Stories were also controlled for equal prominence of the two events (for details, see Chapter 4) so that participants would not be prone to giving a response relevant to whichever event seemed more important to the story.

Production with lead-in

The first production task was a story-based elicitation task involving 12 problem solving scenarios in which participants must work together with a fairy puppet to ask every character their opinion about what they should do next. In this version, the fairy presented participants with a "lead-in", a prompt to help them start their question, as seen in (9):

(9) Fairy: Die Frage beginnt so: was glaubst Du…
The question begins like this: what do you believe…

Participant: **Was** glaubst Du **was** der Fee helfen kann?
What believe you **what** the fairy help can?
What do you believe can help the fairy?

After every character gives their opinion, the story proceeds following a majority rule: they do whatever most characters want to do. This motivates the questions at every juncture, making the task more natural. There were two versions of the task which presented the questions in different orders (questions were arranged such that questions of the same extraction type were not grouped) so that we could determine whether the order of presentation affected performance. The story changes slightly to accommodate this reversal, but the task remains the same. These two orders are referred to as Form A and Form B.

The use of a lead-in is common practice when eliciting complex structures (e.g. Bock & Miller, 1991; Ferreira & Dell, 2000). However, as discussed in Chapter 5, this lead-in is also a potential confound as it adds an extra layer of complexity to the task and makes it even less natural than it already is. For this reason, we added a version of the experiment with no lead-in.

Production with no lead-in

The production task without a lead-in is based on a task used by Schulz (2011) to elicit biclausal productions from L2 speakers of English. We adapted this task for children by creating a game that goes along with a 'translation' task. Again, we have a continuous story with 12 scenarios in which we wish to elicit everyone's opinion, but this time one character, Mork, is an alien who does not understand German. An example of this dialogue is presented in (10), English translation in italics.

(10) Experimenter:	Hm… wer hat den Flügel? Hast Du eine Idee? *Hm… who has the wing? Do you have an idea?*
Participant:	Bauer Hans! *Farmer Hans!*

Experimenter :	Das ist eine tolle Idee. Lass uns herausfinden was Mindy denkt. Kannst Du sie fragen „Wer hat den Flügel?“? *That is a great idea. Let's find out what Mindy is thinking. Can you ask her "Who has the wing?"*
Participant:	Wer hat die Flügel? *Who has the wing?*
Mindy:	Ich denke Doktor Maria hat den Flügel. *I think Doctor Maria has the wing.*
Experimenter:	Okay, Ihr zwei habt wieder zwei verschiedene Ideen, also brauchen wir Morks Meinung. Kannst Du Mindy fragen? *Ok, you have two different ideas again, so we need Mork's opinion. Can you ask Mindy?*
Target:	**<u>Was</u> denkt Mork, <u>wer</u> den Flügel hat?** **What thinks Mork who the wing has** ***Who does Mork think has the wing?***

The scenarios always followed the pattern presented in (10): The Experimenter asked the monoclausal question to the participant, then prompted the participant to ask Mindy the same question. Mindy (voiced by the experimenter) always disagreed with the participant, thus eliciting the need to find out Mork's ideas. Because Mork did not understand German, Mindy had to translate for him, so participants had to ask Mindy to find out Mork's ideas. The task included three practice questions so that participants understood the translation aspect. There were also two filler questions to break up the pattern and which could be used to assist participants if they forgot how the translation task worked.

Of course, these tasks were designed for children, so all adults found them relatively simple, but they also commented on them being enjoyable and sweet. Adults were asked afterwards if any part of the dialogue or task was unnatural, but no adult mentioned any part of the stories as odd or cumbersome in German. This was crucial since they had been designed in English and translated by the native German-speaking RAs.

6.3.4 Data coding and analysis

Comprehension

All responses were coded as either a WSM-like response (1) or not (0). Because the goal of this experiment was simply to ascertain whether the tasks work for eliciting the appropriate interpretation of the scenarios, we will simply be computing the percentage of WSM-like responses out of the whole.

Production

All utterances were transcribed by one of the two RAs who were both native speakers of German. They were then coded by the author for whether they were (1) or were not (0) an example of a SynSM construction, an LD construction, or a SeqQ construction. The author's coding was in turn spot-checked by the RAs and the two were in agreement 100% of the time. This is unsurprising since adults generally produced the target structure. Again, because our goal was simply to ascertain whether these scenarios elicit SynSM or these other constructions types, we present percentages of production types.

6.3.5 Results

Comprehension

Adults performed at ceiling on the comprehension task. One participant missed one filler question, but otherwise responses were unanimous. After the task, adults were asked whether the questions seemed to fit with the scenarios and everyone responded in the affirmative. This suggests we successfully created stimuli which were appropriate for eliciting SynSM constructions in German.

Production with lead-in

Eight participants took part in the production task with the lead-in, Form A. We elicited a total of 96 questions. However, only 20.11% were SynSM. 68.7% were SeqQs and no adult produced an LD question (a grammatical option in at least some dialects according to the linguistic literature on German). 11.2% were a form of parenthetical as in (11). The placement of the auxiliary verb, *kann*, indicates that this is not an embedded clause, but rather a monoclausal question with an aside '*glaubst du'* inserted into it (personal communication, Bettina Braun).

(11) Was glaubst Du, kann die böse Hexe aufhalten?
What think you, can the evil witch stop?
What – do you think – can stop the evil witch? OR
What – in your opinion – can stop the evil witch?

We considered the possibility that these parenthetical structures were happening because in Form A, *was* 'what' questions were elicited first, which allow for the parenthetical structure. Since this construction is possible with *was* questions, participants were perhaps primed to continue using this construction throughout the task. However, in Form B, *who* questions were elicited first which did not tend to elicit this construction (therefore would not prime its use throughout the experiment). We therefore recruited an additional 10 participants to take part in Form B of this task. They produced 120 questions. The only productions were SynSM (16.6%) and SeqQs (83.3%). Across the two groups, this makes an average of 18.4% SynSM constructions and 76% SeqQ constructions in this task.

Production without lead-in

Ten adults participated in the production task without the lead-in, producing 120 questions for analysis. Recall that this task was open-ended and we accepted any structure that was pragmatically appropriate. This resulted in a variety of constructions, but the majority were SynSM (80.3%). Other productions included SeqQs (3.4%) and LD constructions (5.9%). There were also a few copy-constructions e.g., *Wo er denkt, wo der Flügel is?* (*Where does he think where is the wing*?) and accounted for 1.7 % of productions. This is unsurprising as copy-constructions are only viable in some dialects of German (Murphy, 2016). Another .8% were *whether/if* questions such as *Kannst Du Mork fragen, ob er weiß, wer den Weg kennt?* (*Can you ask Mork if he knows who knows the way*). Finally, 7.6% were questions such as (12) which were biclausal, but did not involve wh-movement to the matrix clause:

(12) Mindy, könntest Du bitte Mork fragen, wem wir helfen sollten?
Mindy, can you please Mork ask, who-dative we help should?
Mindy, can you please ask Mork who we should help?

Together, these constructions account for 99.8% of productions. While there were a variety of constructions produced, they were all complex and the majority were SynSM.

6.3.6 Discussion

These data are informative in several ways. In the first place, we have established that German-speaking adults in this southern region of Germany do use SynSM constructions. Furthermore, even though these adults are not speaking to children, the stories are clearly meant for children and many of the stories involve children, so we can also posit that they would use SynSM constructions with children. We can conclude that the comprehension task successfully creates scenarios which elicit SynSM constructions. Finally, the production task without the lead-in successfully elicits SynSM productions from adults. Interestingly, the production task with the lead-in elicits SeqQs from adult German-speakers more frequently than SynSM constructions. The story-lines themselves are similar enough to where it is unlikely this is the result of a pragmatic factor. It is more likely that this is the result of the lead-in. Perhaps the fact that the lead-in itself would always be a complete question (*Was glaubst Du? what do you think?*) biases adults to treat it as a monoclausal question to be followed by another monoclausal question rather than the beginning of a biclausal question. RAs were trained to use a rising tone at the end of the lead-in to indicate it was not, in fact, a complete construction, but required finishing, but participants may have simply construed this to mean they should follow up with something. It is also possible that the use of a puppet (who provided the lead-in) put adult speakers in a situation where they were more likely to use a child-directed register, and therefore produce SeqQs (as suggested by the survey discussed in Chapters 2 & 3).

For the purpose of understanding the acquisition of this structure, we will be able to compare the performance of German-speaking children with the performance by German-speaking adults. This experiment has shown us that German-speaking adults on the one hand respond with WSM-interpretation 100% of the time in our comprehension task and that on the other they use SeqQs and SynSM more frequently than LD or any other structure in our production task. Furthermore, we saw that they were significantly more likely to use SeqQs in the task with the lead-in and significantly more likely to use SynSM in the task without the lead-in.

6.4 Experiment with German-speaking children

This experiment is motivated by three crucial questions. In the first place, we investigated an aspect of the acquisition of German which has yet to be explored. We hope to add to this body of literature by studying an age group slightly older than previous work and by focusing on this complex construction.

The second question this experiment addresses is whether German children will produce SeqQs or SynSM or some other construction during the production tasks. We ask this question because our work with English-speaking children has suggested that the integration of clauses is difficult and children will establish work-around techniques for asking the questions they need to ask. In particular, we have suggested in Chapter 5 that English-speaking children might be asking two questions in a row rather than a biclausal question including a medial-wh phrase, but because English-speaking children are unreliable in their use of subject-auxiliary inversion (Pozzan & Valian, 2017) we cannot make this claim with confidence. Our argument that these are examples of SeqQs would be strengthened if we found evidence of the same work-arounds from a group of children who are reliably good at verb placement. As the review in 6.2 shows, while German-speaking children make other errors, verb placement is something they master early on. Thus, if German-speaking children also rely on work-arounds such as SeqQs, we can more confidently suggest that these productions by English-speaking children are not biclausal questions with medial-wh phrases, but are, instead, instances of SeqQs used as work arounds. We are also interested in whether German-speaking children make any errors (such as blends or disfluencies) which indicate difficulty at the clausal boundary.

Finally, in Chapter 5, we established a relationship between children's performance on the production and comprehension tasks and their WM. If our hypothesis that the errors in English-speaking children's productions and comprehension are the result of processing is correct, then we should see a similar relationship between German-speaking children's performance on these language tasks and their WM.

To summarize, therefore, we hope to establish three things: (1) whether German-speaking children use SynSM in their production and respond appropriately to SynSM questions when asked; (2) whether German-speaking children also use work-arounds such as SeqQs and make errors at the clausal boundary; and (3) whether German-speaking children's performance on production and comprehension tasks is correlated to their WM. Together, the answers to these questions will inform our understanding of language acquisition in German and English as well as the role of processing in acquiring adult-like linguistic performance.

6.4.2 Participants

Forty-eight children were recruited by the BabySprachLabor at the University of Konstanz. They were recruited from the Konstanz area including a preschool on campus. Children who could not answer filler questions or practice questions were excluded from the analysis (n=20, similar to the English-speaking children in Chapter 5). After this exclusion, data from 28 children were analyzed. Children ranged in age from 3;11 to 6;8 (mean 5;6). While we had hoped to collect data for 6 more children, the COVID-19 stay at home order halted data collection in Germany for the foreseeable future.

6.4.3 Procedure

All tasks were presented as stories accompanied by simple PowerPoint animation and images on a DELL XPS laptop computer. They were recorded using an Olympus L100 audio recorder. All children completed a comprehension, production, and WM task. Tasks were balanced for order, but WM was always measured last, for the following reasons. In order for WM to be meaningful across participants, it must be measured at the same point in the task for each participant; obviously children will perform differently at the beginning of an hour-long session than at the end. However, as will be described below, the WM task involves repeating what the experimenter says. Pilot testing revealed that when we started sessions with the WM task, children tended to think all the tasks would be based on repetition and were inclined to continue repeating the experimenter rather than producing relevant constructions. For this reason, the WM task was always done at the end of the sessions.

Children were immediately assigned one of the production tasks (lead-in or no lead-in) and were introduced to the relevant puppets for that task. This allowed children to establish a connection with the characters with whom they were going to interact. Characters then 'went inside the computer' by disappearing behind a screen and appearing in the PowerPoint.

To get a full comparative picture of German-speaking children's abilities, we of course needed to include a comprehension task with SynSM questions as in (12, repeated as 13). However, we also wanted to know if German-speaking children would make similar errors to English-speaking children when faced with questions that include a wh-relativizer (*was*) as in (14). The reader will recall from Chapters 4 and 5 that the comprehension task in English included *how+what* questions (14) and English-speaking children responded as if they were *what*-questions, as in (13). We therefore had two comprehension tasks for children: one followed by questions like (13), one followed by questions like (14).

(13) **Was** hat der böse Stefan Selina erzählt, **was** er stehlen wird?
What has the evil Stefan Selina told, **what** he steal will
What *did Evil Stefan tell Selina (that) he will steal?*

(14) **Wie** hat der böse Stefan Selina erzählt, **was** er stehlen wird?
How has the evil Stefan Selina told, **what** he steal will
How *did evil Stefan tell Sherry* ***what*** *he will steal?*

Similarly, for production, we wanted to establish how German-speaking children would respond to the tasks we had given English-speaking children, we therefore included both the task with the lead-in and without the lead-in.

The study is a mixed-design: every child did one of the two production tasks, one of the two comprehension tasks, and the WM task. The language tasks were presented in one of two orders (Form A and Form B) so we could determine whether there was any effect of order of presentation. Task pairings (which production task with which comprehension task) and ordering (Form A and Form B) were counterbalanced. Each task took roughly twenty minutes per child, though there was variability. Children were encouraged to take a break between each task resulting in each session taking approximately an hour and a half. While this is a long session for people in this age range, children were openly enthusiastic about the game-like tasks, asking whether they could play with Nina and the other characters more.

6.4.4 Materials

Materials are the same as presented in section 6.3.3, but we present brief descriptions here. For further details, we refer the reader to the materials sections of Chapters 4 and 5.

Comprehension

Children were presented a series of short stories followed by a pre-recorded question (recorded by one of the native-speaker RAs). The task included one practice story, two filler stories, and six target stories. While the stories were identical, children were either given SynSM or *wie+was* 'how+what' following them. It is crucial to note that the task was open-ended: children were not given options to choose from, but were allowed to respond freely. They were rewarded for their response (regardless of its accuracy) with a sticker given to them by the friendly puppet, Nina Nilpferd. Nina could come and go from the 'story world' in the computer (using animation) which added an extra level of interest for the children as well as a connection to the stories.

Production with lead-in

The production task which included the lead-in was presented as a continuous story where children had to ask all characters their opinion about what should happen next. The progression of the story depended on finding out everyone's opinion, so children were motivated and the task of asking questions felt natural. The task with the lead-in included a fairy-puppet which children had to work with in order to formulate their questions. This fairy prompted children to begin their questions "*Was glaubst Du…"* thus encouraging a biclausal question. The task elicits 12 questions, controlling for question type (*what* vs. *who*) and extraction type (subject vs. object).

Production without lead-in

The production task without the lead-in was developed to determine whether the presence of the prompt from the fairy was causing children extra difficulty. Two space-travelers (an alien and a racoon) crash land on earth and need the child's help to find all the pieces of their spaceship. Like the production task with no lead-in, the task is one continuous story where we need to ask everyone's opinions about what should happen next. This makes the question felicitous. However, in this story, the racoon, Mindy, must translate for her alien friend, Mork, prompting children to ask Mindy questions like *Who does Mork think can fix the engine*?. There were three practice questions to help children understand the relatively complex idea of translation, as well as 2 filler questions which targeted simpler *where*-questions to help children get the idea of what was required. There were 12 target questions, controlled for extraction type (subject vs. object), but all were *who*-questions. This was so we could tell whether children were fronting the contentful wh-phrase or inserting a scope marker (*was* 'what') to the beginning of the structure.

WM task

The WM task was the standardized "Digit Span" task from the Differential Ability Scales-II (DAS-II) (Elliot, 2007) used with English-speaking children (Chapter 5). We measured both Forward Digit Span (FDS) and Backward Digit Span (BDS) and gave each child a composite score (FDS+BDS). For FDS, the child is asked to repeat a string of numbers, beginning with a string of two digits. If the child repeats these digits in the appropriate order successfully, the experimenter advances to a longer string of digits. This continues until the child is no longer successful. The child's "base" is established as the longest string of digits where the child was successful at least four times. Their "ceiling" is the longest string of digits where they were successful no more than once. Base and ceiling scores are used to compute an "ability score". For the BDS, the task is the same except the child must repeat the string of digits in the opposite order to the experimenter. The composite scores consisted of BDS ability score and FDS ability score. This was to give us a complete picture of the child's WM ability as suggested by Wilde, Strauss, & Tulsey, (2004).

6.4.5 Data Coding and Analysis

Children's questions and responses were transcribed and translated by three paid RAs. Two were the RAs recruited locally by the BabySprachLabor who conducted the experimental tasks and were native speakers of German. The third was a PhD student in the German department at JHU who is a fluent second-language speaker of German with extensive translation experience. Her transcriptions were spot-checked by the native speaker Research Assistants to ensure accuracy and were deemed accurate 100% of the time. Native speakers and the translator judged and coded constructions as adult-like (1), meaning a grammatical sentence of German, or non-adult-like (0). Constructions were then coded by the author to indicate whether they were (1) or were not (0) instances of various types of construction, including, but not limited to SynSM (a complete list is given in Table 1). The author's coding was spot-checked by another native speaker of German with an overall agreement rate of 92% and a 100% agreement rate in the category of SynSM constructions.

Table 1. *Construction Types, examples from child productions*

Construction Type	Example (s)	Translation
SynSM	Was glaubst Du, wer die gute Fee ist?	Who do you think is the good fairy?
SeqQs	Was glaubst Du? Wer ist die gute Fee?	What do you think? Who is the good fairy?
LD	Wen glaubst Du den- der[46] uns helfen kann?	Who do you think can help us?

[46] This is an example of the type of 'swap' children do at the clausal boundary, which we will discuss below. It should be further noted that this is not a typical LD construction, which would require a *dass* complementizer. *Der/den* is used in relative clauses.

Parenthetical	Was - glaubst Du - sollen wir suchen?	What (do you think) should we look for?
"Can you ask…" constructions[47]	Kannst Du ihn fragen, wo der Flügel ist? Mindy, fragst Du mal Mork, was… was er denkt.	Can you ask him where the wing is? Mindy, ask Mork what…what he thinks.
Infinitival Clause	Wen möchte Mork zum Essen einladen?	Who does Mork want to invite to dinner?
Monoclausal Question	Was denkt Mork? OR Wen sollen wir einladen?	What does Mork think? OR Who should we invite

Recall that our goals include (1) surmising whether German-speaking children use SynSM in their productions and respond appropriately to SynSM questions when asked; (2) establishing whether German speaking children also use work-arounds such as SeqQs and make errors that resemble syntactic blends and other disfluencies in their production; and (3) determining whether German-speaking children's performance on production and comprehension tasks is correlated to their WM. For the first goal, we computed by-subject means for each production type. For the second, we coded for whether constructions included (1) various types of errors or not (0) and whether they were (1) or were not (0) examples of other construction types. We will report on the by-subject-mean productions for these various types of construction.

We also will establish whether productions include disfluencies. We counted five types of overt disfluencies only when they appeared at the clausal boundary. These

[47] It should be noted that these were not always mutually exclusive with other question types. For example *Kannst Du Mork fragen, was er denkt, wer den Weg weiß?* includes a *can you ask Mork* frame, but it includes a SynSM structure within it *who does he think knows the way*. These were coded as instances of both, but are not included as a 'work-around' since they do not simplify the construction.

included: (1) an audible "umm", (2) a restart (defined as returning to the beginning of either the structure or the embedded clause), (3) stutter (defined as at least one repetition of the wh-phrase at the clausal boundary, as in *Was denkt Mork* ***wer…wer*** *den Flügel hat?*), (4) an error on the medial wh-phrase, and (5) swapping the medial wh-phrase (as in *Was denkt Mork,* ***wo wir… wer*** *den Motor reparieren kann?*). These disfluencies were chosen based on McDaniel et al. (2010).

To address the third goal, these data were then submitted to a mixed-effects logistic regression model (e.g. Krusxhke, 2015) using the lme4 package version 3.2.2 (Bates, Maechler, Bolker, & Walker, 2015) in the R statistical analysis environment (R Core Development Team, 2015). The model included WM as a fixed effect. Each of the response types (Comprehension) and construction types as well as whether or not the structure contained a disfluency (Production) were tested for whether WM were a significant predictor. The model also included age as a fixed effect, which was centered and scaled as well as Form (A vs. B). The random effects always included participant and trial number. The model was run with the maximal random effects that would converge (Barr, Levy, Scheepers, & Tily, 2013) which minimally included random intercepts for participant and item. WM, Form, and Age were used as a random slope for participants when the model would converge. This model will tell us whether WM is a significant predictor of the various responses and construction types we are interested in and therefore whether they are the result of immature processing mechanisms.

6.4.6 Results

Comprehension

Recall that children were divided into groups that heard the same stories, but one group heard a SynSM question at the end as in (15), while the other group heard a question with a medial wh-phrase as in (16).

(15) **Was** hat Stefan Selina erzählt, **was** er stehlen wird?
What did Steven tell Sherry (what) he will steal?

(16) **Wie** hat Stefan Selina erzählt, **was** er stehlen wird?
How did Stevan tell Sherry what he will steal?

Twelve children were in the SynSM group, 16 were in the wh-relativizer group. Recall that this imbalance was the result of the University of Konstanz shutting down due to the COVID-19 crisis. We will describe the findings from the two groups separately first and then compare them.

SynSM questions

In response to SynSM questions, children gave adult-like responses 83% of the time. The most common incorrect response was an answer to the embedded clause question (e.g. what Steve *will* steal in (15)), which accounted for 13.9% of responses. Together, these account for 96.9% of responses given.

Questions containing wh-relativizers

In response to questions containing wh-relativizers, there were more varied responses. The largest group of responses were adult-like: 55.1% of children's responses were correct responses to the *how* question. However, like the English-speaking children in Chapter 5, the most frequent incorrect response was, in fact, a WSM-like interpretation: 23.1% of children's responses were responses to *what* Steve *said* he was going to steal. 6.4% of responses were responses to the embedded clause (what Steve was going to steal) and 3.8% were incorrect responses to the *how* question. 6.3% of responses were *how* responses that were not either of the methods described, generally this means the child made up their own answer or gave an ambiguous response such as *He said it*. Together these responses accounted for 94.7% of responses children gave.

Correlations with WM

Four children were removed from the analyses of WM because they did not complete the WM task or because their scores were more than two standard deviations away from the mean (1 from SynSM task, 3 from wh-relativizer task). This leaves twenty-four children whose data we included in the following models (11 in SynSM; 13 in the wh-relativizer condition). One of the goals of this experiment was to find out if German-speaking children responded appropriately to these two types of questions. It was crucial, therefore to establish whether children performed equally well on both tasks. We submitted the data to a model asking whether which task the participant took part in predicted an adult-like response. There was no significant difference in adult-like performance between the two tasks ($p>.1$). In other words, while the percentages reported above do indicate more adult-like responses in the SynSM task, the difference between adult-like performance in each task is not significant. German-speaking children performed equally well on the SynSM comprehension task and the wh-relativizer comprehension task.

Next, we established whether the number of WSM-like responses in the wh-relativizer task were significantly different to those in the SynSM task. This will crucially allow us to determine whether German-speaking children can differentiate between the two structure types or if there are significant WSM-like responses in either case. We submitted the data to a model asking whether WSM-like responses were predicted by task. There was a difference between performance on the different forms within the wh-relativizer group: Form A elicited SynSM responses 33% of the time compared to Form B's 14%. However, upon further investigation, we saw that this difference was driven entirely by two children, who both gave SynSM responses most of the time and both happened to have Form A. If those two are excluded, performance on the two forms was identical. We therefore removed Form as a fixed effect from this model. Fixed effects, therefore, included task and age (centered and scaled). Random effects included participant and item with random slopes for task and age. We found that children were significantly more likely to give a WSM-like response in the SynSM task than the wh-relativizer task ($p<.001$). German-speaking children gave SynSM responses to SynSM questions and not to other complex construction types.

We also wanted to establish whether WM predicts performance on these tasks. There were not enough instances of errors for WM to predict any particular error type. Therefore, to that end, we asked whether correct responses were predicted by WM. There were too few participants for the model to converge in each task individually, however, when we look at the group as a whole, WM is a predictor of adult-like performance (SynSM response in the SynSM condition; correct 'how' response in the wh-relativizer condition) across the two tasks ($p=.01$)). This means that when all participants are considered, regardless of task, adult-like responses are predicted by WM scores to a significant level. The results of our model are shown in Table 2.

Table 2: *Logistic regression results: WM as a predictor of Adult-like comprehension*

	B	Error	z-value	Pr(>|z|)
BothWM	0.04	0.01	2.6	0.009*
Scale(Age)	-0.41	0.44	-0.94	0.34
Form	0.95	0.75	1.28	0.2

* denotes significance

It is worth noting that age was not a significant predictor of performance. This is similar to our findings from Chapter 5.

Comparison with English Comprehension

Finally, recall that this German investigation stems from a question about errors made by English-speaking children who give WSM-like responses to wh-relativizer questions. We therefore submitted the English and German data to a model asking whether language predicted adult like-responses. We found no significant difference between performance by the two groups (p>.1). Children speaking both languages performed better in the task with a single true wh-phrase (LD for English-speakers; SynSM for German-speakers) than the wh-relativizer task. Furthermore, both English and German-speaking children mistakenly responded to medial wh-phrases in the wh-relativizer condition:16.1% in English, as reported in Chapter 5; 23.1% in German. This difference is not statistically significant (p=.35). Fig. 1 illustrates the performance of children in each language group on the comprehension task with the wh-relativizer.

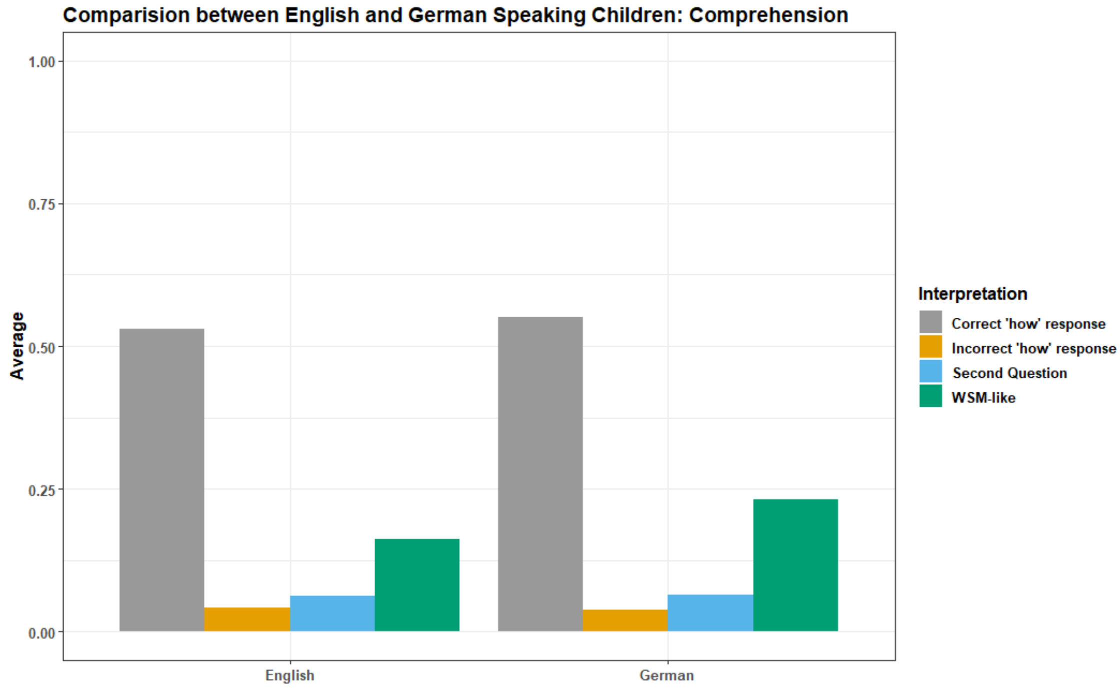


Fig. 1 Average response types for English and German children in the *how+what* condition. In both languages correct responses occur significantly more than other responses. There was no significant difference between the two languages on any response type.

As is clear from Fig. 1, children in each language group performed in the same way: the majority of responses were correct, but the mistake which was next most common was a WSM-like response. There are more WSM-like responses in the German group and slightly more correct 'how' responses in the German-speaking groups, but neither difference is significant (p>.1 in both). Overall, performance on this task was essentially identical in the two language groups.

Production

Recall that in production, there were two different tasks: one that included a lead-in or prompt and one that did not. Sixteen children participated in the task with the lead-in, producing 192 questions and 12 participated in the task without the lead-in, producing 144 questions. Recall, the imbalance between groups was a result of the University of Konstanz shutting down for the COVID-19 crisis. Across both tasks, a total of 21 utterances were excluded for various reasons (generally, the experimenter had to give the child the utterance to prevent frustration). This is comparable to the number of exclusions reported for English-speaking children in Chapter 5. It is also worth pointing out that of the remaining 315 utterances, only 18 are monoclausal, indicating that both tasks worked for eliciting complex constructions from children.

Unlike German-speaking adults, child participants showed no significant difference in their use of SynSM or SeqQs in the two tasks. In the task with the lead-in, 55%[48] of productions were SynSM structures and 35% were SeqQs. This is compared to the adult averages of 18.4% SynSM and 76% SeqQs. In the task with no lead-in 50% of children's productions were SynSM structures and 11% were SeqQs. This is compared to the adult averages of 80.3% SynSM and 3.4% SeqQs. The difference between adult and child performance will be discussed.

The only LD structures appeared in the task with the lead-in (1.6%). Similarly, the parenthetical construction described in the adult study only appeared in the task with the lead-in (3.6%). In both groups, children produced work-around structures; they included questions like (17-19), which were semantically appropriate, but less complex structures than we were eliciting. (17) consists of a matrix yes-no question and an embedded wh-

[48] Percentages in this section are actually by-subject means, the average of participant averages. This is not exactly the same as a percentage because not every participant successfully completed every trial, but gives us a more accurate picture of the overall performance of the group.

question. (18) is a monoclausal question with a complex verb structure[49]. (19) is a monoclausal question.

(17) target: *Was denkt Mork, wer die schöneren T-Shirts hat?*
Who does Mork think has the best shirts?

Production: Mindy, kannst Du fragen Mork[50], welches das schönere T-Shirt ist?
Mindy, can you ask Mork, which the prettiest T-shirt is?
Mindy, can you ask Mork which is the prettiest t-shirt?
participant 50, age: 5;10

(18) target: *Was denkt Mork, wen wir einladen sollen?*
Who does Mork think we should invite?

Production: Wen möchte Mork zum Essen einladen?
Who wants Mork to dinner to invite?
Who does Mork want to invite to dinner?
participant 22, age: 4;10

(19) target: *Was denkt Mork, wen wir einladen sollen?*
Who does Mork think we should invite?

Production: Wen sollen wir einladen?
Who should we invite
Who should we invite?
participant 42, age: 6;1

In the task with a lead-in, these work-arounds only accounted for 5.4% of productions. They were much more prevalent in the task without the lead-in, accounting

[49] Wurmbrandt (1991) does not consider *möchten* 'want' either as a full lexical or "restructuring" verb, but says it contains properties of both. Fully lexical entails the structure would be biclausal; restructuring entails the structure would be monoclausal. "Restructuring" is thus understood to be less syntactically complex than a fully lexical verb taking a true CP complement.

[50] Note that this utterance is ungrammatical because the child has swapped the word order of the verb *fragen* and Mork. This is an example of what Clahsen (1982) found: errors in verb-placement of the matrix clause rather than the embedded one.

for 33.9% of all productions. This is unsurprising since in the task with no lead-in, the child was allowed to ask in any way they preferred.

Finally, it should be noted, that the majority of children's productions were adult-like, meaning they were grammatical sentences of German (regardless of what type of construction they were). In the task with the lead-in 87.9% of productions were judged grammatical by the translators. In the task without the lead-in 76.5% were judged grammatical. Errors children did make were most frequently the use of the wrong morphological case or number. In (20) for example, *glauben* should take the dative case *wem* rather than accusative case *wen*.

(20) Mindy, Fragst Du Mork, *wen wir glauben sollten?
c.f. Mindy, Fragst Du Mork, wem wir glauben sollten?
Mindy, you ask Mork who we should believe

Comparison with English Production

There were several crucial similarities and differences between the German-speaking children's productions and the English-speaking children's productions. Results of our mixed effects logistic regression model indicate, first, that there is no difference between the two language groups in the proportion of biclausal productions. By this we mean, the English-speaking children attempted LD constructions at the same rate as the German-speaking children attempted SynSM. This is an argument against the hypothesis that English-speaking children are using a German-like grammar: if they were, all other things being equal, we would expect the same approximate percentage of SynSM in their productions. Furthermore, as would be predicted if we presume that children learning either language have already established the target-like grammar of complex questions correctly, German-speaking children produced significantly more SynSM constructions than English-speaking children (p=.0007). Finally, as mentioned above, German-speaking children produced fewer non-adult-like constructions (p<.01). However, as we will discuss this is probably (at least in part) because any SeqQs produced by English-speaking children would have counted as medial wh-phrases and thus would not count as adult-like.

Fig. 2 illustrates the comparison between English and German-speaking children's performances on the task without the lead-in. The "complex adult" constructions, in gray, indicate those which were judged as adult-like attempts at LD in English and adult-like attempts at SynSM in German. The green bars indicate the use of any type of grammatical work-around *except* SeqQs (e.g. infinitival clauses). In German, the blue bar indicates use of SeqQs. Recall that in German we can definitively call a construction an example of SeqQs because we know from previous work (e.g. Clahsen, 1982) that German-speaking children correctly place verbs in embedded clauses. In English, however, we cannot definitively call some constructions examples of SeqQs because English-speaking-children are not as reliably good at refraining from subject-auxiliary inversion in embedded clauses (Pozzan & Valian, 2017). Thus, in English, the blue striped bar indicates a construction which included a medial-wh phrase and thus *might* be an example of SeqQs or an example of SynSM (under a Parametric Acquisition Hypothesis) or an example of a syntactic blend, as discussed in Chapter 5.

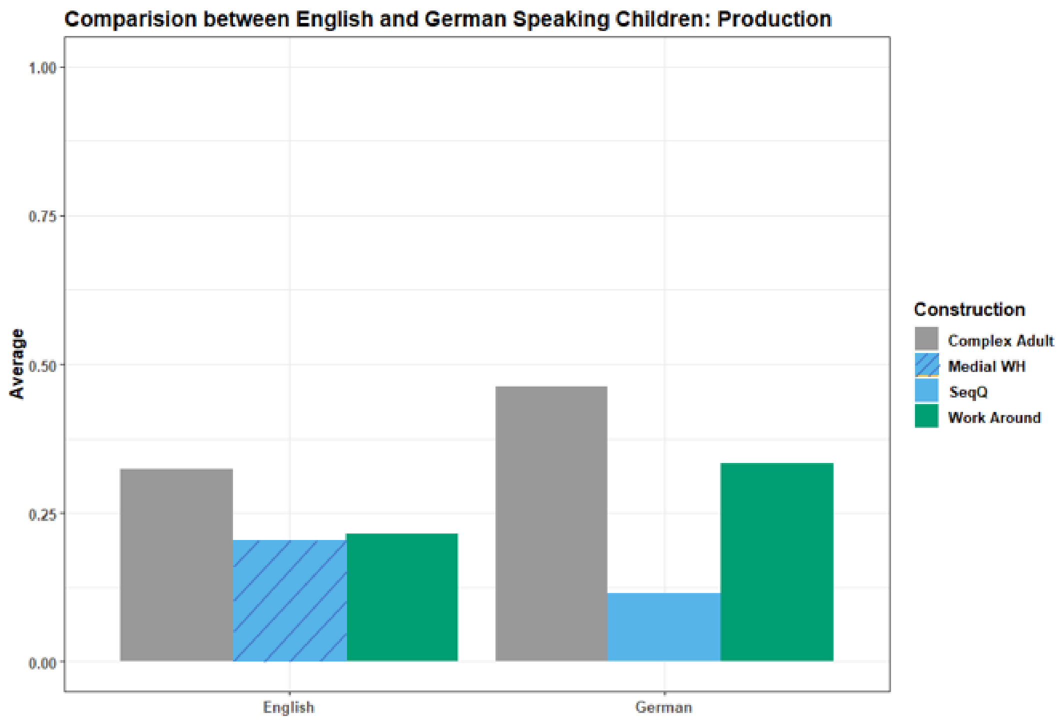


Fig. 2 By Subject Mean production of construction type by language. There were significantly more complex-adult productions in German than in English, but there were not significant differences between other construction types. Crucially there is no significant different between medial-wh productions in English and SeqQ productions in German.

Fig. 2 shows clearly the difference between English and German speaking children's adult-like complex constructions. However, the comparison to focus on is the two blue bars. We know these indicate SeqQs in German. It is reasonable to assume that a task which elicits SeqQs in one language should do so in another (where SeqQs are equally possible). In English, we have constructions with medial wh-phrases which we cannot definitively call SeqQs, but which appear at a similar rate to SeqQs in German. The difference between these occurrences is not significant. We see therefore, a comparable number of complex productions (more of which are adult-like in the German group, which we discuss below), a comparable number of work-around constructions, and a comparable number of productions with medial wh-phrases (which are definitively SeqQ constructions in German). This is a crucial finding which we will discuss thoroughly, below.

Correlations with WM

As in our discussion of the relationship between comprehension an WM, four children were removed from this analysis (two from each group, leaving 14 in the group with the lead-in, 10 in the group without). Two were removed because they did not successfully complete the WM task, two were removed because their WM score was more than 2 standard deviations below average (no child was more than two standard deviations above average). Children's performance on the WM task on average was high (average composite WM of 203), this would be in the 50th percentile for even the oldest children tested, according to Elliot (2007). However, one difference between the German and English-speaking children's WM scores was that there was a much smaller range among the German-speaking participants. While English-speaking children's WM score ranged from a composite score of 109 to 311 (standard deviation of 47), the German-speaking children's scores ranged from a composite of 136 to 264 (standard deviation of 39). This is important because less variability in the scores means finding a correlation will necessarily be less likely. This finding is possibly the result of fewer participants (24 in German compared to 32 in English), which was a result of the fact that we had to halt data collection in Germany.

We ran several mixed-effects logistic regression models to determine whether WM predicted either adult-like productions, SynSM, SeqQs, or work-arounds. None of the construction types produced were significantly predicted by WM. This is possibly because the German-speaking children's WM scores were less varied than the English-speaking children's scores and all were higher than the average reported for the age group, as discussed above (Elliot, 2007). However, German-speaking children's questions were also mostly adult-like (>75% in each task). Given the high rate of adult-like productions and the relatively minimal variability in WM scores, it is unsurprising that no correlations were found. However, the Limited Processing Hypothesis suggests not only that WM should predict adult-like responses, but that WM should predict errors. We found relatively few errors of syntax, so we must therefore turn to another type of error: disfluencies at the clausal boundary. The Limited Processing Hypothesis predicts that disfluencies at the clausal boundary should be predicted by WM score.

Disfluencies at the clausal boundary

Recall that we coded for five types of disfluency: (1) an audible "umm", (2) a restart (defined as returning to the beginning of either the structure or the embedded clause), (3) stutter (at least one repetition of the wh-phrase at the clausal boundary (4) an error on the medial wh-phrase, and (5) swapping the medial wh-phrase. These disfluencies were chosen based on McDaniel et al. (2010). Across all productions (from the task with the lead-in and without), there were 316 utterances in total. 94 of these included at least one form of disfluency at the clausal boundary (29.7%). The most frequent disfluency found was a restart where the child started the clause over. This occurred in 47 utterances:18 utterances included a repeat of the wh-phrase at the clausal boundary; 23 included an audible 'umm' and 7 utterances included an error on the wh-phrase. Finally, 41 (13%) included a 'wh-swap' at the clausal boundary; meaning the child first uttered one wh-phrase and then changed it to be another. All of these disfluencies point to difficulty at the clausal boundary.

There was no significant difference between groups in the appearance of disfluencies at the clausal boundary. We therefore collapsed across groups for this discussion, though analyses were run for the tasks separately as well, with the same results. We investigated whether the presence or absence of any clausal boundary error was predicted by WM. We found that disfluencies at the clausal boundary were significantly affected by children's composite WM score (p=.05) The results of our mixed effects logistic regression model are presented in Table 3 and the correlation is illustrated in Fig. 3. Again, it is worth noting that WM is a better predictor of disfluencies than Age, which is not a significant predictor.

Table 3: *Logistic regression results: WM as a predictor of disfluencies in production*

	B	Error	z-value	Pr(>|z|)
BothWM	-0.01	0.01	-1.9	0.05*
Scale(Age)	0.03	0.22	0.14	0.89

*indicates significance

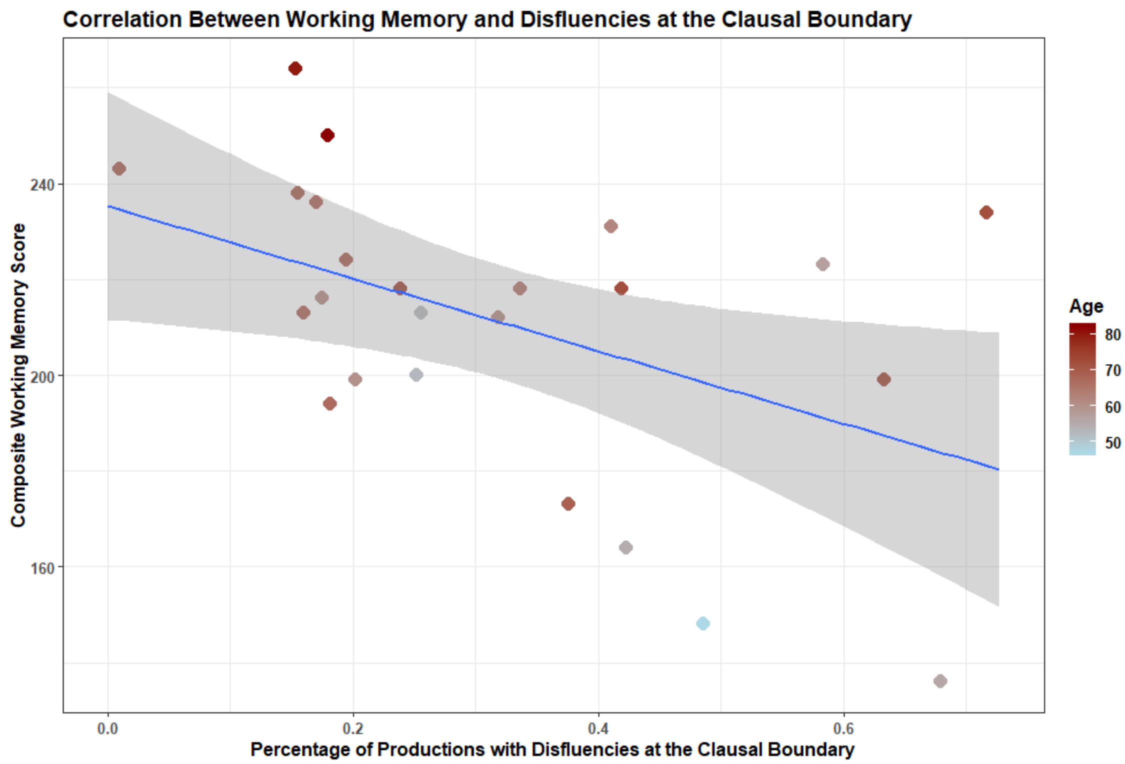


Fig.3 The relationship between composite WM score and disfluencies at the clausal boundary. There was no significant effect of age (which is represented in months by color) (p>.1).

As is apparent in Fig. 3, the higher a participant's WM score, the fewer disfluencies were found at the clausal boundary. Lower WM scores result in more disfluencies. This is precisely as predicted by the Limited Processing Hypothesis.

In summary, the results of the production task indicate no relationship between the type of constructions children produced (SynSM, SeqQs, etc) or in whether those productions are adult-like. The latter is likely the result of a ceiling affect, since over 75% of constructions were judged as adult-like by native speakers. However, there is a strong relationship between WM scores and the presence of a disfluency at the clausal boundary such that the higher the child's WM, the less likely they are to make an error.

6.4.7 Discussion

Comprehension

The comprehension task showed several things. In the first place, German-speaking children do respond to SynSM structures correctly 83% of the time. Second, we have seen that, like their English-speaking counterparts, German-speaking children give a greater variety of responses to questions containing wh-relativizers, but still perform relatively well (55% accuracy on a task that is open-ended). German-speaking children also give WSM-like responses to these questions about 23% of the time, which is not significantly different to the 16.1% of this error type given by English-speaking children.

The results of our logistic regression models indicate a significant relationship between children's adult-like responses and their WM in both tasks. This means that a child with higher WM is more likely to give an adult-like response, regardless of the task. This is precisely what is predicted by the Limited Processing Hypothesis: it is not a question of whether the child knows how to navigate a particular structure, but instead a question of how well her processing mechanisms can handle that structure.

We showed that there is no difference in the proportion of adult-like responses of English and German speaking-children, which means that non-adult-like behavior in response to these questions is not related to the target grammar. Children speaking both languages perform essentially like adults when there is one true wh-phrase (whether the construction is LD or SynSM) and make errors to a similar extent when responding to a question with a wh-relativizer. This suggests that there is something difficult about the wh-relativizer itself or its placement within the clause. One suggestion is that the presence of any wh-phrase at the clausal boundary is confusing for children, but that would also predict errors in the SynSM condition, which we did not find. The processing model of comprehension suggested by Phillips & Ehrenhofer (2015) offers a potential alternative explanation. As discussed in Chapter 5, the model suggests that interpreting a question requires the child to recognize a wh-phrase and hold it in memory as they search for a relevant gap for it to fill (which establishes its grammatical function, subject vs. object). They must simultaneously maintain the matrix wh-phrase in the scope position and they must not allow the relativizer (which shares a form with wh-phrases used in questions) to interfere. In other words, when responding to a question such as *How did Steve tell Sherry what he stole,* the child must hold *how* in their memory as the wh-phrase with highest scope and the one to answer. They will come to the wh-relativizer *what* and they have to keep that relativizer from interfering with the true wh-phrase. We suggest that children speaking either language who have a lower WM have greater difficulty maintaining *how* as the wh-phrase to answer.

A crucial follow-up to this experiment, which would further confirm whether the difficulty was the result of *any* wh-phrase at the clausal boundary or whether relativizers are particularly challenging, would be to investigate the use of SynSM which used a contentful wh-phrase besides *was.* All of our comprehension questions in the SynSM condition were *was*-questions, meaning that the contentful-wh phrase was always the same as the SMer. We had to use *was*-questions in order to establish a comparison with the LD questions we had asked English-speaking children. It would be worth investigating whether children performed in a similar way to the wh-relativizer condition if the contentful wh-phrase were different from the SMer. We would predict that performance on any WSM constructions should require less processing than a question with a wh-relativizer because (regardless of how WSM is analyzed) there is a single wh-phrase to track and answer, whereas in the questions containing a wh-relativizer, there are two. We leave this to future work.

Production

In production, we saw no difference between SynSM productions in the two tasks. There were more SeqQs used in the task with the lead-in than the task without the lead-in, but this difference was not significant. Children's performance showed a significantly different pattern to the adult German-speakers. Adults performed differently in the two tasks, producing many more SeqQ constructions with the lead-in than without it, while children perform identically in both tasks. As mentioned in Section 6.3.6, we postulate that the difference between Adult performance on the two tasks must result from the use of the lead-in. It is possible that the use of the puppet to provide the lead-in made the adults feel they should use their child-directed register, which would result in SeqQ productions, according to our findings in Chapter 3. It is possible that children either are not sensitive to this difference in Register or that the presence of the puppet does not cause the same shift in Register as it does in adults. In either case, these tasks show an unexpected difference between adult and child performance.

In neither task did WM significantly predict any construction type. WM did predict disfluencies at the clausal boundary: the greater your working memory, the less likely you were to make an error at the clausal boundary. It should be noted that this is distinct from the English data. While we performed a similar analysis of disfluencies in English, we found no such correlation in Chapter 5. However, as discussed in the General discussion below, we believe that this is because disfluency errors in English surfaced as syntactic blends (a mixture of two constructions) which included a medial wh-phrase and were therefore counted as a medial wh-production rather than a disfluency per se. Furthermore, in Chapters 4 and 5 we found that disfluencies tended to predict non-adult productions. In other words, the English-speaking children who produced more disfluencies were also the children who tended to produce non-adult-like constructions. In the future it will be worthwhile to investigate this relationship more fully.

We showed that German and English-speaking children produce comparable numbers of biclausal constructions (though more of these are adult-like in German-speakers - a point we will return to in the General Discussion), a comparable number of work-around constructions, and a comparable number of constructions which included a medial wh-phrase (which are definitively SeqQs in German). We take this as further assurance that at least some of these productions by our English-speaking participants were in fact SeqQs. It seems unreasonable to suggest that German-speaking children should produce SeqQs at this rate, but English-speaking children do not produce them at all.

However, not all examples of medial-wh phrase productions in English can be explained as instances of SeqQ constructions. In Chapter 5, we suggested that the errors English-speaking children make, resulting in medial-wh phrase production, are either examples of SeqQs or syntactic blends which are the result of immature processing mechanisms. We have seen now that German-speaking children also make errors at the clausal boundary in their productions. The underlying cause is the same for both languages: integrating clauses is difficult and only children with higher processing abilities will navigate the challenge without error at this age. These errors at the clausal boundary support the idea that children have smaller planning periods than adults, put forth by McDaniel et al. (2010). Though the exact manifestations of the errors are different, errors occur at the clausal boundary in both English and German-speaking children.

Finally, both groups of children often use work-around constructions to ask these questions in a simpler way. These include SeqQs, infinitival phrases, "*can you ask Mork…*" constructions, monoclausal questions, in both languages and parentheticals (specific to German). This use of simpler constructions suggests a preference for constructions that can avoid or minimize the structural complexity of a fully biclausal construction.

Before continuing to the General Discussion, let us return to the specific questions posed in the intro which are relevant to this chapter.

1) Do German-speaking adults in the Konstanz area use SynSM in their productions? Will they use them in 'child-like' tasks? The results of our test with German-speaking adults suggest that adults in the area around University Konstanz do indeed use SynSM in their productions. Furthermore, yes, they use them in these tasks. We designed our tasks based on the results of the survey reported in Chapter 2, so this further confirms that we have established the relevant pragmatics for SynSM constructions. One unexpected result was that German-speaking adults used SeqQs more frequently in the task which included the lead-in, which leads us to the conclusion that something about the lead-in prompts SeqQs. The task without the lead-in was, therefore, all the more crucial for eliciting SynSM constructions.

2) Do German-speaking children use SynSM in their productions? Yes, German-speaking children do use SynSM in their productions. Overall, children made relatively few errors in production outside of those we described as disfluencies, but there were errors in morphological case, which is unsurprising given the complexity of the case system (including their actual forms) in German.

Do they respond to SynSM questions like adults? Yes, children's responses to SynSM questions were adult-like 83% of the time. There was no effect of age: even the youngest children performed well on this version on of the task. Children did, however, make similar errors to English-speaking children when asked a question with a wh-relativizer: they responded as if the question were a SynSM question. We have suggested that this is the result of children failing to maintain the true wh-phrase as the one to respond to when faced with a second wh-item.

3) Do German-speaking children, like English-speaking children, use simpler constructions (work-arounds) to avoid complex SynSM constructions? Yes, German-speaking children use work-around alternatives, including SeqQs, "*Can you ask…*" constructions, parenthetical questions, and infinitival constructions. Generally, these work-arounds were judged as adult-like (over 75% in each production task) even though they were not the target structure. This means that they were always semantically appropriate and correct uses of the relevant constructions.

4)**Are German-speaking children's performance on production and comprehension tasks correlated with their WM abilities?** Yes, in both comprehension and production, we found a significant relationship between performance and WM. In comprehension, we found that adult-like responses were predicted by WM in across tasks (SynSM and wh-relativizer): children with higher WM were more likely to respond with the adult-like response. In production, we did not find a relationship between WM and any production type, but we did find a relationship between WM and average disfluencies at the clausal boundary. This further supports the hypothesis that children plan their productions in smaller phrases than adults, resulting in more disfluencies at the clausal boundary, as suggested by McDaniel et al. (2010).

6.6 General Discussion

Chapter 6 tested a final prediction of the Limited Processing Hypothesis: if comprehension and production of biclausal questions is difficult for English-speaking children because of processing limitations, the same should be true for children speaking another language, even if that language includes SynSM as a viable option. This chapter has successfully shown that German-speaking children use SynSM appropriately and respond to SynSM questions like adults, but they also make similar types of mistakes to English-speaking children. This finding is not only novel, it holds implications for our understanding of the relationship between processing and language learning. We argue that limitations on immature processing mechanisms lead children speaking either English or German to misinterpret questions containing wh-relativizers and to produce disfluencies at the clausal boundaries of SynSM questions.

Before considering our findings in greater detail along with why they support the Limited Processing Hypothesis, let us turn once again to the Parametric Acquisition Hypothesis and consider whether it could explain the findings we discuss here. This hypothesis was brought forward to explain the production of questions with medial-wh phrases by English-speaking children when they attempted to produce an LD question and their WSM-like responses to questions containing wh-relativizers. The prior suggestion was that these errors might stem from children accessing UG, recognizing SynSM as a possible construction and using it rather than LD. This could result either from a 'mis-set' parameter or it could be the result of competing grammars as suggested by Legendre et al. (2002) or Yang (2002). Regardless of the specific model, the presumption is that the child's grammar is imperfect in that it initially admits all constructions licensed by UG, including some that the child will eventually retain while discarding others, based on positive evidence in the language-particular input they are exposed to. We would thus expect some syntactic creativity to surface in German, perhaps leaving the wh-phrase in-situ (as is viable in Hindi) or using LD when adult German-speakers did not (as is viable in English). However, we did not find errors of this type. In comprehension, German-speaking children might still be establishing which question word was the scope-marker and, because WSM is in their input we might expect more WSM-like interpretations. On the other hand, if we consider the production data and the data from the WSM comprehension task, German-speaking children seem to have already established *was* as an SMer. Therefore, children should not show evidence of more difficulty in responding to questions containing wh-relativizers than the English-speaking children. The errors we see in comprehension have a straightforward processing explanation, as do the disfluencies which appear in production.

Of course, it could be the case that German-speaking children, for whatever reason, have already established the adult-grammar when English-speaking children of the same age have not and, thus, the only errors we see in German are processing errors, while the errors we find in English are (even partially) the result of an immature grammar. This is not impossible, but if it were so, we would need to explain why the findings between German and English-speaking children are so similar and why errors in both languages, in both tasks, are correlated with WM. These similarities across languages will be discussed further below, but in summary: the errors in question are not best explained as a result of Parametric Acquisition.

The similarity between comprehension performance by English and German-speaking children is striking. Children in the SynSM comprehension task performed with a high degree of accuracy, exactly like their English-speaking counterparts performed on the LD comprehension task. In both languages, questions containing wh-relativizers caused difficulty for the children. While most responses were adult-like, many children in both language groups gave WSM-like responses to these questions. This suggests that the difficulty children face with such questions does not arise either from their target grammar, but from something they have in common: their immature processing mechanisms. We have suggested that the wh-form of the relativizer used in both languages results in interference with the matrix wh-phrase. Only children with relatively mature processing mechanisms successfully keep the matrix wh-phrase in their memory as the one to respond to.

In production, again we see striking similarities between the two language groups, but we will delve into each construction type more deeply. In the first place, both groups attempt biclausal constructions at a similar rate. This suggests, first, that both groups know what the adult grammar requires. Second, this suggests that the pragmatics of the scenarios have successfully elicited biclausal questions most of the time. One difference is that German-speaking children are significantly more likely to produce an adult-like construction than the English-speaking children. One possible explanation for this which deserves further exploration is that SynSM constructions are inherently less burdensome on the processing system than LD constructions. This could be for several reasons. Perhaps inhibiting a medial wh-phrase at the clausal boundary (as must be done in LD given a standard cyclic movement analysis whereby the wh-phrase transits via the medial CP on its way to the left periphery of the matrix clause) is more taxing, a suggested by Grolla & Lidz (2017). Furthermore, if a clause is a major planning unit as suggested by (e.g.) McDaniel (2015), producing or parsing a wh-phrase before the clause it relates to could add a layer of processing difficulty. It is also possible that moving a single wh-phrase twice is more difficult than moving two wh-phrases a single time, especially since, as we suggest in Chapter 3, moving an SMer generated in SpecAgrOP to SpecCP will be a shorter distance than moving a wh-phrase from the embedded specCP to the matrix specCP. Either of these accounts could explain why SynSM seems to be easier for German-speaking children than LD is for English-speaking children. Both of these explanations would be worth pursuing in future work.

We have also seen a comparable number of work-around constructions in English and German-speaking children. This finding is a particularly interesting result of doing the task without a lead-in. We found that in both language groups, children will sometimes try to use simpler constructions rather than biclausal questions. These work-arounds do not seem to be related to WM ability, but, as we discussed in Chapter 5, children do appear to be consistent with their use of work-arounds. In other words, a child who uses an infinitival clause once is likely to do so again. This consistency is possibly the result of a non-adult-like constraint ranking (as suggested in Chapter 5) which results in a preference for simpler constructions. This, too, is worthy of future work.

Finally, and most crucially, we found that the number of questions with medial-wh phrases produced by English-speaking children is comparable to the number of SeqQs produced by German-speaking children. While it is not possible to say definitively that English-speaking children are using SeqQs, it is possible to do so with German-speaking children. We consider this external evidence for our claim in Chapter 5 that at least some of the productions by English-speaking children which included medial wh-phrases are examples of SeqQs. While not all of them can be explained this way, a large proportion can be (77%). It seems reasonable to assume that if German-speaking children sometimes produced SeqQs in these elicited production types, English speaking children would do the same.

Generally speaking, therefore, we have shown that English and German-speaking children's behavior on these tasks is remarkably similar in many ways. We have also shown that German-speaking children's performance on comprehension and production are both very good, but that their errors are predicted by their WM. Both of these findings are supportive of the Limited Processing Hypothesis.

6.7 Conclusion

To summarize the most crucial points, Chapter 6 has investigated German-speaking children's performance on comprehension and production tasks involving the SynSM construction, which (to our knowledge) has not previously been undertaken. We have found that they behave like adults in many ways: they respond appropriately to SynSM constructions and use SynSM constructions in their productions. However, there are also several ways in which these children's performance was not adult-like: they respond to questions containing wh-relativizers as if they were SynSM constructions and they tend to show disfluencies at the clausal boundary in productions. They also respond differently to the task with the lead-in than adults do, using more SynSM constructions. We have learned that children's WM is a significant predictor of their performance on each task and that they perform similarly to English-speaking children. These findings support the Limited Processing Hypothesis which suggests that children's errors in responding to and producing these complex structures are not the result of immature grammar, but rather the result of immature processing mechanisms.

Chapter 7: Conclusion

This dissertation has examined the phenomenon of WSM both in adult grammars where it is grammatical (German, Hindi, Hungarian) as well as its supposed use in child English as an example of Syntactic Creativity. The novel contributions of this work (both theoretical and empirical) fall into two critical categories: (1) theoretical analysis of (adult) WSM and (2) the role of processing in language acquisition. Chapters 2 and 3 were devoted to the first, while Chapters 3 through 5 were devoted to the second. Together, this work not only contributes to our understanding of the use of WSM, but also the reason for cross-linguistic variation in its use, and, perhaps most substantially, to our understanding of the role of processing in children's non-adult-like production and comprehension. Below, we provide a summary of the contributions of individual chapters including answers to specific questions formulated in the introductory chapter, before concluding with a look to the future.

Chapter 2

Our novel results end up posing a renewed challenge to previous accounts for WSM. While the Direct Dependency Account (DDA, e.g. McDaniel, 1989) accounts for the syntactic phenomenon, it falls short in explaining why the semantic account given for Hindi cannot map to it. Likewise, the Indirect Dependency Account (IDA, e.g. Dayal, 1994), couched in semantic analyses, cannot account for German. Neither can account for Hungarian. We suggest that previous work has, in fact, been examining two distinct phenomena, one syntactic and one semantic, which sometimes align, but not always, resulting in cross-linguistic variation that neither account can fully explain. Our experimental findings from a cross-linguistic survey conducted on-line support the hypothesis that SemSM surfaces as different syntactic constructions cross-linguistically. In other words, a given meaning (semantic and pragmatic features or properties) may surface as different syntactic constructions (including what has been called SemSM in this dissertation). We identified several semantic and pragmatic factors which affect the construction preferred by native speakers. These included the contrastive topic status of the subject, the resolved/unresolved status of the QUDs, and the (child- vs. adult-directed) Register of the production. Depending on subtle combinations of these factors, native speakers of Hindi, German, Hungarian, and English chose different constructions as optimal. While English and Hungarian were less variant (essentially choosing LD and SynSM, respectively, in every context), both German and Hindi showed alternation in which construction was optimal in a given situation. Thus, the variation found is not only cross-linguistic (meaning, for example, a given context might elicit the use of SynSM in German, but not in Hindi) but also within a language. In other words, even within German and Hindi, there is variation in what construction is optimal depending on the context.

These findings are directly relevant to the following question posed in the introduction: **What pragmatic and semantic features are relevant to children learning languages with and without WSM?** We learned in this chapter that not all situations elicit WSM equally in every language. Contrastive Topic, QUD and Register are all relevant to its use and children will have to learn the subtle differences between contexts which render SynSM optimal.

Chapter 3

Chapter 3 builds on the claim from Chapter 2, that previous work had been examining two separate, superficially similar phenomena: SynSM and SemSM. While these phenomena sometimes overlap as in Hindi, for example, where SemSM can map directly onto SynSM (and often does), this is not always the case. In German, for instance, SemSM frequently surfaces as SeqQs, as our survey in Chapter 2 reveals. We posited that SemSM is universal (following Dayal, 2000) and that different languages use different syntactic constructions to express it. In Chapter 3, we demonstrated that by reranking a few universal constraints, we can model the appropriate syntactic output for each semantic input in English, German, Hindi, and Hungarian.

This unified analysis is possible because the particular OT Syntax framework we use (as initially described in Legendre et al., 1995; 1998), which, by its very architecture, exploits the interface of syntax and semantics/pragmatics. Because OT Syntax allows for the semantic input to affect the syntactic output, this analysis illustrates how Input-Output Faithfulness and Markedness constraints conspire to yield a variety of optimal strategies across languages. The success of this effort suggests not only the universality of SemSM, but is a strong argument for OT syntax and the particular architecture we adopt, as well as a testament to its ability to explain such complex cross-linguistic syntactic variation. Finally, it supports our hypothesis that the phenomenon known as WSM is neither exclusively syntactic nor semantic, but rather the result of the complex relationship between form and meaning.

These findings address the following question posed in our introduction: **What cross-linguistically valid analysis of WSM should we adopt that sheds light on the**

mapping between semantics/pragmatics and syntax in complex questions? The analysis we adopt is couched in OT Syntax. We consider SynSM as a possible syntactic output which is sometimes (indeed, often) the optimal output for SemSM, which we associate with the resolved/unresolved status of QUDs involved. However, we have also shown that other aspects of semantics/ pragmatics are also relevant to the use of SynSM, including CT and Register. Cross-linguistic variation in the rankings of violable constraints results in variation in optimal syntactic outputs.

Chapter 4

Having established the adult grammar, we turned to the question of whether the appearance of medial-wh phrases in children's productions and their response to wh-relativizers in comprehension tasks constitute an example of Syntactic Creativity. In other words, in the context of an investigation of syntactic acquisition which assumes access to UG, are these errors the result of immature grammar or immature processing mechanisms? In Chapter 4, we conducted a novel continuous story production task eliciting biclausal questions with the use of a lead-in. We also conducted a questions-after-stories task which allowed us to distinguish between WSM-like interpretations and responses to the embedded clause question. We found that the improvements we made to the methodology reduced the instances of these errors from what was previously reported, but did not eliminate them. We also conducted a third, within-subjects experiment where children did both the comprehension and production tasks and examined whether or not a child who produced medial wh-phrases also showed evidence of WSM-like interpretations. This experiment revealed no correlation between production and comprehension errors, either across the entire group or within individual children. This does not support the hypothesis that child-WSM is a true example of Syntactic Creativity. We conclude Chapter 4 with the suggestion that these errors are, instead, the result of immature processing mechanisms, allowing for asymmetrical performance on the two tasks within an individual.

Chapter 4 addressed the following questions outlined in Chapter 1:

Does children's performance change when we make methodological improvements to the tasks? We found that children's performance did change from what had been previously reported. We found many fewer examples of WSM-like interpretations than previous work reported (e.g., de Villiers and Roeper, 1995). However, errors in both production and comprehension persisted.

Are medial wh-productions and WSM-like interpretations correlated? No, we did not find evidence of a correlation across participants. This was a crucial finding of Chapter 4 and our first solid evidence that did not support the Parametric Acquisition Hypothesis. Indeed, of all the findings, this might be the hardest for the Parametric Acquisition Hypothesis in its original Chomskyan parameter-based construal to account for. If a child is making these errors (for example) in production because she believes her target grammar includes WSM, it would be difficult to explain why she would not also show evidence for this error in her comprehension.

If there is no group relationship, are there exceptional children who show evidence of WSM in both production and comprehension? No, even those children who produced many questions with medial-wh phrases or gave many responses that indicated WSM-like interpretations were not necessarily children who made many of the other error. In other words, even a child who produced many medial-wh phrases did not necessarily give any WSM-like responses. In fact, we had multiple examples of children who made many errors in one task and none in the other.

Chapter 5

Chapter 4 cast doubt on the idea that these errors were examples of Syntactic Creativity, so we turned to alternative explanations for these errors. Chapter 5 directly tested the Limited Processing Hypothesis by investigating whether children's errors of production and comprehension were correlated with their WM-scores, which we used as a proxy for general processing mechanisms (Baddeley, 2017). Thus, our primary, novel finding from this chapter is that both production of medial wh-phrases and giving WSM-like responses to questions containing wh-relativizers in English are predicted by a child's WM. We have made suggestions for the direct cause of processing load in each task based on models of comprehension (Phillips & Ehrenhofer, 2015) and production processing mechanisms (e.g. Dell, 1986). In comprehension, we suggest that WSM-like responses are the result of the child failing to keep the true wh-phrase in memory and preventing the wh-relativizer to interfere. Thus, they end up giving a response to the wh-relativizer as if it had scope over the entire question. In production, we suggest that questions with medial-wh-phrases are one of two things: 1) SeqQs used as a work-around structure or (2) a syntactic blend (Garrett, 1980) of an LD construction and an SeqQ construction that results in something that resembles SynSM. While the empirical data remains somewhat inconclusive we return below to this proposal in light of unambiguous German results (Chapter 6).

The results presented in Chapter 5 are important for several reasons. First, they provide direct evidence in support of the Limited Processing Hypothesis. Second, they replicate our findings from Chapter 4 which showed no correlation between the production and comprehension tasks. Third, we have added to the body of literature suggesting WM is relevant to both production and comprehension. Fourth, this study provides (as far as we are aware), the first instance of evidence of syntactic blends in monolingual children. Finally, we used another novel task to elicit productions without a lead-in. This allowed us to observe what children would do in the absence of the lead-in (inclusion of a lead-in has been a standard feature of production tasks in the previous literature). We found several meaningful differences between the errors produced during this task and the errors produced in Chapter 4 (and in all other reported versions of this experiment). While we found the error persisted, indicating medial-wh production generally was not caused by this aspect of the methodology, we found that children produced many fewer copy-constructions (using the same wh-phrase in the matrix clause and at the clausal boundary) in the absence of a lead-in. In other words, the questions with medial-wh phrases produced in this experiment primarily began with *what* and used *who* medially. We also found that children were creative in their use of work-around constructions, which we were less likely to see when we gave them a lead-in in Chapter 4.

The following questions posed in our introduction are answered in Chapter 5:

Does children's performance change when we make methodological improvements to the tasks? We replicated our findings from Chapter 4. We again found fewer errors (particularly of comprehension) than previous literature, but the errors persisted. We also replicated the finding that there was no correlation between production and comprehension errors.

And when we control for relevant pragmatic and semantic features? In Chapter 5, we added another character to the comprehension task such that all of the questions could be considered as having CT subjects. This was to see if children would be sensitive to this aspect of the input which Chapter 2 suggested was relevant to adult-speakers of WSM-languages. However, there was no difference in English-speaking children's performance from Chapter 4.

Are children's performance on production and comprehension tasks correlated with their WM? Yes, crucially, we found that children's performance on production and comprehension tasks were correlated with their WM scores. We argue that this is evidence for the Limited Processing Hypothesis.

Do children's errors resemble the errors adults make when their processing systems are overloaded? In other words, in production could these errors be syntactic blends? In comprehension, could their non-adultlike responses be the result of interference from other lexical items? Yes, children's errors in production and comprehension resemble these classic error-types described in the adult psycholinguistics literature. We have argued that in production, children's medial-wh phrase productions are likely examples of syntactic blends and that in comprehension their WSM-like responses are in fact the result of interference from the wh-relativizer.

Chapter 6

Finally, Chapter 6 tested another prediction of the Limited Processing Hypothesis. If medial-wh phrase production and WSM-like interpretations are the result of limited processing, then we should expect children of this same age to face similar limitations regardless of what language they are learning. We should see errors in production and comprehension even if the language includes SynSM as a viable construction. We ran our experiments with German-speaking children and adults in Konstanz, Germany in order to test this prediction. We found, first, that German-speaking adults behaved as expected based on our survey in Chapter 2: they used primarily SynSM and SeqQs in their productions and responded to SynSM questions appropriately. Crucially, this verified that our stimuli elicited the SynSM productions and WSM-responses. Our major findings, however, were the results of doing these experiments with German-speaking children. We wish to highlight three findings, in particular. First, German-speaking children performed remarkably well over all, giving more adult-like productions than English-speaking children did. This is an important finding to add to the general understanding of the acquisition of German. Second, like the English-speaking children, adult like-performance in both the comprehension and production tasks were predicted by the children's WM abilities. This suggests that, like English-speaking children, the higher a child's WM, the less likely they were to make an error that rendered the sentence ungrammatical. Finally, we found crucial similarities between German and English-speaking children's errors in both comprehension and production which we enumerate below.

In the comprehension task which asked a question with a wh-relativizer (*How did Steve tell Sherry what he was gonna steal?*) German-speaking children's performance was essentially identical to the English-speaking children. Both groups gave mostly adult-like responses, but both groups also made many more errors in this task than in the LD comprehension task (English-speakers) or the SynSM comprehension task (German-speakers). The primary error, which was made by both groups at a similar rate, was to give a WSM-like interpretation to a question with a wh-relativizer. This finding is particularly telling because it indicates that – regardless of the language of input – children are making the same error in response to the same type of questions. They are giving the wh-relativizer (which appears medially) the scope of the true wh-phrase. This is consistent with what we suggested at the end of Chapter 5: that the errors of comprehension are the result of interference by the wh-relativizer. Children with lower WM fail to hold the true wh-phrase in mind and fail to inhibit interference from the wh-relativizer, which has similar characteristics to a true wh-phrase. In both English and German, only children with more adult-like WM are able to maintain the matrix wh-phrase as the one to answer and give adult-like responses. The fact that both groups behaved in the same way is strong evidence for the Limited Processing Hypothesis.

We also found similarities in performance on the production tasks. In the first place, we saw that both English and German-speaking children mostly tried to produce biclausal questions. This suggests that, in both groups, generally children know that the adult-like response is a biclausal question. A crucial difference we discussed in Chapter 6 is that

German-speaking children were more likely to produce an adult-like, target-like question, meaning a SynSM construction with no errors. We have suggested that it is possible that SynSM is, in fact, less taxing on the processing system, but this finding would need to be tested further. Furthermore, we found that German-speaking children frequently showed signs of disfluency at the clausal boundary, such as stutters and restarts. These disfluencies were significantly correlated with WM: children with better WM made fewer disfluencies at the clausal boundary. In the second place, we found that both groups of children attempted work-around productions at a similar rate. In other words, the strategy, so to speak, of using another, simpler construction was apparent in both groups. Finally, we found that the number of questions which included medial wh-phrases produced by English-speakers is comparable to the number of SeqQs produced by German-speakers. We consider this further evidence of our claim that at least some of the questions with medial wh-phrases produced by English-speaking children are simply SeqQs.

Overall, our findings from Chapter 6 can be summarized as follows. We have found that German-speaking children respond appropriately to SynSM constructions and use SynSM in their productions, frequently in an adult-like way. However, they also make errors which show their striking similarities to English-speaking children: they respond to questions containing wh-relativizers as if they were SynSM constructions and they tend to show disfluencies at the clausal boundary in productions. Children's WM is a significant predictor of adult-like performance on both production and comprehension tasks. These findings support the Limited Processing hypothesis which suggests that children's errors in production and comprehension of these complex structure is the result of immature processing mechanisms.

Chapter 6 has addressed two questions mentioned in our introduction:

What types of errors, if any, do children make when their target grammar includes WSM (German)? German-speaking children make errors in both production and comprehension which are informative to our primary questions. In production, they often show evidence of disfluency at the clausal boundary (in the form of restarts, 'umm's, and stutters). These disfluencies are correlated with WM score. In comprehension, children make few errors in response to SynSM constructions, but behave similarly to English-speaking children when asked questions containing a wh-relativizer. They often give WSM-like responses to these questions just like English-speaking children do.

Are the errors made by German-speaking and English-speaking children similar? Yes, as stated above, the errors made by German- and English-speaking children are quite similar. In production, both children show evidence of disfluencies at the clausal boundary. In comprehension both groups are good at biclausal questions (LD in English; SynSM in German), but make similar errors in response to questions containing relative clauses. In both cases, the error they tend to make is to give a WSM-like response: they treat the wh-relativizer as the contentful wh-phrase to answer. We have argued that in both languages this is the result of children failing to maintain the true wh-phrase as the one to answer when faced with interference from the wh-relativizer, which is superficially similar to a wh-phrase. We have argued that this similarity is strong evidence for the Limited Processing Hypothesis.

7.2 Novel Contributions

The novel contributions of this work are theoretical, methodological, as well as empirical. We have added to our understanding of the semantic analyses of WSM and, as far as we are aware, have addressed the question of what role the pragmatics of context play in the use of SynSM for the first time. We have contributed to the WSM literature by suggesting that research focused on semantic analyses and syntactic analyses have been analyzing two similar, but distinct phenomena. Furthermore, our resulting OT analysis crucially utilizes input-output faithfulness constraints and could not work without them. This is a strong argument for the particular OT framework proposed in Legendre et al. (1995, 1998) as opposed to other OT syntax frameworks which do not incorporate input-output faithfulness as a major aspect of the evaluative component of the grammar. We have also contributed to theories of processing in both comprehension and production in that our results support models of production (e.g., Dell, 1986; Bock 1987) and comprehension (Phillips and Erehnhofer, 2015) put forth by previous literature. Finally, as far as we are aware, we are the first to suggest that these errors might be examples of syntactic blends due to processing overload, which supports Coppock (2010)'s suggestion that this is the case for adults.

Methodologically, this largely experimental work includes several innovations. In the first place, we greatly reduced the number of errors in the comprehension task by balancing emphasis on the two relevant QUDs, illustrating (once again, after Thornton, 1995) that children are sensitive to this aspect of the semantics. Our two production tasks are also novel in that they engage children in a single, continuous story and render the asking of these questions much more felicitous. Finally, our adaption of Schulz (2011)'s translation task for bilingual adults to a task for monolingual children has proven critical for both our findings in German and English. It allows us to see the variety of productions children use and reduces the number of copy constructions dramatically (e.g., 37% of medial-wh productions resembled copy constructions in our task without a lead-in, compared to 82 % in our task with a lead in Chapter 4) which necessarily changes our understanding of the phenomenon.

Empirically, we have collected adult data which inform our understanding of when SynSM constructions are used and added to previous findings documenting children's production and comprehension in both English and German. In English, we have shown that this production error persists longer than previously thought (through age 6;7) and that if children are not given a lead-in, they produce fewer copy-like constructions and more work-arounds. We have documented for the first time, as far as we are aware, German-speaking children's ability to produce and comprehend these constructions.

7.3 Future Directions

As always in research, our work has opened as many questions as it has answered. Before closing, we wish to outline several avenues we hope to pursue in the future. First, in further pursuit of the question of variation in the use of SynSM, we hope to expand our survey to include more participants and improve the survey in various ways. Second, there are several variations on the experiments we wish to pursue. For instance, if these errors are the result of limited processing in children, it stands to reason that we should be able to elicit these errors from adults who are given another task to complete in the meanwhile (a handicap). While the task might have to be very taxing to elicit such errors, the prediction of the Limited Processing Hypothesis is clear: if a mature processing system is overloaded enough, adults should make similar errors to children. Furthermore, we would like to conduct our experiments with children learning a language which does not allow WSM, but which does allow one to distinguish between SeqQs and a medial-wh production. If we find a similar proportion of SeqQ production, this would further affirm our claim that those productions with medial wh-phrases found in English are often SeqQs.

The work in German also needs to be expanded. In the first place, simply expanding the number of participants would be worthwhile. We also want to follow up with a version of the comprehension task which asks a SynSM question where the contentful wh-phrase is *who.* This would help us to establish whether children truly understand the nature of the SMer or if they allow it to interfere with their response types. We also wish to pursue the nature of the work-around productions. Children in both languages produced them frequently and consistently. More work which elicits complex structures without the use of a lead-in is required to further understand this general strategy. Finally, we wish to expand to other production types which carry a large processing load to see if our findings extend to other error types. All of these ideas will help us better understand the relationship between processing and grammatical competence.

7.4 General Conclusion

There is one initial question we have not yet answered directly, which is the primary driving question of this research: **Are children's errors in production and**

comprehension tasks the result of incorrect parametric choices or are they due to grammar-external factors such as experimental artifact or processing difficulties? Every chapter has addressed this question. Chapters 2 and 3 established our take on the grammar of WSM-languages and addressed the cross-linguistic variation seen. They established the relevant pragmatic contexts for eliciting such productions and judgments from children and taught us what to look for in children's behavior. Chapter 4 addressed the possibility that these errors were the result of experimental artifact and showed that, indeed, English-speaking children produce medial-wh phrases and respond to questions containing wh-relativizers as if they were SynSM, as previous research suggests, even with improvements to the methodology. However, Chapter 4 also showed that children's performance on these two tasks was not correlated, meaning that a child who produces medial-wh phrases is not necessarily also a child who responds to wh-relativizers as if they were SynSM. This is not what would be predicted by the Parametric Acquisition Hypothesis. Then, in Chapter 5 we directly tested the Limited Processing Hypothesis by testing children's WM as well as their production and comprehension abilities and we found a strong correlation: these errors are made by children with lower WM scores, suggesting that these errors could be the result of immature processing abilities. Finally, in Chapter 6, we tested the Limited Processing Hypothesis again by conducting our studies with German-speaking children. We found that, like English-speaking children, German children's WM scores predicted their adult-like performance on both comprehension and production tasks. Crucially, we also found striking similarities between children's performance on the production and comprehension tasks across languages. In production, both language groups showed evidence of difficulty integrating clauses. For English-speakers, we argue that it manifests in the form of a syntactic blend. For German-speakers, it surfaces as a disfluency at the clausal boundary. In comprehension, both groups performed well when there was only one contentful wh-phrase in the question (LD for English, SynSM for German) and both groups made similar mistakes when the question included a wh-relativizer. We have argued that the similarities in their errors are evidence for the Limited Processing Hypothesis.

And so, in response to the question ***Are children's errors in production and comprehension tasks the result of incorrect parametric choices or are they due to grammar-external factors such as experimental artifact or processing difficulties?,*** we respond that our findings show no evidence that these errors result from incorrect parametric choices, and much evidence that they are due to immature processing mechanisms. The children we have worked with seem to have a generally adult-like grasp of the relevant grammars as couched in traditional UG-based parametric terms, but are limited in their performance by their immature processing mechanisms.

Appendices

Appendix A: Production Task Story

Designed to elicit 12 questions (3 who-obj; 3 who-subj; 3 what-obj; 3 what-subj). This is an abbreviated version of the script, when told to children, there is more repetition. Also, as the story is interactive, children sometimes asked questions and this meant there would be stops and restarts. The child is given a non-functioning microphone and told that that is our means of communicating with characters inside the computer.

Children are introduced to Batman, Kermit, and Dora who go "inside" the computer to go on and adventure. When they arrive, they are greeted by the Pink Fairy who says:

"Welcome to the fairy world! Here everything is very fair and we will always do whatever two out of the three of you want to do. Does that make sense? [allow child to respond] What would you like to do first?"

(to child) Can you ask everybody what they want to do? Let's start with Kermit. Can you say "Kermit, what do you want to do? (Child asks) Great. Kermit says "I want to pick some flowers". How about Batman? Can you ask him? (Child asks) Great! Batman says "I want to pick an apple."

The fairy interrupts to say "Be careful! There's a scary witch who lives here. She owns this apple tree! If you pick on of her apples, she might get angry."

Hm. Now that we know that, let's ask Dora. Can you ask Dora?

Child asks

Dora says "I'm not scared of a witch! I want to pick an apple" So, because two out of the three of them want to do it, they decide to pick an apple. But look! As soon as Dora picks the apple the witch appears! She turns to the Pink Fairy and says "How dare you tell these people about my apple tree! I'm going to send you to the real world as punishment! Poof!" (The fairy appears as a puppet) for the child to talk to.

"Oh no," says the fairy. "What will your friends do now that I'm not there to help them! Let's watch and find out."

The witch says "For taking my apple, Dora, you must be punished! I'm going to put you under a spell! Poof! You just try talking!"

Dora mumbles, and the witch cackles.

"Haha! There are two rules for my spell. Rule number 1, Dora can only talk when (CHILD) asks her a question. Rule number 2, CHILD and the Pink Fairy have to work together to make that question."

And the witch flies off angrily. (to child) That's really complicated! Dora can only talk when you ask a question, but you have to work *with* the fairy to make the question. I wonder what that means… I have an idea! How about, the fairy can tell you how the question starts, and then you say the whole question since you have the microphone? (allow for further explanation if needed) Let’s try one just to see if that’s what she meant. How about this: I happen to know that someone in the fairy world thinks green is the best color. Let’s find out who! Let’s ask Dora “Who thinks green is the best color?”, so you have to let the fairy tell you the beginning and then you say the whole question, ok?

Fairy “It starts like this! ‘Who thinks…’”

(Child asks- experimenter helps if necessary)

Dora says “Kermit the frog thinks green is the best color because he is green!”

Nice that worked! Let’s try another one. This time we’ll ask everyone… Let’s find out… who thinks apples are really important. Can you ask Kermit “who thinks apples are really important?” (Child asks)

Kermit says “I think it’s the witch who thinks apples are important”.

Great! Now can you ask Batman? (Child asks)

Batman says, “I think it’s the Pink Fairy who thinks apples are really important”

Ok! They disagree, so let’s ask Dora. Remember, to ask Dora, you have to work with the Pink Fairy.

Pink fairy says “It starts like this “Who thinks…”

Child asks.

Dora says “I think it’s the witch who thinks apples are really important because when I took one of her apples, she put a spell on me!”

Awesome, now we know how to help Dora talk! But these guys need some help, don’t they? We should ask someone for help. Who should we ask? Can you ask Kermit “who should we ask?”

Child: who should we ask?

Kermit: I believe we should ask a good wizard.

E: Good! Now can you ask batman?
Child: Who should we ask?

Batman: I believe we should ask a unicorn.

E: Ok! Let’s ask Dora

Fairy: It starts like this “Who do you believe…”

Child (1): Who do you believe we should ask?

Dora: I believe we should ask a wizard because wizards are very wise.

So the friends go looking for a wizard. Soon they come to a cave. Maybe the good wizard lives here, but maybe the scar witch lives here. Who will we meet here? Can you ask Kermit "Who will we meet here?"

Child: who will we meet here?
Kermit: I believe we will meet the good wizard here.

E: Good! Now ask batman.

Child: Who will we meet here?
Batman: I believe we will meet the good wizard here.

Ok, let's ask Dora.

Fairy: It starts like this "Who do you believe…"

Child (2): Who do you believe we will meet here?

D: I believe we will meet the good wizard in the cave because wizards love caves!

They go into the cave and who do they find? The good wizard! Kermit says "Excuse me! My friend is under a spell? Can you break the spell and help us to go home?"

The wizard says "That is a very powerful spell. To break it you will need three things. Frist, a magic ruby. Second, the help of a good fairy. And third an enchanted object like a fairy wand or a dragon's egg. I like helping, so I will give you my magic ruby." And he gives the ruby to Kermit. "Then you will need to find the fairy tree in the middle of the forest where the fairies live. On the way, you will come to a place with two roads. One goes to the flower field where the fairy wands grow and one goes to the dragon mountain where the dragons lay their eggs. You will need to choose what to look for. Good luck!"

So the friends head off down the path. Then a bluebird appears! "Don't go this way," she says " The forest is on fire and you will get hurt!" But then a friendly troll appears. "Don't trust her," he says "She is really the witch in disguise!" Uh-oh! Who should we trust?? Can you ask Kermit "who should we trust?"

Child: Who should we trust?

Kermit: I believe we should trust the friendly troll.

Good, now can you ask Batman?
Child: Who should we trust?

Batman: I believe we should trust the bluebird.

Ok, let's ask Dora!

Fairy: It starts like this "who do you believe…"

Child (3): Who do you believe we should trust?

Good job! Dora says " I believe we should trust the friendly troll because he seems really nice"

When Dora says that, the bluebird reveals she was really the bad fairy in disguise and she flies off angrily! Whew!

The troll says "I am glad you trusted me! I will help you some more. In the road ahead is a big snake. He won't let you get by, but don't worry. He is easy to trick. Just play some music and he will fall asleep. I have a guitar and tambourine. I will give you one of them if you decide what you want to play."

Hm. What should we play? Can you ask Kermit "what should we play?"

Child asks.

Kermit says, "I believe we should play the guitar."

Nice, now can you ask Batman?
Child asks.

Batman says "I believe we should play the tambourine."

Hm. They disagree, but maybe Dora can break the tie! Remember, to ask Dora we work with the pink fairy.

The pink fairy says "It starts like this "What do you believe…"

Child (4): What do you believe we should play?

Dora jumps and says "I believe we should play the tambourine because tambourines jingle and snakes love jingling!"

So, the troll gives the tambourine to the Kermit. The friends thank the troll and head off down the path.

Pretty soon they come to the big snake, but Kermit plays the tambourine and the snake falls asleep and the friends sneak on by. Then they come to the place with two roads, one that goes to the flower field where the fairy wands grow and one that goes to the dragon mountain where the dragons lay their eggs. We need to choose what to look for: a fairy wand or a dragon's egg. What should we look for? Can you ask Kermit "what should we look for?"

Child: What should we look for?
Kermit: I believe we should look for a fairy wand.

Good, now can you ask Batman?
Child: What should we look for?

Batman: I believe we should look for a dragons egg.

Ok, you know what to do, let's ask Dora!

Fairy: It starts like this: What do you believe?

Child (5): What do you believe we should look for?

Good job! Dora says "I believe we should look for a dragon's egg, because that will be big and easy to find."
So the friends head off down the path toward the dragon mountain. But then a big bear appears in the middle of the road!
"Grrr!" says the bear. "I won't let you by until you give me something I want to eat!"

What does the bear want to eat? Can you ask Kermit "What does the bear want to eat?"

Child asks.

Kermit says, "I believe the bear wants to eat a cookie."

Nice, now can you ask Batman?
Child asks.

Batman says "I believe the bear wants to eat an apple."

Hm. They disagree, but maybe Dora can break the tie! Remember, to ask Dora we work with the pink fairy.

The pink fairy says "It starts like this "What do you believe…"

Child (6): **What do you believe the bear wants to eat?**

Dora jumps and says "I believe the bear wants to eat an apple because apples grow on trees and bears live in the woods!"

So now we know what to feed the bear, but someone has to feed the bear… No one wants to feed the bear… Who should feed the bear? Can you ask Kermit, "Who should feed the bear?"

Child asks.

Kermit says, "I believe Batman should feed the bear."

Nice, now can you ask Batman?
Child asks.

Batman says "I believe Kermit should feed the bear."

Hm. They disagree, but maybe Dora can break the tie! Remember, to ask Dora we work with the pink fairy.

The pink fairy says "It starts like this "Who do you believe…"

Child (7): Who do you believe should feed the bear?

Dora jumps and says “I believe Kermit should feed the bear because he is the biggest!”

So, Kermit tries his best to be brave and he takes the apple and gives it to the bear, but the bear says “Grr! I don’t like apples!” and he chases them off down the path.

So they decide, they’re going to look for a fairy wand instead. Soon they come to the flower field where the fairy wands grow. They see one really big flower with fairy wands growing at the top! They’re really high! Who can get the wand? Can you ask Kermit “who can get the wand?”

Child asks.

Kermit says, “I believe I can get the wand.”

Nice, now can you ask Batman?
Child asks.

Batman says “I believe I can get the wand.”

Hm. They disagree, but maybe Dora can break the tie! Remember, to ask Dora we work with the pink fairy.

The pink fairy says “It starts like this “Who do you believe…”

Child (8): **Who do you believe can get the wand?**

Dora jumps and says “I believe Batman can get the wand because Batman can fly!”

So Batman flies to the top of the sunflower and picks a wand and flies back down.

Now all we need is to find the fairy tree in the middle of the woods! The friends keep going till they find it. When they get there, they see two blue fairies.

Fairy: Uh-oh! Two blue fairies! There’s only one real blue fairy. One must be the witch in disguise!

Uh-oh! Who is the good fairy?? Can you ask Kermit “who is the good fairy”?

Child: Who is the good fairy?

Kermit: I believe the fairy by the mushrooms is the good fairy.

Ok! Let’s ask Batman.

Child: Who is the good fairy?

Batman: I believe fairy in the treetops is the good fairy.

Ok! Let’s ask Dora.

Fairy: The question starts “who do you believe…”

Child (9): Who do you believe is the good fairy?

Dora: I believe the fairy in the treetops is the good fairy because she is flying a lot better! Just then the fairy by the mushrooms reveals she's really the witch! She chases the good fairy and catches her in a magic rope!

"Please help me!" says the blue fairy!

What can help the fairy? Can you ask Kermit "what can help the fairy?"

Child: What can help the fairy?

Kermit: I believe the magic wand can help the fairy.

How about Batman, can you ask him?

Child: What can help the fairy?

Batman: I believe the magic ruby can help the fairy.

Let's ask Dora!

Fairy: It starts like this "What do you believe…"

Child (10): What do you believe can help the fairy?

Dora: I believe the ruby can help the fairy because the wizard said we would need it!

So, batman throws the ruby to the fairy and she uses it to become really small and escape the witch's rope. But then the witch starts flying around everywhere. We've got to stop her! What can stop the witch? Can you ask Kermit "What can stop the witch?"

Child: What can stop the witch?

Kermit: I believe a rope can stop the witch.

Good, now ask Batman!
Child: What can stop the witch ?

Batman: I believe a net can stop the witch.

Ok, let's ask Dora!
Fairy: It starts like this " What do you believe… "

Child (11): What do you believe can stop the witch?

Dora: I believe a net can stop the witch because a net covers a lot of space.

So batman throws a net over the witch and it knocks her to the ground.
"Uhg! You caught me!"

Ooh- this witch is not happy.

Kermit says "Blue Fairy, now that we helped you and stopped the witch, can you break the spell on Dora and send us home?"

But the blue fairy says “I can’t. I don’t have enough magic to break the spell and send you home and bring my fairy sister back to the fairy world! Maybe if you can make the witch happy she will help me.

Ooh. What can make the witch happy? Can you ask Kermit “what can make the witch happy?”

Child: What can make the witch happy?

Kermit: I believe a lollipop can make the witch happy.

Good! How about Batman?
Child: What can make the witch happy?

Batman: I believe a chocolate bar can make the witch happy.

Ok, let’s ask Dora.

Child (12): What do you believe can make the witch happy?

Dora: I believe a chocolate bar can make the witch happy because chocolate always makes me happy.

So batman takes a chocolate bar and gives hit to the witch and he gives her a big hug and now she feels so much better! Look at her big smile!

“Thank you for making me feel better,” says the witch. “Give the magic wand to the fairy and together we can send you home.”

So the witch and the fairy work together and send everybody home!

The End

Appendix B: Production Error Types found in Experiment 1.

This appendix presents the extended types of productions given by children in Experiment 1. Note: percentages do not sum to 100 as some production types overlap (e.g., that-trace extraction and ‘believe in’)

Production Type	Example Target:	Example of Production Error:	By - Sub means (converted to %)
Believe in	What do you believe can stop the witch?	What do you believe in that … something can stop the witch?	0.3%
Changed Extraction	What do you believe we should play?	What do you believe can make the snake fall asleep?	8.9%
Changed verb	Who do you believe can get the wand?	Who do you believe can pick the wand?	24.4%
Changed Auxiliary	Who do you believe can get the wand?	Who do you believe should get the wand	23.1%
D-linking	Who do you believe is the good fairy	Who do you believe … which fairy is the good fairy?	1.7%
Drops_Auxiliary	What do you believe we should look for?	What do you believe what we look for?	10.1%

Failed to Extract WH	Who do you believe should feed the bear?	Who do you believe that you should feed the bear?	.9%
Infinitival_clause	Who do you believe we should ask?	Who do you believe to ask to break the spell?	5.6%
Subject/aux_inversion_error (without 2Q)	What do you believe we should play?	what do you believe should we play?	0.6%
Overt Complementizer *can be grammatical or ungrammatical	What do you believe can make the witch happy?	what do you believe that can make the witch happy?	34.9%
Two Question Words	What do you believe the bear wants to eat?	What do you believe what does the bear wanna eat?	22.5%
Target-Like Production	What do you believe we should play?	What do you believe we should play?	35.9%

Appendix C: Prosodic Analysis

A full investigation of potential prosodic differences using our current data is not feasible (and is otherwise beyond the scope of the present paper). However, we can consider preliminary evidence from the single participant whose recordings were of good enough quality to investigate and who produced both questions with medial wh-phrases and questions with *that* at the embedded clause boundary. This will allow us to compare wh-questions with medial wh-phrases to wh-questions which necessarily include embedded clauses. We compared prosodic attributes of her questions which reliably included embedded clauses (*believe+that,* n=8) and questions which were ambiguous between an embedded vs. an independent wh-phrase (*believe+what,* n=3). Using PRAAT (Boersma, 2001), we examined two aspects of the child's prosody which could be relevant. First, we measured the time between the release of the [v] of the word *believe* and the beginning of the next word in each utterance. In *believe+that* questions, this time averaged .41 seconds, while the average gap in *believe+what* questions was .26 seconds. Furthermore, the range of pause duration was large across trials: .09 seconds to .43 seconds in *believe+what* questions. This dramatic range in so few utterances is unreliable at best, however it does suggest a difference between these types of utterances.

Second, we computed the duration of the [i] vowel in the word *believe*. Previous work (e.g. Cooper & Paccia-Cooper, 1980) has suggested that syntactic properties can result in lengthening of lexical items. We thus hypothesize that there could be a difference in the length of the vowel at the end of a clause compared to at the end of a question. On average, the duration of [i] in *believe+that* constructions was .35 seconds, while the duration of the vowel in *believe+what* construction was .42 seconds. Though the small number of data points means this data cannot be considered significant, we did find differences in each of these categories between this participant's *believe+that* and *believe+what* utterances, suggesting *believe+what* questions are distinct from questions with embedded clauses and therefore likely not instances of WSM. A thorough investigation of the prosody is needed in the future to confirm this merely suggestive evidence.

Appendix D: Comprehension Task Stimuli

8 stories (6 target; 2 fillers) designed to be followed by biclausal questions.

Past tense matrix verb (3 target; 1 filler)

Lewis & Sally

This is Lewis and this is Sally. They are having a party tonight with some of their friends, but they don't know what food they should serve. Lewis thinks he could pick some fruit in his new basket and they could have that for dessert. Sally thinks that's a great idea, but she's going to stay home and get the rest of the party ready. She still needs to know what Lewis picks, though, so she can decide what dessert to make. She gives Lewis a walkie-talkie and asks him to use that to tell her what he picks. Lewis says ok and heads off into the woods.

Pretty soon Lewis comes to the pond and sees a big patch of strawberries. That will make a great dessert! So Lewis picks a strawberry and then he picks another one and soon he has a whole basket full of strawberries. He takes out his walkie-talkie to tell Sally what he picked, but he drops it! It falls into the pond! Oh no! Now Lewis can't use the walkie-talkie to tell Sally what he picked! What's he gonna do?

Then he remembers that there is a phone at the ice-cream store nearby. He can use the phone to tell Sally what he picked. So Lewis heads toward town. On the way he gets hungry so he eats a strawberry… and then he eats another one… and by the time he gets to town, Look! He only has one strawberry left! They can't have one strawberry for dessert at their party! Fortunately, Lewis sees this lady selling apples. That will make a good dessert, so Lewis buys some apples. But, he's so embarrassed that he ate all the strawberries that he decides he's going to pretend he picked the apples. He puts them in his basket, so it looks like he picked them. When he gets to the phone and Sally says "Hello," Lewis is still so embarrassed he ate all the strawberries he picked that he says "Hi Sally, I picked lots of apples in the apple orchard."

"That's great!" says Sally. "I will make an apple pie. Hurry home!"

So, Lewis hurries home with his basket of apples. When he gets there, Sally has already made a pie crust, so they put the apples in crust and the pie in the oven and then they have a great party with all of their friends.

That's the end. Are you ready for your question?

How+what (experimental): **How did Lewis tell Sally what he picked?**

What+that (control): **What did Lewis tell Sally that he picked?**

Emily and Sarah

This is Emily. She is throwing a surprise birthday party for her friend Sarah. She has all the decorations ready, but she still hasn't bought a present. Emily knows that Sarah really likes sports, so she decides to go to a sports store. She thinks Sarah would like either a soccer ball or a colorful beachball. She looks for a beachball first, but look: they are sold out! There are no more beachballs, so Emily can't buy a beachball for Sarah. She decides to look some more and soon she sees a soccer ball! That's perfect. Sarah has wanted a soccer ball for a long time. So Emily buys a soccer ball for Sarah and now she is all ready for the birthday party.

But the party is not till Saturday and today is only Thursday. That means tomorrow is Friday and both girls have to go to school. There's no talking allowed in school, but before school starts Sarah comes over to Emily and whispers "I saw you at the sports store yesterday! What did you buy?"

Uh-oh! Emily can't tell Sarah what she bought, that will ruin the birthday surprise. She decides to tell Sarah she bought something else. She starts to whisper, but the teacher says "Girls! Go back to your desks! No talking in school."

So the girls go back to their desks, but Emily still needs to answer Sarah's question. She decides to write her a note. She takes a piece of paper and on it she writes:

"Hi Sarah, I bought a colorful beachball! Come over to my house on Saturday and I will show you."

Then, while the teacher is helping another kid, Emily sneaks her note over to Sarah and then she sneaks back to her desk. Sarah reads the and nods that she can come over on Saturday. Perfect!

Saturday comes and the girls' friends all go over to Emily's house and get the food and decorations ready and when Sarah walks through the door they shout "Surprise! Happy Birthday Sarah!"

Sarah is *so* surprised. She was not expecting that at all. The friends have a great party and Sarah loves all her new presents. It was a great surprise birthday party.

That's the end. Are you ready for your question?

How+what (experimental): **How did Emily tell Sarah what she bought?**

What+that (control): **What did Emily tell Sarah that she bought?**

Jane and Jeremy

This is Jane and this is Jeremy and they are on a field trip with their school. Their teacher wants them to take a picture of a rainforest bird. Jeremy says that his favorite bird in the world is a toucan. He's really excited that he might get to see one. Jane says she would also like to see a toucan, but a parrot would also be cool. She doesn't really care. Jeremy says "ok, why don't you go down by the bananas and I'll go down by the river. We can use these special rainforest radios to tell each other what we see. You can let me know if you see a toucan. So, the friends go in different directions looking for rainforest birds.

Jane goes down by the banana grove and when she gets there, look what she sees! A beautiful toucan. It's so pretty and Jane is so excited that she shouts "WOW!". But when she shouts, it scares the toucan and he flies away. Jane tries to go after him to see where he lands, but he is too fast and gets away. Oh no! Jane scared away the toucan before Jeremy could see it! She decides she won't tell him what she saw. She doesn't want to hurt his feelings.

But just then, the radio buzzes! It's Jeremy. He says "Jane! I heard you shout! What did you see?"

Jane really doesn't want to tell Jeremy she saw a toucan and scared it away, so she decides she'll tell him she saw something else. She says "I saw…" but Jeremy interrupts and says "Wait, these radios are too loud. We are scaring the birds. Let me come to the banana grove so we can whisper.

When Jeremy gets to the banana grove he leans in and whispers "Remember we still need a picture of a rainforest bird, so whisper! What did you see?"

Jane really doesn't want to hurt Jeremy's feelings, so she decides to tell him she saw something else. She leans in and whispers, "I saw a parrot. Let's keep looking for toucans."

The friends keep looking for toucans until it is time to leave. They go back to the bus and look what they see! A beautiful toucan is waiting for them! Jeremy takes a picture and even though the toucan flies away, he will always have a picture to remember seeing his favorite bird.

That's the end! Are you ready for your question?

How+what (experimental): **How did Jane tell Jeremy what she saw?**

What+that (control): **What did Jane tell Jeremy that she saw?**

Past tense filler:

Elisha

This is Elisha and her favorite thing in the world are dinosaurs! She needs to write a report for school, so she decides she'll do it about dinosaurs. Now, it would be best if she could see a real dinosaur, but there are no more dinosaurs anymore. She decides her next best option will be to talk to her friend, Professor Siders, and his robot, Clank, to see if they know anything about dinosaurs.

She asks the professor, but he says he doesn't know much about dinosaurs, but he *did* just finish making a time machine. They can go back in time to see some dinosaurs! So they jump in the time machine and go whirring back in time to the time of dinosaurs.

When they get there and they jump out of the machine, Clank, Elisha, and Professor Siders realize there is so much to see, they will need to split up to explore more. Elisha says "Clank, you go down by the field of strange trees with weird fruit. The professor and I will go down by the pond and later, we'll tell each other what we see." So the friends go in different directions looking for dinosaurs.

When Elisha and Professor Siders get down to the pond look what they see! A nest of dinosaur eggs!

"What kind of dinosaur made those eggs?" asks Elisha. "And what does that dinosaur eat?"

"A brontosaurus made those eggs," says professor Siders. "But I don't know what they eat."

"I bet a brontosaurus would eat either leaves or fruit or maybe both," says Elisha.

Just then, they hear Clank the robot shout "You guys, come quick! There's a brontosaurus! And he's eating!" So, Elisha and professor Siders run to the field of strange trees with the weird fruit. When they get there they see Clank the robot, and a *big* brontosaurus.

Clank says “You guys missed it, but this brontosaurus ate LOTS of leaves.”

Just as he finishes talking, the volcano erupts! Quick everyone run back to the time machine!

The friends run back to the time machine and jump inside- Clank, Elisha, and professor Siders – and they go whirring back to the present, just in time. When they get home, everyone agrees it was amazing to see a brontosaurus, but Elisha is the most excited because now she can write the best report ever!

That’s the end! Are you ready for your question?

Experimental: **Who told Elisha what the brontosaurus ate?**

Control: **Who told Elisha that the brontosaurus ate leaves?**

Future tense matrix verb (3 target ; 1 filler)

Evil Steve and Detective Sherry

This is Detective Sherry and she has solved many mysteries against her rival, Evil Steve, but she has never caught him. Today, the Queen’s crown is on display at the museum and Sherry is pretty sure that Steve will try to steal it.

Meanwhile, over at Steve’s office, he *is* thinking of ways to steal the crown! But, he knows that Sherry is really smart and will expect that. He decides he’s going to trick her by telling her he will steal something else. But how can he tell Sherry something without getting caught? Steve has an idea. He has just invented a TV-machine that can put him on people’s televisions even if their TV isn’t on. He will use the TV machine to tell Sherry what he will steal.

Pretty soon, over at Sherry’s office, they TV turns on! There’s Evil Steve on the TV. He laughs and says “Foolish Detective Sherry—” but it cuts off and the TV goes black again. “What was Evil Steve doing on my television!?” thinks Detective Sherry.

Meanwhile, at Steve’s office, you can see his TV machine is broken. There is smoke everywhere and wires coming out of it. Steve can’t use the TV machine to tell sherry anything. He decides he will write her a letter. He’ll deliver it himself and watch from the window to be sure she reads it.

Pretty soon, a letter arrives at Sherry’s door. She’s looking at the letter, so she doesn’t see Steve watching at the window! She opens the letter and on it she reads “Foolish Detective Sherry, I am going to steal the Queen’s diamond ring and there’s nothing *you* can do to stop me! Hahaha! Evil Steve”

“Oh no! I’ve got to protect the ring!” Shouts Sherry and she runs to the Queen’s palace. When she gets there, though, the guard says “no, the queen’s ring is fine. It’s here on her dresser where it belongs.”

“Steve must have been tricking me!” thinks Detective Sherry. “I bet he told me this to distract me from the crown!” So she runs to the museum where the crown is on display.

When she gets there, she meets a guard and he says “The queens crown is fine! It’s inside on a pedestal right where it belongs. Go see for yourself!” So Sherry goes inside, but when she get’s to the pedestal where the crown should be, there’s nothing there! “That guard looked mighty familiar,” thinks Sherry “and what was in that bag he was carrying? I bet that was Evil Steve in disguise!” So, she runs back outside and stops the guard. She makes him take off his hat and empty his bag. In it, she finds the Queen’s crown and Evil Steve’s outfit and mask.

“I caught you Evil Steve,” says Detective Sherry. “You are going to jail for a long time.”

That’s the end! Are you ready for your question?

How+what (experimental): **How did Evil Steve tell Detective Sherry what he was gonna steal?**

What+that (control): **What did Evil Steve tell Detective Sherry that he was gonna steal?**

Grandma and Lizzie

This is Lizzie. She is having a game night tonight with her friends. Lizzie’s grandma comes into her room and says “I’m going to the store to buy some food so I can make something tonight for you and your friends.”

“Thanks, Grandma. That’s so nice!” Says Lizzie. And Grandma leaves, but then Grandma realizes she doesn’t know *what* grandma is gonna make. She *hopes* Grandma will make yummy, yummy chocolate cake, but Grandma really likes to make yucky, yucky broccoli casserole. Grandma is not good at making broccoli casserole, but Lizzie has never told her that because she doesn’t want to hurt her feelings. She wants to know if that’s what Grandma will make tonight, so that she can warn her friends.

Lizzie runs to the bus stop, but when she gets there, Grandma is already on the bus. Lizzie shouts, “Grandma! What are you gonna make tonight?” On the bus, Grandma could hear Lizzie’s question, but she can’t shout on the bus, that’s not nice, so Grandma decides she will draw a picture and hold it in the window to tell Lizzie what she’s going to make. She starts to draw, but before she can finish, the bus drives away, so Grandma can’t tell Lizzie that way.

Lizzie is disappointed. She really wants to know what Grandma will make! But then her phone beeps. It’s a text message from Grandma it says “Hi Lizzie, I heard your question. I am going to make broccoli casserole tonight. Love Grandma”

Oh no- Lizzie is disappointed they won’t have chocolate cake, but at least she can warn her friends. When her friends come over and Grandma is in the kitchen, Lizzie says “Guys, my grandma is going to make broccoli casserole and it's going to be yucky, but we have to pretend we like it so we don't hurt her feelings." Lizzie's friends agree. They don't want to hurt Grandma's feelings. But then in comes Grandma with chocolate cake! She just wanted it to be a surprise. The friends eat Grandma's yummy, yummy chocolate cake and then finish their game while Grandma cleans up. It was a great game night.

That's the end! Are you ready for your question?

How+what (experimental): **How did Grandma tell Lizzie what was gonna make?**

What+that (control): **What did Grandma tell Lizzie that she was gonna make?**

Tom and Katie

This is Tom and this is Katie and they are at the beach playing a pirate game! They are getting instructions from a real, live pirate, Pirate Jack. Pirate Jack says that he has hidden somewhere on the island *either* a pirate hat *or* a pirate hook. Whoever finds one of those things will get a pirate prize. He says there is a clue hidden somewhere near a pile of boulders. Whoever finds the clue will probably win. Tom and Katie run in different directions looking for a pile of boulders.

Tom comes to a pile of boulders first. He sees a piece of paper. It's a map! It shows a pirate hat next to two palm trees. So now Tom knows what the winner will find and where to find it! They'll find a pirate hat next to two palm trees. But Tom needs to be sure Katie doesn't get there first. He decides to leave her a fake clue. He takes his finger and writes in the sand "The winner will find a pirate hook next to the boats," but a big wave comes and washes Tom's message away. Tom can't tell Katie that way. He looks around and finds a bottle. Tom can leave Katie a message in a bottle! That's a great pirate clue! He takes piece of paper and on it he writes "the winner will find a pirate hook next to the boats," and he leaves a message in a bottle for Katie.

Tom sees Katie coming and hides behind the boulders. Katie picks up the bottle and she thinks it's a real clue, so she runs off in the direction of the boats. Tom is happy his trick worked and *he* runs off in the direction of the palm trees. When he gets there, look what he finds! A pirate hat! He picks it up and takes it back to Pirate Jack. Pirate Jack puts the pirate hat on his head and gives Tom a gold coin for winning the pirate game.

That's the end! Are you ready for your question?

How+what (experimental): **How did Tom tell Katie what the winner would find?**

What+that (control): **What did Tom tell Katie that the winner would find?**

Future tense filler:

Alex

This is Alex. In the forest near Alex's house, there's a haunted house! In the haunted house live lots of ghosts and mummies. Alex's friends think the ghosts and mummies must be mean and scary, but Alex is sure they must be nice and friendly. He thinks if he could meet a ghost or a mummy and introduce them to his friends everyone would know that ghosts and mummies are nice, so Alex heads off into the woods to find the haunted house.

On the way, Alex meets a farmer. He asks the farmer "Do you know where the haunted house is? Do you know what I will meet there?"

The farmer says, "I don't know what you will meet there because I can't see the future, but I do know the haunted house is on a tall hill in the middle of the woods. If you really want to know what you will meet there, you should ask the witch, Esmerelda. She has a tall pointy hat and long green hair. She lives in a clearing nearby. Sometimes she can see the future in her crystal ball. Go find her! Good luck!"

Alex says thank you to the farmer and goes to the clearing. He sees the witch, Esmerelda, with her tall pointy hat and her long green hair and he asks her "What will I meet in the haunted house?" Esmerelda looks deep into her crystal ball and says "Aha! You will meet a mummy in the haunted house!" Alex is excited about that and hurries to the tall hill. When he goes inside, he doesn't see anything… but then a ghost appears! But he flies away before Alex could meet him. Alex is disappointed, but then in comes a friendly mummy!

"I am excited to meet you," says the mummy. "My friend the ghost is very shy, but I love making new friends."

"I'm excited to meet you too," says Alex. "Would you like to meet my friends? I want to prove to them that ghosts and mummies are nice."

"Of course, I would love to meet your friends," says the mummy. "Bring them to the font yard of my haunted house."

So the next day, Alex brings his friends to the front yard of the haunted house and they all get to meet the friendly mummy and now everybody knows that ghosts and mummies are nice.

That's the end! Are you ready for your question?

***Experimental:* Who did told Alex what he would meet?**

***Control:* Who told Alex that he would meet a mummy?**

Appendix E: Production Error Types found in Experiment 3

This appendix presents the extended types of productions given in Experiment 3. Again, these do not sum to 100% because a single utterance can contain multiple errors.

Answer Type	Example Target:	Example of Production Error:	Percentage
Believe in	What do you believe can stop the witch?	What do you believe in that … something can stop the witch?	0%
Only ever asked a monoclausal question (only accepted after multiple attempts at biclausal)	Who do you believe is the good fairy?	Who do you believe is right?	11.2%
Follow up questions	Who do you believe we can trust?	Who do you believe? The bird or the fairy?	5.4%
Changed Extraction	What do you believe we should play?	What do you believe can make the snake fall asleep?	5.4%
Changed embedded clause verb	Who do you believe can get the wand?	Who do you believe can pick the wand?	22.5%
Changed matrix clause verb	What do you believe will make the witch happy?	What do you think that will make the witch happy?	1.2%
Changed Auxiliary	Who do you believe can get the wand?	Who do you believe should get the wand	13.3%

D-linking	Who do you believe is the good fairy	Who do you believe … which fairy is the good fairy?	.8%
Drops Auxiliary	What do you believe we should look for?	What do you believe what we look for?	3.3%
Failed to Extract WH	Who do you believe should feed the bear?	Who do you believe that you should feed the bear?	2.5%
Infinitival clause	Who do you believe we should ask?	Who do you believe to ask to break the spell?	6.25%
Subject/auxiliary inversion error (without 2Q)	What do you believe we should play?	what do you believe should we play?	2.08%
Overt Complementizer *can be grammatical or ungrammatical	What do you believe can make the witch happy?	what do you believe that can make the witch happy?	11.67%
Two Question Words	What do you believe the bear wants to eat?	What do you believe what does the bear wanna eat?	15.41%
Target-Like Production	What do you believe we should play?	What do you believe we should play?	46.67%

Appendix F: Production task story with no lead-in

This is the story associated with Experiment 4 designed to elicit biclausal questions from children without the use of a lead-in. Target structures are indicated in bold. Children's responses are hypothetical. Regardless of the idea they present when they choose an option, Mindy disagrees, thus ensuring we need to ask about Mork's ideas.

These are two friends going on an adventure! This is an alien, but this is a racoon! She has been visiting other planets in her spaceship and now she wants to earth to her friend. Uh-oh! On the way something goes wrong and the ship crashes! (The racoon puppet is removed from behind the computer)

E(xperimenter): Can you say hello and ask her name?

C(hild): Hello! What's your name?

Mindy: Hello! My name is Mindy! This is my friend, but he doesn't speak English. When you want to talk to him, tell me what to find out and I'll ask him.

E: Let's ask about his name can you say "What's his name?"

C: What's his name?

Mindy: I'll ask what he wants us to call him (alien exchange). He says you can call him Mork!

E: Let's get to know each other. Do you have a favorite color?

C: responds

E: Nice! Why don't you ask Mindy "What's your favorite color?"

C: What's your favorite color?

Mindy: My favorite color is pink!

E: Now let's find out about Mork- can you ask Mindy about Mork?

C: What’s his favorite color?

Mindy: I’ll ask him! (alien exchange) He says his favorite color is brown.

E: Nice! Why don’t we find out what they want to see on earth. Can you ask Mindy “What do you want to see?”

C: What do you want to see?

Mindy: I want to see a restaurant.

E: Let’s find out about Mork! Can you ask Mindy “What does Mork want to see?”

C: What does Mork want to see?

Mindy: I’ll ask him! (alien exchange) He says he’d like to see some fish! Our spaceship broke into lots of pieces. Can you help us look?

E: I think we can, don’t you? Let’s see, where should we look? There’s a pond over here… and a forest over here… Do you have an idea?

C: The forest!

E: The forest is a good idea. Let’s find out about Mindy. Can you ask Mindy “Where should we look?”

C: Where should we look?

Mindy: I think we should look by the pond.

E: Ok, you have two different ideas. Let’s find out about Mork. Can you ask Mindy about Mork? (if needed) Can you say “Where does Mork think we should look?”

C: Where does Mork think we should look?

Mindy: I’ll ask him! (alien exchange) He says he thinks we should look by the forest!

E: So they go to the forest and look what they find! The nose of the ship! Can you place that on your ship? Good! They keep going and they come across a farmer.

Mindy: Excuse me! We are looking for the pieces of our ship have you seen something that might be part of it?

Farmer Bob: I did see something that looked like a wing! I fell over there. Either by Farmer John's Farm or by Dr. Mary's hostpital.

E: Hm… who has the wings? Do you have an idea?
C: Dr. Mary!

E: That's a great idea! Let's find out about Mindy- can you ask her "Who has the wings?"

C: Who has the wings?

Mindy: I think Farmer John has the wings.

E: ok- you have two different ideas, so let's find out about Mork. Can you ask Mindy?

C: Who does Mork think has the wings?

Mindy: I'll ask him! (alien exchange) He says he thinks Farmer John has the wings.

E: So they go to Farmer John's farm and look what they see! A wing! Stick that on your spaceship. Good job! But look! The wing broke Farmer Johns fence when it fell! We should fix it! Who can fix the fence? We know Dr. Mary lives close by and we just met Farmer Bob, so he's not far off either. Do you have an idea?

C: I think Farmer Bob.

E: Good idea! Let's find out about Mindy can you ask her "who can fix the fence?"

C: Who can fix the fence?

Mindy: I think Dr. Mary can fix the fence.

E: Ok, let's find out about Mork. Can you ask Mindy?

C: Who does Mork think can fix the fence?

Mindy: I'll ask him! (alien exchange) He thinks Farmer Bob can fix the fence.

E: So they go find Farmer Bob again. When they get there, he says "I found this map in the road! Is it yours?" It sure is! Put that in the spaceship! Farmer Bob, we need help fixing Farmer John's fence. Can you help?

Farmer Bob: I can, but I will need a hammar and I don’t have one. Teacher Ben might have a hammer… and so might Jenn the mechanic…

E: Hm. Who has a hammer? Do you have an idea?

C: I think Jenn the mechanic.

E: Good idea! Can you ask Mindy “who has a hammer?”

C: Who has a hammer?

Mindy: I think Teacher Ben has a hammer.

E: Ok, let’s find out about Mork. Can you find out about Mork?

C: Mindy, Who does he think has a hammer?

Mindy: I’ll ask him! (alien exchange) He thinks Jenn the mechanic will have a hammer.

E: So they go find Jenn the mechanic. She has a hammer and also says she has found a strange chair that doesn’t go in any of the cars she is fixing! It belongs in the spaceship! Jenn says to come back if we need more help because she can fix anything. The friends take the hammer back to Farmer John’s farm. They give the hammer to Farmer Bob and he uses it to fix the fence.

Mindy: Farmer Bob, Farmer John. Have you seen anything else that might be part of our spaceship?

Farmer John: Yes, I saw something that looked like an engine fall over there. Either by the gazebo or by the pond.

E: Hm where is the engine… do you have an idea?
C: The pond!

E: Good idea- let’s ask Mindy “Where is the engine?”
C: Where is the engine?
Mindy: I think the engine is by the gazebo.

E: Hm. Let's find out about Mork can you ask Mindy "Where does Mork think the engine is?"
C: Where does Mork think the engine is? (Filler 1)

Mindy: I'll ask him! (alien exchange) He says he thinks it's over by the pond.

E: So, the friends go look. They find the engine, but it's badly broken. Who should we ask for help? Farmer Bob fixed the fence… maybe we should ask him… but then Jenn the mechanic said she could fix anything… Do you have an idea?

C: Jenn!

E: good idea. How about Mindy? Can you ask her "Who should we ask"?

C: Who should we ask?

Mindy: I think we should ask Farmer Bob.

E: Ok! Let's find out about Mork!
C: Who does Mork think we should ask?

Mindy: I'll ask him! (alien exchange) He thinks we should ask Jenn the mechanic.

E: So the friends return to Jenn's shop. When they get there they ask her if she can help fix the engine.

Jenn: I sure can, but I'm really busy. Maybe you can help me! Teacher Ben needs a new steering wheel and Farmer Bob needs four new tires. If you can help one of them, I can work on the engine.

E: Who should we help? Teacher Ben or Farmer Bob? Do you have an idea?

C: I think Teacher Ben.

E: Great idea. Can you ask Mindy "Who should we help"?

C: Who should we help

Mindy: I think we should help Farmer Bob.

E: Ok! How about Mork?

C: Who does Mork think we should help?

Mindy: I'll ask him! (alien exchange) He thinks we should help farmer Bob.

E: They go find Farmer Bob and he says he needs four new tires. Do you see four new tires on the screen? (Child points) Good job! But we can't put the tires on. We better give them to someone who can… Hm. Who should we give the tires to? There's Jenn, but she's really busy… and there's also her assistant, Rosie. Do you have an idea?

C: I think Rosie!

E: Good! How about Mindy. Can you ask her "who should we give the tires to?"

C: Who should we give the tires to?

Mindy: I think we should give the tires to Jenn the mechanic.

E: Ok! Let's ask Mindy about Mork. Can you ask her?

C: Who does he think we should give the tires to?

Mindy: I'll ask him! (alien exchange) He thinks we should give the tires to Rosie.

E: So they give the tires to Rosie who says she'll put them on the tractor, but also that she found another wing of the spaceship! Here you go! Jenn the mechanic comes and says "Guys, that engine will take a long time to fix. You'd better go explore a bit while I fix it" The friends decide they should see the things they want to on earth! Remember Mindy wanted to see a restaurant. Let's go eat lunch. The friends head down the path when they meet Dr. Mary and Sue the tennis player. Let's invite one of them to lunch. Who should we invite?

C: Dr. Mary

E: Very nice. Let's find out about Mindy. Ask her "who should we invite".

C: Who should we invite?

Mindy: I think we should invite Sue the tennis player.

E: Ok- let's find out about Mork!

C: Who does he think we should invite?

Mindy: I'll ask him! (alien exchange) He thinks we should invite Dr. Mary.

Dr. Mary: There are two places to eat. The diner and the café. Where should we eat?

E: Do you have an idea?
C: The café!

E: Let's ask Mindy. Can you ask "Where should we eat"?

C: Where should we eat?

Mindy: I think we should eat at the diner.

E: Ok, let's find out about Mork. Can you ask Mindy?

C: Where does Mork think we should eat?

Mindy: I'll ask him! (alien exchange) He thinks we should eat at the diner.

Dr. Mary: ok, but I'm not totally sure how to get there from here. Teacher Ben is nearby and so is Jeremy the student. Who knows the way?

C: I think Teacher Ben!
E: Good, let's find out about Mindy. Can you ask "who knows the way?"

C: Who knows the way?
Mindy: I think Jeremy the student knows the way.

E: Ok, let's ask about Mork. Can you ask Mindy?

C: Who does he think knows the way?

Mindy: I'll ask him! (alien exchange) He thinks Teacher Ben knows the way.

E: So they go to the school yard and ask Teacher Ben how to get to the diner. He says it's down this road by the super market! Let's go. So they go to the Diner and when they get there, they see Farmer Bob and Farmer John eating lunch. Farmer Bob says "Good choice coming here! You should get the burger. It's the best" Farmer John says "No, get the macaroni! That's the best". Hm. Who should we listen to? Do you have an idea?

C: I think Farmer John!

E: Ok, let's find out about Mindy. Can you ask "who should we listen to?"

C: Who should we listen to?

Mindy: I think we should listen to Farmer Bob.

E: Ok, let's find out about Mork. Can you ask Mindy about Mork?

C: Who does he think we should listen to?

Mindy: I'll ask him! (alien exchange) He thinks we should listen to Farmer Bob.

E: So the friends get a burger and share it. Now they have seen a restaurant. Do you remember what Mork wanted to see? Fish! Let's go to the aquarium. When they get there, they realize they only have time to visit one exhibit: the jelly fish or the sharks. Who should we see?

C: The sharks!

E: Good idea. Let's find out about Mindy. Ask "Who should we see?"

C: Who should we see?

Mindy: I think we should see the jellyfish.

E: Ok, let's ask about Mork!

C: Who does Mork think we should see?

Mindy: I'll ask him! (alien exchange) He says he thinks we should see the sharks.

E: Ok- let's take a tour. There are two tour guides. One with a shark hat and one with a fish hat. Who can show us the sharks?

C: The one with the shark hat!

E: Ok, let's ask Mindy! "Who can show us the sharks"

C: Who can show us the sharks?

Mindy: I think the one in the fish hat can show us the sharks.

E: Ok, let's ask about Mork!

C: Who does he think can show us the sharks?

Mindy: I'll ask him! (alien exchange) He thinks the one with the shark hat can show us the sharks.

E: So they go with the guy in the shark hat and have a great visit looking at the sharks. So cool! On their way out of the aquarium, the see a guy selling shirts with turtles and a girl selling shirts with dolphins. Who has a better shirt?

C: The girl!

E: Let's find out about Mindy. Can you ask "Who has a better shirt"?

C: who has a better shirt?

Mindy: I think the boy has a better t-shirt.
E: Ok, let's find out about Mork.

C: Who does he think has a better t-shirt?

Mindy: I'll ask him! (alien exchange) Mork says he thinks the boy has the better t-shirt.

E: So Mork buys a t-shirt from the aquarium to remember his visit. The friends go back to Jenn's shop and when they get there she says "Here's the engine! It's all fixed!" Great! Take that and put it inside the ship! Now look, we can put the ship in the computer! It's all fixed (places the ship behind the computer as it appears on screen). The friends jump inside and off they go. Thank you for helping them continue their adventures!

Appendix G: Production and Comprehension Tasks in German

Production with lead-in

Produktionsaufgabe Geschichte mit Einführung

Den Kindern werden Batman, Kermit und Dora vorgestellt, die „in" den Computer auf ein Abenteuer gehen. Sobald sie im Computer angekommen sind, werden sie von der pinken Fee begrüßt.

Fee: Willkommen im Feenreich! Hier ist alles sehr gerecht und wir machen immer das, was zwei von euch drei machen wollen. Macht das Sinn? –Kind reagiert- Was möchtet ihr als Erstes machen? Kannst du[51] alle fragen, was sie machen wollen? Fangen wir mit Kermit an. Kannst du sagen „Kermit, was möchtest du machen?“?

K: Kermit, was möchtest du machen?

Fee: Sehr gut!

Kermit: Ich möchte ein paar Blumen pflücken.

Fee: Was ist mit Batman? Kannst du ihn fragen?

K: Was möchtest du machen?

Fee: Super!

Batman: Ich möchte einen Apfel pflücken.

Fee: Achtung! Hier wohnt eine böse Hexe und ihr gehört der Apfelbaum. Wenn du einen von ihren Äpfeln pflückst, könnte sie sauer werden! Hm… Da wir das jetzt wissen, fragen wir Dora, was sie machen will. Kannst du Dora fragen?

K: Was möchtest du machen?

Dora: Ich habe keine Angst vor der Hexe. Ich möchte einen Apfel pflücken.

Fee: Da zwei von drei Leuten einen Apfel pflücken möchten, haben sie beschlossen, einen Apfel zu pflücken. Aber schau mal! Sobald Dora den Apfel gepflückt hat, erscheint die Hexe!

Hexe (zur Fee): Wie kannst du es wagen, diesen Leuten von meinem Apfelbaum zu erzählen? Als Strafe sende ich dich in die echte Welt. Poof! –die Fee erscheint als Puppe-

Fee: Oh nein! Was werden deine Freunde nur ohne meine Hilfe machen? Schauen wir mal und finden es heraus.

Hexe: Dora, du hast einen meiner Äpfel gepflückt, dafür musst du bestraft werden! Ich werde dich verfluchen. Poof! Versuch doch mal zu reden.“

Dora murmelt und die Hexe lacht.

Hexe: Haha! Es gibt zwei Regeln für meinen Fluch. Die erste Regel ist, daß Dora nur reden kann, wenn du (das Kind) ihr eine Frage stellst. Die zweite Regel ist, daß du (das Kind) und die pinke Fee zusammen die Frage formulieren müsst.

E: Und dann fliegt die böse Hexe davon. –zum Kind- Das ist sehr kompliziert! Dora kann nur reden, wenn du ihr eine Frage stellst, aber du musst zusammen mit der Fee die Frage erstellen. Ich frage mich, was das heißt… Ich hab‘ eine Idee! Wie wär’s, wenn die Fee den Anfang von der Frage sagt, und da du ein Mikrofon hast, kannst du die ganze Frage stellen. –Zeitpunkt für weitere Erklärung, falls nötig- Lass uns eine Frage ausprobieren, um herauszufinden, ob die Hexe es so gemeint hat. Wie wär’s damit: Ich weiß zufällig, daß jemand in der Feenwelt denkt, daß grün die beste Farbe ist. Lass uns herausfinden, wer das

[51] We note that German would typically require *Du* with a capital ‘D’, however, the native-speaking German RAs said that lower-case *du* was more natural here because they were reading the text to a child.

ist! Lass uns Dora fragen „Wer denkt, daß grün die beste Farbe ist?“. Also du lässt die Fee jetzt den Anfang der Frage stellen und du sagst dann die ganze Frage, okay?

Fee: Die Frage beginnt so: Wer glaubt,…“

Kind: Wer glaubt, daß grün die beste Farbe ist? –dem Kind wird geholfen, falls nötig-

Dora: Kermit der Frosch denkt, daß grün die beste Farbe ist, weil er selber grün ist.

E: Super! Das hat funktioniert. Lass uns das noch einmal versuchen. Dieses Mal werden wir alle fragen, wer denkt, daß Äpfel sehr wichtig sind. Kannst du Kermit fragen „Wer denkt, daß Äpfel sehr wichtig sind?“?

K: Wer glaubt, daß Äpfel sehr wichtig sind?

Kermit: Ich glaube, die Hexe glaubt, daß Äpfel sehr wichtig sind.

E: Super! Kannst du jetzt Batman das Gleiche fragen?

K: Wer glaubt, daß Äpfel sehr wichtig sind?

Batman: Ich glaube, daß die pinke Fee glaubt, daß Äpfel sehr wichtig sind.

E: Okay, sie sind sich nicht einig. Lass uns Dora fragen. Denk daran, daß wenn du Dora etwas fragst, du mit der pinken Fee zusammenarbeiten musst.

Fee: Es beginnt so: Wer glaubt,…

Kind: Wer glaubt, daß Äpfel wichtig sind?

Dora: Ich glaube, daß die Hexe denkt, daß Äpfel wichtig sind, weil sie mich verflucht hat, als ich einen ihrer Äpfel genommen habe.

E: Sehr gut. Jetzt wissen wir, wie wir Dora reden helfen können. Aber diese Leute brauchen ein bisschen Hilfe, oder nicht? Schauen wir mal, ob wir jemanden finden, der uns helfen kann. Die drei Freunde gehen den Weg entlang. Aber sie kommen nicht weit, weil sie auf eine blaue Fee treffen, die in einem Käfig gefangen ist. Ihre Augen sind verbunden und sie kann nichts sehen.

Blaue Fee: Bitte helft mir! Die Hexe hat mich verflucht und ich kann nicht entkommen. Ich brauche ein magisches Objekt, um zu entkommen. Es ist etwas kleines, das glänzt.

E: Hm. Siehst du etwas, das klein ist und glänzt? –Kind darf suchen- Gut gemacht! Kermit sieht sich um und findet einen Rubin. Das ist klein und glänzend, also hebt er es auf. Batman findet ein Fläschchen mit Feenstaub. Das ist auch klein und glänzend, also hebt er es auf. Was kann der blauen Fee helfen? Kannst du Kermit fragen „Was kann der Fee helfen?“?

K: Was kann der Fee hFeen?

Kermit: Ich glaube, der magische Rubin kann der Fee helfen.

E: Gut, kannst du jetzt Batman fragen?

K: Was kann der Fee helfen?

Batman: Ich glaube, der Feenstaub kann der Fee helfen.

E: Hm, sie sind sich nicht einig. Aber vielleicht weiß es ja Dora. Denk daran, daß du zusammen mit der Fee die Fragen stellen musst.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K(target 1): Was glaubst du, was der Fee helfen kann?

Dora: Ich glaube, der Feenstaub kann der Fee helfen, weil Feen Feenstaub für ihre Magie benutzen.

E: Also nimmt Batman den Feenstaub und gibt ihn der blauen Fee. Sie nimmt den Feenstaub und lässt damit die Augenbinde und den Käfig verschwinden.

Blaue Fee: Vielen Dank! Wie kann ich euch helfen?

Kermit: Kannst du den Fluch von Dora aufheben und uns nach Hause bringen?

Blaue Fee: Das kann ich nicht, der Fluch ist zu stark. Nur die böse Hexe oder der gute Zauberer können diesen Fluch brechen. Ich glaube, du suchst besser nach dem guten Zauberer.

E: Also bedanken sich die Freunde bei der blauen Fee und gehen den Weg weiter. Aber plötzlich erscheint die böse Hexe! Sie fliegt überall herum und lässt sie nicht weiter. Wir müssen sie aufhalten! Was kann die Hexe stoppen? Kannst du Kermit fragen „Was kann die Hexe stoppen?“?

K: Was kann die Hexe stoppen?

Kermit: Ich glaube, ein Seil kann die Hexe aufhalten.

E: Gut, kannst du jetzt Batman fragen?

K: Was kann die Hexe stoppen?

Batman: Ich glaube, ein Netz kann die Hexe aufhalten.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 2): Was glaubst du, was die Hexe aufhalten kann?

Dora: Ich glaube, ein Netz kann die Hexe aufhalten, weil ein Netz sehr groß ist.

E: Also schmeißt Batman ein Netz über die Hexe und hält sie damit fest. Die Hexe fällt auf den Boden und ihr Besen zerbricht.

Hexe: Ugh! Du hast mich geschnappt und meinen Besen kaputt gemacht.

E: Oh nein, sie ist sehr sauer. Ich frage mich, was sie wieder glücklich macht. Kannst du Kermit fragen „Was kann die Hexe glücklich machen?“?

K: Was kann die Hexe glücklich machen?

Kermit: Ich glaube, ein Lollipop macht die Hexe wieder glücklich.

E: Gut, kannst du jetzt Batman fragen?

K: Was kann die Hexe glücklich machen?

Batman: Ich glaube, eine Tafel Schokolade macht die Hexe wieder glücklich.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 3): Was glaubst du, was die Hexe glücklich machen kann?

Dora: Ich glaube, eine Tafel Schokolade macht die Hexe glücklich, weil Schokolade mich immer glücklich macht.

E: Also nimmt Batman eine Tafel Schokolade und gibt sie der Hexe und gibt ihr einen Kuss auf die Wange, damit sie sich besonders fühlt. Die Hexe isst die Schokolade und schau mal! Sie hat ein riesengroßes Lächeln auf dem Gesicht.

Hexe: Danke, jetzt fühle ich mich besser.

Kermit: Da du jetzt glücklich bist, kannst du Doras Fluch aufheben und uns nach Hause schicken?

Hexe: Das kann ich nicht machen. Ihr habt meinen Besen zerbrochen und das hat meine Magie zerstört. Ihr braucht die Hilfe von der guten Fee. Damit sie euch helfen kann, braucht die Fee einen Zauberstab. Ihr müsst zum Blumenfeld gehen, wo die Zauberstäbe wachsen und dort müsst ihr einen pflücken. Aber seid vorsichtig! Es gibt gute, aber auch böse Feen, die euch vielleicht austricksen wollen, damit ihr einen Zauberstab klaut. Viel Glück.

E: Die Freunde bedanken sich bei der Hexe und gehen den Weg weiter, bis sie zu dem Blumenfeld kommen, wo die Zauberstäbe wachsen. Sie sehen die Zauberstäbe, aber sie sind ganz weit oben, an der Spitze einer hohen Sonnenblume. Wer kann den Zauberstab holen? Kannst du Kermit fragen „Wer kann den Zauberstab holen?“?

K: Wer kann den Zauberstab holen?

Kermit: Ich glaube, ich kann den Zauberstab holen.

E: Gut, kannst du jetzt Batman fragen?

K: Wer kann den Zauberstab holen?

Batman: Ich glaube, ich kann den Zauberstab holen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K(target 4): Was glaubst du, wer den Zauberstab holen kann?

Dora: Ich glaube, Batman kann den Zauberstab holen, weil er fliegen kann.

E: Also fliegt Batman zu der Spitze der Sonnenblume und pflückt den Zauberstab und fliegt wieder zurück. Genau in dem Moment, in dem er den Boden berührt, erscheinen zwei grüne Feen.

Pinke Fee: Oh-oh. Zwei grüne Feen! KIND, es gibt nur EINE echte grüne Fee. Eine von den beiden muss eine böse Fee sein, die sich verkleidet hat.

Fee bei der großen Blume: Gib mir den Zauberstab und ich werde euch helfen, den Zauberer zu finden.

Fee bei der kleinen Blume: Gib mir den Zauberstab und ich werde den Fluch aufheben.

E: Welche ist die gute Fee? Kannst du Kermit fragen „Welche ist die gute Fee?"?

K: Welche ist die gute Fee?

Kermit: Ich glaube, die Fee bei der großen Blume ist die gute Fee.

E: Gut, kannst du jetzt Batman fragen?

K: Welche ist die gute Fee?

B: Ich glaube, die Fee bei der kleinen Blume ist die gute Fee.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K(target 5): Was glaubst du, welche die gute Fee ist?

Dora: Ich glaube, die Fee bei der großen Blume ist die gute Fee, weil sie von dem Zauberer wusste.

E: Also schmeißt Batman den Zauberstab der Fee bei der großen Blume zu. Sie fängt ihn und benutzt ihn um zu zeigen, daß die Fee bei der kleinen Blume die böse Fee war. Die böse Fee fliegt beleidigt davon.

Grüne Fee: Danke, daß ihr mir geholfen habt! Ich hatte euch versprochen zu helfen, also sage ich euch, daß der Zauberer im Wald in der Nähe eines Teichs wohnt. Er braucht übrigens ein verzaubertes Objekt, wie zum Beispiel einen magischer Ring oder ein Drachenei, damit er den Fluch aufheben kann. Eins davon müsst ihr zuerst finden. Wenn ihr diesen Weg weitergeht, dann kommt ihr an einen Ort, mit zwei Straßen: eine Straße führt zu dem Drachenberg, wo die Drachen ihre Eier legen und eine Straße führt zu dem verzauberten Wald, in dem magische Ringe auf Bäumen wachsen. Ihr müsst euch entscheiden, welchen Weg ihr wählt. Ich gebe euch diese Karte, die funkelt, wenn ihr **entweder** in der Nähe der Dracheneier **oder** der magischen Ringe seid. Viel Glück!

E: Eine Karte erscheint in Batmans Händen. WOW! Die Freunde bedanken sich bei der Fee und gehen den Weg weiter. Aber plötzlich erscheint ein großer Bär in der Mitte der Straße.

Bär: Grrr! Ich werde euch nicht vorbei lassen, außer ihr gebt mir etwas, daß ich essen kann/ will.

E: Oh-oh. Jemand muss den Bären füttern. Aber niemand will das machen. Wer sollte den Bären füttern? Kannst du Kermit fragen „Wer sollte den Bären füttern?"?

K: Wer sollte den Bären füttern?

Kermit: Ich glaube, Batman sollte den Bären füttern.

E: Gut, kannst du jetzt Batman fragen?

K: Wer sollte den Bären füttern?

Batman: Ich glaube, Kermit sollte den Bären füttern.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K(target 6): Was glaubst du, wer den Bären füttern sollte?

Dora: Ich glaube, Kermit sollte den Bären füttern, weil er der Größte ist.

E: Okay, jetzt wissen wir, wer den Bären füttern sollte, aber wir wissen noch nicht, was der Bär möchte. Kannst du Kermit fragen „Was möchte der Bär essen?"?

K: Was möchte der Bär essen?

Kermit: Ich glaube, der Bär möchte einen Keks essen.

E: Gut, kannst du Batman fragen?

K: Was möchte der Bär essen?

Batman: Ich glaube, der Bär möchte einen Apfel essen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 7): Was glaubst du, was der Bär essen möchte?

Dora: Ich glaube, der Bär möchte einen Apfel essen, weil Äpfel auf Bäumen wachsen und Bären leben im Wald.

E: Also versucht Kermit sein Bestes, mutig zu sein und nimmt den Apfel und gibt ihn dem Bären. Aber der Bär sagt:

Bär: Grr! Ich mag keine Äpfel.

E: Und dann jagt er die Freunde den Weg entlang. Sie entkommen dem Bären und schau mal! Sie kommen zu einer Stelle mit zwei Straßen. Eine Straße führt zu dem Drachenberg, wo die Drachen ihre Eier legen und eine Straße führt in den verzauberten Wald, wo die magischen Ringe auf Bäumen wachsen. Wonach sollen wir schauen? Kannst du Kermit fragen „Wonach sollen wir schauen?“?

K: Wonach sollen wir schauen?

Kermit: Ich glaube, wir sollten nach dem magischen Ring schauen.

E: Gut, kannst du jetzt Batman fragen?

K: Wonach sollten wir schauen?

Batman: Ich glaube, wir sollten nach einem Drachenei schauen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 8): Was glaubst du, wonach wir schauen sollten?

Dora: Ich glaube, wir sollten nach einem magischen Ring schauen, weil er klein und leicht zu tragen ist.

E: So machen sich die Freunde auf den Weg zum verzauberten Wald. Auf dem Weg treffen sie auf einen Troll.

Troll: Seid vorsichtig, Freunde. In der Straße vor euch ist eine große Schlange. Sie wird euch nicht vorbei lassen, aber keine Sorge. Man kann sie leicht austricksen. Ihr müsst einfach ein bisschen Musik spielen, und dann schläft sie ein. Ich habe eine Gitarre und ein Tamburin. Ihr könnt eins davon haben, wenn ihr euch entschieden habt, was ihr spielen wollt.

E: Hm. Was sollen wir spielen? Kannst du Kermit fragen „Was sollen wir spielen?“?

K: Was sollen wir spielen?

Kermit: Ich glaube, wir sollten Gitarre spielen.

E: Gut, kannst du jetzt Batman fragen?

K: Was sollen wir spielen?

Batman: Ich glaube, wir sollten Tamburin spielen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was denkst du,…

K(target 9): Was denkst du, was wir spielen sollten?

Dora: Ich glaube, wir sollten Tamburin spielen, weil Tamburine klimpern und Schlagen lieben Klimpern!

E: Also gibt der Troll Kermit die Tamburine. Die Freunde bedanken sich bei dem Troll und gehen den Weg weiter. Nach kurzer Zeit treffen sie auf die Schlange, aber Kermit spielt auf dem Tamburin und die Schlange schläft ein, sodass sich die Freunde vorbei schleichen können. Dann beginnt die Karte in Batmans Hand zu leuchten und zu funkeln. Sie müssen ganz nah bei dem magischen Ring sein! Schau mal, der Baum funkelt auch. Kermit schaut sich den Baum ganz genau an und sieht ein, zwei, drei magische Ringe. Er nimmt einen und geht zurück zur Gruppe. Jetzt müssen wir nur noch den Zauberer finden. Als die Freunde weiter gehen, erscheint eine grüne Fee.

Grüne Fee: Freunde, geht nicht diesen Weg entlang! Der Wald brennt und ihr werdet euch verletzen!

E: Aber dann erscheint auch noch ein Rotkehlchen.

Rotkehlchen: Glaubt ihr nicht! Sie ist eine böse Fee, die sich verkleidet hat. Sie will, daß ihr euch verlauft.

E: Wem sollen wir glauben? Kannst du Kermit fragen „Wem sollen wir glauben?“?

K: Wem sollen wir glauben?

Kermit: Ich glaube, wir sollten dem Rotkehlchen glauben.

E: Gut, kannst du jetzt Batman fragen?

K: Wem sollen wir glauben?

Batman: Ich glaube, wir sollten der grünen Fee glauben.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 10): Was glaubst du, wem wir glauben sollen?

Dora: Ich glaube, wir sollten dem Rotkehlchen glauben, weil die grüne Fee nicht den Zauberstab hat, den wir ihr vorhin gegeben haben.

E: Die grüne Fee zeigt, daß sie eine böse Fee ist, die sich verkleidet hat und fliegt sauer davon.

Rotkehlchen: Ihr seid so nah dran, den Zauberer zu finden. Ihr müsst einfach den Weg weiter gehen.

E: Also gehen die Freunde den Weg weiter. Aber auf einmal verschwindet er! Jetzt wissen sie nicht, wohin sie gehen müssen. Sie müssen jemanden nach Hilfe fragen. Da drüben sind ein paar grüne Feen und dort ist ein Einhorn. Wen sollen wir fragen? Kannst du Kermit fragen „Wen sollen wir fragen?“?

K: Wen sollen wir fragen?

Kermit: Ich glaube, wir sollten die grünen Feen fragen.

E: Gut, kannst du jetzt Batman fragen?

K: Wen sollen wir fragen?

Batman: Ich glaube, wir sollten das Einhorn fragen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage beginnt so: Was glaubst du,…

K (target 11): Was glaubst du, wen wir fragen sollen?

Dora: Ich glaube, wir sollten das Einhorn fragen, weil Einhörner sehr weise sind.

E: Die Freunde gehen zu dem Einhorn.

Kermit: Entschuldigung, weißt du, wo wir den Zauberer finden können?

Einhorn: Ja klar! Ihr seid sehr nah. Ihr müsst einfach den lila Blumen folgen, bis ihr zu dem Teich kommt. Aber Vorsicht! Dort leben viele böse Feen.

E: Die Freunde bedanken sich bei dem Einhorn und folgen den lila Blumen, bis sie den Teich erreichen. Dort sehen sie eine Höhle. Vielleicht lebt dort der Zauberer, aber vielleicht leben dort auch die bösen Feen. Wen werden wir hier treffen? Kannst du Kermit fragen „Wen werden wir hier treffen?“?

K: Wen werden wir hier treffen?

Kermit: Ich glaube, wir werden hier den guten Zauberer treffen.

E: Gut, kannst du jetzt Batman fragen?

K: Wen werden wir hier treffen?

Batman: Ich glaube, wir werden hier die bösen Feen treffen.

E: Sie sind sich nicht einig. Aber vielleicht kann Dora helfen. Denk daran, zusammen mit der pinken Fee kannst du Dora fragen.

Pinke Fee: Die Frage fängt so an: Was glaubst du,…

K (target 12): Was glaubst du, wen wir hier treffen werden?

Dora: Ich glaube, wir werden hier den guten Zauberer treffen, weil die grüne Fee uns gesagt hat, daß er in der Nähe von einem Teich lebt.

E: Die Freunde gehen in die Höhle und wen sehen sie? Den guten Zauberer!

Kermit: Kannst du bitte Doras Fluch aufheben?

Zauberer: Habt ihr das verzauberte Objekt dabei, das ich brauche?

Kermit: Ja klar! Hier, der magische Ring!

E: Der Zauberer nimmt den Ring und benutzt seine Magie, um die pinke Fee zurück in die Feenwelt zu schicken und die Freunde hier her.

Batman: Du warst ein Held!

Kermit: Das hast du ganz toll gemacht!

Dora: Danke, daß du mir geholfen hast, zu reden!

E: Das war's! Gut gemacht!

Production No lead-in

Geschichte

Das sind zwei Freunde, die gemeinsam auf ein Abenteuer gehen. Das ist ein Alien und das ist ein Waschbär. Sie hat bereits andere Planeten mit ihrem Raumschiff besucht und jetzt möchte sie auf die Erde zu ihrem Freund/ ihrer Freundin. Oh-oh! Auf dem Weg geht etwas schief und das Schiff macht eine Bruchlandung! (Die Waschbär-Puppe wird von hinter dem Computer weggenommen.)

Experimentleitung (E): Kannst du Hallo sagen und nach ihrem Namen fragen?

Kind (K): Hallo! Wie heißt du?

Mindy: Hallo! Ich heiße Mindy. Das ist mein Freund, aber er spricht kein Deutsch. Wenn du mit ihm reden willst, sag mir, was du wissen willst und ich werde ihn fragen.

E: Lass uns seinen Namen herausfinden. Kannst du sagen: „Wie heißt du?"?

K: Wie heißt du?

Mindy: Ich werde ihn fragen, wie wir ihn nennen sollen. –Alien Sprache- Er sagt, daß du ihn Mork nennen kannst.

E: Lass uns ein bisschen besser kennenlernen. Was ist deine Lieblingsfarbe?

K: -antwortet-

E: Sehr schön! Warum fragst du nicht Mindy „Was ist deine Lieblingsfarbe?"?

K: Was ist deine Lieblingsfarbe?

Mindy: Meine Lieblingsfarbe ist pink.

E: Und jetzt Mork. Kannst du Mindy wegen Mork fragen?

K: Was ist Morks Lieblingsfarbe?

Mindy: Ich frage ihn mal kurz. –Alien Sprache- Er sagt, daß seine Lieblingsfarbe braun ist.

E: Toll! Lass uns jetzt herausfinden, was die Beiden auf der Erde sehen wollen. Kannst du Mindy fragen „Was wollt ihr hier sehen?"?

K: Was wollt ihr hier sehen?

Mindy: Ich möchte ein Restaurant sehen.

E: Mal schauen, was Mork sehen will. Kannst du Mindy fragen „Was will Mork hier sehen?“?

K: Was will Mork hier sehen?

Mindy: Ich werde ihn kurz fragen! –Alien Sprache- Er sagt, daß er gerne ein paar Fische sehen möchte. Unser Raumschiff ist in viele einzelne Teile kaputt gegangen. Kannst du uns helfen, sie zu suchen?

E: Ich glaube das können wir, oder was denkst du? Mal schauen, wo sollten wir zuerst schauen? Da drüben ist ein Teich, und da drüben ist ein Wald. Was denkst du?

K: Im Wald.

E: Im Wald ist eine gute Idee. Schauen wir mal, was Mindy denkt. Kannst du Mindy fragen „Wo sollen wir nachschauen?“?

K: Wo sollen wir nachschauen?

Mindy: Ich denke wir sollten bei dem Teich nachschauen.

E: Okay, ihr Beide habt zwei unterschiedliche Ideen. Schauen wir mal, was Mork dazu sagt. Kannst du Mindy wegen Mork fragen? –wenn nötig- Kannst du sagen „Was denkt Mork, wo wir nachschauen sollten?“?

K: Was denkt Mork, wo wir nachschauen sollten?

Mindy: Ich frage ihn! –Alien Sprache- Er sagt, er denkt, wir sollten im Wald schauen.

E: Also gehen Mindy und Mork in den Wald und schau mal, was sie gefunden haben! Die nose von dem Schiff! Kannst du sie an dem Schiff befestigen? Super! Nun gehen sie weiter und treffen auf einen Landwirt/ Bauer.

Mindy: Entschuldigen Sie! Wir suchen nach Teilen von unserem Raumschiff, haben Sie etwas gesehen, daß vielleicht ein Teil davon sein könnte?

Farmer Bob: Ich habe etwas gesehen, daß wie ein Flügel aussah. Es ist dort hinten herunter gefallen. Entweder bei dem Bauernstall von dem Bauer Hans, oder bei dem Krankenhaus von Doktor Maria.

E: Hm… wer hat den Flügel? Hast du eine Idee?

K: Doktor Maria.

E: Das ist eine tolle Idee. Lass uns herausfinden was Mindy denkt. Kannst du sie fragen „Wer hat den Flügel?“?

K: Wer hat die Flügel?

Mindy: Ich denke Bauer Hans hat den Flügel.

E: Okay, ihr zwei habt wieder zwei verschiedene Ideen, also brauchen wir Morks Meinung. Kannst du Mindy fragen?

Kind (target 1): Was denkt Mork, wer den Flügel hat?

Mindy: Ich werde ihn fragen. –Alien Sprache- Er sagt er denkt Bauer Hans hat den Flügel.

E: Also gehen sie zu dem Bauernstall von Bauer Hans und schau mal was sie sehen! Ein Flügel! Den kannst du an das Raumschiff stecken. Gut gemacht! Aber schau mal! Der Zaun von Bauer Hans ist kaputt gegangen, als der Flügel herunter fiel. Wir sollten ihn reparieren. Aber wer kann ihn reparieren? Wir wissen, daß Doktor Maria in der Nähe wohnt und wir haben gerade eben Bauer Bob getroffen, also wird er noch in der Nähe sein. Was denkst du?

K: Bauer Bob.

E: Gute Idee. Mal schauen, was Mindy denkt. Kannst du sie fragen „Wer kann der Zaun reparieren?"?

K: Wer kann den Zaun reparieren?

Mindy: Ich denke Doktor Maria kann den Zaun reparieren.

E: Okay, finden wir heraus, was Mork davon hält. Kannst du Mindy fragen?

K (target 2): Was denkt Mork, wer den Zaun reparieren kann?

Mindy: Ich frage ihn kurz! –Alien Sprache- Er denkt, Bauer Bob kann dabei helfen, den Zaun zu reparieren.

E: Also gehen Mindy und Mork los, um Bauer Bob wieder zu finden. Als sie ihn treffen, sagt er „Ich habe diese Karte auf der Straße gefunden! Gehört sie euch?" Natürlich gehört sie uns. Mach sie in unser Raumschiff. Bauer Bob, kannst du uns helfen, den Zaun von Bauer Hans zu reparieren?

Farmer Bob: Ja das kann ich, aber ich brauche einen Hammer und ich habe keinen. Der Lehrer Ben könnte einen Hammer haben… so wie die Mechanikerin Jenny.

E: Hmm. Wer hat den Hammer? Hast du eine Ahnung?

K: Jenny die Mechanikerin.

E: Gute Idee! Kannst du Mindy fragen "Wer hat den Hammer?"?

K: Wer hat den Hammer?

Mindy: Ich denke der Lehrer Ben hat einen Hammer.

E: Okay, lass uns herausfinden, was Mork denkt. Kannst du Mindy fragen?

K (target 3): Mindy, was denkt Mork, wer den Hammer hat?

Mindy: Ich werde ihn fragen. –Alien Sprache- Er denkt, daß die Mechanikerin Jenny einen Hammer hat.

E: Mindy und Mork machen sich auf den Weg, Jenny die Mechanikerin zu finden. Sie hat einen Hammer und erzählt, daß sie einen merkwürdigen/ seltsamen Sitz gefunden hat, der in keins der Auto passt, die sie repariert. Der Sitz gehört in das Raumschiff! Jenny sagt, daß wir immer zu ihr kommen können, wenn wir Hilfe brauchen, weil sie alles reparieren kann. Die Freunde nehmen den Hammer und bringen ihn zum Bauernhof von Bauer Hans. Sie geben Bauer Hans den Hammer und er kann damit den Zaun reparieren.

Mindy: Bauer Bob, Bauer Hans. Haben Sie noch etwas gesehen, daß vielleicht Teil unseres Raumschiffes sein könnte?

Bauer Hans: Ja, ich habe etwas gesehen, daß aussah wie ein Motor und es ist dort drüben hingefallen. Entweder bei dem Pavillon oder bei dem Teich.

E: Hmm… wo ist der Motor? Was denkst du?

K: Beim Teich.

E: Sehr gut. Lass uns Mindy fragen „Wo ist der Motor?"

K: Wo ist der Motor?

Mindy: Ich denke der Motor ist bei dem Pavillon.

E: Okay, lass uns herausfinden, was Mork denkt. Kannst du Mindy fragen „Was denkt Mork, wo der Motor ist?"?

K (Filler 1): Was denkt Mork, wo der Motor ist?

Mindy: Ich frage ihn! –Alien Sprache- Mork sagt, er denkt, daß der Motor bei dem Teich ist.

E: Die Freunde gehen zum Teich und finden den Motor! Aber der Motor ist stark beschädigt. Wen sollen wir nach Hilfe fragen? Bauer Bob hat den Zaun repariert. Vielleicht sollten wir ihn fragen. Aber die Mechanikerin Jenny sagte, daß sie alles reparieren kann. Hast du eine Idee?

Kind: Jenny.

E: Gute Idee. Was sagt Mindy dazu? Kannst du sie fragen „Wen sollen wir fragen?"?

K: Wen sollen wir fragen?

Mindy: Ich finde, wir sollten Bauer Bob fragen.

E: Okay, und nun fragen wir nach Mork!

K (target 4): Was denkt Mork, wen wir fragen sollen?

Mindy: Ich werde ihn fragen! –Alien Sprache- Er denkt, wir sollten die Mechanikerin Jenny fragen.

E: Die Beiden gehen zurück zu Jennys Shop. Als sie dort sind, fragen sie Jenny, ob sie helfen kann, den Motor zu reparieren.

Jenny: Natürlich kann ich euch helfen, aber leider habe ich sehr viel zu tun. Vielleicht könnt ihr mir helfen! Lehrer Ben braucht ein neues Lenkrad und Bauer Bob braucht vier neue Reifen. Wenn ihr Einem helfen könnt, dann kann ich an dem Motor arbeiten.

E: Wem sollten wir helfen? Lehrer Ben oder Bauer Bob? Was denkst du?

K: Lehrer Ben.

E: Gute Entscheidung. Kannst du Mindy fragen „Wem sollen wir helfen?“?

K: Wem sollen wir helfen.

Mindy: Ich denke wir sollten Bauer Bob helfen.

E: Gut. Wie sieht es mit Mork aus?

K (target 5): Was denkt Mork, wem wir helfen sollen?

Mindy: Ich frage kurz. –Alien Sprache- Er denkt, wir sollten Bauer Bob helfen.

E: Sie machen sich auf den Weg zu Bauer Bob, und er erzählt ihnen, daß er vier neue Reifen brauchen. Siehst du vier neue Reifen auf dem Bildschirm? –Kind zeigt darauf- Gut gemacht! Aber leider können wir die Reifen nicht drauf machen. Wir geben sie lieber jemandem, der das kann. Aber wem sollen wir die Reifen geben? Jenny vielleicht, aber sie hat keine Zeit, aber sie hat eine Assistentin, Rosie. Was denkst du?

K: Rosie.

E: Gut. Was ist mit Mindy? Kannst du sie fragen „Wem sollen wir die Reifen geben?“?

K: Wem sollen wir die Reifen geben?

Mindy: Wir sollten der Mechanikerin Jenny die Reifen geben.

E: Okay. Hat Mork eine Idee? Kannst du Mindy fragen?

K (target 6) : Was denkt Mork, wem wir die Reifen geben sollen?

Mindy: Ich werde ihn kurz fragen. –Alien Sprache- Er denkt, wir sollten Rosie die Reifen geben.

E: Also geben sie Rosie die Reifen. Rosie sagt, daß sie die Reifen auf den Traktor machen wird, und auch, daß sie einen weiteren Flügel vom Raumschiff gefunden hat. Perfekt! Die Mechanikerin Jenny kommt und sagt „Leute, es wird sehr lange dauern, bis der Motor repariert ist. Ihr geht lieber noch ein bisschen auf Erkundungstour, bis ich ihn repariert habe.“ Die Freunde beschließen, die Dinge anzusehen, die sie auf der Erde anschauen wollten. Erinnerst du dich, daß Mindy ein Restaurant sehen wollte? –Kind antwortet- Lass uns Mittagessen gehen. Die Freunde gehen den Weg zum Restaurant entlang, bis sie auf Doktor Maria und Susanne die Tennisspielerin treffen. Lass uns eine von beiden zum Essen einladen. Aber wen sollen wir einladen?

K: Doktor Maria.

E: Alles klar. Lass uns herausfinden, was Mindy denkt. Frag sie bitte „Wen sollen wir einladen?"

K: Wen sollen wir einladen?

Mindy: Wir sollten die Tennisspielerin Susanne einladen.

E: Und Mork?

K (target 7): Was denkt Mork, wen wir einladen sollen?

Mindy: Ich frage schnell nach. –Alien Sprache- Er denkt, wir sollten Doktor Maria einladen.

Doktor Maria: Es gibt zwei Orte, an denen man essen kann. Ein Burgerrestaurant und ein Café. Wo sollen wir essen?

E: Hast du eine Idee?

K: Im Café.

E: Lass uns Mindy fragen. Kannst du sie fragen „Wo sollen wir essen gehen?"?

K: Wo sollen wir essen gehen?

Mindy: Ich finde, wir sollten im Burgerrestaurant essen.

E: Und Mork?

K (Filler 2): Was denkt Mork, wo wir essen gehen sollen?

Mindy: Ich frage ihn! –Alien Sprache- Er denkt, wir sollten im Burgerrestaurant essen.

Dr. Maria: Alles klar. Aber ich bin mir nicht sicher, wie wir dort hinkommen. Der Lehrer Ben und der Schüler Jonas sind beide in der Nähe. Wer kennt wohl den Weg?

K: Lehrer Ben.

E: Gut, fragen wir Mindy. Kannst du fragen „Wer kennt den Weg?"?

K: Wer kennt den Weg?

Mindy: Ich denke der Schüler Jonas kennt den Weg.

E: Und Mork?

K (target 8): Was denkt Mork, wer den Weg kennt?

Mindy: Ich frage ihn schnell. –Alien Sprache- Er denkt der Lehrer Ben kennt den Weg.

E: Sie gehen nun zum Schulhof und fragen Lehrer Ben, wie man zum Burgerrestaurant kommt. Er sagt, daß wir die Straße entlang müssen und das Burgerrestaurant neben dem Supermarkt ist. Also los geht's. Als sie beim Burgerrestaurant ankommen, treffen sie Bauer Bob und Bauer Hans, die gerade zu Mittag essen. Der Bauer Bob sagt „Gute Entscheidung hier herzukommen! Ihr solltet die Burger ausprobieren. Das sind die besten Burger in der Stadt." Bauer Hans sagt „Nein, nehmt die Makkaroni! Das ist das Beste hier." Auf wen sollen wir jetzt hören? Was denkst du?

K: Bauer Hans.

E: Okay, und jetzt mal schauen, was Mindy dazu sagt. Kannst du sie fragen „Auf wen sollen wir hören?"?

K: Auf wen sollen wir hören?

Mindy: Wir sollten auf Bauer Bob hören.

E: Und Mork?

K (target 9): Was denkt Mork, auf wen wir hören sollen?

Mindy: Ich frage ihn. –Alien Sprache- Er denkt, wir sollten auf Bauer Bob hören.

E: Die Freunde entscheiden sich, einen Burger zu kaufen und zu teilen. Jetzt haben sie ein Restaurant gesehen. Weißt du noch, was Mork sehen wollte? Fische! Lass uns zu einem Aquarium gehen. Die Freunde kommen beim Aquarium an, und merken, daß sie nur für eine Ausstellung Zeit haben. Entweder Quallen oder Haie. Was sollen wir uns anschauen?

K: Haie.

E: Gute Idee. Lass uns herausfinden, was Mindy sehen will. Frag sie „Was sollen wir uns anschauen?"

K: Was sollen wir uns anschauen?

Mindy: Ich denke, wir sollten uns die Quallen anschauen.

E: Okay, und jetzt fragen wir Mork.

K (target 10) Was denkt Mork, was wir uns anschauen sollten?

Mindy: Ich frage ihn! –Alien Sprache- Er sagt, er denkt, wir sollten uns die Haie anschauen.

E: Okay. Dann beginnen wir mal die Führung. Es gibt zwei Führungsleiter. Der Eine trägt eine Hai-Mütze und der Andere trägt eine Fisch-Mütze. Wer kann uns die Haie zeigen?

K: Der Eine mit der Hai-Mütze.

E: Okay, lass uns Mindy fragen. „Wer kann uns die Haie zeigen?"

K: Wer kann uns die Haie zeigen?

Mindy: Ich denke der Eine mit der Fisch-Mütze kann uns die Haie zeigen.

E: Okay, lass uns wegen Mork fragen.

K (target 11): Was denkt Mork, wer uns die Haie zeigen kann?

Mindy: Ich frage ihn! –Alien Sprache- Er denkt, der eine mit der Hai-Mütze kann uns die Haie zeigen.

E: Also leitet der Mann mit der Hai-Mütze die Führung und die Freunde haben sehr viel Spaß dabei, die Haie anzuschauen. Wie cool! Als sie das Aquarium verlassen, sehen die Freunde einen Mann, der T-Shirts mit Schildkröten, und eine Frau, die T-Shirts mit Delphinen verkauft. Wer hat die schöneren T-Shirts?

K: Die Frau.

E: Lass uns herausfinden, was Mindy dazu sagt. Kannst du sie fragen „Wer hat die schöneren T-Shirts?"?

K: Wer hat die schöneren T-Shirts?

Mindy: Ich denke, der Mann hat die schöneren T-Shirts.

E: Okay, lass uns herausfinden, was Mork dazu sagt.

K (target 12): Was denkt Mork, wer die schöneren T-Shirts hat?

Mindy: Ich werde ihn schnell fragen. –Alien Sprache- Mork sagt, er denkt der Mann hat die schöneren T-Shirts.

E: Als Erinnerung an das Aquarium kauft sich Mork ein T-Shirt mit Schildkröten von dem Mann. Die Freunde gehen zurück zu Jennys Shop und Jenny sagt „Hier ist der Motor! Alles ist repariert." Super! Nimm den Motor und mach ihn zurück in das Raumschiff. Schau mal, wir können das Schiff in den Computer machen! Es ist nicht mehr kaputt –Schiff erscheint auf dem Bildschirm und wird hinter dem Computer platziert- Die Freunde steigen in das Raumschiff ein und fliegen los. Vielen Dank für deine Hilfe! Ohne dich hätten ihr Abenteuer nicht weitergehen können.

Comprehension

Lilly (Übungsgeschichte)

Das ist Lilly. Sie hat Besuch von ihrem Freund Robert. Schau mal, Robert hat ihr eine Überraschung mitgebracht! / Lilly geht zu dem Geschenk, um es zu öffnen / und findet darin ein Kätzchen. Lilly freut sich sehr über das Geschenk und liebt ihr neues Kätzchen. /

Wer hat Lilly eine Überraschung gebracht?

Alex (Filler)

Das ist Alex. In dem Wald in der Nähe von seinem Haus gibt es / ein Geisterhaus. In diesem Geisterhaus gibt es / viele Geister und Mumien. Die Freunde von Alex denken, daß die Geister und Mumien böse und gruselig sind, aber Alex ist sich sicher, daß sie nett und freundlich sind. Er denkt sich, daß wenn er einen Geist oder eine Mumie treffen könnte und sie seinen Freunden vorstellen könnte, dann wüsste jeder, daß Geister und Mumien nett sind. /Also macht sich Alex auf den Weg in den Wald, um das Geisterhaus zu finden.

Auf dem Weg trifft Alex einen Bauern. / Er fragt den Bauern „Weißt du wo das Geisterhaus ist? Weißt du was ich dort finden werde?" / Der Bauer sagt „Ich weiß nicht, was du dort finden wirst. Aber ich weiß, / daß das Geisterhaus auf einem hohen Hügel in der Mitte vom Wald ist. Wenn du wirklich wissen willst, was du dort finden wirst, solltest du / die Hexe Esmeralda fragen. Sie trägt einen langen spitzen Hut und hat lange grüne Haare. / Sie kann mit ihrer Kristallkugel in die Zukunft schauen. / Sie lebt auf der Lichtung, hier in der Nähe."

Alex bedankt sich bei dem Bauern und geht zu der Lichtung. / Er sieht die Hexe Esmeralda mit ihrem langen spitzen Hut und ihren langen grünen Haaren. / Er fragt sie „Was werde ich in dem Geisterhaus finden?" Esmeralda schaut tief in ihre Kristallkugel und sagt „Aha! / Du wirst eine Mumie im Geisterhaus treffen."

/ Alex freut sich und läuft schnell zu dem hohen Hügel. / Er betritt das Geisterhaus und als er hineingeht, sieht er gar nichts. / Aber dann erscheint ein Geist, / welcher gleich wieder wegfliegt, bevor Alex etwas sagen konnte. Alex ist enttäuscht, / aber dann kommt eine nette Mumie in das Haus hinein. „Hallo!" sagt die Mumie. „Mein Freund, der Geist, ist sehr schüchtern, aber ich lerne gerne neue Leute kennen." „Es freut mich dich kennenzulernen" sagt Alex. „Würdest du gern meine Freunde treffen? Ich möchte ihnen beweisen, daß Geister und Mumien nett sind." „Natürlich" sagt die Mumie. „Bringe sie zu dem Garten vor dem Geisterhaus."

/ Also bringt Alex am nächsten Tag seine Freunde zu dem Garten vor dem Geisterhaus und sie lernen alle die freundliche Mumie kennen. Jetzt wissen sie alle, daß Geister und Mumien nett sind.

Q:Wem hat Alex erzählt, wen er treffen würde?

Tom und Katharina

Das ist Tom. / Er und Katharina sind am Strand und spielen ein Piratenspiel. Sie bekommen Anweisungen von einem echten, lebendigen Piraten, dem Pirat Jakob. / Der Pirat Jakob sagt, daß er irgendwo auf der Insel entweder einen Piratenhut oder einen Piratenhaken versteckt hat. Wer eine der beiden Sachen findet, bekommt einen Piratenpreis. / Er sagt auch, daß ein sprechender Papagei namens Georg Hinweise geben kann und / daß ein weiterer Hinweis in der Nähe von Felsbrocken versteckt ist. / Wer einen Hinweis findet, wird wahrscheinlich gewinnen. / Tom und Katharina rennen in unterschiedliche Richtungen und suchen nach Hinweisen.

/ Nach kurzer Zeit trifft Katharina Georg den sprechenden Papagei. Sie sagt „Georg. Wonach soll ich schauen? Was wird der Gewinner finden?" / Georg sagt „Piratenhut, Piratenhut!" Perfekt! Katharina rennt weg und sucht nach einem Piratenhut. /

/ Zur gleichen Zeit findet Tom die Felsbrocken. Er sieht ein Stück Papier. / Es ist eine Karte! Darauf sieht man einen Piratenhut neben zwei Palmen. / Jetzt weiß Tom, was der Gewinner finden wird und wo er es finden wird. Aber er muss sicher gehen, daß Katharina nicht zuerst dort ankommt. Er beschließt, ihr einen falschen Hinweis zu hinterlassen. / Er schreibt mit seinen Fingern in den Sand „Der Gewinner wird einen Piratenhaken neben den Booten finden" / aber eine große Welle kommt und / wäscht Toms Nachricht weg. Also muss sich Tom etwas anderes ausdenken. / Er schaut sich um und findet eine Flasche. Tom kann Katharina eine Nachricht in der Flasche hinterlassen. Das ist ein toller Piratenhinweis. / Er nimmt ein Stück Papier und darauf schreibt er „Der Gewinner wird einen Piratenhaken neben den Booten finden“ / und dann lässt er die Nachricht für Katharina, in der Flasche.

/ Tom sieht, wie Katharina kommt und versteckt sich hinter den Felsen. / Sie hebt die Flasche auf und denkt, daß es ein echter Hinweis ist, / also rennt sie in die Richtung der Boote. / Tom ist glücklich, daß sein Trick funktioniert hat und rennt in die Richtung der Palmen. / Tom kommt bei den Palmen an und schau mal was er findet. / Einen Piratenhut! Er hebt ihn auf und / bringt ihn zu dem Piraten Jakob. / Der Pirat zieht den Hut an und gibt Tom eine Goldmünze als Belohnung. Tom hat das Piratenspiel gewonnen.

Wie+was (Wh-Relativizer): Wie hat Tom Katharina gesagt, was der Gewinner finden wird?

Was+was(WSM): Was hat <u>Tom</u> Katharina gesagt, was der Gewinner finden wird?

Emma und Sarah

Das ist Emma. Sie plant eine Überraschungsgeburtstagsfeier für ihre Freundin Sarah. / Aber sie hat noch kein Geschenk für Sarah gekauft. Emma weiß, daß Sarah sehr gerne Sport macht, / also entscheidet sie sich, zu einem Sportgeschäft zu gehen. Sie denkt, daß Sarah entweder einen Fußball oder einen bunten Strandball möchte. / Sie schaut zuerst nach einem Strandball, aber schau mal! Sie sind alle ausverkauft. Es gibt keine Strandbälle mehr, also kann Emma keinen Strandball für Sarah kaufen. / Sie entscheidet sich, sich noch ein bisschen länger umzusehen und sie findet schnell einen Fußball. / Das ist sehr gut. Sarah wollte schon seit sehr langer Zeit einen Fußball. / Also kauft Emma einen Fußball für Sarah und jetzt ist alles fertig für die Geburtstagsfeier. /

Aber die Feier ist erst am Samstag und heute ist erst Donnerstag. Das bedeutet morgen ist Freitag und die Mädchen müssen noch in die Schule gehen. / In der Schule ist Reden nicht erlaubt, aber bevor der Unterricht begonnen hat, / geht Sarah zu Emma und flüstert „Was hast du gestern gemacht?“ / Emma flüstert „Ich war...“/ doch die Lehrerin sagt „Mädels! Zurück zu euren Tischen. In der Schule wird nicht geredet.“

Die Mädchen gehen zurück zu ihren Tischen, aber Emma muss Sarah noch sagen, daß sie etwas anderes gekauft hat. Sie entscheidet sich, ihr eine Nachricht zu schreiben. / Sie nimmt ein Stück Papier und schreibt sie: „Hi Sarah, ich habe einen bunten Ball für den Strand gekauft. Wenn du am Samstag zu mir nach Hause kommst, kann ich ihn dir zeigen.“ /

Während die Lehrerin einem anderen Kind hilft, / schleicht Emma mit ihrem Zettel zu Sarah und / dann zurück an ihren Tisch. Sarah liest die Nachricht und nickt Emma zu. Das heißt, sie kann am Samstag vorbeikommen. Perfekt!

/ Es ist Samstag und alle Freunde gehen zu Emma nach Hause und helfen das Essen und die Dekoration vorzubereiten. / Als Sarah in das Haus hereinkommt, rufen die Kinder „Überraschung! Alles Gute zum Geburtstag Sarah!“ Sarah ist so überrascht. Das hat sie überhaupt nicht erwartet. / Die Freunde haben sehr viel Spaß bei der Feier und Sarah freut sich über ihre schönen Geschenke. Das war eine tolle Überraschungsfeier.

Wie+was(WH-relativizer): Wie hat Emma Sarah gesagt, was sie gekauft hat?

Was+was(WSM): Was hat <u>Emma</u> Sarah gesagt, was sie gekauft hat?

Jana und Jonas

Das sind Jana und Jonas. Sie sind auf einem Ausflug mit ihrer Schule. / Die Lehrerin möchte, daß sie ein Foto von einem Regenwaldvogel machen. / Jonas sagt, daß sein absoluter Lieblingsvogel ein Tukan ist. Er ist sehr aufgeregt, weil er vielleicht Einen sehen wird. / Jana sagt, sie würde auch gerne einen Tukan sehen, aber ein Papagei wäre auch sehr cool. Es macht für sie keinen Unterschied. / Jonas sagt „Okay, Jana wieso gehst du nicht zu den Bananen und / ich gehe dann runter zum Fluss. / Wir können diese speziellen Regenwaldradios benutzen, um uns zu sagen, was wir sehen.“ / Also gehen die Freunde in unterschiedliche Richtungen und schauen nach Regenwaldvögeln.

/ Jana geht zu der Stelle, wo die Bananen wachsen und als sie dort ankommt, sieht sie einen wunderschönen Tukan. / Er ist so schön und Jana ist so aufgeregt, daß sie ganz laut „WOW“ ruft. / Aber ihr Rufen hat den Tukan erschreckt und er fliegt weg. / Jana versucht, ihm hinterherzugehen, aber er ist zu schnell und sie verliert ihn. Oh nein. Jana hat den Tukan verscheucht, bevor Jonas ihn sehen konnte. Sie beschließt, daß sie Jonas nicht erzählen wird, was sie gesehen hat. Sie will seine Gefühle nicht verletzen. /

Jonas nimmt das Radio und sagt „Jana, hast du einen Tukan gesehen?“ / Jana möchte Jonas wirklich nicht erzählen, daß sie einen Tukan gesehen und verscheucht hat, also sagt sie „Ich habe…“ aber / Jonas unterbricht sie und sagt „Warte, diese Radios sind viel zu laut. Damit werden wir nur die Vögel verscheuchen. Ich komme zu dir und dann können wir flüstern.

/ Als Jonas bei Jana ankommt, flüstert er „Denk dran, wir müssen noch ein Foto von einem Vogel aus dem Regenwald machen, also musst du flüstern. Was hast du gesehen?“ / Jana möchte die Gefühle von Jonas nicht verletzen, deswegen sagt sie ihm, daß sie etwas anderes gesehen hat. Sie lehnt sich zu Jonas und flüstert „Ich habe einen Papagei gesehen. Lass uns weiter nach Tukanen schauen.“ Die Freunde schauen weiter nach Tukanen bis es Zeit ist, zurück zu fahren. / Sie gehen zurück zum Bus und schau mal was sie sehen! Ein wunderschöner Tukan wartet auf sie. / Jonas macht ein Foto, und / obwohl der Tukan schnell wieder weg fliegt, hat Jonas ein Foto und kann sich immer daran erinnern, daß er seinen Lieblingsvogel gesehen hat.

Wie+was(WH-Relativizer): Wie hat Jana Jonas gesagt, was sie gesehen hat?

Was+was(WSM): Was hat <u>Jana</u> Jonas gesagt, was sie gesehen hat?

Elisa (Filler)

Das ist Elisa. Sie liebt Dinosaurier über alles. Sie soll für die Schule einen Bericht schreiben und sie entscheidet sich, / daß sie ihn über Dinosaurier schreibt. Dafür wäre es am besten, wenn sie einen echten Dinosaurier sehen könnte, aber es gibt keine mehr. / Ihre nächstbeste Option ist jetzt, daß sie mit ihrem Freund, Professor Schmid, und seinem Roboter, Clank, redet, und fragt, ob sie etwas über Dinosaurier wissen. Sie fragt ihren Professor, aber er sagt, daß er nicht viel über Dinosaurier weiß. / Aber er ist gerade damit fertig geworden, eine Zeitmaschine zu bauen. / Sie können in der Zeit zurückreisen, um ein paar Dinosaurier anzuschauen. / Also gehen sie in die Zeitmaschine und reisen in die Zeit der Dinosaurier.

/ Als sie dort ankommen und / aus der Maschine aussteigen, bemerken Elisa, Professor Schmid und Clank, daß es so viel zu sehen gibt und deswegen beschließen sie, sich aufzuteilen. / Elisa sagt „Clank, du gehst zu dem Feld mit den komischen Bäumen, die seltsame Früchte haben. / Der Professor und ich gehen runter zum Teich und nachher erzählen wir uns gegenseitig, was wir gesehen haben.“ Die Freunde gehen in unterschiedliche Richtungen um Dinosaurier zu finden.

/ Als Elisa und Professor Schmid den Teich erreichen, entdecken sie ein Nest voll mit Dinosauriereiern. / „Welche Art von Dinosaurier hat diese Eier gelegt?“ fragt Elisa. „Und was frisst dieser Dinosaurier?“ / „Ein Brontosaurus hat diese Eier gelegt.“ Sagt Professor Schmid. „Aber ich weiß nicht, was sie fressen.“ / „Ich wette ein Brontosaurus frisst entweder Blätter oder Früchte, oder beides“, sagt Elisa. / Genau in dem Moment, hören sie Clank, den Roboter, rufen „Hey Freunde, kommt schnell hier her! Da ist ein Brontosaurus. Und er frisst.“ / Elisa und Professor Schmid rennen zu dem Feld mit den komischen Bäumen, die seltsame Früchte haben. Als sie dort ankommen sehen sie Clank, den Roboter, und einen sehr großen Brontosaurus. / Clank sagt „Ihr habt gerade verpasst, wie dieser Brontosaurus total viele Blätter gefressen hat.“

/ Kurz nachdem Clank das sagt, bricht der Vulkan aus. / Rennt alle schnell zurück zur Zeitmaschine! / Die Freunde rennen zurück zur Zeitmaschine / und springen hinein, und dann fliegen sie genau zur richtigen Zeit zurück in die Gegenwart. / Als sie zuhause ankommen, stimmen alle zu, daß es großartig war, einen Brontosaurus zu sehen, und Elisa freut sich am Meisten, daß sie jetzt den besten Bericht überhaupt schreiben kann.

Wer hat Elisa gesagt, was der Brontosaurus gefressen hat?

Oma und Lisa

Das ist Lisa. Sie trifft sich heute Abend mit ihren Freunden für einen Spieleabend. / Lisas Oma kommt in ihr Zimmer und sagt „Ich gehe jetzt zum Supermarkt, um etwas Essen zu kaufen, damit ich etwas für dich und deine Freunde machen kann.“ „Danke Oma, das ist sehr lieb“ sagt Lisa. / Die Oma geht und dann fällt Lisa ein, daß sie nicht weiß, was Oma zu Essen machen wird. / Sie hofft, daß Oma ihren super leckeren Schokoladenkuchen macht, / aber Oma macht gerne diesen ekligen Gemüseauflauf. Oma ist nicht gut darin, einen Gemüseauflauf zu machen, aber Lisa hat ihr das noch nie gesagt, weil sie ihre Gefühle nicht verletzen will. Lisa möchte wissen, ob Oma das heute Abend machen wird, damit sie ihre Freunde warnen kann.

/ Lisa rennt zur Bushaltestelle, aber als sie dort ankommt, ist Oma schon im Bus. / Lisa ruft „Oma! Was wirst du heute Abend machen“ / Oma kann die Frage im Bus hören, aber sie möchte nicht rufen, weil das nicht nett ist. Also beschließt Oma, daß sie ein Bild malen wird und es gegen das Fenster hält, damit sie Lisa sagen kann, was sie machen wird. / Sie beginnt zu malen, aber bevor sie fertig ist, fährt der Bus los, und Oma muss sich etwas anderes überlegen. /

Lisa ist will wirklich wissen, was Oma machen wird. / Dann klingelt auf einmal ihr Handy. Lisa hat eine Nachricht von ihrer Oma erhalten, in der steht „Hi Lisa, ich habe deine Frage gehört. Ich werde heute Abend Gemüseauflauf machen. Liebe Grüße Oma“. /

Oh nein. Lisa ist enttäuscht, daß sie heute Abend keinen Schokoladenkuchen haben werden, aber immerhin kann sie ihre Freunde warnen. / Als ihre Freunde ankommen, ist Oma in der Küche und Lisa sagt / „Freunde, meine Oma macht einen Gemüseauflauf, der nicht gut schmeckt, aber wir müssen vorgeben, daß er gut schmeckt, damit wir ihre Gefühle nicht verletzen.“ Lisas Freunde sind damit einverstanden. / Aber dann kommt Oma in das Zimmer mit einem Schokoladenkuchen. Sie wollte Lisa damit überraschen. Die Freunde essen den super leckeren Schokokuchen / und spielen dann ganz viele Spiele.

Wie+was(Wh-relativizer): Wie hat Oma Lisa gesagt, was sie machen wird?

Was+was(WSM): Was hat <u>Oma</u> Lisa gesagt, daß sie machen wird?

Luis und Sandra

Das sind Luis und Sandra. Heute Abend feiern sie mit ihren Freunden, aber sie wissen nicht, welches Essen sie anbieten sollen. / Luis denkt, er könnte ein paar Früchte pflücken und diese könnten sie dann als Nachtisch essen. / Sandra denkt, daß das eine gute Idee ist, aber sie bleibt zu Hause und bereitet alles für die Party vor. Aber sie muss trotzdem wissen, welche Früchte Luis pflückt, damit sie weiß, welchen Nachtisch sie machen soll. / Sie gibt ihm ein Walkie-Talkie und sagt, daß er ihr damit sagen soll, was sie pflückt. Luis sagt „Okay" und macht sich auf den Weg. /

Luis geht zum Teich und sieht ein großes Beet voller Erdbeeren. Das wäre ein toller Nachtisch! /Also pflückt Luis eine Erdbeere, und noch eine und ganz schnell ist sein kompletter Korb voll mit Erdbeeren. / Er holt sein Walkie-Talkie aus seiner Tasche, um Sandra zu sagen, was er gepflückt hat. Aber er lässt es fallen! / Es fällt in den Teich. Oh nein! Jetzt kann Luis sein Walkie-Talkie nicht benutzen, um Sandra zu sagen, was er gepflückt hat. /Was macht er jetzt nur? Dann erinnert sich Luis, / daß es ein Telefon in der nahegelegenen Eisdiele gibt. Er kann das Telefon benutzen, um Sandra zu sagen, was er gepflückt hat. Also geht er los zur Eisdiele.

/ Luis bekommt auf dem Weg zur Eisdiele Hunger und isst eine Erdbeere… und dann isst er noch eine… und bis er endlich an der Eisdiele angekommen ist, ist nur noch eine Erdbeere in seinem Korb. Aber sie können doch nicht nur eine Erdbeere als Nachtisch machen. / Zum Glück entdeckt Luis eine Frau, die Äpfel verkauft. Äpfel eignen sich gut als Nachtisch, also kauft er welche. Es ist ihm so peinlich, daß er alle Erdbeeren gegessen hat, deswegen gibt er vor, er hätte die Äpfel gepflückt. / Er legt die Äpfel in seinen Korb, und jetzt sieht es so aus, als hätte er sie selbst gepflückt. / Er ruft Sandra an / und sagt „Hi Sandra, / ich habe ganz viele Äpfel vom Apfelbaum gepflückt." „Das ist super!" sagt Sandra. / „Ich mache einen Apfelkuchen. Komm schnell Heim."

/ Luis geht schnell nach Hause mit seinem Korb voller Äpfel. Als er ankommt, hat Sandra bereits den Kuchenboden gemacht, also machen sie die Äpfel auf den Kuchenboden und den Kuchen in den Ofen / und dann haben sie eine tolle Party mit all ihren Freunden.

Wie+was(Wh-realtivizer): Wie hat Luis Sandra erzählt, was er gepflückt hat?

Was+was(WSM): Was hat <u>Luis</u> Sandra gesagt, was er gepflückt hat?

Der böse Stefan und Ermittlerin Selina

Das ist die Ermittlerin Selina. Sie hat bereits viele Fälle von ihrem Feind, dem bösen Stefan, gelöst, aber sie hat ihn noch nie erwischt.

/ Zur gleichen Zeit ist Stefan in seinem Büro, / und denkt darüber nach, wie er die Krone klauen kann. Aber er weiß, daß Selina sehr schlau ist und ihn erwarten wird. Er beschließt, sie auszutricksen, indem er ihr sagt, daß er etwas anderes klauen wird. Aber wie kann er Selina etwas sagen, ohne daß er dabei erwischt wird? / Stefan hat eine Idee. / Er hat neulich eine Fernsehmaschine erfunden, mit der er in die Fernseher von anderen Personen gezaubert wird, auch wenn ihre Fernseher nicht eingeschaltet sind. Er wird die Fernsehmaschine benutzen, um Selina zu sagen, was er stehlen wird.

/ Kurz darauf geht in Selinas Büro der Fernseher an! Sie sieht den bösen Stefan und er lacht und sagt „Leichtsinnige Ermittlerin Selina-„ aber Stefan wird unterbrochen / und der Fernseher wird wieder schwarz. „Was hat der böse Stefan in meinem Fernseher gemacht?" denkt sich Ermittlerin Selina.

/ Gleichzeitig sieht man Stefans kaputte Fernsehmaschine. Überall ist Rauch und Kabel kommen aus der Maschine heraus. Stefan kann seine Fernsehmaschine nicht mehr benutzen, um Selina irgendwas mitzuteilen. / Er beschließt, ihr einen Brief zu schreiben. Er wird ihn persönlich vorbeibringen und vom Fenster aus beobachten, daß sie ihn auch wirklich liest.

/ Nach kurzer Zeit erhält Selina einen Brief. Sie schaut ihn sich an, und so sieht sie nicht, wie Stefan sie am Fenster beobachtet. / Sie öffnet den Brief und darin steht „Leichtsinnige Ermittlerin Selina, ich werde den Diamantring der Königin klauen und es gibt nichts, das du dagegen machen kannst! Hahaha! Stefan" / „Oh nein! Ich muss den Ring beschützen!" ruft Selina / und rennt zu dem Palast der Königin. / Als sie dort ankommt, sagt eine Wache „Nein, der Ring der Königin ist sicher. Er ist hier auf der Kommode, genau dort, wo er sein sollte." „Stefan hat mich bestimmt reingelegt!" denkt sich die Ermittlerin Selina. / „Ich wette, das hat er nur gesagt, um mich von der Krone abzulenken." / Also rennt sie zu dem Museum, in dem die Krone ausgestellt wird.

/ Als sie im Museum ankommt, / trifft sie auf einen Wachmann und er sagt „Die Krone der Königin ist sicher! Sie ist drinnen auf einem Podest, genau dort, wo sie sein sollte. Du kannst selber nachschauen." / Also geht Selina in das Museum und als sie an dem Podest ankommt, auf dem die Krone sein sollte, sieht sie, daß dort nichts ist! / „Der Wachmann sah sehr bekannt aus" denkt sich Selina „und was war in der Tasche, die er dabei hatte? Ich bin mir sicher, daß das der böse Stefan war, der sich verkleidet hat!" / Selina rennt raus und hält den Wachmann auf. / Sie bringt ihn dazu, seinen Hut auszuziehen und seine Tasche zu leeren. / In der Tasche findet Selina die Krone der Königin und die Klamotten und die Maske von Stefan.

„Ich habe dich erwischt böser Stefan" sagt die Ermittlerin Selina. / „Du wirst für eine lange Zeit ins Gefängnis gehen."

Wie+was(wh-relativizer): Wie hat der böse Stefan der Selina erzählt, was er stehlen wird?

Was+was(WSM): Was hat <u>der böse Stefan</u> Selina erzählt, was er stehlen wird?

References

Adams, A.M & Gathercole, S.E. (2000). Limitations in working memory: implications for language development. *International Journal of Language and Communication Disorders*, *35* (1), 95-116.

Ambridge, B., Rowland, C., Theakston, A., & Tomasello, M. (2006). Comparing different accounts of inversion errors in children's non-subject wh-questions: "What experimental data can tell us?" *Journal of Child Language, 33*(3), 519–557.

Atkinson, E., Wagers, M., Lidz, J., Phillips, C., & Omaki, A. (2018). Developing incrementality in filler-gap dependency processing. *Cognition, 179*, 132–149.

Baddeley, A. (2017). Modularity, working memory and language acquisition. *Second Language Research. 33*(3), 299-311.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language, 56*(1), 65–85.

Baker, M. (2005). Mapping the terrain of language acquisition. *Language Learning and Development, 1*(1), 93–129.

Bakovic, E. & Keer, E. (2001). Optionality and Ineffability. In. G. Legendre, J. Grimshaw, &S. Vikner (Eds.) *Optimality Theoretic Syntax*. 97-112. MIT Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Beck, S. & Berman, S. (1996). Wh-Scope Marking: Direct vs. Indirect Dependency. In U. Lutz & G. Müller (Eds.), *Papers on Wh-Scope Marking*. 59-83

Behrens, H. (2006). The input-output relationship in first language acquisition. *Language and Cognitive Processes, 21*, 2-24.

Bhatt, R., Yoon, J. (1992). On the composition of COMP and parameters of V2. In: *Proceedings of the Tenth WCCFL*, 41-53. CSLI, Stanford CA.

Blunter, R. (2000). Some aspects of Optimality Theory in interpretation. *Journal of Semantics, 17*, 189-216.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry, 21*(4), 491–504.

Bock, K., & Miller, C. A. (1991). Broken agreement. *Cognitive Psychology, 23*(1), 45–93.

Bock, J.K. (1987b). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the psychology of language, 3*, 337-390.

Bock, K & Ferreira, V.S. (2014). Syntactically Speaking. In M. Goldrick, V. Ferreira, & M. Miozzo (Eds), *The Oxford Handbook of Language Production*. 1-34.

Bock, J.K. & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology, 23*, 45-93.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International, 5*(9/10), 341–345.

Borer, H. (1984). Restrictive relatives in Modern Hebrew. *Natural Language & Linguistic Theory, 2*(2), 219–260.

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

Bruening, B. (2004). Two types of wh-scope marking in Passamaquoddy. *Natural Language and Linguistic Theory, 22*(2), 229–305.

Bruening, B. (2006). Differences between the wh-scope-marking and wh-copy constructions in Passamaquoddy. *Linguistic Inquiry, 37*(1), 25–49.

Buring, D. (1997b). *The Meaning of Topic and Focus – The 59th Street Bridge Accent*. Routledge, London.

Büring, D. (2003). On D-trees, beans, and B-accents. *Linguistics and Philosophy 26*:511–545

Carnie, A. (2013). *Syntax: A Generative Introduction -- Third Edition.* Wiley & Blackwell Publishing.

Choi, H. (1996). *Optimizing Structure in Context: Scrambling and Information Structure*. PhD Dissertation. Stanford University.

Chomsky, N. (1973). Conditions on transformations. In S. Anderson & P. Kiparsky (Eds.), *A festschrift for Morris Halle* (pp. 232–286). New York: Holt, Rinehart & Wilson.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, the Netherlands: Foris.

Chomsky, N. (1991) Some Notes on the Economy of Derivation. In Freidin (ed,), *Principles and Parameters in Comparative Grammar*, (pp. 417-454).

Chomsky, N. (1995). *The Minimalist Program*. MIT Press.

Chomsky, N. (2001). Derivation by phase. In M. Kenstowicz (Ed.), Ken Hale: *A life in language,* (1–52).Cambridge, MA: MIT Press.

Chomsky, N., & Lasnik, H. (1977). Filters and control. *Linguistic Inquiry, 8*(3), 425–504.

Chung, S. (1991). Functional heads and proper government in Chamorro. *Lingua, 85*(2–3), 85–134.

Cole, P., & Hermon, G. (1998). The typology of wh-movement, wh-questions in Malay. *Syntax, 1*(3), 221– 258.

Coppock, E. (2010). Parallel grammatical encoding in sentence productions: Evidence from syntactic blends. *Language and Cognitive Processes, 25*(1), 38-49.

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry, 40*(3), 446–486.

Corsi, P. M. (1972). Human memory and the medial temporal region of the brain. *Dissertation Abstracts International, 34* (2), 819B. (University Microfilms No. AAI05-77717).

Clahsen, H. (1982). *Spracherwerb in der Kindheit: Eine Untersuchung zur Entzicklung der Syntax bei Kleinkindern [Language Acquisition in Childhood: A study of the Acquisition of Syntax in Children)*, Narr, Tübingen.

Clahsen, H. (1993/1994). Functional Categories in Early Child German. *Language Acquisition. 3*(4), 395-429.

Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. Dowty, L. Kartunnen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 94–128). Cambridge: Cambridge University Press.

Cunnings, I. & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes, 28,* 188-219.

Dayal, V. (1994). Scope marking as indirect wh-dependency. *Natural Language Semantics, 2*(2), 137–170.

Dayal, Veneeta. (2000). Scope Marking: Cross-Linguistic Variation in Indirect Dependency. In Uli Lutz, Gereon Müller, & Arnim von Stechow (eds.), *Wh-Scope Marking*. (pp.156-193).

Delage, H. & Frauenfelder, U.H. (2019). Relationship between working memory and complex syntax in children with Developmental Language Disorder. *Journal of Child Language*, 1-33.

de Villiers, J., Kotfila, J., & Klein, M. (2019). Parsing, pragmatics, and representation: Children's comprehension of two-clause questions. In T. Ionin & M. Rispoli (Eds.), *Three streams of generative language acquisition research*. (pp. 85–105). Amsterdam: John Benjamins.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*, 1037–1060.

de Villiers, J., & Roeper, T. (1995). Relative clauses are barriers to wh-movement for young children. *Journal of Child Language, 22*(2), 389–404.

de Villiers, J., Roeper, T., Bland-Stewart, L., & Pearson, B. (2008). Answering hard questions: Wh- movement across dialects and disorder. *Applied Psycholinguistics, 29*(1), 67–103.

de Villiers, J., Roeper, T., & Vainikka, A. (1990). The acquisition of long-distance rules. In L. Frazier & J. de Villiers (Eds.), *Language processing and language acquisition* (pp. 257–297). Dordrecht, the Netherlands: Kluwer .

Dell, G.S. (1986). A Spreading-Activation Theory of Retrieval in Sentence Production. *Psychological Review, 93*(3), 283-321.

Dell, G.S. & Reich, P.A. (1981). Stages in Sentence Production: An Analysis of Speech Error Data. *Journal of Verbal Learning and Verbal Behavior, 20,* 611-629.

Demirdache, H., & Oiry, M. (2007). On the felicity condition for long-distance questions in L1 acquisition. In H. C. Nulton, S. Kulatikake, & I. Woo (Eds.), *Proceedings of BUCLD 31*, 184–195. Somerville, MA: Cascadilla Press.

Dickey, M. W. (1996). Constraints on the sentence processor and the distribution of resumptive pronouns. In M. W. Dickey & S. Tunstall (Eds.), *Linguistics in the laboratory* (pp.157–192). Amherst: University of Massachusetts.

Diessel, H. (2004). *The acquisition of complex sentences.* Cambridge: Cambridge University Press.

Diessel, H., & Tomasello, M. (2000). Why complement clauses do not include a that-complementizer in early child language. *Annual Meeting of the Berkeley Linguistics Society, 25*(1), 86–97.

Elliot, C.D. (2007). *Differential Ability Scales – Second Edition*. San Antonio, TX: Pearson.

Fanselow, Gisbert & Mahajan, Anoop. (2000). Towards a Minimalist Theory of Wh-Expletives,Wh-Copying, and Successive Cyclicity. In Uli Lutz, Gereon Müller, & Arnim von Stechow (eds.), *Wh-Scope Marking.* (pp. 195-230).

Fay, D. (1982). Substitutions and splices: A study of sentence blends. In A. Cutler (Ed.), *Slips of the tongue and language production*. (pp.163-195).

Felser, C. (2004). Wh-copying, phases, and successive cyclicity. *Lingua, 114*(5), 543–574.

Ferreira , F. & Henderson, J.M. (1991). Recover from misanalyses of garden-path sentences. *Journal of Memory and Language, 30,* 725-745.

Ferreira, F., & Swets, B. (2005). The production and comprehension of presumptive pronouns in relative clause “island” contexts. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 263–278). Mahwah, NJ: Lawrence Erlbaum Associates.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology, 40*(4), 296–340.

Fodor, J. D. (1998). Parsing to learn. *Journal of Psycholinguistic Research, 27*(3), 339–374.

Franck, J., Soare, G., Frauenfelder, U. H., & Rizzi, L. (2010). Object interference in subject–verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language, 62*(2), 166–182.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language & Linguistic Theory, 5*(4), 519–559.

Frazier, L. (2015). Do null subjects (mis-) trigger pro-drop grammars? *Journal of Psycholinguistic Research, 44*(6), 669–674.

Friedmann, N., Belletti, A., and Rizzie, L. (2009) Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua,119,* 67-88.

Garrett, M. F. (1980). Levels of processing in sentene production. In B.L. Butterworth (Ed.), *Language production, Volume I: Speech and talk.*177-220.

Gavruseva, E., & Thornton, R. (2001). Getting it right: Acquisition of whose-questions in child English. *Language Acquisition, 9*(3), 229–267.

Geçkin, V., Thornton, R., & Crain, S. (2017). Children’s interpretation of disjunction in negative sentences: A comparison of Turkish and German. *Language Acquisition, 25,* 1–16.

Gerken, L. (1996). Prosody’s role in language acquisition and adult parsing. *Journal of Psycholinguistic Research, 25,* 345–356.

Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax, 7*(1), 55–78.

Goldrick, M., Putnam, M., & Schwarz, L. (2016). Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition. 19*(5). 857-876.

Goro, T., & Akiba, S. (2004). The acquisition of disjunction and positive polarity in Japanese. In V. Chand, A. Kellehar, A. J. Rodriguez, & B. Schmeiser (Eds.), *Proceedings of the 23rd west coast conference of formal linguistics*. 251–264. Somerville, MA: Cascadilla Press.

Grimshaw, J. (1993). *Minimal Projection, Heads, and Optimality*. Rutgers University manuscript.

Grimshaw, J. (1997). Projection, Heads, and Optimality. *Linguistic Inquiry 28* (3), 373-422.

Grolla, E., & Lidz, J. (2018). A performance account for medial wh-questions in child English. InA. B. Bertolini & M. J. Kaplan (Eds.), *Proceedings of the 42nd annual Boston University conference on language development.* 289–302. Somerville, MA: Cascadilla Press.

Gualmini, A., Hulsey, S., Hacquard, V., & Fox, D. (2008). The question-answer requirement for scope assignment. *Natural Language Semantics, 16*(3), 205–237.

Hamblin, C. L. (1973). Questions in Montague English, *Foundations of Language 10*, 41-53.

Hamburger, H., & Crain, S. (1982). Relative acquisition. In S. Kuczaj (Ed.), *Language Development, 5*, 245–274. Hillsdale, NJ: Lawrence Erlbaum.

Herburger, E. (1994). *A semantic Difference between Full and Partial Wh-Movement in German*. Paper presented at the Annual LSA Meeting, Los Angeles.

Horvath, Julia. (1995). Partial Wh-Movement and Wh-"Scope-Markers". In István Kenesei (ed.), *Levels and Structures,* (pp. 69-124). Szeged: JATE. (Approaches to Hungarian 5.)

Horvath, J. (1997). The status of 'wh-expletives' and the partial wh-movement construction of Hungarian. *Natural Language & Linguistic Theory, 15*(3), 509–572.

Horvath, J. (1998). Multiple wh-phrases and the wh-scope-marker strategy in Hungarian interrogatives. *Acta Linguistica Hungarica, 45*(1), 31-60.

Horvath, Julia. 2000. On the Syntax of "Wh-Scope Marker" Constructions: Some Comparative Evidence. In Uli Lutz, Gereon Müller, & Arnim von Stechow (eds.), *Wh-Scope Marking.* 271-316.

Hurewitz, F., Brown-Schmidt, S., Thorpe, K., Gleitman, L. R., & Trueswell, J. C. (2000). One frog, two frog, red frog, blue frog: Factors affecting children's syntactic choices in production and comprehension. *Journal of Psycholinguistic Research, 29*(6), 597–626.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Hyams, N., & Wexler, K. (1993). On the grammatical basis of null subjects in child language. *Linguistic Inquiry, 24*(3), 21–459.

Jager, J. (2005). *Kids' Slips: What young children's slips of the tongue reveal about language development.* Psychology Press.

Jakubowicz, C., & Strik, N. (2008). Scope-marking strategies in the acquisition of long distance wh-questions in French and Dutch. *Language and Speech, 51*(1–2), 101–132.

Johnson, M.K., McMahon, R.P., Robinson, B.M., Harvey, A.N., Hahn, B., Leonard, C.J., Luck, S.J., Gold, J.M. (2013). The relationship between working memory capacity and broad measures of cognitive ability in healthy adults and people with schizophrenia. *Neuropsychology. 27*(2), 220-229.

Kam, C. L. H., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology, 59*(1), 30–66.

Kandybowicz, J. (2006). Comp-trace effects explained away. In D. Baumer, D. Montero, & M. Scanlon (Eds.), *Proceedings of the 25th west coast conference on formal linguistics*. 220–228. Somerville, MA: Cascadilla Proceedings Project.

Kazanina, N., & Phillips, C. (2001). Coreference in child Russian: Distinguishing syntactic and discourse constraints. In A. H. J. Do, L. Dominguez, & A. Johansen (Eds.), *Proceedings of the 25th Boston University conference on language development*. 413–424. Somerville, MA: Cascadilla Press.

Kiss, K.E. (1998). Identificational focus versus information focus. *Language, 74*(2), 245-273.

Kroch, A. (1981). On the role of resumptive pronouns in amnestying island constraint violations. In R. Hendrick, C. Maseh, & M. Miller (Eds.), *Papers from the 17th regional meeting of the Chicago Linguistic Society*.125–135. Chicago, IL: Chicago Linguistic Society.

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R and BUGS (2nd ed.)*. New York: Academic Press.

Labelle, M. (1990). Predication, wh-movement, and the development of relative clauses. *Language Acquisition, 1*(1), 95–119.

Lahiri, Utpal. (2002). On the Proper Treatment of “expletive-wh” in Hindi. *Lingua, 112*, 501-540.

Larson, R. (1988). On the double object construction. *Linguistic Inquiry, 1*, 335-391.

Lassotta, R., Omaki, A., & Franck, J. (2016). Developmental changes in misinterpretation of garden-path wh- questions in French. *The Quarterly Journal of Experimental Psychology, 69*(5), 829–854.

Legendre, G., Wilson, C., Smolensky, P., Homer, K., & Raymond, W. (1995). “Optimality and wh-extraction,” in J. Beckman, L. Walsh Dickey, and S. Urbanczyk (Eds), *Papers in Optimality Theory* (pp. 607-636) (University of Massachusetts Occasional Papers,18). Amherst, MA: GLSA Publications.

Legendre, Géraldine, Paul Smolensky, and Colin Wilson. 1998. When is Less More? Faithfulness and Minimal Links in Wh-Chains. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis, and D. Pesetsky (eds), *Is the Best Good Enough? Optimality and Competition in Syntax*. (pp. 249-289).MIT Press.

Legendre, G. (2001). An Introduction to Optimality Theory in Syntax. In. G. Legendre, J. Grimshaw, &S. Vikner (Eds.) *Optimality Theoretic Syntax.* (pp. 1-20). MIT Press.

Legendre, G., Vainikka, A., Hagstrom, P., & Todorova, M. (2002). Partial constraint ordering in child French syntax. *Language Acquisition, 10*(3), 189–227.

Legendre, G., Hagstrom, P., Chen-Main, J., Tau, L., & Smolensky, P. (2004). Deriving output probabilities in child Mandarin from Dual-Opimization grammar. *Lingua, 114*, 1147-1185.

Legendre, Géraldine.2010. The Neutralization approach to ineffability in syntax. In C. Rice & S. Blaho (Eds.) *Modeling Ungrammaticality in Optimality Theory.* Equinox. 237-266.

Legendre, G., Putnam, M., de Swart, H., Zaroukian, E. (eds.).(2016a). *Optimality-theoretic syntax, semantics, and pragmatics: From uni-to bidirectional optimization.* Oxford: Oxford University Press.

Levelt, W.J. M, (1989). *Speaking: From Intention to Articulation*. MIT Press. Cambridge, Massachusetts.

Lewis, S., Hacquard, V., & Lidz, J. (2017). “Think” pragmatically: Children’s interpretation of belief reports. *Language Learning and Development, 13*(4), 1–23.

Lidz, J., & Gagliardi, A. (2015). How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics, 1*, 333–353.

Lutz, U., & Müller, G. (Eds.). (1996). *Papers on wh-scope marking*. Stuttgart, Germany: Universitat Stuttgart.

Lutz, U., Müller, G., & von Stechow, A. (Eds.). (2000). *Wh-scope marking.* Philiadelphia, PA: John Benjamins.

Lutken, C.J. and Omaki, A. (2017). Production Comprehension Asymmetry in Children’s Medial Wh-Questions. *Boston University Conference on Language Development 42.* Oral Presentation.

Lutken, C.J., Legendre, G., and Omaki, A. (2020). Syntactic Creativity in Children’s Wh-Question. *Cognitive Science.* 44(7). E12849.

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrene Erlbaum Associates.

Mahajan, Anoop. (2000). Towards a Unified Treatment of WH-Expletives in Hindi and German. In Uli Lutz, Gereon Müller, & Arnim von Stechow (eds.), *Wh-Scope Marking.* 317-332

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T. J., Xu, F., & Clahsen, H. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development, 57*(4), 1–178.

McCloskey, J. (1990). Resumptive pronouns: A-binding and levels of representation in Irish. In R. Hendrick (Ed.), *Resumptive pronouns at the interface* (pp. 199–248). Philadelphia, PA: John Benjamins.

McDaniel, D. (1989). Partial and multiple wh-movement. *Natural Language and Linguistic Theory, 7*(4), 565–604.

McDaniel, D., Chiu, B., & Maxfield, T. L. (1995). Parameters for wh-movement types: Evidence from child English. *Natural Language and Linguistic Theory, 13*(4), 709–753.

McDaniel, D., McKee, C., Cowart, W., & Garrett, M. F. (2015). The role of the language production system in shaping grammars. *Language, 91*(2), 415–441.

McDaniel, D., McKee, C., & Garrett, M. F. (2010). Children's sentence planning: Syntactic correlates of fluency variations. *Journal of Child Language, 37*(1), 59–94.

McKee, C., & McDaniel, D. (2001). Resumptive pronouns in English relative clauses. *Language Acquisition, 9*(2), 113–156.

Murphy, A. (2016). What copying (doesn't) tell us about movements: Remarks on the derivation wh-copying in German. In K. Barnickel, M. Guzm an, J. Hein, S. Korsah, A. Murphy, L. Paschen, Z. PuVskar, & J. Zaleska (Eds.), *Replicative processes in grammar* (pp. 149–188). Linguistische Arbeits Berichte. Leipzig, Germany: Institut fur Linguistik, Universitat Leipzig.

Müller, Gereon. (1997). Partial Wh-Movement and Optimality Theory. *The Linguistic Review, 14*, 249-306.

Musolino, J., & Lidz, J. (2006). Why children aren't universally successful with quantification. *Linguistics, 44*(4), 817–852.

Newmeyer, F. J. (2005). Possible and probable languages: A generative perspective on linguistic typology. New York: Oxford University Press.

Omaki, A., Davidson White, I., Goro, T., Lidz, J., & Phillips, C. (2014). No fear of commitment: Children's incremental interpretation in English and Japanese wh-questions. Language Learning and Development, 10 (3), 206–233.

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition, 22*(2), 158–192.

Orfitelli, R., & Hyams, N. (2012). Children's grammar of null subjects: Evidence from comprehension. *Linguistic Inquiry, 43*(4), 563–590.

Pearl, L., & Sprouse, J. (2013). Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition, 20*(1), 23–68.

Pesetsky, D. (1997). Optimality Theory and Syntax: Movement and Pronounciation. In D. Archangeli and D.T. Langendoen (eds), *Optimality Theory, An Overview*. (134-170). Blackwell Publishers.

Phillips, C. &Erehnhofer, L. (2015) The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism. 5*(4). 409-453.

Pinker, S. (1989). *Learning, development, and conceptual change. Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: The MIT Press.

Poeppel, D. & Wexler, K. (1993). The Full Competence Hypothesis of Clause Structure in Early German. *Language, 69,* 1-33.

Pozzan, L., & Valian, V. (2017). Asking questions in child English: Evidence for early abstract representations. *Language Acquisition, 24*(3), 209–233.

Prince, A. & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar.* Oxford: Blackwell. (1st edn. 1993. Technical report CU-CS-696-93, University of Colorado at Boulder and Technical report ROA 537, Rutgers University, New Brunswick, NJ.)

R Core Development Team. (2015). R: *A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rawlins, K. (2015). A Case of Compositional Scope Marking in English. Johns Hopkins

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica 27*, 53-94.

Repp, Sophie. 2009. Defining 'contrast' as an information-structural notion in grammar. *Lingua 120*, 1333-1345.

Riggle, J. (2004). *Generation, Recognition, and Learning in Finite State Optimality Theory*. Ph.D. thesis, University of California, Los Angeles.

Riggle, J. (2009). *Generating Contenders.* Rutgers Optimality Archive, 1044.

Riemsdijk, Henk van. 1983. Correspondence Effects and the Empty Category Principle. In Y. Otsu et al. (eds.), *Studies in Generative Grammar and Language Acquisition,* 5-16. Tokyo.

Rizzi, L. (1982). *Issues in Italian syntax*. Dordrecht: Foris Publications.

Rizzi, L. (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.

Rizzi, L. (1996). Residual verb second and the Wh-criterion. In A. Belletti & L. Rizzi (Eds.), *Parameters and functional heads: Essays in comparative syntax* (pp. 63–90). New York: Oxford University Press.

Roberts, C. (2012). Information structure in discourse: Towards and integrated formal theory of pragmatics. *Semantics & Pragmatics, 5*(6), 1–69.

Roeper, T., & de Villiers, J. G. (2011). The acquisition path for wh-questions. In J. G. de Villiers & T. Roeper (Eds.), *Handbook of generative approaches to language acquisition* (pp. 189–246). Dordrecht, the Netherlands: Springer.

Sabel, J. (2000). Expletives as Features. In. R. Billerey et al. (Eds.) *WCFL 19 Proceedings.* 411-424. Somerville, MA: Cascadilla Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Samek-Lodovici, V., & Prince, A. (2002). Fundamental properties of harmonic bounding. *RuCCS-TR, (71).*

Schulz, B. (2011). Syntactic creativity in second language English: Wh-scope marking in Japanese–English interlanguage. *Second Language Research, 27*(3), 313–341.

Schütze, C., & Sprouse, J. (2014). Judgment data. In R. Podesva & D. Sharma (Eds.), *Research Methods in Linguistics* (pp. 27-50). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139013734.004

Sells, P. (1984). *Syntax and semantics of resumptive pronouns*. Retrieved from ProQuest Digital Dissertations, AAI8500135.

Shatz, M. & Gelman, R. (1973). The Development of Communication Skills: Modifications in the Speech of Young Children as Function of Listener. *Monographs of the Society for Research in Child Development. 38*(5). 1-38.

Shlonsky, U. (1992). Resumptive pronouns as a last resort. *Linguistic Inquiry, 23*(3), 443–468.

Smolensky, P. (1996b). On the comprehension/production dilemma in child language. *Linguistic Inquiry, 27,* 720-731.

Snedeker, J. (2013). Children's sentence processing. In R. P. G. von Gompel (Ed.), *Sentence processing* (pp.189–220). New York: Psychology Press.

Snyder, W. (2007). Child language: The parametric approach. New York: Oxford University Press. Stepanov, A. (2000). Wh-scope marking in Slavic. *Studia Linguistica, 54*(1), 1–40.

Stepanov, A., & Stateva, P. (2006). Successive cyclicity as residual wh-scope marking. *Lingua, 116*(12), 2107–2153.

Sturt, P. (2007). Semantic Re-Interpretation and garden path recovery. *Cognition, 105*, 477-488.

Stowe, L. (1986). Parsing wh-constructions: Evidence for online gap location. *Language and Cognitive Processes, 1*(3), 227–245.

Stromswold, K. (1995). The acquisition of subject and object wh-questions. *Language Acquisition, 4*(1–2), 5–48.

Szabolcsi, A. (1983). The possessor that ran away from home. *The Linguistic Review, 3*(1), 89–102.

Thornton, R. J. (1990). *Adventures in long-distance moving: The acquisition of complex wh-questions*. Retrieved from Doctoral Dissertations, AAI9109855.

Thornton, R. (1995). Referentiality and wh-movement in child English: Juvenile d-linkuency. *Language Acquisition, 4*(1–2), 139–175.

Thornton, R., & Crain, S. (1994). Successive cyclic movement. In T. Hoekstra & B. Schwartz (Eds.) *Language acquisition in generative grammar* (pp. 215–252). Philadelphia, PA: John Benjamins.

Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry, 15*(1), 103–129.

Ure, J., & Ellis, J. (1977). Register in descriptive linguistics and linguistic sociology. *Issues in sociolinguistics*, 197-243.

Valian, V. (1990). Null subjects: A problem for parameter-setting models of language acquisition. *Cognition, 35*(2), 105–122.

Valian, V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition, 40* (1–2), 21–81.

Viau, J., Lidz, J., & Musolino, J. (2010). Priming of abstract logical representations in 4-year-olds. *Language Acquisition, 17*(1–2), 26–50.

von Stechow (2000). Partial Wh-Movement, Scope Marking, and Transparent Logical Form. In Uli Lutz, Gereon Müller, & Arnim von Stechow (eds.), *Wh-Scope Marking*.447-478.

Wagner, K.R. (1974). *Die Sprechsprache des Kindes. Teil1: Theorie un Analyse*. Düsseldorf: Präger.

Wagner, Michael. (2012). Contrastive Topics Decomposed. *Semantics & Pragmatics. 5* (8), 1-54.

Wagner, L., Greene-Havas, M., & Gillespie, R. (2010) Development in Children's Comprehension of Linguistic Register. *Child Development. 81*(6).

Wilde, N.J., Strauss, E. & Tulsky, D.S. (2004) Memory Span on the Wechsler Scales, *Journal of Clinical and Experimental Neuropsychology, 26*(4), 539-549.

Willis, C & Gathercole, S.E. (2001). Phonological short-term memory contributions to sentence processing in young children. *Memory, 9*(4-6). 349-363.

Wilson, C. (2001). Bidirectional Optimization and the theory of anaphora. In: Legendre, G., Grimshaw, J., Vikner, S. (Eds) *Optimality Theoretic Syntax*. MIT Press, Cambridge. MA. 465-507.

Wode, H. (1971). Some stages in the acquisition of questions by monolingual children. *Word*, *27*(1-3), 261-310.

Wurmbrand, S. (1998). *Infinitives.* PhD Dissertation. MIT

Yang, C. (2002). *Knowledge and learning in natural language*. New York: Oxford University Press.

Yang, C. (2004). Universal grammar, statistics or both? *Trends in Cognitive Sciences, 8*(10), 451–456.

Carolyn Jane Lutken

E-mail : clutken2@jhu.edu

Phone : 662-638-5858

Website: cjlutken.com

Education

Johns Hopkins University: Ph.D. Cognitive Science, defended Dec. 9, 2020
Dissertation Title: *Cross-Linguistic Investigations of 'Syntactic Creativity' Errors in Children's WH-Questions*
Primary Advisor: Geraldine Legendre, Ph.D.

JHU Summer Teaching Institute: May 2019 – *An intensive 3-day course on teaching methods, philosophies, and techniques. Put on by the JHU Teaching Academy, this Institute's goal is to explore the benefits of active learning, ongoing assessment, and responsiveness to diversity.*

Newcastle University: Graduated with Distinction
Master of Arts in Linguistics and Language Acquisition 2013

Whitman College:
Graduated Cum Laude 2010
Bachelor of Arts in English Literature
Minor in French Language

University of Nantes, France - one semester studying abroad 2008

Professional Experience

Undergraduate Teaching Experience

Upcoming: Summer 2021	*Psycholinguistics* Johns Hopkins University
Currently: Spring 2021	*Psycholinguistics* Johns Hopkins University
Spring Intersession 2020	*Learning from their Mistakes: Experimental Approaches to Language Learning* Johns Hopkins University

Guest Lectures

Spring 2020	Guest Lecture: *First Language Acquisition* For *Julia Yarmolinskaya,* Johns Hopkins University
Fall 2019	Guest Lecture: *Experimental Methods in Psycholinguistics* For *Bettina Braun,* University Konstanz

Fall 2019	Guest Lecture: *Language Development* For *Celia Litovsky*, Goucher College Prison Program
Spring 2019	Guest Lecture: *Bilingualism in the Brain* For *Julia Yarmolinkaya*

Teaching Assistantships

Spring 2018	Bilingualism | *Julia Yarmolinskaya* Guest lecture: *Bilingual Acquisition* Cognitive Science Department, Johns Hopkins University
Fall 2017	Language and Mind | *Colin Wilson* Guest Lecture: *Language Processing* Cognitive Science Department, Johns Hopkins University
Spring 2017	Bilingualism | *Julia Yarmolinskaya* Cognitive Science Department, Johns Hopkins University
Fall 2016	Language and Mind | *Julia Yarmolinskaya* Guest Lecture *Introduction to Syntax* Cognitive Science Department, Johns Hopkins University
Fall 2015	Cognition | *Colin Wilson* Cognitive Science Department, Johns Hopkins University

Research Interests

Cognitive Development
General cognitive and processing mechanisms underlying language and language learning
First language acquisition of syntax, particularly complex structures
Cross linguistic variation in acquisition and structure use
Syntax and the Syntax-Semantics Interface

Grants

NSF-Doctoral Dissertation Research Improvement Grant (Co-PI): *Cross-Linguistic Investigations of Syntactic Creativity Errors in Children's Wh-questions*. G. Legendre (PI). BCS-1853297. 2019-2020. Amount (DC): $11,453

International Research Experience

Sept.-Nov. 2020	Conducted Research at the University of Konstanz with Bettina Braun at BabySprachLabor. This research was part of the DDRIG listed above.

Sept. -Dec. 2013 Research Assistant at Newcastle University: *How does linguistic input in infancy inform language change?* | Joel Wallenberg (PI)

Publications

2020 Lutken, C.J., Legendre, G., and Omaki, A. (2020). Syntactic Creativity in Children's Wh-Question. *Cognitive Science*. 44(7). E12849.

In Preparation

Lutken, C.J., Legendre, G. Pragmatic Constraints on the use of Wh-Scope Marking vs. Long Distance constructions cross-linguistically.

Lutken, C.J., Legendre, G. A cross-linguistic analysis of Scope Marking at the syntax/semantics interface using Optimality Theory.

Planned

Lutken, C.J. & Legendre, G. Cross-Linguistic evidence for processing difficulties in complex question formation: Evidence from German. (To be submitted to *Language Acquisition*)

Lutken, C.J. & Legendre, G. Correlation between medial wh-production and working memory ability: a processing explanation of a classic error. (To be submitted to *Language Acquisition*)

Research Presented at Professional Meetings

Invited Seminar and Conference Presentations

2019 Lutken, C.J. and Legendre, G. Syntax or Something Else? What causes systematic errors in children's production and comprehension of complex questions? Newcastle University, United Kingdom. Colloquium Talk.

2019 Lutken, C.J. and Legendre, G. Errors as Evidence: What children's syntactic errors can tell us about how they learn grammar. University of Konstanz, Germany. Invited Talk.

2019 Lutken, C.J. and Legendre, G. An Introduction to Optimality Theory: Survival of the Fit Enough. University of Pisa, Italy. Invited Talk.

2018 Lutken, C.J. Syntactic Creativity: what we can learn about language acquisition from children's errors? Stevenson University. Invited talk.

<u>Refereed by Abstract or Paper Submission:</u>

2021 Lutken, C.J. & Legendre, G. A Cross-linguistic comparison in children's interpretations of complex questions. Linguistics Society of America 95. Oral Presentation

2020 Lutken, C.J., & Legendre, G. Immature Syntax or Processing? What causes "scope marking errors" in English-Speaking 5-year-olds? Boston University Conference on Language Development 45. Boston, MA. Poster

2019 Lutken, C.J., Omaki, A., and Legendre, G. What do you think what's the cause of children's errors in biclausal questions? Boston University Conference on Language Development 44. Boston, MA. Poster.

2019 Lutken, C.J. and Legendre, G. An Optimality Theory analysis of Scope Marking at the syntax/semantics interface. Mid-Atlantic Colloquium of Studies in Meaning 8. New York, NY. Poster.

2019 Lutken, C.J. Legendre, G., and Omaki A. Syntax or something else: what is the source of systematic errors in children's complex questions? Society for Research in Child Development 19. Boston, MA. Poster.

2018 Lutken, C.J. and Omaki, A. Children's Medial Wh-questions: Asymmetry in comprehension and production. University of Delaware Linguistics and Cognitive Science Student Conference. Oral Presentation.

2017 Lutken, C.J. and Omaki, A. Production Comprehension Asymmetry in Children's Medial Wh-Questions. Boston University Conference on Language Development 42. Oral Presentation.

2017	Lutken, C.J. and Omaki, A. Imbalance in English Speaking Children's Production and Comprehension of Wh-Questions. Linguistics Society of America Summer Institute. Lexington, KY. Poster.
2017	Lutken, C. J. and Omaki, A. What do you think why American children produce Russian wh-questions?. CUNY Sentence Processing Conference. Boston, MA. Poster.

Other Honors

2010- Fulbright Recommended Scholar to the English Teaching Assistantship Program in France (one of 55 students nation-wide)

Girl Scout Gold Award

Other Relevant Skills

English – Native
French – Proficient

Other Teaching Experience

Jan.-May 2014	Tutor in all core subjects for 8^{th} grade student with Autism
Oct. 2010-May 2011	English Teaching Assistant in Dunkirk, France in two Middle schools
Sept. – Dec. 2008	Native Speaker- taught English lessons average of 6 hours a week at an elementary school in Nantes, France