ABSTRACT

Title of Dissertation: ISLAND CONSTRAINTS: WHAT IS THERE FOR CHILDREN TO LEARN?

Mina Robinson Hirzel
Doctor of Philosophy, 2022

Dissertation directed by: Professor Jeffrey Lidz
Department of Linguistics

Associate Professor Ellen Lau
Department of Linguistics

This dissertation presents behavioral studies that target the early syntactic representations of *wh*-movement during infancy and early childhood. Previous studies show that by 20 months-old, infants represent *wh*-movement and use this knowledge to respond to *wh*-questions during language comprehension tasks (Gagliardi 2012; Gagliardi et al., 2016; Seidl et al., 2003). Studies probing the nature of early representations of *wh*-movement show that by around 4 years-old, children represent island constraints (e.g., de Villiers et al., 1990; de Villiers & Roeper, 1995a, 1995b; Fetters & Lidz, 2016; Goodluck et al., 1992). It remains unclear how knowledge of *wh*-movement develops. What is the source of this 'empirical gap' between the onset of knowledge of *wh*- movement, and the observation that children respect island constraints? One possibility is that knowledge of island constraints is a component of

Universal Grammar (e.g., Chomsky 1965, 1986; Hornstein & Lightfoot 1981). In this case, the 'empirical gap' in the linguistic abilities of infants compared to young children isn't indicative of their linguistic knowledge, but rather the difficulties with testing infants and toddlers on complex syntax. Another possibility is that knowledge of island constraints is acquired via experience (e.g., Pearl & Sprouse, 2013). In this case, the 'empirical gap' reflects a knowledge gap, and there's no evidence for knowledge of island constraints during infancy because it has yet to be acquired. Experiment 1 shows that by 19 months-old, infants have knowledge of *wh*-movement, and use this knowledge during language comprehension. Results are consistent with recent work which shows that 18 month-olds, but not 17 month-olds, know that *wh*-phrases co-occur with gap positions in *wh*-object questions (Perkins & Lidz, 2021). Experiment 2 shows that 3 year-olds respect locality constraints on *wh*-movement in *wh*- questions, and Experiment 3 shows that adults behave as expected on this task. Experiments 4 and 5 test children and adults on locality constraints on *wh*- movement in relative clauses, but these results are inconclusive (likely due to difficulties with moving the task online). The results of Experiment 3 raises challenges for learning hypotheses of island constraints which emphasize the role of linguistic experience. Learning models which propose that linguistic experience is the key factor in the acquisition of island constraints must consider these behavioral results when estimating the amount of data that the learner needs to solve the acquisition problem. These behavioral results are consistent with the hypothesis that knowledge of island constraints is innate, but further work is needed to close the 'empirical gap' between the onset of knowledge of *wh*-movement and the onset of knowledge of island constraints.

ISLAND CONSTRAINTS: WHAT IS THERE FOR CHILDREN TO LEARN?

by

Mina Robinson Hirzel

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2022

Advisory Committee:
Professor Jeffrey Lidz, Chair
Associate Professor Ellen Lau, Co-chair
Associate Professor Yi Ting Huang, Dean’s Representative
Associate Professor Alexander Williams
Professor Emeritus Norbert Hornstein

Preface

This dissertation is built from collaborative research. Chapter 4 reports on collaborative work with Jeff Lidz. The study in Chapter 3 was done in collaboration with Laurel Perkins and Jeff Lidz. This work has been supported in part by the National Science Foundation (NRT NRT-DESE #1449815).

Acknowledgements

This work has benefited from many conversations with many wonderful people over the years. A big thank you to my advisors, Jeff Lidz and Ellen Lau for the hours and hours you put into helping me refine these experiments and think about the bigger questions. Jeff, it was so much fun to brainstorm new experiments with you! Thank you for always keeping your door open to talk about anything acquisition and for looking through countless drafts of everything. Ellen, thank you for bringing your perspective to this work over the years – my thinking and my scientific approach has changed so much from our discussions. Thank you for being so supportive and adaptive, and for helping me find a path forward when I was really, really stuck. I'm happy we could have meetings outside with the cicadas crawling everywhere.

Thank you to the rest of my committee, Alexander Williams, Norbert Hornstein and Yi Ting Huang, for lending your expertise to this work and your comments and positive energy. Alexander, these experiments work because of your keen eye to smooth out the jagged edges so that infants and children don't get bamboozled and lose control – thank you for all your suggestions, especially the ones that were jokes in disguise.

Thank you to the linguistics faculty at the University of Maryland, many of whom were my teachers and mentors throughout the years. I benefitted greatly from comments and feedback from Peggy Antonisse, Tonia Bleam, Naomi Feldman, Howard Lasnik, Paul Pietroski, Colin Phillips, Valentine Hacquard and Omer Preminger. Thank you to Omer for gently reminding me about various syntactic phenomena that I would inevitably forget as the semesters went by. A special thank you to Valentine for teaching me how to write and for your valuable feedback on these experiments, and for your tireless support of my work.

Thank you to Alan Munn and Cristina Schmitt for getting me into the glamorous world of linguistics! I had such a great time working in the language acquisition lab with you at Michigan State, I didn't want it to end. Thank you for all the energy you put into being amazing teachers and mentors, and for your continued support.

Thank you to Tara Mease, friend and Lab Manager. Thank you for your kind and honest and thoughtful and thorough feedback over the years on so many topics, and for all your work that has moved these studies forward. This would have been so incredibly difficult (and much less enjoyable) without you.

Thank you to my incredible RAs in the language acquisition lab, who put so much of their time into supporting this work. Thank you to Mariam Aiyad, Sterling Mullenix, Jiabao Zhou, Becca Lebovitz, Gabby Hochmuth, William Simpson, Abby Kameny, Lilly Righter and Alexander Shushunov. Mariam, I was so excited when you agreed to help me film these stimuli, it was a lot of fun and I'm so glad that we were able to work together on other projects as well. Sterling, thank you for all your help with stimuli editing and reorganization, and for your efforts getting the study online and collecting data. It's not easy to be an RA for a dissertating student, and I'm happy that we could work together. Thank you for all you've done to keep moving this project ahead. Jiabao, thank you for your work on this study when it was in its early stages, I always remember your enthusiasm when I'm presenting on it.

Thank you to Kim Kwok for your tireless work keeping the department running and for all the last minute fixes that you pulled for me, with no questions asked.

Thank you to the students at the Maryland language acquisition lab for all your feedback and support, Julie Gerard, Mike Fetters, Rachel Dudley, Laurel Perkins, Adam Liter, Hisao Kurokami, Dan Goodhue, Jack Yuanfan Ying, Imane Bou-Saboun and Yu'an Yang. Rachel, thank you for always advocating for my work, for giving me an appreciation for the finer things, and for sharing so much wisdom when I was a first year student…when I was writing, all your words came flooding back to me. Thank you Laurel for being so supportive of me, for always being there to bounce ideas off of and workshop the next new study. This work is vastly improved thanks to your contributions, at all stages. Yu'an your positive energy is contagious, and I admire you so much for building your research program from scratch. Thank you for being there to the bitter end!

Thank you to so many students, post docs, lab managers and Baggett Fellows at the Maryland Linguistics Department. This work has been touched by so many of you, and I really loved wandering around the department and chatting and thinking about linguistics together. Thank you to Paulina Lyskawa, Nick Huang, Suyoung Bae, Allyson Ettinger, Jeff Green, Kasia Hitczenko, Ted Levin, Craig Thorburn, Aura Cruz Heredia, Jackie Nelligan, Maxime Papillon, Annemarie van Dooren, Chris Baron, Jon Burnsky, Masato Nakamura, Hanna Muller, Anton Malko, Lara Ehrenhofer, Zoe Ovans, Phoebe Gaston, Chia-Hsuan Liao, Zoe Schlueter, Mike McCourt, Jad Wehbe, Rachel Dudley, William Matchin, Luisa Seguin, Clara Cuonzo, and Jéssica Viana Mendes. Thank you Zoe for the adorable and delicious dissertation cake.

Maxime, thank you for all the late-night linguistics, teaching, and writing consultations in the Solarium – and the chats about nothing in particular. Since I met you, sometimes I cook without measuring any of the ingredients. For me, the Solarium lost its spark after you left.

Thank you Annemarie for bringing so much joy to the days we spent together in the department and at conferences/workshops, no matter what happens I always end up laughing when we're together.

Thank you to my amazing cohort, Sigwan Thivierge, Anouk Dieuleveut, Aaron Doliana, Rodrigo Ranero and Tyler Knowlton (SMAART). I still remember that first year with you all so clearly. I have had so many linguistics conversations with each of you about this work as it developed, and I love to think back on all the breakfasts, late nights, trips and parties we experienced together. I admire you all so much, as brilliant linguists and as dear friends. Sig, thank you for the hours watching fireflies twinkle in the swale, the hours spent watching our houseplants grow new leaves, and the hours walking home and sitting on the bus together. Anouk, thank you for sharing your baking gifts and your love of nature with me, and for finally convincing me to come to the pool. Aaron, thank you for showing me a career path forward that I didn't know I could take, for grabbing an espresso with me whenever, and for sharing your innovative ideas with me. Rodrigo, thank you for telling it like it really is (it's a talent), for all the late night scary movies, and for your hilarious stories. Tyler, thank you for sharing a desk clump with me at the office for years, and for your unmatched "Yes, and!" energy and willingness to jump into the next big project. Even if the next big project is making an encyclopedic poster of office mugs.

Thank you to the philosophy boys for taking me under your wing when I needed a place to stay, Andrew Knoll, Chris Vogel, Brendan Ritchie and Christian Tarsney. Answering your roommate-wanted ad is by far my best Craigslist transaction – I loved living at Queensbury House with you all. Thank you to my covid roommates Paulina, Rodrigo, Sig, Ted, and Jad at Hill House for keeping it together and getting creative with entertainment while stuck at home. Thanks Tedrigo for making breakfast a full-on experience…I still hear the Suspiria Theme sometimes when I have my morning coffee. Paulina, thank you for your amazing birds-eye perspective on all things, and for helping me start a garden in the front yard, my sweet escape from everything. Thank you to my post-covid roommates Luisa, Clara, Jéssica and Imane for bringing amazing energy into the house (and into my life!) during my dissertating year, and for being good friends to me even though I only stayed for a semester.

Thank you to my parents, Andy and Tamara Hirzel for your continued support during school (all these years of school!). Thank you Papa, I got my love of languages from you and your endless word games. Thank you to my Mom and my sisters Grace and Sophie for listening to my linguistics ramblings and being excited for me not because you found it interesting, but because I was excited. Your love and support means so much to me. Thank you to my Aunt Amy and Uncle Michał, who always wanted to know what I was up to at school, and to Grandpa Rudy, who never lost track of how many years I had to go.

Jad, thank you for your kind and sincere feedback and support as I finish this

document. You are a light in my life, and I'm proud to share this moment with you. Thank you for being there for me during some very difficult times these past years, and for creating some amazing memories with me. You bring out the best in me, and I love you dearly.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Each discovery about the complexity and richness of the linguistic system invites questions about how children acquire linguistic knowledge. The more abstract the linguistic discovery, the more difficult these questions are to answer. Island constraints are a classic case. For example, (1) is a *wh*- question, and (2) shows that *wh*- movement holds from embedded clauses. Example (3) shows that there are certain limitations on the application of *wh*- movement, called *island constraints* (Ross, 1967).

1. What did Mariam hug?
2. What did he say [$_{CP}$ that Mariam hugged]?
3. *What did Sterling investigate the statement [$_{CP}$ that Mariam hugged]?

How do children acquire knowledge that movement from the clause *that Mariam hugged* is allowed in (2), but not in (3)? It's not feasible for children to rely on their linguistic experience to acquire syntactic constraints because their experience would have to include negative evidence, or information about utterances that the grammar does not generate (namely, a labeled, ungrammatical sentence). Knowledge of island constraints is unconscious linguistic knowledge, and untrained speakers are often surprised when confronted with island-related observations about their own language, so it's untenable that parents provide these kinds of linguistic data points to their children (and even if they did, children would likely ignore it (e.g., Pinker, 1989)). But even if children did have access to this information, direct observation can only confirm or disconfirm generalizations about particular linguistic expressions. Conditions on syntactic rules can only be confirmed or disconfirmed by grammars (e.g., Chomsky, 1973, 1981).

Since island constraints cannot be determined on the basis of linguistic experience alone, it must be the case that children themselves are responsible for islanc constraints. In other words, children are innately disposed to acquire grammars with island constraints. This is an example of *an argument from the poverty of the stimulus* (Chomsky, 1976; Hornstein & Lightfoot, 1981; Lidz, 2018; Lidz & Lasnik, 2017; Pearl, 2021). It is irrefutable that children acquire grammars by virtue of some abstract properties of the mind, and this dissertation is committed to the hypothesis that at least some of these innate properties are fundamentally *linguistic properties* (i.e., Universal Grammar).

From this theoretical perspective, it is not the syntactic constraints that children acquire, but instead the linguistic form that *wh-* movement takes in their language, as well as the form of the islands themselves (e.g., Otsu, 1981). Consider sentence (4a), which is remarkably similar to (4b). Despite the surface similarities of the two examples, the syntactic structure of (4a) and (4b) are different. *Wh-* movement reveals this structural difference: when *the bear* is moved from the embedded clause in (4a), an island violation obtains (5). When *the bear* is moved from the embedded clause in (4b), the resulting *wh-* question is acceptable (6). Adult speakers know that it's possible to extract *the bear* in (4a) but not (4b) because they know these strings have different structures. For example, they also know that *girl* does not select for clausal complements, but *statement* does.

4.

a. Sterling investigated the statement that Mariam hugged the bear

b. Sterling told the girl that Mariam hugged the bear

c. NP VP NP CP NP VP NP

5. *What did Sterling investigate the statement that Mariam hugged?

6. What did Sterling tell the girl that Mariam hugged?

It's easy to simulate what it would be like to lack this knowledge (i.e., put yourself in a child's situation) by playing Lila Gleitman's "Great Verb Game" and replacing these words with nonsense (7). The lexical information here disambiguates the structure – without it, it's impossible to determine whether (7) has the structure of (7a) (compare to 4a), which blocks extraction, or (7b) (compare to 4b), which allows extraction.

7. Sterling gorped the blicket that Mariam hugged the bear

   a. [$_{TP}$ [$_{NP}$ Sterling] [$_{VP}$ gorped [$_{NP}$ the blicket [$_{CP}$ that [$_{TP}$ [$_{NP}$ Mariam] [$_{VP}$ hugged [$_{NP}$ the bear]]]]]]]

   b. [$_{TP}$ [$_{NP}$ Sterling] [$_{VP}$ gorped [$_{NP}$ the blicket] [$_{CP}$ that [$_{TP}$ [$_{NP}$ Mariam] [$_{VP}$ hugged [$_{NP}$ the bear]]]]]

Children must learn to represent the syntax of these sentences, and only then will it be possible to determine whether they have acquired knowledge of island constraints. If children are tested for this linguistic knowledge prematurely, then they might appear to have selected a grammar without island constraints, when in fact they have simply mis-parsed the sentence. Similarly, island constraints might appear to vary cross-linguistically – not because some human grammars have different conditions on syntactic rules, but because the syntactic analysis of the construction is not what it seems.

Some acquisition models reject the premise that *no* linguistic data is available which would guide children to acquire island constraints. Pearl & Sprouse (2013) devises a learner which makes use of *indirect negative evidence* to induce island constraints by determining that structures which are completely absent from the linguistic input (i.e., have a probability of zero) are ungrammatical. This learner begins with several “innate” linguistic biases, but these biases are weak enough to accommodate cross-linguistic variation of island constraints. This means that the island constraints which the learner acquires are closer to the list of linguistic generalizations than they are to conditions on syntactic rules. This dissertation contributes to more rigorous learning models of island constraints through the study of the early syntactic representations of children and infants. Certain aspects of the Pearl & Sprouse (2013) learning model have been estimated based on the results of behavioral research, like the quantity and quality of the linguistic input. The input quantity is calculated based on the age at which children have demonstrated knowledge of island constraints in experimental settings – around 5 years-old (e.g., de Villiers et al., 1990; de Villiers & Roeper, 1995). However, there has been no targeted effort to identify whether children younger than 4 years-old have acquired knowledge of island constraints. It’s possible that the age range used to calibrate the training data set for the model does not reflect reality, and is an overestimate of the quantity of linguistic data made available to the learner.

Given concerns that the training data set is already relatively small (Phillips, 2013), it’s reasonable that an even smaller linguistic input might affect the learning outcomes of the model, or the kinds of linguistic biases required for the model to succeed. Chapter 4 presents a behavioral experiment which shows that 3 year-olds respect island constraints, which indeed suggests that the (Pearl & Sprouse, 2013) learner overestimates the input quantity. In ongoing work, a similar experiment is used to probe for knowledge of island constraints in 2 year-olds, with the goal of determining the earliest age at which knowledge of island constraints is acquired. As previously discussed, knowledge of syntax is a somewhat limiting factor for testing knowledge of island constraints during infancy: if infants cannot accurately parse the sentence, then there's no hope of observing the effects of knowledge of island constraints (see Otsu, 1981 for relevant discussion). Since children acquire verb argument structure between 1 and 2 years-old, testing children younger than 2 for knowledge of island constraints is somewhat precarious.

In Chapter 3, I address the empirical validity of a precondition on learning island constraints for the Pearl & Sprouse (2013) learner: children accurately parse *wh*-questions. Previous findings suggest that children acquire knowledge of *wh*-movement around 18 to 20 months-old (Gagliardi et al., 2016; Perkins & Lidz, 2021; Seidl et al., 2003). Chapter 3 shows that 19 month-olds parse the *wh*- phrase as the direct object in questions like *What is she washing with the sponge?* These results confirm that infants have acquired knowledge of *wh*- movement early enough that they might make use of *wh*- argument questions to learn island constraints as suggested in (Pearl & Sprouse, 2013).

Taken together, these results suggest that children acquire knowledge of island constraints over the course of about a year or so – this is less than half the time (meaning, half the linguistic input) estimated in the Pearl & Sprouse (2013) model.

These results raise questions about whether the model would succeed in a shorter timeline, and can be used to develop more rigorous learning models of island constraints.

To summarize, this dissertation explores the idea that determining *when* children have mastered some aspect of syntax can inform models of *how* they acquire that knowledge. In particular, this dissertation presents behavioral findings that 19 month-olds have acquired knowledge of *wh*-movement, and that 3 year-olds have acquired knowledge of island constraints. Behavioral research in the field of language acquisition often involves testing children for their grammatical competence, and concluding that children at age X months-old have knowledge of Y syntactic construction – but that's where it often stops. In light of recent experimental and modeling work concerning the acquisition of *wh*- movement, the time is right to shift toward a theory of how island constraints are acquired, as one small step toward a comprehensive theory of the acquisition of syntax. Existing learning models of island constraints (Pearl & Sprouse, 2013) provide a valuable opportunity for exploring how the linguistic input affects the learning outcome. Keeping the learner as realistic as possible will help us determine whether to pursue these kinds of models further, or look elsewhere.

The structure of this dissertation is as follows. Chapter 2 provides a brief linguistic overview of *wh*- movement and island constraints. Chapter 3 presents a behavioral study of early representations of *wh*- movement and predictive sentence processing at 19 months-old. Chapter 4 presents a behavioral study of early representations of island constraints at 3 years-old. Chapter 5 summarizes and reviews the Pearl & Sprouse (2013) learning model. Chapter 6 concludes.

Chapter 2: Overview of *Wh*-movement and Island Constraints

Chapters 3 and 4 present behavioral experiments which target syntactic representations of *wh*- movement and island constraints at early stages of language acquisition. This chapter provides a brief overview of these syntactic phenomena. *Wh*- questions in English (e.g., *Who did Mariam hug?*) are an example of *wh*- movement, a syntactic dependency which can span multiple clauses via smaller movements. In (8), the phrase *who* represents the direct object, and is syntactically related to *hug* via *wh*- movement.

8. Who did Mariam hug?

Wh- movement is said to be *unbounded* because there is no upper limit on the number of clause boundaries that a *wh*- movement relation can cross. For example, (9) shows that *wh*- movement holds across one clause boundary, (10) shows *wh*- movement across two clause boundaries, and (11) shows *wh*- movement across three clause boundaries (and so on). There is grammatical potential for *wh*- movement to cross infinite clause boundaries. For overviews of *wh*- movement, see Cheng & Corver, 2006; Citko, 2014; Richards, 2014; Sprouse & Hornstein, 2013.

9. [$_{CP}$ What$_1$ did Sterling say [$_{CP}$ _$_1$ that Mariam hugged _$_1$] ?
10. [$_{CP}$ What$_1$ did Jiabao think [$_{CP}$ _$_1$ he said [$_{CP}$ _$_1$ that Mariam hugged _$_1$]]] ?
11. [$_{CP}$ What$_1$ did Jiabao say [$_{CP}$ _$_1$ she thought [$_{CP}$ _$_1$ he said [$_{CP}$ _$_1$ that Mariam hugged _$_1$]]]] ?

If the intermediate [Spec,CP] (the landing site for *wh*- movement) is occupied by another syntactic element, then *wh*- movement is blocked (Chomsky, 1973, 1977). For example, some *wh*- complementizers (e.g., *whether*) occupy this position. Example (13) shows that *whether* blocks *wh*- movement which is otherwise allowed (12).

12. Jiabao remembered whether Mariam hugged the brown bear
13. *Which bear did Jiabao remember whether Mariam hugged?
14. Jiabao remembered that Mariam hugged the brown bear
15. Which bear did Jiabao remember that Mariam hugged?

Most indirect questions involve *wh*- movement, which leaves the intermediate [Spec,CP] occupied and blocks long-distance movement out of the *wh*- complement. Example (16) shows that *wh*- movement across a *when* complement is blocked, and example (16) shows *wh*- movement across a *how* complement is blocked. As a result, it's not possible to use (16c) to express the following coherent thought: *Jiabao remembers when Mariam hugged some bear…which bear was it?*. Example (16c) does not have an interpretation because it's ungrammatical. The same is true of (17c).

16.
   a. Jiabao remembered that Mariam [[hugged the brown bear] [at the party on Thursday]]
   b. Jiabao remembered when$_1$ Mariam [[hugged the brown bear] [_$_1$]]
   c. *Which bear$_2$ did Jiabao remember when$_1$ Mariam [[hugged _$_2$] [_$_1$]] ?

17.
   a. Jiabao remembered that Mariam hugged the brown bear tightly
   b. Jiabao remembered how$_1$ Mariam [[hugged the brown bear] [_$_1$]]
   c. *Which bear$_2$ did Jiabao remember how$_1$ Mariam [[hugged _$_2$] [_$_1$]] ?

Because the landing site of *wh-* movement is [Spec,CP], and long-distance movement proceeds as a series of smaller movements, *wh-* movement itself blocks long-distance movement (Chomsky, 1973). Metaphorically speaking, *wh-* movement creates an *island* for *wh-* movement (Ross, 1967). This generalization extends beyond *wh-* complements to explain why relative clauses block *wh-* movement. In (18a), *who won the brawl* is a subject relative clause, and is part of the direct object of like (*the hockey player who won the brawl*). Example (18b) shows that the subject relative clause blocks extraction of its direct object *the brawl*. In (19a), *that Rachel refurbished* is an object relative clause; a null operator (*op*) relates [Spec,CP] of the clause to the direct object gap. Example (19b) shows that the object relative clause blocks extraction of the temporal modifier *years ago*.

18. Subject relative clause

    a. Luisa liked the hockey player [$_{CP}$ who$_1$ [$_{TP}$ _$_1$ won the brawl]]

    b.* What$_2$ did Luisa like the hockey player [$_{CP}$ who$_1$ [$_{TP}$ _$_1$ won _$_2$]] ?

19. Object relative clause

    a. Imane appreciated the red chair [$_{CP}$ *op*$_1$ that Rachel refurbished _$_1$ years ago]

    b.* When$_2$ did Imane appreciate the red chair [$_{CP}$ *op*$_1$ that Rachel refurbished _$_1$ _$_2$] ?

Unlike example (18a), the embedded clause in (19a) does not feature a *wh-* phrase, but the salient linguistic properties that it shares with (18a), in particular its island sensitivity, suggest that it shares the same syntax as (18a). Island constraints have been a powerful tool for diagnosing the syntactic similarities between superficially different linguistic constructions since their discovery (Ross, 1967).

Chomsky (1977) leverages island sensitivity to argue that several well-known and superficially diverse clausal constructions are instances *wh*-movement (see (20) and (21), for example).

20. Comparatives

   a. Anouk has more leaves [$_{CP}$ *op*$_1$ than Aaron rakes _$_1$ from his yard each autumn]

   b.* When$_2$ does Anouk have more leaves [$_{CP}$ *op*$_1$ than Aaron rakes _$_1$ from his yard _$_2$] ?

21. Topicalization

   a. [$_{CP}$ [This book]$_1$, Rodrigo lent me _$_1$ four years ago]

   b.* [$_{CP}$ When$_2$ [this book]$_1$, did Rodrigo lend me _$_1$ _$_2$] ?

The observation that *wh*- movement blocks *wh*- movement is referred to as the *wh*- island constraint (Chomsky 1973). This generalization can be expanded even further, to any linguistic structure which has the effect of forcing *wh*- movement to proceed in *one fell swoop*. Ross (1967) observes that complex NPs, clausal adjuncts, Subjects, Coordination and Left-branching structures are all islands for *wh*-movement (for overviews of island constraints, see Boeckx, 2012; Sprouse & Hornstein, 2013). The Complex NP (Noun Phrase) Constraint is the generalization that Noun Phrases which contain sentences or clauses (for example, *the claim that Jéssica loves a vegan cooking challenge*) block *wh*- movement (Ross 1967). In example (22a), [$_V$ claimed] takes a clausal complement. Example (22b) shows that it is possible to extract the direct object *a vegan cooking challenge* from the complement of *loves* in the embedded clause. In example (23a), [$_N$ claim] takes a clausal complement, but example (23b) shows that it is not possible to extract the direct object from the complement of *loves* in the embedded clause. The embedded clause is identical in (22) and (23), but its relation to the matrix clause is different. In most cases, NP blocks *wh*- movement.

22. [$_{V}$ claim]

a. Jad recalled that Jéssica claimed [$_{CP}$ that she loves a vegan cooking challenge]

b. What did Jad recall that Jéssica claimed [$_{CP}$ that she loves _] ?

23. [$_{N}$ claim]

c. Jad recalled the claim [$_{CP}$ that she loves a vegan cooking challenge]

d.* What did Jad recall the claim [$_{CP}$ that she loves _] ?

Clausal adjuncts are another case where *wh*- movement is blocked (Huang, 1982; Ross, 1967). For example, when the clause *that we finally printed the nametags* is introduced as a VP-complement (24a), the direct object *the nametags* is free to move to the top of the matrix clause (24b). When the same clause is introduced as a clausal adjunct (e.g., *before we finally printed the nametags*) (25a), then movement of *the nametags* is blocked (25b).

24.

a. Sig remembered [$_{CP}$ that we finally printed the nametags [$_{CP}$ after making coffee]]

b. What$_{1}$ did Sig remember [$_{CP}$ _$_{1}$ that we finally printed _$_{1}$ [$_{CP}$ after making coffee]] ?

25.

a. Sig remembered [$_{CP}$ that we made coffee [$_{CP}$ before we finally printed the nametags]]

b.* What$_1$ did Sig remember [$_{CP}$ that we made coffee [$_{CP}$ before we finally printed _$_1$]] ?

When a clause is introduced as the subject, *wh-* movement is blocked (Huang, 1982; Ross, 1967). For example, when *that Rodrigo loves to watch movies* is a VP-complement (26a), it's possible to move the direct object of the embedded clause to the matrix clause (26b). When the same clause *that Rodrigo loves to watch movies* is introduced as the subject (26c), then it's not possible to extract the direct object.

26. Sentential Subject Condition

a. It is obvious that Rodrigo loves to watch movies

b. What is it obvious that Rodrigo loves to watch _ ?

c. That Rodrigo loves to watch movies is obvious

d. *What is that Rodrigo loves to watch _ obvious?

The sentential subject condition, clausal adjunct constraint, complex NP constraint and *wh-* island constraint are not the only locality constraints on *wh-*movement, but they have been highlighted here because they have been used to test the outcome of a prominent learning model of island constraints (Pearl & Sprouse, 2013), which is introduced and discussed in Chapter 5. Chomsky (1973) proposed the *Subjacency Condition* to unify these four island constraints, which stipulates that *wh-*movement cannot cross more than one bounding/cyclic node, where bounding nodes are TP and NP. Since subjects are NPs, *wh-* movement from a subject crosses NP and TP. *Wh-* movement from clausal adjuncts and *wh-* complements crosses the TP of the embedded clause and the TP of the matrix clause, and *wh-* movement from complex

NPs crosses the dominating NP and the matrix TP. The effect of this principle is that the gap of the moved element must be contained by a clause *subjacent* to its landing site.

A complete picture of constraints on *wh*- movement is more nuanced than the examples included here. Ross (1967) observed that *wh*- movement is subject to other constraints which can involve syntactic domains that are smaller than a clause, like coordinate structures (e.g., **What did Sterling like to make videos and throw _ ?* Answer: pottery) and the left-most branch of an NP (e.g., **Whose does Jiabao like to eat _ hamburgers?* Answer: Shake Shack's). Many languages also impose a ban on extraction from PP, and instead require that the entire PP undergo *wh*- movement (English is an unusual language in that it allows either).

The experiments in Chapter 4 use a small island to test 3 year-olds for knowledge of island constraints: NP-adjuncts. In general, *wh*- movement from DP is blocked (e.g., Bach & Horn, 1976; Davies & Dubinsky, 2003). For example, (27) has at least two interpretations: (a) Mariam used the blanket as a means to give the bear a hug (*with a blanket* is a VP modifier), and (b) the bear who received the hug from Mariam is the one holding the blanket (*holding a blanket* is an NP modifier). When *the blanket* is moved (28), only the interpretation (a) survives. This ambiguous example further illustrates that island violations result in the loss of an interpretation. If the baseline sentence (27) is ambiguous, then the transformed sentence (28) will still have at least one interpretation – just not the structure-meaning pairing blocked by the island. Note that the *wh*- cleft in (29) allows for both interpretation (a) and (b), and in general deletion of the linguistic material which contains the island but leaves the *wh*- phrase (e.g., sluicing) ameliorates island violations (e.g., Fox & Lasnik, 2003; Merchant, 2008; Rottman & Yoshida, 2013).

27. Mariam hugged the bear with a blanket

    a. Mariam [$_{VP}$ [$_{VP}$ hugged [$_{DP}$ the bear]] [with a blanket]]

    b. Mariam [$_{VP}$ hugged [$_{DP}$ the bear with a blanket]]

28. What did Mariam hug the bear with _ ?

29. Mariam hugged the bear with the blanket. Which blanket was it?

This dissertation work is committed to the hypothesis that island constraints are grammatical constraints. Alternatively, reductionist accounts pursue extra-grammatical explanations for island constraints (e.g., Ambridge & Goldberg, 2008; Deane, 1991; Erteschik-Shir, 1973; Givón, 1979; Goldberg, 2013; Hofmeister & Sag, 2010; Kluender & Kutas, 1993). These proposals tend to focus on explaining *weak islands* as well as other factors which influence the acceptability of *wh*- movement. For weak islands, extracting an argument tends to be more acceptable than extracting an adjunct (e.g., *wh*- islands) (e.g., Cinque, 1990; Szabolcsi, 2006; Szabolcsi & Den Dikken, 1999). *Wh*- movement is also sensitive to the intervention of negation (Cinque, 1990), quantificational adverbs (Obenauer, 1985), factive and manner-of-speaking verbs (e.g., Liu et al., 2019) and similarity-based intervention (e.g., Boeckx, 2008; Rizzi, 1990), and definiteness (e.g., compare *Who did she steal a picture of?* to *Who did she steal the picture of?*) (e.g., Davies & Dubinsky, 2003).

This chapter gives a brief overview of island phenomena and their treatment in generative syntax. This dissertation is committed to the hypothesis that island constraints are syntactic in nature, but introduces island constraints in a descriptive way, barely touching on modern treatment of island constraints. The purpose of including this discussion is to provide context for the language acquisition experiments in the following chapters. These techniques are suited for testing infants and children for knowledge of descriptive generalizations, but cannot go further to differentiate syntactic theories.

Chapter 3: 19 month-old infants represent *wh*-movement

3.1 Introduction

An important step in acquiring knowledge of *wh*- movement is learning that some linguistic elements are syntactically related at a distance, and that the displaced *wh*-phrase leaves a gap. For some learning theories, this knowledge is a precondition for acquiring knowledge of locality constraints on *wh*-movement (e.g., Pearl and Sprouse 2013), or constraints on how the *wh*-movement rule is applied. These models rely on behavioral results to shape their estimates of the language input (both its qualitative and quantitative properties) and the learner (its linguistic and conceptual capacities). The more accurate these estimates are, the more convincing the learning model is. The current chapter is primarily concerned with the early emergence of knowledge of *wh*-object questions during language acquisition.

Using a modified preferential-looking task (Hollich et al., 1999; Lidz et al., 2017; Spelke, 1976), this experiment tests the hypothesis that 19 month-old infants represent *wh*- movement. The results show that infants parse *what* as the direct object of the verb in *wh*-object questions which contain a novel Noun Phrase (e.g., *What is she pulling with the wug?*). Previous results show that 19 month-old infants mis-parse *the wug* as the direct object in intransitive sentences like *She's pulling with the wug*. When the sentence is transitive (e.g., *She's pulling that thing with the wug!*), these same infants correctly parse *the wug* as the object of the instrument (Lidz et al., 2017). In this experiment, we familiarize infants either with sentences which have been shown to induce this parsing error (e.g., *She's pulling with the wug!*) or with *wh*-object questions (e.g., *What is she pulling with the wug?*). Infants who have learned to represent *wh*-movement will represent *wh*-object questions as transitive, which will allow them to represent the *wh*-phrase as the direct object and learn that *the wug* refers to the instrument.

Instead of measuring how infants respond to the *wh*- object question, this experiment measures what infants have learned that *the wug* refers to, and use that as a window into how they represented the *wh*- object question. This approach successfully isolates the contribution of syntactic knowledge on the comprehension of *wh*-questions. Identifying the contribution of the syntax on language comprehension is difficult for experiments which rely solely on infants' responses to *wh*- questions, as many non-syntactic factors contribute to infant behavior in response to a question.

The results of this experiment show that infants mistakenly learn that *the wug* refers to the pullee in the intransitive condition (e.g., *She's pulling with the wug*), but successfully learn that *the wug* refers to the puller in the *wh*- object question condition (e.g., *What is she pulling with the wug?*, indicating that they have parsed *what* as the direct object of the sentence. Because sentence processing mechanisms are resistant to linguistic revision (in both children and adults), we suggest that infants parse *what* as the direct object **before** they parse the rest of the sentence (i.e., …*with the wug*). The alternative – that infants parse *the wug* as the direct object and then revise their analysis to accommodate the *wh*- phrase – is untenable given well-known difficulties with revision cited in the sentence processing literature (Omaki & Lidz, 2015; Snedeker, 2013; Trueswell et al., 1999). This result is consistent with the hypothesis that even the earliest sentence processing mechanisms are guided by grammatical knowledge.

3.2 Background

The topic of *wh*- movement has historically received a lot of attention in the field of language acquisition for several reasons. Primarily, the linguistic representations involved (e.g., local relations which hold across an arbitrary distance, or non-locally) are sufficiently abstract to be of interest to generative linguists. From a practical standpoint, children produce *wh*- questions from an early age, giving the impression that young children might understand these constructions. Moreover, *wh*- questions are common in child-directed speech, so it's unlikely that the linguistic environment will be a limiting factor for the acquisition of *wh*-movement. From a practical standpoint, *wh*- questions are an effective way to get a response, to these constructions are well-suited for the kinds of interactive experiments that work well for testing young children.

Most behavioral experiments which probe for knowledge of *wh*-movement use *wh*- questions and have largely focused on language comprehension in children ages 3 and older (e.g., Cairns & Hsu, 1978; Radford, 1994; Roeper & Villiers, 2011). More recently, there has been a shift toward studying whether infants represent *wh*-questions as syntactic dependencies. This experiment is targeting the onset of this linguistic knowledge, so this section reviews literature on the early acquisition of *wh*- questions.

Even before they can form a linguistic response themselves, infants get exposure to *wh*- questions in the input. In a study of child-directed speech drawn from the Suppes dataset (1 child ages 1;11-3;3) (Suppes 1974), the Brown dataset (Adam (ages 2;10-5;2), Eve (ages 1;6-2;3) and Sarah (ages 2;3-5;1) (Brown 1973)) and the Valian dataset (21 children ages 1;10-2;8) (Valian 1991), Pearl & Sprouse (2013) finds that 20,923/101,838 (~20%) of utterances in speech to children ages 1 to 5 years-old contained *wh*- argument questions, and most of those are *wh*- object questions (76.7% *wh*- object questions versus 12.8% *wh*- subject questions). This estimate is high compared to other analyses of child-directed speech. Studies focused on *wh*- questions directed toward infants yield smaller proportions of *wh*- questions relative to other linguistic constructions. For example, in a small study of 10 mother-infant pairs, only 10% (59/680) of utterances during free-play situations contained *wh*- questions (Broen, 1972). Another small study finds 13% of parent utterances during free-play (9 parent-infant pairs with infants younger than 2) contained *wh*-questions (Newport et al., 1977). A third study finds that about 17% of parent utterances (7 parent-infant pairs with infants around 2 years-old) recorded in the home contained *wh*- questions (Furrow et al., 1979).

The results of studies of larger corpora (with Pearl & Sprouse, 2013 being the outlier) estimate that the number of *wh*-questions is less than 10% of total utterances. Since studies of child language corpora opt to increase the size of the data set by including a wider age range, the larger percentage of *wh*-questions observed in smaller corpora likely reflects a qualitative difference in how the age of the child affects the ways that parents talk to them (for related discussions, see Huttenlocher et al., 1998, 2002). In a study of 5 children (ages 1 to 5 years-old), only 4.2% (5,994/143,353) of utterances contained a filler-gap dependency with an argument *wh*-phrase (Atkinson, 2016: 67). Note that this group contains *wh*- argument questions as well as other constructions, like relative clauses and embedded questions, so it necessarily *overestimates* the number of *wh*-argument questions in child-directed speech. This makes the discrepancy between this result and the Pearl & Sprouse (2013) estimate more surprising. Of particular interest to this study, 2.4% (3,504/143,252) of utterances contained a *what* question with a direct object gap, and >1% (1,204/143,252) of utterances contained a direct object gap with an overt PP (e.g., *What is she pulling with the wug?*). Note that the asymmetry between subject and object filler-gap dependencies observed in Pearl & Sprouse is also observed in this study (26.7% (1,490/5,573) of argument filler-gap dependencies contained subject gaps versus 62.9% (3,504/5,573) object gaps[1]) Interestingly, the opposite pattern has been reported for relative clauses, with subject relatives more common than object relatives (Tabor et al., 1997).

In a study of the quantity and distribution of *wh-* dependencies (*wh-* questions, relative clauses, embedded questions, etc.) in child ambient speech (26 children, ages ranged from 1 to 6 years-old) revealed that 10% (14,427/146,363) of lines contained overt *wh-* dependencies (Omaki, 2010: 214). Since this analysis includes constructions other than *wh-* questions, like *wh-* adjunct questions, *wh-* argument questions must constitute less than 10% of the input (*wh-* dependencies containing argument *wh-*phrases make up 8% (11,639/146,363) of the input). Only 2% (305/14,427) of *wh-*dependencies were long distance movement across a finite clause. To summarize, studies of the input indicate that children do get exposure to *wh-* questions, with *wh-*argument questions making up anywhere from 4.2% to 20% of child-directed utterances (although most estimates hover around 10%).

Early child productions of *wh-* questions give a rough estimate of linguistic knowledge but should still be considered if the goal is to converge on the onset of

[1] The missing 10.4% (579/5,573) contained prepositional object gaps.

knowledge of *wh*- movement. Children's language production tends to lag behind their comprehension (e.g., Gerken 1987, 1990, 1996; Hirsh-Pasek and Golinkoff 1996; Shipley et al., 1969 and many others). This generalization is commonly used to guide estimates of the onset of some linguistic knowledge. If a child uses some construction, then it must be the case that they have acquired grammatical knowledge of that construction. While this may be true in some cases, it's not obvious that it must be true for every child utterance. One study of 12 children (ages 1 to 6) finds that children begin to use *wh*- argument questions (i.e., *who*, *what* or *which* questions) just after their second birthday, on average . Children used their first *wh*- subject question at a mean age of acquisition of 2;5.2 (years;months) (age range: 1;8.5 to 3;8.2) and their first object question at a mean age of 2;3.4 (age range: 1;9.8 to 2;8.8) (Stromswold 1995: 26). For *what* questions with direct object gaps, first use was observed at a mean age of acquisition of 2;3.6 (years;months) (age range: 1;9.8 to 2;9.2) (Stromswold 1995: 28). Questions were counted as *wh*-object questions "if the most likely answer to the question was an object NP." (Stromswold 1995, page 21). The observation that children begin to use *wh*-object questions earlier than *wh*-subject questions tracks neatly with the observation that adults use *wh*-object questions more frequently than *wh*-subject questions in child-directed speech. For a thorough discussion of alternative explanations for this asymmetry, see Stromswold (1995).

Note that the first-use measure does an excellent job at ensuring that the earliest possible datapoints are considered, but it doesn't give any information regarding how children tend to use *wh*- questions. Another study of 7 children (ages 2-3) uses a quantitative measure and reports that 2;2 is the average age at which children begin to use *what* questions which contain a verb. Most of the *wh-* questions that 2 year-olds use are very simple: 28% lack a verb entirely. Only 12% (959/7,887) contained 'descriptive verbs' (e.g., *jump, eat, bring…*), while most (60%) contained 'pro-verbs' (*be, do, happen* or *go*) (Bloom et al., 1982; reprinted in 1991). Only 5% (397/7,887) were *what* questions which contained a 'descriptive verb'. These are distinct from the *wh-* object questions tested in the current study (e.g., *What is she pulling _ with NP?*). Because very few *wh*-questions contain verbs, Bloom (1982) suggests that these early questions are learned as routines. A more recent study shows that (like adults) most (62.2%) of the argument filler-gap dependencies that children produce have direct object gaps (1,544/2,481), and only 19.8% (491/2,481) have subject gaps (Atkinson 2016: 70). Only 1% (1,597/158,194) of children's utterances are *what* questions with a postverbal gap (i.e., direct object or prepositional object gaps).

To summarize, children begin using *wh-* object questions just after they turn 2 years-old, and these tend to be *what* questions. Because most of their first *wh-* questions use verbs without clear postverbal gap positions, or lack the verb entirely, any knowledge of *wh-* movement that these children have is effectively masked by their production system (alternatively, as Bloom suggested, they don't have the knowledge at all). If 2 year-olds represent *wh-* movement, then it's likely that they have done the work of acquiring this knowledge between 1 and 2 years-old. For this reason, studies of the onset of knowledge of *wh-* movement focus on how infants react to *wh-* questions, as opposed to what they produce.

The first comprehension study of *wh-* movement aimed at children younger than 2 was designed to test whether infants knew enough about *wh-* questions to look more toward the answer to the question than toward the referent of another DP in the sentence (Seidl, Hollich & Jusczyk 2003). It is natural for infants to look at the referent when they hear its linguistic label, but will they be able to look at the answer to a *wh*-question? For example, in *wh*- subject questions like *What hit the keys?*, the answer to the question is whatever object was the hitter in a hitting event. Looking toward the hitter in response to the question indicates awareness that a question has been asked and awareness that the question was not about *the keys* – even though *the keys* are explicitly mentioned in the sentence, while *what* is a pronoun.

This study used a preferential-looking paradigm (Hirsh-Pasek & Golinkoff 1996) to test 20 month-olds, 15 month-olds and 13 month-olds on *wh*- subject questions, *wh*- object questions, and *where* questions (control condition). Question type is manipulated as a within-subjects factor, so each infant heard all three question types. The *where* questions (e.g., *Where are the keys?*) were included as a control condition: the answer to the *where* question is also the explicitly realized DP *the keys*, unlike for the argument questions. Infants watched simple events, like an image of a book ricocheting off of an image of a set of keys. After the event, infants saw a split-screen display with the book and the keys, one on each side of the screen. Infants were tested on either a subject, object or *where* question (depending on the trial).

The results show that 13 month-olds do not look reliably more toward the correct answer in any of the conditions, suggesting that 13 month-olds have not yet acquired knowledge of *wh*- questions. 15 month-olds look reliably more toward the correct answer in response to *wh*- subject questions and *where* questions, but look roughly equally at both objects in response to *wh*- object questions (they show a slight preference for the correct answer, but it's not significant). 20 month-olds look reliably more toward the correct answer in every condition. To summarize, performance improves as a function of age. The authors suggest that 20 month-olds' responses are driven by knowledge of *wh-* movement, and conclude that this knowledge develops sometime between 13 and 20 months-old.

Work following up on this study carefully considers the hypothesis that 15 month-olds have knowledge of *wh-* movement, but this knowledge is masked in the *wh-* object condition of Seidl et al., (2003) because the task is both too demanding and too strange (Gagliardi, Mease & Lidz 2016). They point out that the within-subjects design in the Seidl et al., (2003) study requires switching from one topic to the next, the materials were somewhat abstract, and the questions themselves were infelicitous because there were never alternative answers to consider. For example, had a third object swooped in and nearly hit the book, then the question *What **hit** the book?* would have been better supported by the experimental context. See Gagliardi (2015) for a detailed critique of Seidl et al., (2003).

To address these issues, Gagliardi et al., (2016) tested 15 month-olds and 20 month-olds on *wh-* object and *wh-* subject questions in a modified version of Seidl et al., (2003). Question type was presented as a between-subjects factor to require less topic shifting on the part of the infant. Infants were familiarized to videos in which a cat bumped a dog, who in turn bumped another cat. During test, infants saw a split-screen display of the cats. A question like *Which cat did **the dog bump**?* is supported by this context, because there is a cat which the dog bumped, and there is another cat which the dog didn't bump.

Gagliardi et al., (2016) reports that 15 month-olds look more toward the agent in the subject condition (about 60% looks-to-agent) than the object condition (about 40% looks-to-agent) – but only in the second half of the experiment. The first half of the Gagliardi et al., (2016) experiment shows the opposite result, but the authors suggest that infants are not able to demonstrate their knowledge in the short time of the first three trials – which is why, they conclude, this knowledge does not reveal itself in the 15 month-old population in the Seidl et al., (2003) study. Turning to the 20 month-olds, Gagliardi et al., (2016) observes that infants look more toward the agent in the subject condition (about 60%) than the object condition (about 30%), but only in the second half of the experiment, and only after the second test question. This is one instance where infant looking patterns in response to the first test question do not seem to be driven by their understanding of the question, but instead reflect their arbitrary baseline preferences. This same pattern obtains in the current study.

In a second experiment, Gagliardi et al., (2016) tests 15 month-olds and 20 month-olds for knowledge of subject and object relative clauses. Again, 15 month-old looking patterns are consistent with knowledge of *wh*-movement: in the second half of the experiment, in response to the second test utterance, infants look more toward the agent (about 55%) in the subject relative condition (e.g., *Show me the dog that bumped the cat*) than in the object relative condition (about 40%) (e.g., *Show me the dog that the cat bumped*). The 20 month-olds, on the other hand, fail to show an effect of syntax (about 50% looks-to-agent in the subject relative condition compared to about 45% looks-to-agent in the object relative condition).

In a follow-up study, Perkins & Lidz (2019) reproduce the findings that 15 month-olds succeed at understanding subject and object *wh*-questions as well as subject and object relative clauses using a streamlined version of this design. They further find that vocabulary size influences infant performance. Taken together, these three studies are consistent with the hypothesis that 15 month-olds represent *wh*-movement. However, this conclusion would fail to explain why the 20 month-olds fail in the relative clause experiment. To explain how it could be that the younger infants succeed where the older ones fail, Gagliardi et al., (2016) pursues the hypothesis that the younger 15 month-olds actually do not represent *wh*- movement (aligning with the results of the original Seidl et al., (2003) study). Instead, 15 month-olds succeed in these tasks by noticing that the string (e.g., *the cat bumped*) is missing an obligatory argument, and "search the discourse context for a referent that could fill this thematic structure" (Gagliardi 2015: 238). 15 month-olds succeed at this comprehension task without representing the relation between the *wh*- phrase and the gap in the syntax. The 20 month-olds have acquired knowledge of *wh*- movement, but fail to apply this knowledge during the relative clause experiment without a clear signal (e.g., a *wh*-phrase). A third experiment tests 20 month-olds on relative clauses like *Show me the dog who bumped the cat* and *Show me the dog who the cat bumped*, which use the *wh*-phrase *who* to signal the relative clause more clearly than *that*. Results show that 20 month-olds look more toward the agent in the subject condition than the object condition. This finding supports the hypothesis that 20 month-olds, but not 15 month-olds, have knowledge of *wh*-movement (Gagliardi et al., 2016: Experiment 3).

The hypothesis that 15 month-old infants succeed at these tasks armed with only partial knowledge of *wh*- movement is an intriguing way to capture their success in light of the 20 month-olds' failure (Gagliardi, Mease and Lidz 2016: Experiment 2), but it raises questions about whether these experimental methods are suitable for determining if knowledge of *wh-* movement drives success at any age. The performance contrast between Experiment 2 and Experiment 3 is suggestive, but additional evidence that 20 month-olds represent *wh-* movement would make the argument more compelling.

A recent study side-steps the difficulties of testing infants' interpretations of *wh-* argument questions by using a listening technique, instead of prompting infants to respond to *wh-* questions (Perkins 2019, Perkins & Lidz 2021). In this infant-controlled listening time task, infants were first familiarized with blocks of transitive sentences while they watched a soothing video of abstract, rotating shapes. This gave infants the opportunity to grow accustomed to the experiment and discover that they could control the duration of the trial by looking away from the shapes for at least two seconds. During the test phase of the experiment, infants heard alternating blocks comprised of lists of grammatical and ungrammatical sentences. The clause type (declarative versus *wh*-question) was manipulated as a between-subjects factor. In the *wh*-condition, grammatical blocks had object gaps (30a), but the ungrammatical blocks did not (30b).

30. *Wh-* question condition
    a. Hey, which monkey should the giraffe kiss?[2] Wow, which cow should the lion tickle? Yay, which tiger should the lion hug? …
    b. Hey, which monkey should the giraffe hug her? Wow, which cow should the lion tickle him? Yay, which tiger should the lion hug her? …

[2] These are not the actual sentences used in the study, but very similar.

In the declarative condition, ungrammatical blocks had object gaps 31a), but the grammatical blocks did not (31b).

31. Declarative condition

  a. Hey, A monkey! The giraffe should kiss her. Wow, a cow! The lion should tickle him. Yay, a tiger! The lion should hug her. …

  b. Hey, A monkey! The giraffe should kiss. Wow, a cow! The lion should tickle. Yay, a tiger! The lion should hug. …

In Experiment 1 (focused on younger infants), 14 month-olds did not show any listening preference, but 15 month-olds preferred to listen to object gap trials (30a) and (31b). In Experiment 2 (focused on older infants), 18 month-olds but not 17 month-olds prefer to listen to object gap trials in the *wh*- condition (30a), but overt object trials in the declarative condition (30a). Strikingly, 18 month-olds prefer to listen to grammatical sentences over ungrammatical sentences, which suggests that by 18 months, infants have learned that fronted *wh*- phrases co-occur with object gaps – a crucial step in acquiring *wh*- movement.

These experiments add an important datapoint to the developmental timeline of the acquisition of *wh*- movement by corroborating the hypothesis that the success of 15 month-old infants in comprehension studies is due to partial syntactic knowledge, while the success of 20 month-old infants is due to knowledge of *wh*-movement.

The current study takes a different approach which skirts the complications introduced by interpretation tasks by probing for knowledge of *wh*- movement via infant sentence processing mechanisms. Instead of measuring infants' responses to the *wh*- question of interest, the current experiment measures infants' responses to a follow-up question about the meaning of a novel Noun Phrase used in the *wh*- question of interest. If infants learn the meaning of the novel NP, then they must have parsed the fronted *wh*- phrase as the direct object of the verb. In doing so, they cease to make a robust sentence processing error (Lidz et al., 2017). Before moving on to the details of the experiment, the remaining paragraphs in this section provide some background on *wh*- processing strategies and how they develop.

Wh- processing refers to the processing strategies or mechanisms applied to parse filler-gap dependencies, like *wh*- questions and relative clauses. Listeners make incremental parsing commitments based on various linguistic and extra-linguistic information which is accessible to them in the moment. In the case of English *wh*-questions, the *wh*- phrase is pronounced at the beginning of the sentence, which provides a clear linguistic signal from the start. If the *wh*- phrase is *who*, for example, listeners might anticipate a gap in an argument position (perhaps even a subject gap if *who* is more frequently used in subject questions than object questions). If the *wh*-phrase is *where*, then listeners might anticipate an adjunct gap, and so on. Since the gap position determines the interpretation of the *wh*- phrase, it's to listener's advantage to identify the gap position as quickly as possible. There is ample experimental evidence that adults parse the fronted *wh*- phrase in *wh*- questions *predictively* – before encountering the gap (Stowe 1986, Aoshima, Phillips, & Weinberg, 2004; Lee, 2004, Traxler and Pickering 1996, Boland, Tanenhaus, Garnsey, & Carlson, 1995; Garnsey, Tanenhaus, & Chapman, 1989; Omaki & Schulz, 2011; Phillips, 2006; Pickering & Traxler, 2003; Wagers & Phillips, 2009; Omaki et al., 2010).

Many studies of the development of *wh*- processing are pre-occupied with how children might develop these predictive parsing mechanisms and whether children exhibit the same incremental processing strategies that adults do when parsing *wh*-questions. Language comprehension studies which use offline measures of sentence processing show that 5 year-olds are biased to actively complete *wh*- dependencies at the first verb they encounter. In a question-after-story task, English and Japanese learning children associate *where* with the linearly closest VP when asked questions like (32) or (33) (Omaki et al., 2014).

32. Where did Lizzie {tell/say to} someone that she was gonna catch butterflies?

33. Doko-de Yukiko-chan-wa [*pro* choucho-o tsukameru to] itteta-no?

    Where-at Yukiko-Dim-Top she butterfly-Acc catch Comp was telling-Q

    "Where was Yukiko telling someone that she will catch butterflies?"

English learning children prefer to associate *where* with the linearly closest VP, which is the matrix clause (the location of the event report) (Omaki et al., 2014: Experiment 1).[3] Japanese learning children also prefer to associate *where* with the linearly closest VP, which is the embedded clause (the butterfly-catching location). In a follow-up experiment, children associate *where* with the embedded VP even when that clause already supports another location modifier. These results suggest that 5 year-olds incrementally parse the *wh*-phrase (also see Love (2007); Lasotta et al., (2015)).

[3] When the matrix clause expresses the source of information (e.g., *Sig says that there's leftover creamer*) then the matrix clause interpretation of *wh*-adjunct questions like *Where did Sig say that there's leftover creamer* is generally unavailable. Some verbs (like *say*) are highly correlated with this "parenthetical" use, so the matrix verb affects the likelihood that listeners will entertain the long distance reading of some *wh*-adjunct questions (e.g., de Villiers et al., 2008; de Villiers et al., 1990). Pearson, 2008; Roeper & de Villiers, 1992)), the matrix verb influences how people interpret these questions (for detailed discussion, see (Omaki 2010; Lewis, Hacquard & Lidz, 2012))

In contrast, the results of language comprehension studies which use online measures to test sentence processing mechanisms (the preferred method for testing adults) during child development are less clear. Adult behavior is typically used as a baseline to interpret child behavior. When children are slower to react than adults, is it because they use different processing mechanisms, or is it because they have the same processing mechanisms but these mechanisms just deploy more slowly? In a recent eye-tracking study, for example, the effects of incremental *wh-* processing are detected at 6 years-old, but not at 5 years-old[4] (Atkinson et al., 2018, see also Atkinson 2016). In a visual world paradigm, children listen to pre-recorded *wh-* questions like *Can you tell me what Emily was eating the cake with?* after listening to a short story (34) represented by pictures (Figure 1). The visual display serves as a reminder of what happened in the story, and makes it possible to measure what children are looking at while they're listening to the test question.

34. Hi, my name is Emily. Today I'd like to eat some cake, but I also need to wash the dishes. Hmm, what should I do first? I think I'm gonna eat the cake, and for that I need a fork. Mmm! That cake was yummy. Now it's time to wash the dishes. I'm gonna need to use a sponge. Oh, those dishes are so clean. I did a great job today.

    Wh-question: Can you tell me **what** Emily was eating the cake with?

    Yes/no question: Can you tell me **if** Emily was eating the cake with the fork?

[4] For a complete representation of these findings, see Atkinson (2016) where data is discussed in more detail, and results from a 7 to 8 year-old age group is included.

Figure 1. Sample visual array (Atkinson et al., 2018). Clockwise from the left: a crumb-covered fork, a cake with crumbs where a slice used to be, a sparkling clean stack of dishes, a dirtied sponge, Emily.

Following previous experimental protocol used with adult participants, children's looking preferences are measured at two points: during the verb *eating* and during the direct object *the cake*. If children parse *what* as the direct object of the verb before they encounter the gap position, then they will look more toward the cake during the verb region in the *wh-* question condition than in the yes/no question condition. Even though all the test questions featured prepositional object gaps, filler *wh-*questions featured direct object gaps, so it's reasonable that children might expect a direct object question to pop up during the experiment. If children parse *what* as the object of the preposition before they encounter the gap position, then they will look more toward the fork during the direct object region in the *wh-* question condition than in the yes/no question condition.

Results show anticipatory looks toward the fork during the direct object region in the *wh-* question condition compared to the yes/no question condition, but this experiment fails to find any anticipatory looks during the verb region. Also observed is an uptake in looks toward the cake in the direct object region. This looking pattern is consistent with two hypotheses: (a) children look at the referent of the NPs that they hear and (b) children, like adults, incrementally parse *what* as the direct object before encountering its gap position, but are slow either to commit to the parse or to react to it. As a result, these anticipatory looks aren't detected until the verb is long gone. Further research is needed to understand how sentence processing mechanisms link to looking patterns in young children to better understand these results. The current study uses an offline measure of sentence processing, which avoids the risk associated with on online measures.

3.3 Experiment 1: *Wh-* questions (infants)

The current experiment takes advantage of a known parsing error that is characteristic of infants around 19 months-old (Lidz, White & Baier 2017) to develop a new test for early knowledge of transitivity and *wh*-movement. This section describes the experimental conditions in which this parsing error arises. Using a modified intermodal preferential looking paradigm (Hollich et al., 1999), 16 and 19 month-old infants were familiarized to an event with narration (e.g., a woman using a toy fishing rod to pull a toy train around in circles). During the event, infants heard either a transitive linguistic description (e.g., *She's pulling the wug!*), an intransitive one (e.g., *She's pulling the wug!*) or an uninformative one (e.g., *It's a wug!*). At test, infants saw a split-screen display of the toy fishing rod and the toy train, and they are tested on what they think *the wug* refers to (e.g., *Where's the wug? Which one's the wug?*).

Success in this task requires knowledge of syntax: in the intransitive condition (e.g., *She's pulling with the wug!*), infants must parse *the wug* as the object of the preposition, interpret it as the puller, identify that the puller relation is instantiated by the fishing rod in the video of the pulling event, and learn that *the wug* refers to the fishing rod. In the transitive condition (e.g., *She's pulling the wug!*), infants must parse *the wug* as the direct object of the verb, interpret it as the pullee, identify that the pullee relation is instantiated by the toy train in the video, and learn that *the wug* refers to the train. Results show that 16 month-olds look more toward the puller (i.e., the fishing rod) during the test phase in the intransitive condition than the transitive condition, indicating that they can use verb-argument structure to draw inferences about meaning. This result is consistent with a large body of research in the syntactic bootstrapping literature which shows that knowledge of verb argument structure is acquired before 24 months (Fisher et al., 2020; Hirsh-Pasek & Golinkoff, 1996; Seidl et al., 2003).

Unlike the 16 month-olds, 19 month-olds do not look more toward the puller during the test phase in the intransitive condition compared to the transitive condition. The authors suggest that sentence processing mechanisms mask grammatical competence in the 19 month-old age group: infants parse *the wug* as the direct object, which leads them to interpret *the wug* as the pullee and learn that *the wug* refers to the toy train. 19 month-olds fail to learn that *the wug* refers to the fishing rod because their incrementally-assigned linguistic representation resists revision (Lau & Ferriera 2005, MacDonald et al., 1992, Frazier 1990, Slattery et al., 2003, Fodor & Inoue 1998, Huang

& Ferriera 2021). Several follow-up experiments corroborate this conclusion. A corpus search of child-directed speech reveals that the verbs used in this experiment (*hit, pull, push, throw, tickle, touch, wash, wipe*) are used in transitive sentence frames 70% (1,126/1,617) of the time, compared to intransitive frames containing PP modifiers (2% (37/1,617)). The syntactic bootstrapping literature shows that infants can use their syntactic knowledge of verb argument structure to learn the meaning of novel verbs, which indicates that infants must track this information in their input. If 19 month-olds have accumulated experience with the syntactic distribution of these common verbs, then they might use this frequency knowledge to guide sentence processing mechansims, much like older children (Snedeker, 2013) and adults (Chomsky, 1977).

When the 16 month-olds from Experiment 1 are split into two groups based on the number of words that they produce (as determined by parental report on the MacArthur-Bates Communicative Development Inventory m-CDI (Fenson et al., 2006)), the 16 month-old infants in the high vocabulary group behave like the 19 month-olds, and show no preference for the puller in the intransitive condition compared to the transitive condition. The effect of syntax in the entire 16 month-old group is driven entirely by the behavior of the infants in the low vocabulary group. This is consistent with the hypothesis that the behavior of 19 month-olds is driven by their knowledge of verb distribution frequencies; low vocabulary 16 month-olds don't make the parsing error that their high vocabulary peers do because they lack this frequency information.

In a follow-up experiment, 16 month-olds are tested on how they interpret intransitive sentences using the preposition *on* (e.g., *She's pulling on the wug!*) or a novel preposition (e.g., *She's pulling gub the wug!*) given the same visual materials and design as the original experiment. In the *on* condition, infants should look at the toy train at test if they know what *on* means, and in the *gub* condition they should be at chance because the context allows for a variety of interpretations for the novel preposition. This experiment confirms that infant success in the original experiment is not driven by some low-level knowledge, like arguments which are separated from the verb by some linguistic material are not assigned the pullee relation. Results show that at test, infants look more toward the pullee in the *on* condition than in the *gub* condition. This result indicates that 16 month-old infants have acquired knowledge of the meanings of prepositions, including the prepositions *with* and *on*. This finding supports the hypothesis that 16 month-olds succeed in the original experiment because they have knowledge of verb argument structure, which is consistent with the hypothesis that 19 month-olds have the requisite grammatical knowledge to succeed at these tasks.

Another follow-up behavioral experiment familiarizes 19 month-olds to transitive sentences controlling for the syntactic position of the novel noun phrase (using the same visual materials and design as the original experiment). In the critical condition, the novel NP appears in the direct object position (e.g., *She's pulling* ***the wug*** *with that thing!*), and in the control condition the novel NP appears as the prepositional object (e.g., *She's pulling that thing* ***with the wug!***). At test, 19 month-old infants look more toward the fishing rod when *the wug* was introduced as the prepositional object (critical condition) compared to when it appears as the direct object (control condition). This behavior indicates that when 19 month-old infants' expectations about the transitivity of the frame are met, then they assign the correct syntactic representation to the object of the preposition *with*. This accurate syntactic representation allows them to interpret *the wug* as the puller, and learn that *the wug* refers to the fishing rod.

A follow-up study shows that when tested on novel verbs, 19 month-olds successfully avoid making the parsing error observed in the original study (White & Lidz 2022). The value in this test case is that if the verb is novel, then it's not possible for infants to use its syntactic distribution to guide sentence processing. This experimental manipulation 'levels the playing field' by putting 19 month-olds in the same position as the low vocabulary 16 month-olds from the original experiment. In this study, 19 month-olds are familiarized with the same transitive and intransitive sentences as Lidz, White & Baier (2017: Experiment 1), but with novel verbs instead of actual verbs (e.g., *She's* ***stipping*** *with the wug!*). To learn that *the wug* refers to the fishing rod in the intransitive condition, 19 month-old infants must rely on their knowledge of the meaning of *with.* Results show a behavioral split correlated with productive verb vocabulary: 19 month-olds with high verb vocabulary estimates behave exactly like the 16 month-olds in the original experiment, and look more toward the puller at test in the intransitive condition compared to the transitive condition. Using novel verbs blocks parsing strategies which rely on the syntactic distribution of the verb, which allows them to succeed at this task where the failed before. 19 month-olds with low verb vocabulary estimates continued to make the parsing error observed in the original experiment (Lidz, White and Baier 2017), which suggests that infants with less experience had difficulty with a task that introduced multiple novel NPs in one sentence.

To summarize, 19 month-old infants tend to parse the object of the preposition *with* as the direct object in transitive sentences like *She's pulling with the wug!* because they rely on their knowledge of the verb's syntactic distribution to guide sentence processing. Sentence processing mechanisms are predictive and resistant to linguistic revision by nature (see Omaki, 2010 for a great overview). During sentence processing, speakers rapidly integrate information from different sources to incrementally parse the sentences they hear. This can lead to temporary mis-analyses, which should be revised in light of later information. However, research shows that these revised representations are weak, and easily overwhelmed by initial representations. For example, when adults read sentences like, *While Anna dressed the baby played in the crib* (with no comma), they tend to initially interpret *the baby* as the direct object of *dressed* in the matrix clause. This is an example of an incremental parsing commitment. But, directly after *the baby*, the verb *played* reveals that this syntactic commitment was not the right choice for this sentence: *the baby* is the subject of the matrix clause, not the direct object of *dressed*. Evidence for these incremental commitments come from eye-tracking experiments (for example), which measure the behavioral effects of this realization as the process unfolds (while the sentence is being read). From the comprehension perspective, there is an important question about what kind of information adults retain – or learn – from these sentences (or more generally, garden-path sentences which require syntactic revision). When adults are tested on what they remember about these sentences, they confirm that (a) the baby played in the crib, but they also consistently and incorrectly remember that (b) Anna dressed the baby (Christianson et al., 2001; Ferreira et al., 2002). The fact that adults carry the false belief that Anna dressed the baby suggests that their initial interpretation has persisted. Another way to describe this behavior is to say that sentence processing resists revision.

Unsurprisingly, children's sentence processing mechanisms display this same property, although it affects their behavior in different linguistic contexts than with adults. In one well-known sentence processing study, 5 year-old children respond to requests like *Put the frog on the napkin in the box!* by picking up the toy frog and "hopping" it onto the empty napkin and then into the empty box (Figure X). The correct behavior would be to pick up the frog who's sitting on a napkin, and then put that frog in the empty box (Trueswell et al., 1999).

This "hopping error" reveals that children initially represent *on the napkin* as the destination of *put* instead of waiting to hear the entire sentence to interpret this phrase. If children waited until the end to interpret the command, then they would know that *on the napkin* is a modifier of *frog*, and that *in the box* is the destination of *put.* Note that *put* requires a destination, and adults (and children) resolve this requirement actively during sentence processing (Tanenhaus & Trueswell 1995; Trueswell et al., 1999). This behavior persists in follow-up experiments which make different attempts to help children overcome their initial parsing commitments (Hurewitz et al., 2000; Meroni & Crain, 2003; Weighall, 2008). The behavior of the 19 month-old infants in Lidz et al., (2017) is consistent with these observations. Infants learn that *the wug* refers to the toy train in the intransitive condition (e.g., *She's pulling with the wug*) because they incrementally parse *the wug* as the direct object, and this representation persists. The video events in these experiments are designed to align with two interpretations (i.e., *the wug* instantiates the pullee or the puller relation), so the experimental context intentionally does not provide any cues to infants that they might formed the wrong hypothesis about what *the wug* refers to.

The current experiment probes for knowledge of *wh*-movement at 19 months-old by testing whether a *wh*-object question obviates the parsing error observed in Lidz, White and Baier (2017: Experiment 1). In a modified preferential looking paradigm, infants are familiarized with either *wh*-object questions (e.g., *What is she pulling with the wug?*) or intransitive sentences known to induce a parsing error around 19 months-old (e.g., *She's pulling with the wug!*).

If infants parse *what* as the direct object in the *wh*-question condition (e.g., *What is she pulling with the wug?*), then they will parse *the wug* as the object of the preposition, interpret *the wug* as the pullee, and learn that *the wug* refers to the fishing rod. Given previous findings, the control condition (e.g., *She's pulling with the wug!*) should induce a parsing error at 19 months: infants will parse *the wug* as the direct object, interpret it as the pullee and learn that *the wug* refers to the toy train. Knowledge of *wh*-movement will reveal itself as a difference between the *wh*- condition and the control condition, with more looks-to-instrument in the *wh*- condition than the control condition.

Methods

This experiment uses a modified version (Lidz et al., 2017; Trueswell et al., 2012) of the Intermodal-Preferential Looking Paradigm (Hirsh-Pasek & Golinkoff, 1999; Spelke, 1976). In this paradigm, infants are familiarized to sentences in context: infants watch short videos of events with audio descriptions that introduce them to a novel NPs, and then they're tested on what they have learned the novel NP refers to. For example, infants watch a short video of a woman use a fishing rod to pull a toy train around in circles. During this familiarization video, the clause type used in the event description is manipulated as a between-subjects factor: infants in the *wh*- condition hear *wh*- object questions (e.g., *What is she pulling with the wug?*) and infants in the control condition hear intransitive declarative clauses (e.g., *She's pulling with the wug!*).

At test, infants see a split-screen display of still images of the train and the fishing rod from the familiarization video that they just saw. Their task is to associate only one picture with the following test audio: *Which one's the wug?* This method works because infants tend to *look longer on average* at the image which they associate with the test audio. Their looking preferences indicate which object they think *the wug* refers to, which is contingent on how they parsed the sentences during familiarization.

Materials

Each condition consists of 1 training trial (*throw*) and 6 test trials (*hit, pull, tickle, touch, wash, wipe*) (Table 1). These are the same verbs used in Lidz et al., (2017: Experiment 1), except *push*, which was excluded due to previously cited concerns that the toy bulldozer (which "Bob the Builder" fans will instantly recognize as Muck) used as the instrument in the trial was too interesting and disproportionately drew looks during test regardless of syntactic condition. These verbs were chosen because they are well-suited for infants and easy to visually depict, can be used to express events which support instruments, and sound natural in both transitive and intransitive frames[5]. The same

[5] In the progressive tense.

novel NPs used in Lidz et al., (2017) were used here.

New audio was recorded and edited using Praat and Adobe Audition. The same event videos and test images used in Lidz et al., (2017: Experiment 1) were used in the current experiment.

Table 1. Linguistic and visual stimuli (Experiment 1). Verb-Noun pairs used in the stimuli audio and the objects used as instruments and patients during the event videos.

Verb	Novel Noun	Instrument	Patient
hit	*tam*	ruler	mini traffic cone
pull	*wug*	fishing rod	train
throw	*frap*	cup	ball
tickle	*seb*	feather	mouse puppet
touch	*pint*	pipe cleaner	small pumpkin
wash	*pud*	sponge	toy car
wipe	*tig*	cloth	camera

Design

Each infant saw 7 trials total: 1 practice trial and 6 test trials. The 6 test trials were organized to create two different presentation orders such that the target answer never appeared on the same side of the screen during test for more than two consecutive trials. Approximately half the infants saw order 1, the other half saw order 2. Each trial has the same structure, consisting of two phases: a Familiarization Phase (the video portion) and a Test Phase (the split-screen display).

Familiarization Phase

The purpose of the Familiarization Phase is to introduce infants with a language-event pairing. In each 15 second video, a woman uses a toy as an instrument to manipulate another toy. The videos are cropped and angled such that the toys take up most of the screen, so only the woman's hands are visible. The linguistic stimulus was presented 3 times during the video to give infants the opportunity to parse the novel NP and learn which toy it refers to. The linguistic stimuli (*What is she pulling with the wug?* versus *She's pulling with the wug!*) is manipulated as a between-subjects factor.

The final sentence *Ooh, that's a nice wug!* is the only sentence with uninformative syntax with respect to the meaning of *the wug*, and it's the only deviation from the familiarization in Lidz et al., (2017: Experiment 1). This sentence was added to address concerns that infants would have difficulty shifting topics in the *wh*-condition between familiarization and test. During familiarization, if infants understand the question *What is she pulling with the wug?*, then they will be attending to the toy train – the pullee – because that's the answer to the question. At test, they're asked *Where's the wug?*, which is a question about the instrument (the fishing rod). Since these are young infants, it's possible that this shift from the train to the instrument would be difficult for them to accommodate, and introduce noise into the sample, which has the potential to mask any signal of linguistic knowledge. For example, infants might be drawn toward the train at test because it was so central to the familiarization phase. This behavior would make it impossible to determine that infants parsed the *wh*- object question accurately during familiarization, because it's the opposite of what's predicted. To address this concern, the sentence *Ooh, what a nice wug!* is used last during familiarization to shift infants' focus to the *the wug* at the end of the video. If they have learned that *the wug* is the instrument, then they will shift to the fishing rod. If they have misparsed the sentence and learned that *the wug* is the pullee, then they will attend to the train.

Test Phase

The purpose of the Test Phase is to determine which object infants have learned to associate with the novel NP, which is the window into how they have represented the syntax of the sentences they heard during the Familiarization Phase. A blank (black) screen separates the Test Phase from the Familiarization Phase, lasting for 2 seconds. During this 2 second period, the question *Where's the* NOVEL NOUN*?* is asked once. This is the first of two test question prompts. At the offset of the novel noun, a split-screen display of the two objects from the Familiarization Phase appears, and infants have 2 seconds of silence to look at the images on the screen. This period after the first test question (starting from the onset of the split-screen visual and ending after 2 seconds of silence) is referred to as *Window 1*. Studies which use similar methods report that infant looking patterns after the first test question indicate their baseline looking preferences, and are not an accurate reflection of how they've understood the test question (Gagliardi et al., 2016, Lidz et al., 2017, Perkins & Lidz 2019). In this experiment, looking patterns during *Window 1* are taken to indicate infants' baseline preferences.

Two seconds after the onset of the split-screen display, the second test question *Which one's the* NOVEL NOUN? is asked once. Infants have 3 seconds of silence to look at the images on the screen following this question. This 3 second period after the second test question is *Window 2*. Given the results of previous studies using similar methods, *Window 2* is the region of interest, or the time window during which an effect of syntax will surface (if it exists).

Table 2. Detailed structure of single trial (Experiment 1).

Phase	**Duration**	**Video**	**Audio**
Pre-trial	2 seconds 5 seconds	Blank screen Smiling baby	*silence* Baby giggle
Familiarization	15	Train being	Hey, look at that! (What is)

		seconds	pulled by a fishing rod	she('s) pulling with the wug? Wow! (What) do you see her wiping with the wug? Yay! (What is) she('s) pulling with the wug? Ooh, that's a nice wug!
Test		2 seconds	Blank screen	Where's the wug?
	[window1]	2 seconds	Split screen: train and fishing rod	*silence*
	[window2]	3 seconds		Which one's the wug?

Procedure

Infants were tested in person in a laboratory setting. Before beginning the study, infants played with researchers in the lab waiting area while parental consent was obtained from their guardian. Infants and guardians were escorted to a small sound-proofed testing room, where infants sat on their guardian's lap or in a highchair centered six feet from a 51 inch, wall-mounted television monitor. Guardians were instructed to look away from the screen and to resist talking to their child or trying to divert their attention in any way while the experiment was in session. The experiment lasted less than 5 minutes, and infants who became too restless or fussy were given a 10 to 15 minute break to play in the waiting area before being brought back to finish the experiment.

Infants were recorded during the experiment using a digital camcorder with a sample rate of 30 frames per second mounted above the television monitor. The angle and zoom of the camera was controlled by a trained researcher in another room to ensure that the infants' eyes were fully visible in the recording at all times.

A trained researcher coded recordings offline frame-by-frame for direction of look (right or left) using the SuperCoder program (Hollich, 2008). The researcher was blind to the syntactic condition and coded recordings without audio to eliminate bias.

3.4 Participants

64 19 month-old infants (32 male, mean age=19;25, range=19;2-21;3) were included in the final sample of this study.

3.5 Results

This experiment tests the hypothesis that 19 month-olds represent *wh*-movement. If infants represent *wh*-movement, then in the critical *wh*- object condition (e.g., *What is she pulling with the wug?*) they will parse *what* as the direct object, and infer that the pullee (e.g., the train) is the answer to the question. This will allow them to parse *the wug* as the object of the preposition *with*, interpret it as the puller (or instrument), and learn that *the wug* refers to the fishing rod. During test (i.e., *Where's the wug?*) infants will look more toward the fishing rod compared to the train. In the control condition (e.g., *She's pulling with the wug!*), 19 month-olds will mistakenly parse *the wug* as the direct object, interpret it as the pullee, and learn that *the wug* refers to the train – replicating the parsing error observed in Lidz et al., 2017. At test (i.e., *Where's the wug?*) infants will look more toward the train. If infants look more toward the fishing rod than the toy train at test in the critical condition, then they parsed *what* as the direct object of the verb during familiarization, and therefore represent *wh*-movement. If there is no difference in behavior between the critical and control conditions, then infants did not parse *what* as the direct object of the verb, and therefore do not represent *wh*-movement (or failed to deploy this knowledge during the experiment).

Results show that infants in the critical condition (e.g., *What is she pulling with the wug?*) look reliably more toward the fishing rod than infants in the control condition (e.g., *She's pulling with the wug!*) (Figure 2). We fit a linear mixed-effects regression model of proportion looks to instrument in R using the lme4 package (Bates et al., 2015), including as predictor variables test window, condition, and their interactions. Random intercepts were included for participants and items. Model comparison revealed a significant interaction between test window (baseline vs. test) and condition (control vs. *wh*-question) (χ2 = 3.97, p <0.046). Therefore, 19 month-olds must (a) know the syntax of *wh*-questions and (b) rapidly deploy that knowledge during sentence processing.

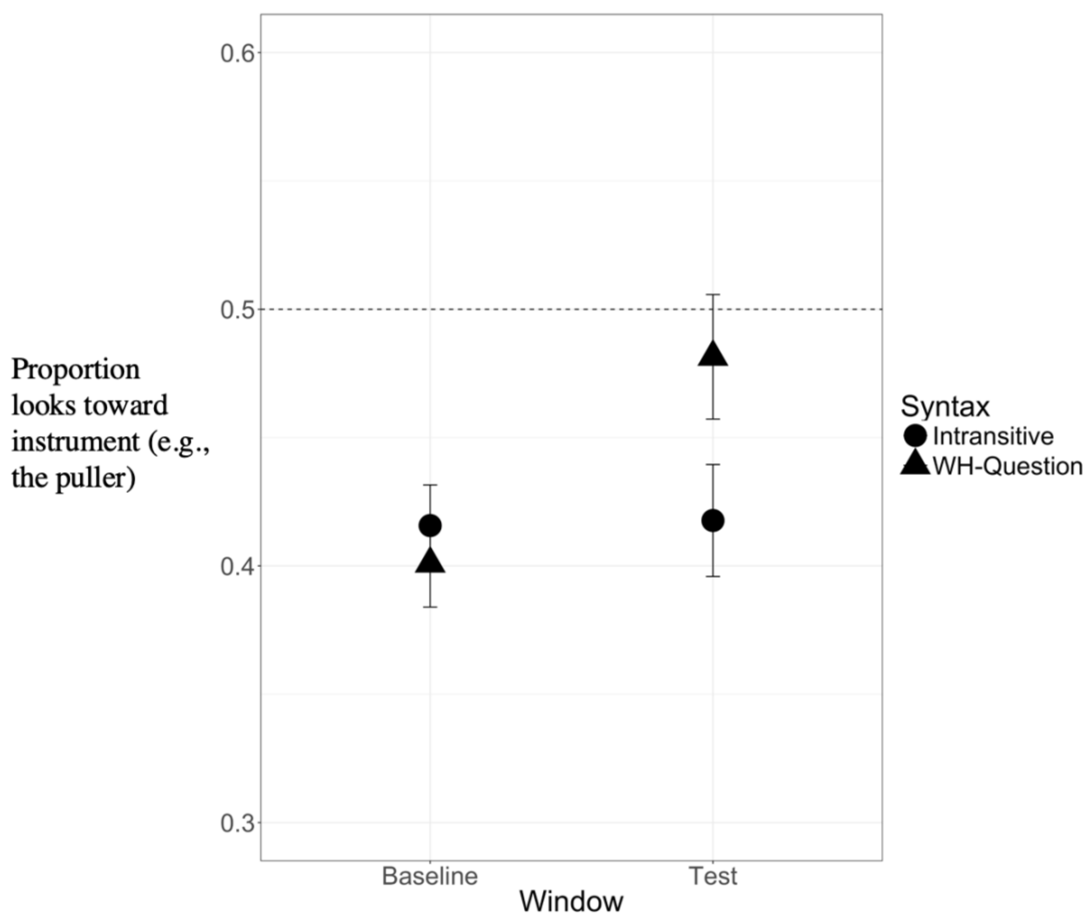


Figure 2. Average infant looking preferences at Baseline (e.g., *Where's the wug?*) and Test (e.g., *Which one's the wug?*). Infant looking preferences shift to the instrument in response to the second test question in the *wh*- question condition, but not the intransitive control condition.

3.6 Discussion

The only difference between the two experimental conditions is the syntax used to introduce the novel NP (*What is she pulling with the wug?* vs. *She's pulling with the wug!*), so knowledge of the syntax of *wh*-object questions is responsible for infants' success. This result shows that by 19 months-old, infants have knowledge of *wh*-movement and can use that knowledge to understand *wh*- object questions during language comprehension tasks. This finding is consistent with the hypothesis that upon acquiring a grammar, children's sentence processing mechanisms develop according to knowledge of grammar, as opposed to developing independently. Without knowing that the *wh*- phrase is related to an upcoming syntactic position, it's not clear that infants would be able to parse it as the direct object in time to satisfy the transitivity of the verb and successfully interpret the novel noun phrase as the instrument.

This interpretation of the results is committed to the following hypothesis about sentence processing procedures. When infants hear a sentence, they automatically assign a syntactic representation as the sentence unfolds. For example, given the simple sentence *She's pulling the gop*, 19 month-olds hear *She…* and represent *she* as the subject. When they hear *…'s pulling…*, they incorporate this into the syntactic structure that they're building as the verb. Because they recognize that *pulling* is a transitive verb, incorporating this fragment entails building a direct object slot. This is what is meant by incremental or predictive parsing. Upon hearing *…the gop*, they use this to fill the direct object slot in their syntactic representation. Returning to the experiment presented in the current chapter, 19 month-olds do the same thing in the control condition. They incorporate *the gop* into their syntactic representation as the direct object, and treat *with* as noise. In the *wh*- question condition, the first thing that infants hear is *what*…. It's possible that at this moment, infants parse *what* as the direction object, but given that infants hear a variety of utterances beginning with *what* that are not direct object questions (e.g., *What a good baby!, What are you doing?*, *What?,* etc.) it cannot be taken for granted that they do so. However, when they hear *…is she pulling…* they incorporate *what* into their syntactic representation as the direct object, filling the direct object slot that they predicted when they parsed *pulling*. From then on, they proceed to incorporate *with the wug* into their syntactic representation.

This is a strong hypothesis, that infants predictively process *wh*- object questions. Because this experiment uses offline measures, the evidence for this hypothesis is indirect, and further study is required to determine the exact timeline of these linguistic predictions. It's possible that infants arrive at the correct interpretation after they've heard the sentence, but critically, something about the syntax of the *wh*-question allows them to overcome this robust sentence processing error made in the intransitive declarative condition. The strong hypothesis raises interesting questions given Atkinson's work on predictive *wh*- processing in older children, which fails to find evidence for predictive processing younger than 6 years-old. Because different linguistic materials and experimental set-ups are used for infants and older children, it's not possible to compare these studies directly. Future work is needed to refine the developmental timeline of the development of predictive *wh*- processing.

In previous experiments (Lidz et al., 2017), when familiarized with an uninformative description, like *It's a gop!*, infants show a preference for the object used as the pullee. This slight preference is revealed in the current study after infants have been prompted with the first (baseline) test question. One possible explanation for this bias is that the objects used as patients happen to be more attention-grabbing than the objects used as instruments. This explanation is unconvincing; it's just as easy to argue that the objects used as instruments are more attention-grabbing during the videos because they are the source of the movement on the screen, so why shouldn't infants show a baseline preference for the instrument? Alternatively, this *patient bias* could be driven by a conceptual structure which comes naturally to the minds of infants when they see these videos in which the pullee is highlighted, or has a privileged role compared to the instrument. This conceptual structure must hook up with the linguistic description. When infants hear the uninformative description *It's a gop!* they align this with their event concept in a very natural way: the gop refers to the object which embodies the privileged role: the train. This hypothesis is consistent with the generalization that transitive verbs often appear in transitive frames. Moreover, it adds some richness to the explanation of why infants seem to be unable to revise their initial interpretive commitments in the control condition (e.g., *She's pulling with the gop!*). If infants' conceptual structure naturally highlights the patient, then a linguistic description which has left out the syntactic argument which encodes the patient relation is an unnatural description, from their perspective. For adults, it's easy to take a different perspective based on linguistic input, but for infants this might be more difficult. The control condition would require that infants take a perspective on the event were the pullee is completely backgrounded (i.e., not explicitly represented in the linguistic frame). It's possible to test this hypothesis. For example, would infants notice if the object that was used as the instrument suddenly changed (say, from a red fishing pole to a blue one)? If neither the conceptual structure nor linguistic representation draw their attention to the instrument, then they would not notice even a dramatic change to it. On the other hand, would a similar change in the object used as the pullee elicit a response?

By focusing on sentence processing, this experiment circumvents many of the issues with the proto question-after-story tasks developed for young infants. However, because the control condition is designed to invoke a processing error characteristic of a particular age group, 19 months-old is the only age appropriate for this experimental paradigm. Another limitation of the experiment is the kinds of wh-questions that it can be used to test. It's difficult to expand it to *wh*-questions which more directly probe the instrument relation, for example, *What is she wiping the gop with?* because it's unclear if infants will be able to accommodate a filled gap, or if they will respond to filled-gap effects uniformly, such that it their looking behavior would reveal a pattern. It has been suggested that comparing *What is she pulling the wug with?* to *What is she pulling with the wug?* would be a better minimal pair for this experiment, since these sentences are closer linguistic matches to each other, and a more natural continuation to the experiments in the original study. This comparison complicates the sentence processing predictions. For example, in the case of *What is she pulling the wug with?*, children should incrementally interpret *what* as the direct object – but then they will encounter *another* direct object in the canonical direct object position: *the wug*. How will these be reconciled? Both *what* and *the wug* cannot be the direct object because it's not possible to have more than one direct object for a single verb. Adult sentence processing studies reveal that encountering a second direct object where the gap should be is disruptive enough that its effects can be detected and quantified. The study of *filled gap effects* has been a useful window into the predictive nature of *wh-* processing. While reading, adults slow down or change their looking pattern when they encounter a filled gap. It's reasonable to expect that this would also cause difficulties for children, but this experiment does use a method which can detect filled gap effects. The inferences that we draw here about developing linguistic representations are indirect, from the way that infants interpret the sentences (as evidenced by what they learn that the gop refers to), and this method is not suitable for observing their behavior in real time.

3.7 Conclusion

This experiment provides important detail about the timeline of the acquisition of early syntax by showing that at 19 months-old, infants use their knowledge of *wh*-movement to understand *wh*-object questions during language comprehension. If children intrinsically represent *wh*-movement and its locality constraints, then it is possible [for linguists] to observe the effects of knowledge of locality constraints at a very early age. This experiment provides empirical support for a core assumption made by learning models of island constraints: like adults, infants use their grammatical knowledge of *wh*-movement to process *wh-* questions incrementally. Despite the puzzling results from *wh-* processing studies on older children, the results of the current experiment suggest that knowledge of grammar influences sentence processing from an early age.

Chapter 4: 3 year-olds respect island constraints

4.1 Introduction

Research shows that 4 year-olds have acquired knowledge of island constraints (de Villiers et al., 1990; de Villiers & Roeper, 1995a, 1995b; Fetters & Lidz, 2016; Goodluck et al., 1992). The experimental results from Chapter 2 show that 19 month-olds represent *wh*-movement, about 2.5 years before their 4th birthday. The question remains, what is the nature of early syntactic representations of *wh*-movement? Do 19 month-olds have knowledge of island constraints, and we (as researchers) just have yet to observe it? Or, is knowledge of island constraints acquired from experience (Pearl & Sprouse, 2013)? This second hypothesis would explain why knowledge of island constraints hasn't been observed in younger children, but it's at odds with conclusions drawn from syntactic theory about the nature of island constraints. Island constraints are cross-linguistically robust, which indicates that children never fail to acquire them (although certain aspects of them are subject to parametric variation). As discussed in Chapter 1, these constraints are abstract in nature, and are underdetermined by the set of sentences generated by the adult grammar, and certainly underdetermined by the input to language acquisition. These observations align with the hypothesis that island constraints are part of Universal Grammar (Berwick et al., 2011; Chomsky, 1965, 1980).

For a proponent of the innate hypothesis of island constraints, the difficulty lies in devising an experiment which would reveal this knowledge. It's not enough for children to represent *wh*-movement and island constraints; children must also represent the clausal relations of the sentence in an adult-like way, and have the ability to apply this knowledge accurately throughout the course of the experiment. A child who demonstrates knowledge of island constraints has only done so after making many discoveries about her language (Otsu, 1981). An experiment which shows evidence for knowledge of island constraints before 4 years-old is consistent with the hypothesis that knowledge of island constraints is innate, but an experiment that fails to show knowledge of island constraints is not evidence against the innate hypothesis. Great care is required to settle the question of how children acquire knowledge of island constraints. The experiments presented in this chapter are one step along the way.

Turning to the question of how children develop knowledge of island constraints, there are two established proposals. One proposal is that knowledge of island constraints is innate, so knowledge of *wh*-movement will appear inseparable from it. Upon acquiring knowledge of *wh*-movement, children already have knowledge of island constraints, and they might be able to demonstrate this knowledge given an understanding of clause structure. This proposal is born in the tradition of generative syntax. The other proposal is that children acquire *wh*-movement, and then they learn island constraints from the *wh*-argument questions in their linguistic input (Pearl & Sprouse, 2013). These proposals make different predictions about when knowledge of island constraints will emerge in development. The role of the input in the first proposal is very small, so even very inexperienced learners – even infants – might respect island constraints, so long as they accurately represent the clause structure of their language. Again, even if infants have knowledge of island constraints, we (as researchers) are not guaranteed to have the tools to measure it successfully. The best we can do is try.

In the second proposal, certain properties of *wh*-argument questions will indicate to learners that their language has island constraints. As a consequence of this learning by induction, there will be a developmental stage during which children have knowledge of *wh*-movement, but have imperfect knowledge of island constraints. ‘Imperfect knowledge’ is used here to mean that children are not aware of some or all island constraints, and not that they are fine-tuning various parametrically determined linguistic properties (e.g., what the bounding nodes are in their language). The strongest evidence for this proposal would be a solid experiment which shows that children at some age have thorough knowledge of *wh*-movement and the clause structure of their language, but they obey no island constraints. If children obey some island constraints but not others, then the warranted conclusions weaken considerably.

This experiment uses a question-after-story task to show that 3 year-olds have acquired knowledge of a locality constraints on *wh*-movement. Following experimental techniques introduced in the 1980s-1990s for testing knowledge of island constraints in children, the test questions in the experiment are designed to be ambiguous for children who have not yet acquired knowledge of the locality constraint, but unambiguous for children who have acquired this knowledge (de Villiers et al., 1990). The major advantage of this technique is that children are never asked questions that have no interpretation (in other words, questions that are nonsense). For example, (35) is ambiguous, and has two relevant interpretations: (35a-b) and (35c-d). In (35a-b), the book is used as an instrument of grabbing (35a) and *with the book* modifies VP (35b). In other words, *with the book* is used in an adverbial way, so (35a-b) is an adverbial interpretation of the PP. In (35c-d), the book is used as a unique identifier of the bear (35c) and *with the book* modifies NP (35d). In other words, *with the book* is used in an adjectival way, so (35c-d) is an adjectival interpretation of the PP.

35. Mariam grabbed the bear with the book

    a. Mariam used the book to grab the bear

    b. Mariam [$_{VP}$ [$_{VP}$ grabbed [$_{NP}$ the bear]] [$_{PP}$ with the book]]

    c. Mariam grabbed the bear who was holding the book

    d. Mariam [$_{VP}$ grabbed [the bear [$_{PP}$ with the book]]]

When *the book* undergoes *wh*- movement, the resulting *wh*- question is unambiguously a question about the instrument of grabbing (36a). Critically, the adjectival interpretation no longer available (36b). Example (37) uses a *wh*- cleft to ask the same question, but avoids using the *wh*- movement rule in the same way as (36b), and both interpretations are available again.

36. Which book did Mariam grab the bear with?

    a. Which book did Mariam [$_{VP}$ [$_{VP}$ grab [$_{NP}$ the bear]] [$_{PP}$ with _]] ?

       Answer: the book that Mariam is holding

    b.* Which book did Mariam [$_{VP}$ grab [the bear [$_{PP}$ with _]]] ?

37. Mariam grabbed the bear with the book. Which book was it?

    a. Mariam [$_{VP}$ [$_{VP}$ grabbed [$_{NP}$ the bear]] [$_{PP}$ with the book]] . Which book was it?

       Answer: the book that Mariam is holding

    b. Mariam [$_{VP}$ grabbed [the bear [$_{PP}$ with the book]]] . Which book was it?

       Answer: the book that the bear is holding

Example (36a) shows *wh*-movement from a VP adjunct, which is unrestricted (Chomsky, 1986). There's a locality constraint on partial extractions from NP: *wh*-movement from NP-adjuncts is blocked. Additionally, it's well-known that *wh*-movement is sensitive to definiteness effects (Ross, 1967). It's worth noting that the locality constraint leveraged here is not a typical exemplar of island constraints – the features that make it an excellent test case for young children and infants make it a poor test case for syntacticians. Mainly, this test question is mono-clausal. The tests developed by syntacticians require enough linguistic material to see the effects of the test – but utterances with multiple clauses pose several problems for testing children and infants. These utterances often require elaborate backstories with multiple characters, which both increase the length of the overall experiment and the difficulty of the task. Multi-clause utterances pose linguistic problems as well; additional testing is required to ensure that children understand the embedding verbs used in the study.

In this experiment, children watch short videos of a woman using a book (for example) to interact with a teddy bear who is holding another book. After the video, half the children hear an unambiguous follow-up question like (36), and the other half hears an ambiguous follow-up question like (37). If children give the same responses regardless of what question they heard, then the experiment has either failed to detect early knowledge of island constraints, or children do not have knowledge of island constraints. If children give different responses to each question type, and they prefer to answer with the book used as the instrument when asked questions like (36), then this behavior is consistent with knowledge of island constraints. Results are consistent with the hypothesis that children respect island constraints.

4.2 Acquisition Background

To test for knowledge of island constraints in children, the most popular method is to measure children's responses to carefully built *wh*- questions (de Villiers et al., 1990; Otsu, 1981). Question-after-story tasks work well for several reasons. One, *wh*-questions are a relatively common instance of *wh*- movement in speech to children (Newport et al., 1977), and have received much attention in comprehension and production studies. Two, these *wh*- questions allow researchers to probe children's linguistic representations indirectly by observing how they answer the question. As mentioned in the Introduction, many studies test children with *wh*- questions which are only ambiguous in the absence of knowledge of island constraints. If children respond the same way to these critical questions as they do questions which are ambiguous regardless of knowledge of island constraints, then the experiment fails. If children respond differently to critical questions than to control questions, the experiment provides evidence for knowledge of island constraints in that age group.

Many early studies tested wider age ranges with fewer children overall, and feature a wide variety of control conditions, making results difficult to compare and interpret. For example, in a study of 5 year-olds that tests for knowledge of the Complex NP Constraint, children listened to illustrated stories (38) and answered follow-up *wh*-questions like (38) (de Villiers & Roeper, 1995a).

38. *Story*: These little girls really wanted to go on a trip on their bikes, but the older girls didn't want them to come. This little girl found a note in the trash about a plan the big girls had to leave for the beach early in the morning! So the little girls got up really early and surprised them!

39. Where did they discover _ [$_{NP}$ the plan to ride *(_)] ?

Answer: in the trash/*to the beach

Without knowledge of island constraints, (39) is ambiguous. It could be a question about the plan, in which case the answer would be *to the beach*. It could also be about the discovery, in which case the answer would be *in the trash*. However, *the plan to ride* is a Complex NP, or an NP which contains a sentence or clause, which blocks *wh*-movement. If children give answers about the location of discovery, then they have interpreted (39) in an adult-like way. If children give answers about the location of riding, then they have violated the Complex NP Constraint. Results show that children most often give answers about the location of discovery. To ensure that this behavior is linguistically driven (and not influenced by extra-linguistic factors), the same children were tested on co-reference interpretations (long-distance syntactic binding relations) into Complex NPs. Children heard brief stories (40) followed by a test question (41). The idea is that children who allow co-reference interpretations into Complex NPs must know that the pronoun is syntactically bound by an antecedent outside of the Complex NP, which demonstrates that they are capable of representing long-distance syntactic relations. But since they don't in the *wh*-question condition, then this must be because knowledge of the Complex NP Constraint is driving their behavior. Interestingly, when the matrix verb is *make* (a 'light' verb) instead of *discover* in (39), children violate the Complex NP Constraint (as do adults (Ross, 1967)). This wrinkle highlights the difficulties that come with testing young children on multi-clausal utterances, and ultimately make the study difficult to interpret.

40. *Story*: Now we have Ernie and Big Bird. And a feather! Show me what

happens when I say:

41. Big Bird overheard a plan to tickle him. Who got tickled? (Child must point to a toy)

    Co-reference: him refers to Big Bird, so Big Bird got tickled.

    Free reference: him refers to Ernie, so Ernie got tickled.

Carefully considering the test question reveals another problem that can arise in experimental contexts. The *wh*- question following the story is clearly some kind of memory test. The story provided the correct answer, which is also the only possible answer. An improved story would feature a location of near-discovery, and a location that the girls considered visiting. This would make it possible to evaluate whether children were paying attention during the story, and make explicit the alternatives under consideration, which is something that children have been shown to have difficulty with (Noveck, 2001).

A set of experiments testing children ages 3 to 5 years-old for knowledge of temporal adjunct islands suffers the same problems (Goodluck et al., 1989, 1992). Children heard stories like (42), followed by *wh*-questions like 43).

42. The fox ran down to the river.

    He ate an icecream cone.

    Then he whistled a tune he'd heard on the radio.

    The fox felt pretty happy.

43. What did the fox eat _ [$_{CP}$ before whistling *(_)] ?

    Answer: the icecream cone/*a tune

The phrase *before whistling* is a temporal adjunct, which is an island for *wh*-

movement. Optionally intransitive verbs (e.g., *eat*, *whistle*) support two interpretations for (43): one in which *what* is the direct object of *eat*, and another in which *what* is the direct object of *whistling*. 3 and 4 year-olds tend to give matrix responses to wh-questions like (43), but it's unclear whether or not knowledge of island constraints drives this behavior. The Control Conditions test children's willingness to assign long-distance interpretations more generally e.g., *Who did the elephant ask before helping?*), and furthermore show different preferences at different ages, making it difficult to draw strong conclusions about knowledge of island constraints.

Similar difficulties surface in studies on relative clause islands (de Villiers & Roeper, 1995b) and embedded questions (de Villiers et al., 1990). For example, children are tested on ambiguous questions like (44), and questions like (44) which contain embedded questions (e.g., *how to paint*). Note that this experiment also contrasted embedded adjunct questions like (45) with embedded argument questions, not shown here. Again, variation in children's answers to the control condition make their answers in the critical condition difficult to interpret.

44. When did the boy say he hurt himself?

    a. When did the boy say _ he hurt himself?

    b. When did the boy say he hurt himself _?

45. Who did Big Bird ask how to paint?

    a. Who_1 did Big Bird ask $\__1$ [how_2 to paint $\__2$]?

    b.* Who_1 did Big Bird ask [how_2 to paint $\__1$ $\__2$]?

In a study of relative clause islands, story (46) features two boys who each drink their own milk, in their own unique way. This provides a natural reason to use a relative clause to pick out one boy in the follow-up question (47) (e.g., the boy who drank with a straw vs. the boy who drank straight from the carton). But, there's an asymmetry here: the *wh-* question lends itself better to be about the manner of drinking than the manner of sneezing. There are no alternatives provided for manner-of-sneezing, and the straw and carton have clear visual correlates. Children prefer to associate the *wh-* phrase with the matrix verb, but it's unclear whether this was driven by knowledge of island constraints or the research materials.

46. These two brothers went to the circus. The clown came and tickled the little boy on the nose with a feather. He sneezed very hard and blew the clown's wig right off! After the circus they were very thirsty and they went to buy some milk. The little boy drank his milk with a straw but the big brother drank his milk straight from the carton.

47. How did the boy drink _ who sneezed *(_) ?

    Answer: with a straw/*very hard

Otsu's (1981) pioneering work on the acquisition of islands uses a slightly different syntactic control: a series of independent syntactic tests was developed to evaluate whether children in each age group had mastered the relevant syntax for a particular island (e.g., the structure of relative clauses). In the island condition, children were told simple stories like (48) with follow up questions like (49). In (49), the potential ambiguity stems from how *with* is related to the rest of the sentence. Is *with* contained by the relative clause (island-violating parse) or is it outside the relative clause?

48. Jim is catching a cat with a net. The cat is climbing up a tree with a ladder.

49. What is Jim catching a cat that is climbing up a tree with?

Answer: a net/*a ladder

Children's performance improves with age: by the time children are 6 years-old, they prefer to associate the PP with the matrix clause, and also perform well on the independent syntactic tests. Children younger than 6 years-old perform poorly on both tasks, and even show a preference to associate the PP inside the relative clause at ages 3 and 4. However, it's difficult to conclude that this behavior is driven by immature knowledge of island constraints, because younger children also fail to demonstrate they have knowledge of relative clauses. Otsu (1981) concludes that only children who had acquired the syntax of *wh*-questions and relative clauses also had knowledge of island constraints, and argues for the weak conclusion that children who hadn't acquired the syntax of relative clauses were unable to demonstrate knowledge of the relative clause island constraint.

Recent work overcomes many of these problems by including alternatives in the story, and a sluice condition as appropriate syntactic control, which allows for the stories to be held constant across both the test and control conditions (Fetters & Lidz, 2016). In a question-after-story task, 4 year-old children watch videos of a person (Mike Fetters) using little toys to act out pre-recorded stories like (50). Children are tested on either relative clause islands (51) and or their sluiced counterparts (52). The ambiguity again rests on whether *with* is contained by the relative clause (island-violating parse) or whether *with* is outside the relative clause.

50. Daddy Turtle and Baby Turtle were out in the woods looking for food, when they spotted a big, yummy onion ring. Daddy Turtle and Baby Turtle were

taking the onion ring home when Troll saw them. Troll wanted the onion ring for himself, so he started to chase Daddy Turtle and Baby Turtle. Daddy Turtle and Baby Turtle carried the onion ring back to their turtle home, but they needed to hide it from Troll!

"I think I know how to cover the onion ring," said Baby Turtle, and he went to grab a leaf. He was about to cover the onion ring with the leaf, but it kept blowing away!

Just then, Crab came by and said, "Troll is coming for you, let me throw you into your home so he can't find you!" Crab got his catapult and threw Baby Turtle back into the turtle home.

"I think I know how to cover the onion ring," said Daddy Turtle, and he got a blanket. Daddy Turtle covered the onion ring with the blanket so Troll couldn't see it.

"Let me throw you into your home so Troll doesn't find you," said Crab, and Crab got a big spoon and threw Daddy Turtle back into the Turtle home. Troll came by and couldn't find the onion ring, so he left. The Turtles got to eat the onion ring for dinner!

51. Can you tell me what Crab threw [$_{NP}$ the turtle that covered the onion ring] with?

    Answer(s): *the blanket, the spoon

52. Crab threw the turtle that covered the onion ring with something. Can you tell me what?

    Answer(s): the blanket, the spoon

A child who hasn't acquired knowledge of island constraints might answer with *the blanket* because Daddy Turtle used a blanket to covered the onion ring in the story. If they answer with *the leaf*, then they weren't paying attention or couldn't remember what happened because Daddy Turtle almost used the leaf to cover the onion ring, but didn't in the end. In the *wh*-question condition, 4 year-olds prefer to answer *wh*-questions with an instrument of throwing (e.g., the spoon). In the sluice condition (52), children prefer to answer with the instrument of covering (e.g., the blanket). Since the only difference between the two conditions is the presence of an island in (51), but not (52), Fetters & Lidz, (2016) conclude that children's behavior is driven by knowledge of island constraints.

A critical review of these experiments reveals that a valid question-after-story task testing for knowledge of islands must have (1) alternatives answers, and a syntactic control which (ideally) is different from the *wh*-question only in the dimension of islandhood. The following section explains how the current experiment meets both criteria while staying baby-friendly.

4.3 Experiment 2: *Wh*- questions (children)

This experiment is designed to test 3 year-olds for knowledge of locality constraints on *wh*-movement. The design and stimuli take infant preferential-looking studies as inspiration to maximize the chances that it will be effective for testing younger children, but relies on the question-after-story task and ambiguous *wh*-questions used in previous research focused on older age groups.

In this question-after-story task, the critical *wh*-question is unambiguous for children who respect locality constraints, but ambiguous for those who don't. This experiment takes advantage of simple ‘telescope ambiguities’ to package a small, simple ‘island’ into a mono-clausal *wh*- question suitable for testing very young learners. Recall that the sentence *Mariam hugged the bear with the blanket* is at least two-ways ambiguous: under one interpretation, the bear is holding onto the blanket (= the adjectival interpretation). Under the other interpretation, Mariam is using the blanket as an instrument to hug the bear (= the adverbial interpretation). Critically, a *wh*-question formed by fronting the object of the preposition *with* (e.g., *Which blanket did Mariam hug the bear with?*) only has the adverbial interpretation, and can only be answered with Mariam’s blanket. The adjectival interpretation does not survive *wh*-movement. Note that the adjectival interpretation survives in the sluiced counterpart (e.g., *Mariam hugged a bear with something. What was it?*) which we take as a diagnostic for islandhood (e.g., (Ross, 1969)). So, the adjectival interpretation of the prepositional phrase isn’t available for the *wh*- question *Which blanket did Mariam hug the bear with?* because it is blocked by a locality constraint.

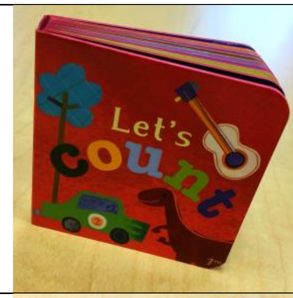
Looking ahead to the experimental stimuli, if a *wh*- question like *Which blanket did she hug the bear with?* is asked in a situation where there are two blankets – one held by the bear, and another one used by Mariam to hug the bear – then both interpretations of the *wh*- question are supported by the context – it’s the syntax of the *wh*- question which restricts the interpretations.

As a syntactic control, children see the same context but hear a truly ambiguous question: e.g., *She hugged the bear with the blanket. Which blanket was it?* If children choose the blanket held by the bear in response to the ambiguous question, this indicates that the context supports this interpretation. It is crucial that children show a preference for the bear's blanket in the control condition. If children prefer the bear's blanket in both the control and critical conditions, then it's possible that they consider both test utterances ambiguous, and do not respect island constraints. If children prefer the bear's blanket in the control, but not the critical condition, then it's likely that their knowledge of locality constraints affected their interpretations of the test question in the critical condition. But if children prefer Mariam's blanket in the control condition and the critical condition, it's unclear if knowledge of locality constraints influences their preference in the critical condition (see Experiment 3 for further discussion). If children prefer Mariam's blanket in the control condition and the bear's blanket in the critical condition, this would be very mysterious.

Materials

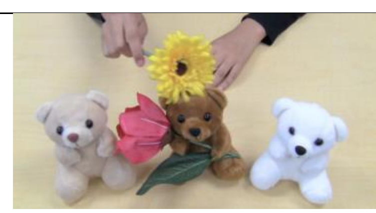
To avoid a situation where children prefer Mariam's blanket in the control condition, stimuli are designed to draw attention toward the object held by the bear and support the NP adjunct interpretation of the ambiguous sentence *She hugged the bear with the blanket* (Figure X). For example, the verbs used in this study (*hug, touch, pick, grab, move,* and *hold*) were easy to depict visually, but were also selected because adult speakers prefer to interpret *with* PPs as instruments at much lower rates for these verbs compared to other transitive verbs (e.g., *hit*) (Trueswell and Snedeker, 2004). The objects chosen for the study are blankets, books and flowers, which are non-canonical instruments, but are also familiar to 30 month-old infants (Frank et al., 2016). Each trial featured a pair of objects, which differed along shape and color dimensions to keep them distinct. Since the test portion of the trial requires children to choose one of the objects, it was important for them to be distinct without being drastically different in size so that they would occupy approximately the same amount of space on the test screen. Objects that were primary colors (red, yellow or blue) were used because they are easy to distinguish from each other. Because there were six verbs and 3 objects, each pair of objects was reused once during the experiment (Table X).

Table 3. Verb and object pairings (Experiment 2).

Verb	Object1 label	Object1 image	Object2 label	Object2 image
grab/move	blue book		red book	
hold/hug	yellow blanket		blue blanket	
pick/touch	red flower		yellow flower	

Each verb was depicted with a unique event. Because of the (intentionally) odd instruments and the fact that Mariam is interacting with small, stuffed bears, the events have a pretend-play feel to them (Table X). Note that each event was filmed twice: once with Object1 as the instrument, and once with Object2 as the instrument. The color of the instrument was manipulated as a between-subjects factor. See Figure X for a storyboard of the internal trial structure.

Table 4. Verbs, event descriptions and event depictions (Experiment 2).

Verb	Event Description	Event	
grab	Mariam uses the book to grab the bear and slide it		
move	Mariam uses the book to lift the bear off the table and set it down		
hold	Mariam uses the blanket to hold the bear by the ear		
hug	Mariam uses the blanket to give the bear a hug		
pick	Mariam uses the flower to bop the bear on the head		
touch	Mariam uses the flower to stroke the bear on the head		

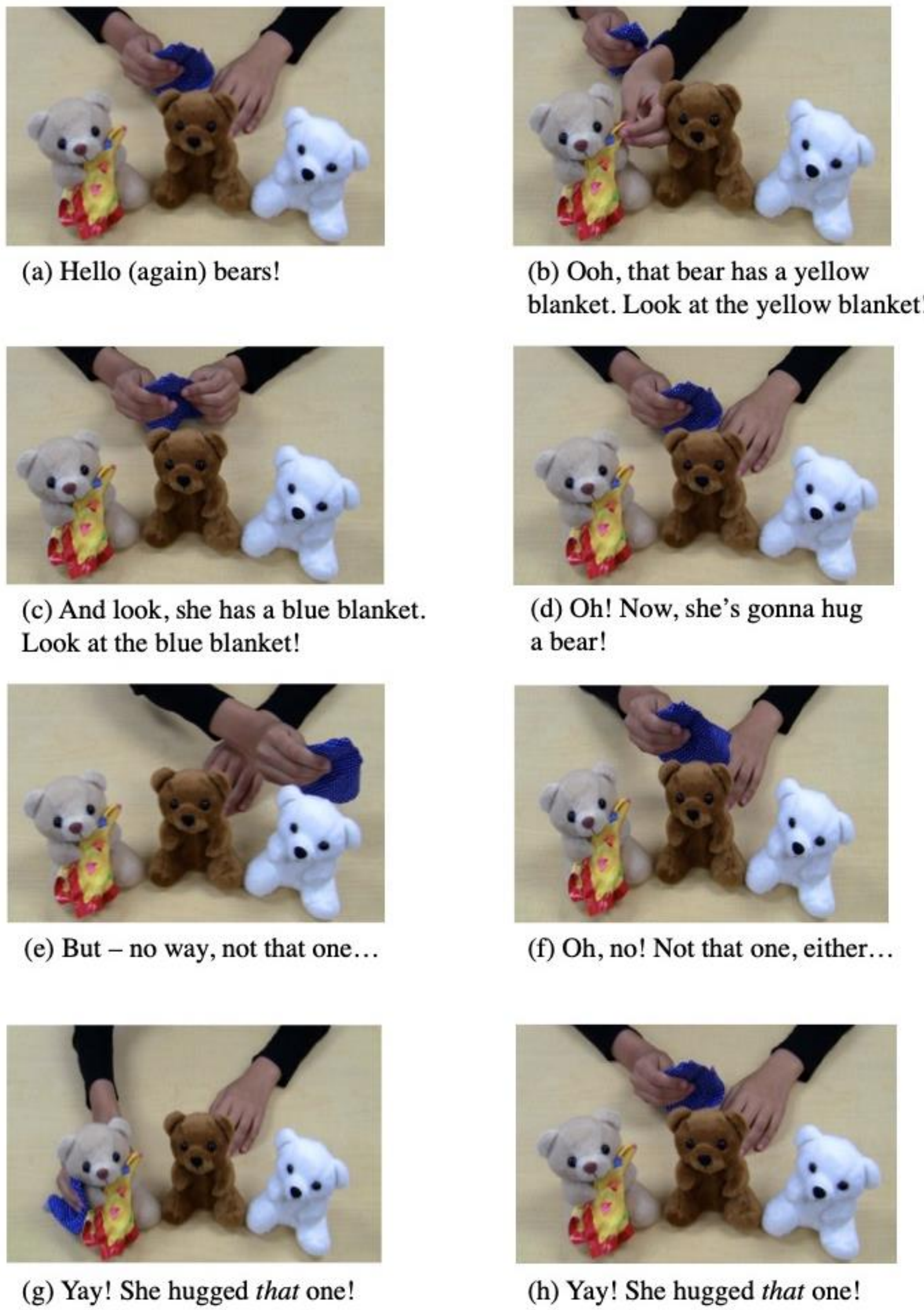


Figure 3. Storyboard of the internal structure of a *hug* trial (Experiment 2).

Table 5. Detailed explanation of Storyboard structure (Experiment 2).

(a)	This greeting brings children’s attention back to the screen at the beginning of the trial, and makes it clear that the bears are what matter most in the video.
(b)	Mariam reaches across and gently wiggles the blanket that the bear is holding, while the narrator (author) repeats its label twice (e.g., yellow blanket). This gives children two chances to hear the object label, and to recognize the blanket as a (tiny) blanket, etc. Note that without this explicit linguistic labelling routine, even adults found it difficult to remember which blanket was used for what purpose once the test phase of the experiment begins (result obtained during piloting).
(c)	Mariam withdraws to her own blanket, wiggling it while the narrator labels it in the same way as before.
(d)	Children notified about what’s coming.
(e)	The first bear is rejected.
(f)	The second bear is rejected. Bears are rejected one-by-one to make it particularly salient that Mariam is after the bear who’s holding an object. The presence of the other two bears, and the fact that one can be uniquely identified by being a blanket-holder, makes it appropriate to use a PP modifier to point to a particular bear during test. Critically, this gives supports the island-violating interpretation of the test question in the island condition.
(g)	Mariam finally gives a little one-handed hug to this lucky bear by using the blanket that she’s holding.
(h)	Mariam returns her hands to their starting position, and the familiarization video ends, transitioning to black screen before the test audio begins.

When the familiarization video ends, children see 2 seconds of black screen, after which they hear the test question. At the offset of the test question, children see the two objects from the trial presented as a split screen (Figure X). The side of the screen that the objects appeared on was balanced as a between-subjects factor due to concerns with interference from previous trials placing too much burden on children’s memory. This is not meant to be a memory test, but a linguistic test.

Figure 4. View of a *hug* test trial (Experiment 2).

In the critical (*wh*-question) condition, children heard a test question like (53), and in the control condition, children heard a test question like (54). (54) is ambiguous in the adult grammar, but maintains as much of the syntax and word order of (53) as possible without containing *wh*-movement from an island domain.

53. Which blanket did she hug the bear with?
54. She hugged the bear with the blanket. Which blanket was it?

The video narration was pre-recorded by the author using child-directed speech. The pause between the verb and its direct object was altered to keep the duration consistent across both conditions, so as not to introduce prosodic bias. To see what children found salient about the experiment given only the familiarization video, it was essential to keep the test question in the control condition as 'ambiguous' as possible. Otherwise, the control condition wouldn't be a measure of what was made salient in the story, but instead would be a measure of children's sensitivity to the prosody of ambiguous sentences. At the offset of the test question, an image of each object that was used during familiarization appears on opposite sides of the screen, remaining there for 3 seconds. After 1 second of black screen, the next familiarization video begins automatically.

Note that definiteness plays a role in the acceptability of extraction from NPs. For example, sentences like *Who did she find the picture of?* are degraded compared to sentences like *Who did she find a picture of?* Extraction from a definite NP is worse than extraction from its indefinite counterpart. Definite DPs were used in the current experiment to first establish whether children are sensitive to locality constraints, before testing whether they are sensitive to the effects definiteness on extraction from DPs, which is left to future work.

Design

Each child saw 6 video trials (one for each verb) edited together into one video less than 5 minutes long. Recall that the side that the instrument appeared on and the color of the instrument were balanced as a between-subjects factor. In addition, two orders of trial presentation were created, making 4 experimental videos per syntactic condition (island vs. control), or 8 in total. Each child saw only one video.

Procedure

All testing was done at local pre-schools by the author. In a quiet room, children sat next to the experimenter and were given a brief description of the task (55).

55. “Hello! Today, I have some special videos for you. They’re videos about *these three bears* [gestures at bears on the screen]. Aren’t they cute? Okay, so…you’re going to watch some videos, and at the end of each video, there’ll

be a question. Your job is to point to your answer to the question on the screen – like this! [points at screen]. Are you ready to give it a try? Yeah, let's do it!"

Children found the task very easy. If they hesitated after the test question, which only ever happened during the first trial, the experimenter paused the video to allow the child to make their choice before the next trial started. The experimenter looked toward the child during the test trial to avoid giving them accidental looking cues by staring at the laptop screen, which was placed in front of the child. After the child chose an object, the experimenter recorded their response with pen and paper. Most children responded by pointing at an object on the screen, but some preferred to yell it out (e.g., "Yellow!" or "Blue book!"). In cases where parental permission was granted, the session was video-recorded using the test laptop, or a separate video camera.

Predictions

This study tests the following two hypotheses: 3 year-old children have mature knowledge of island constraints, or they do not. The independent variable is islandhood: the critical (*wh*- question) condition has *wh*-movement and therefore the potential for extraction from an island (56), and the control condition does not have *wh*- movement in the relevant utterance (e.g., *She hugged the bear with the blanket*) (57). The dependent variable is children's responses to (56) and (57).

56. Which blanket did she hug the bear with?
57. She hugged the bear with the blanket. Which blanket was it?

If 3 year-olds don't have mature knowledge of island constraints, then they will behave the same regardless of syntactic condition: both (56) and (57) will have the same interpretations, so their answers to these questions will be largely determined by what they attended to during the familiarization video. Again, it's possible that children have knowledge of island constraints, and this experiment fails to detect it. In this case, follow-up work is needed.

If 3 year-olds have mature knowledge of island constraints, then they will prefer to choose the object used as the instrument during the video in the critical (*wh*-question) condition (55). The experiment makes it possible to eliminate alternative explanations for this behavior by including a syntactic control in which children see the same videos, but instead hear an ambiguous utterance during test. If children behave differently in the control condition than in the critical (*wh*-question) condition in the right direction, then this experiment allows us to conclude that children have mature knowledge of island constraints.

Participants

22 children (mean age: 3;7;23; age range: 2;8;6 to 4;2;25) participated in this study.

Results

Children in the critical (*wh*-question) condition chose the blanket used as an instrument significantly more than children in the control condition ($t(20) = -2.42$, $p = .017$) (Figure 5). An individual response was obtained for each child by averaging their responses in every trial. Individual responses were averaged to generate one average preference score, plotted in Figure 5 as the proportion of trials that children pointed toward the object used as an instrument during familiarization.

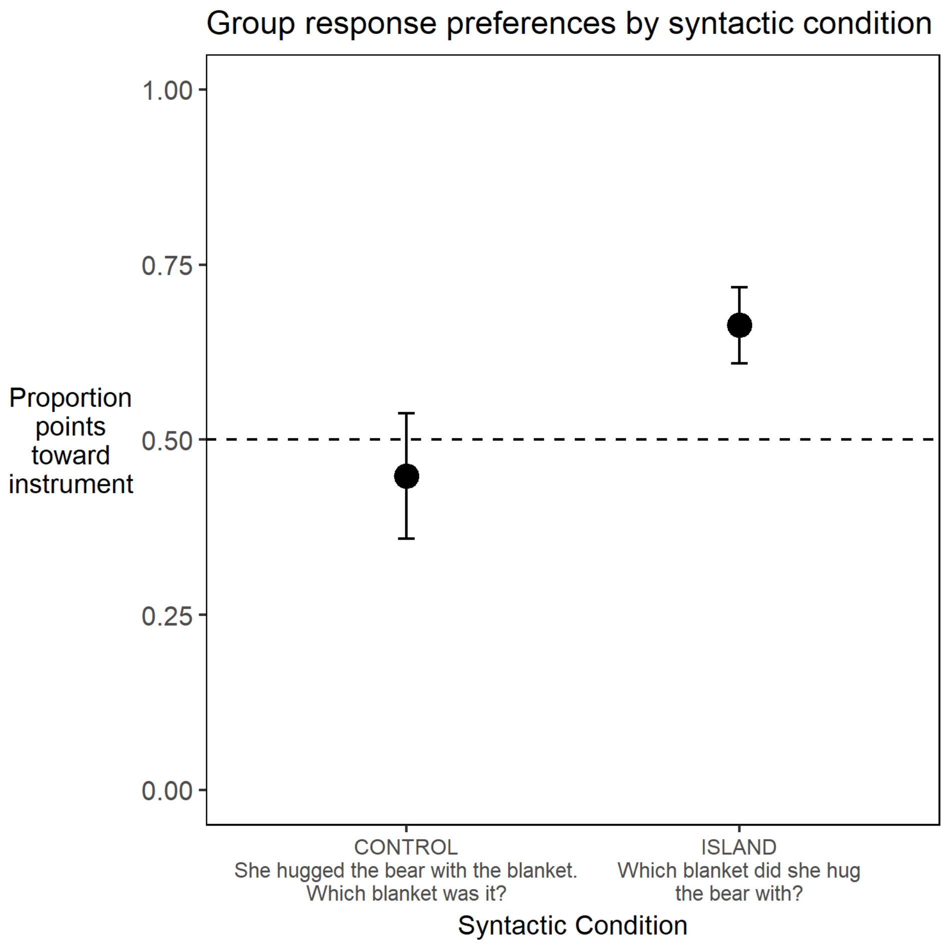


Figure 5. Average preference score for the object which was used as an instrument during the video in each condition. Preference score is calculated as the average proportion of the total responses for each syntactic condition (Experiment 2).

This behavior pattern is best explained by the hypothesis that children have acquired knowledge of island constraints, because the presence of an island is the only difference between the two conditions. Behavior in the control condition indicates that the test question allowed for children to answer with the blanket held by the bear, which confirms that it's an appropriate comparison for behavior in the island condition.

Plotting the individual child responses reveals that most of the individual children in the island condition prefer to choose the blanket used as an instrument (at least 4 out of 6 trials). In the control condition, however, children's behavior either reveals no preference, or a preference for the blanket held by the bear (Figure 6).

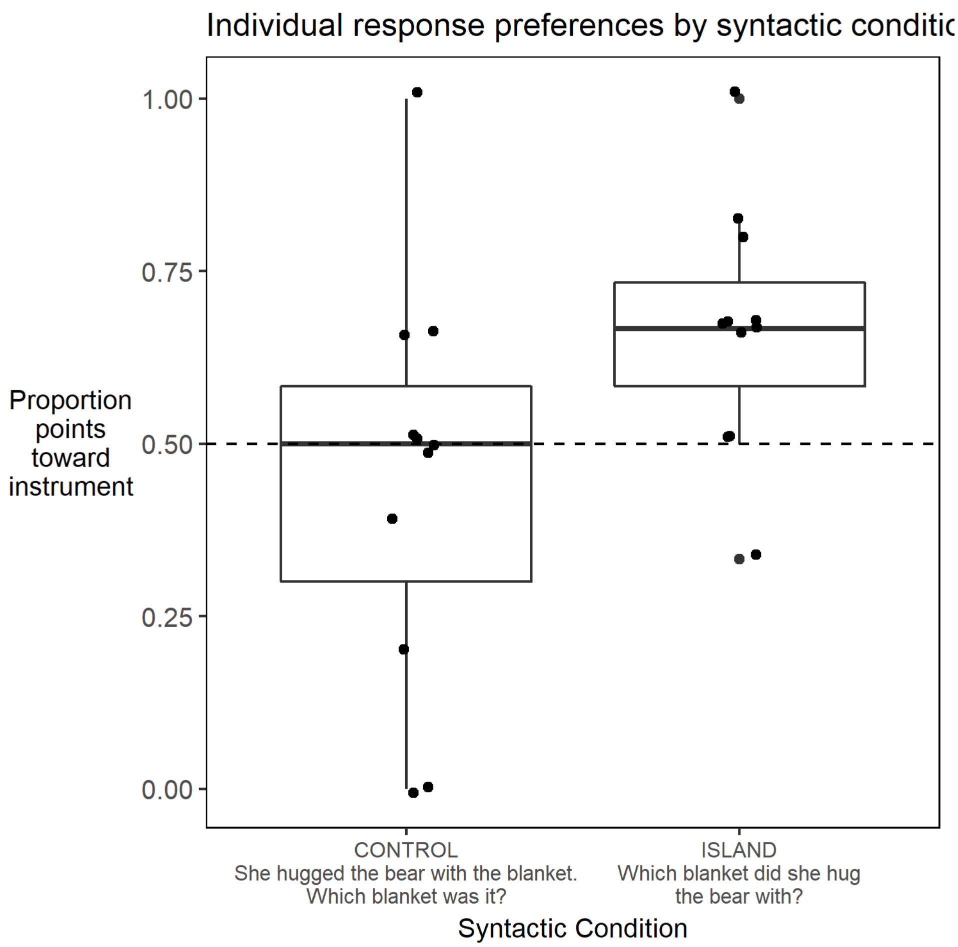


Figure 6. Boxplot of the average preference score for the object which was used as an instrument during the video in each condition. Preference score is calculated as the average proportion of the total responses for each syntactic condition (Experiment 2).

A learning hypothesis for island constraints (e.g., Pearl & Sprouse, 2013) predicts that children's performance will improve as they get older. A 2 (Syntax: island versus control) by 2 (Age: younger versus older) between-subjects ANOVA reveals that performance doesn't improve as a function of age. The main effect of syntax on responses was significant, F = 4.46, p = .04. The main effect of age group on response wasn't significant F = 2.83, p = .10, and there is no interaction (F = .18, p = .66) (Figure 7) and (Figure 8).

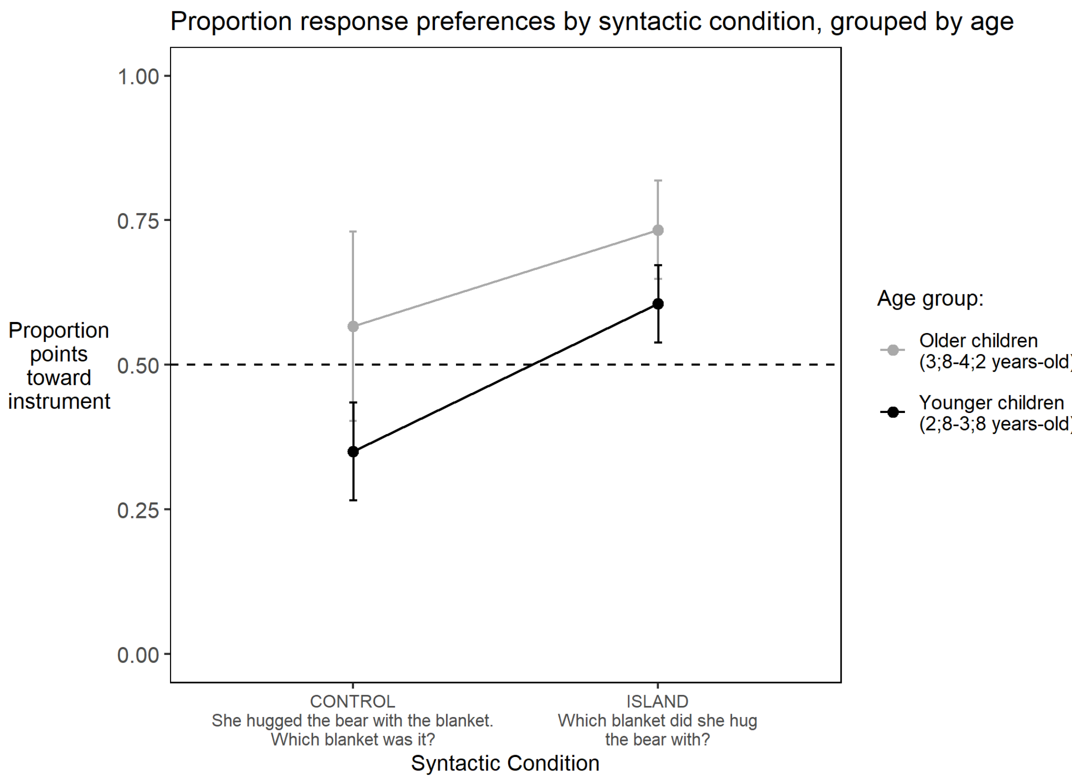


Figure 7. The average preference score for the instrument during the video in each condition, as a function of age. Age groups were established using a median split. Preference score is calculated as the average proportion of the total (Experiment 2).

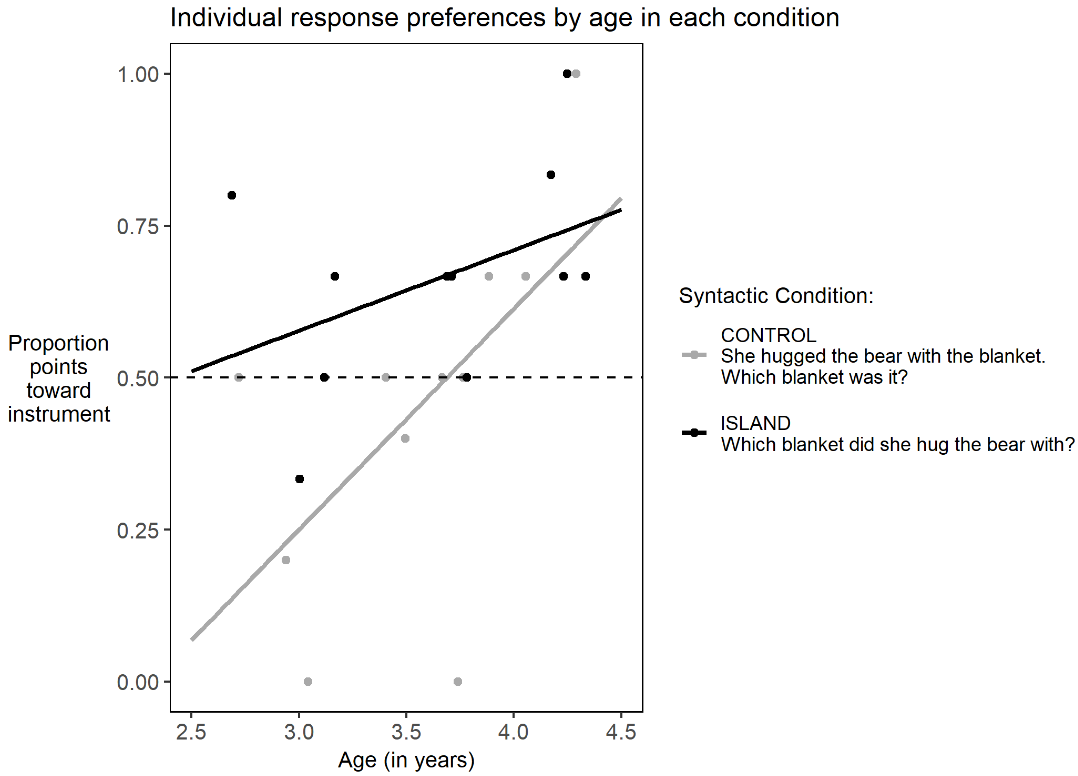


Figure 8. Average preference score for the object which was used as an instrument during the video in each condition, as a function of age. Each point represents an individual child (Experiment 2).

Discussion

This question-after-story task tests 3 year-olds for knowledge of locality constraints on *wh*-movement. Results show that 3 year-olds prefer to answer with the object used as the instrument in the critical condition, and their response preference in the critical condition is significantly different from their response preference in the control condition. This result supports the conclusion that 3 year-olds respect locality constraints on *wh*-movement.

Although the result is significant and the difference between conditions is in the right direction, there is considerable variation in the children's responses. This amount of variation is typical for children younger than 4 years-old; in longitudinal studies variation tends to decrease as age increases. On top of this expected variation, children are given an ambiguous prompt in the control condition. It's reasonable to expect that this would add variability to the response profile. Without testing adult (mature) speakers, it's not possible to claim with certainty that there isn't something about the experiment itself introducing this variability. Adult results would confirm that this experiment is an effective test for knowledge of locality constraints on *wh*-movement, and confirm that the adjectival interpretation is preferred in the control condition (an important prerequisite to interpreting behavior in the critical condition. Experiment 2 tests adult participants on the same experiment, with minor changes to the experimental procedure (i.e., online presentation).

One lingering concern is that the critical and control conditions, although minimally different in the most important respect (islandhood) are too dissimilar in other ways. For example, in the *wh*-question, *the bear with* appears at the end of the sentence, which ultimately resulted in *the bear with* being pronounced more slowly than in the control condition, which can also be perceived as longer pauses between the words. Because prosodic information is linked to syntactic structure, it could be that children's responses are somehow affected by these low-level prosodic differences between the two conditions. To control for prosodic differences, a second experiment was developed (Experiment 3) in which children either hear either relative clauses or the control condition from Experiment 1 during test.

4.4 Experiment 3: *Wh*- questions (adults)

The purpose of this experiment is to ensure that adult participants behave as expected on the question-after-story task introduced in Experiment 1. Adult responses serve to (a) verify baseline preferences in the ambiguous control condition and (b) confirm that the critical (*wh*-question) condition is an effective way to probe for knowledge of locality constraints. Experiment 2 tests adult participants on the same syntactic conditions as the children from Experiment 1 (58-59).

58. *Critical Condition:* Which blanket did she hug the bear with?

59. *Control Condition:* She hugged the bear with the blanket. Which blanket was it?

Materials

The videos, images and audio are identical to those used in Experiment 1. Testing limitations imposed by COVID-19 made it impossible to test adult participants in person, so the experiment was adapted to an online survey, developed using Qualtrics software. Due to limitations with Qualtrics, videos were embedded into the survey via Youtube links. Since adults were unsupervised during the experiment, additional measures were taken (e.g., training trials, filler trials, written instructions) to confirm whether adult participants were watching the videos and listening to the test questions when they gave their answers.

Each test trial begins with a familiarization video (Figure 9). Participants were instructed to watch the video, and after the video ended, a button appeared at the bottom of the screen which allowed them to advance to the test question, which was presented on the next page of the survey.

Figure 9. Example familiarization video page in Qualtrics Survey. The button to advance to the next page does not appear at the bottom of the page until the video ends.

On each test question page, a still image of the end of the training video served as a reminder of what had happened during the familiarization video (Figure 10). The test audio clip was positioned underneath the image. To keep the experiment as close to Experiment 2 as possible, test questions were audio only and were not written anywhere on the screen. Adults listened to the same recordings as the children did to ensure that the prosody was consistent between experiments. Participants used radio buttons to select their answer to the test question. Once an answer was selected, participants were able to advance to the next page.

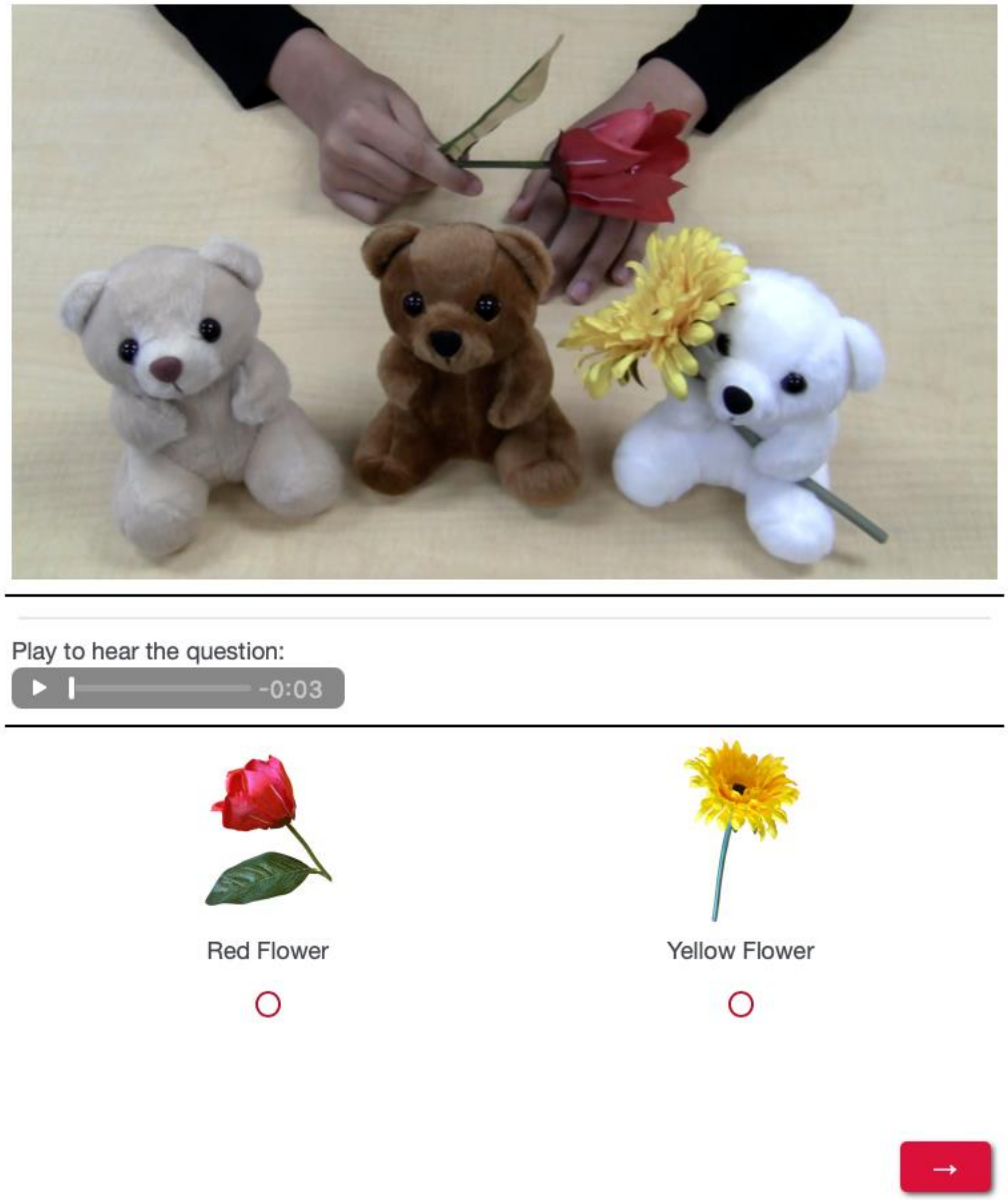


Figure 10. Example test question page in Qualtrics Survey (Experiment 3).

Three filler questions were presented at random throughout the experiment to check whether participants listened to the test questions. At the earliest, a filler question could appear after the first test trial (60-62).

60. Which one’s the blue/yellow blanket?

61. Which one’s the blue/red book?

62. Which one's the red/yellow flower?

All filler questions were multiple choice (Figure 12). Like the test trials, each question had two images to choose from: a blue blanket and a yellow blanket, a blue book and a red book, and a red flower and a yellow flower. These images are the same images used during test. The color of the object (blanket/book/flower) in the question and the position of the objects on the screen was randomly determined. These questions are designed to be easy for participants so long as they listen to the question audio, making these questions an effective way to exclude participants who weren't completing the survey as intended (i.e., playing the audio to hear the question).

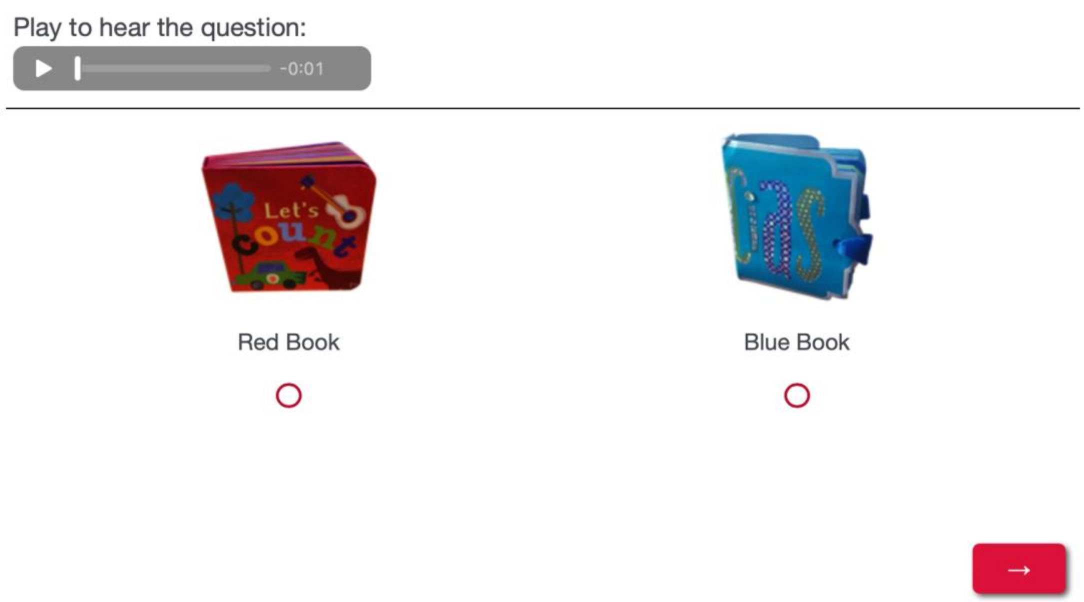


Figure 11. Example filler question (Experiment 3).

To help participants familiarize themselves with the experiment before starting the test trails, each experiment started with one training trial, which appeared directly after moving on from the experiment instructions. This trial was designed to give adult participants the impression that the questions in this experiment had correct and incorrect answers, which was a concern given the ambiguous nature of the control condition. The training question was always *Which one's the blue/yellow blanket?* Like the filler trials, the color word in the question and the order of the objects on the screen was randomly determined. If participants selected the correct answer, then they moved on to the first test trial. In case participants selected the incorrect answer, when they tried to advance the survey they got the following error message, highlighted in red: "Incorrect! Please listen to the question." Participants were not able to advance past the training question without selecting the correct answer to the question. The training trial is designed to give participants a chance to interact with the survey and discover how it works before moving on to test trials.

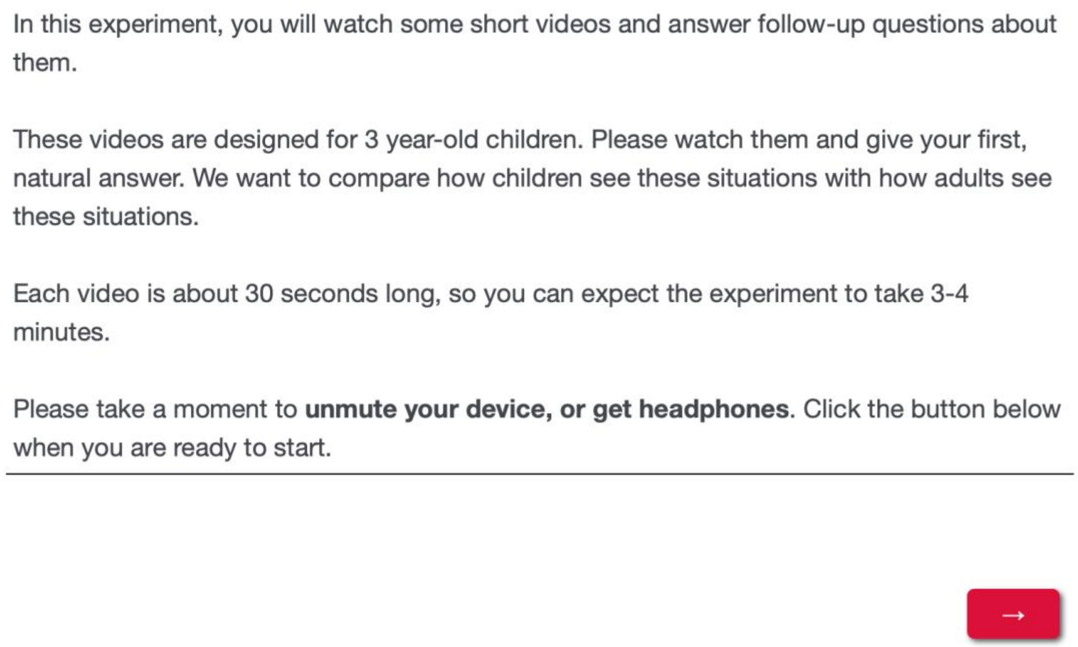
Design

The experiment structure consists of an instruction page, 1 training trial, 3 filler trials and 6 test trials. The *hug* test trial always directly follows the training trial (due to constraints imposed by the Qualtrics randomization software). The five remaining test trials are presented in random order. Filler questions were distributed such that they never appeared twice in a row, and had to appear before the last test trial. Participants saw at most 2 test trials in a row before getting a filler trial.

Participants were assigned to a syntactic condition at random, but the number of participants in each condition at any point during data collection was kept roughly even (control condition: n=16; critical condition: n=16). Like in Experiment 2, the color of the object that was used as the instrument was manipulated as a between-subjects factor, and participants were assigned to a color condition randomly. For example, half (n=16) of the participants saw the blue blanket used as the instrument throughout the experiment, and half (n=16) of participants saw the yellow blanket used as the instrument throughout the experiment.

Procedure

Participants used an invitation link to access the Qualtrics survey and completed it on their own time. Participants took an average of 8-9 minutes to complete the survey (range = 6-16 minutes). The first page of the survey is the consent form, followed by a short demographic questionnaire (initials, gender, age, and approximate age at which they learned English). Participants had the option to end the survey immediately (without answering any further questions) after reading the consent form. Basic information about the participants' browser and device were collected to check whether the same person did the experiment twice under a pseudonym, although this never happened. After submitting their demographic information, participants read instructions before advancing to the training question (Figure 12). Participants' responses were automatically recorded.



In this experiment, you will watch some short videos and answer follow-up questions about them.

These videos are designed for 3 year-old children. Please watch them and give your first, natural answer. We want to compare how children see these situations with how adults see these situations.

Each video is about 30 seconds long, so you can expect the experiment to take 3-4 minutes.

Please take a moment to **unmute your device, or get headphones**. Click the button below when you are ready to start.

→

Figure 12. Instructions page for Qualtrics Survey (Experiment 3).

Participants

Adult participants (n=32, mean age=28, females=22) included in the final sample were native speakers of American English. Three participants were excluded from analysis for being older than 2.5*IQR above the third quartile (31.25+(2.5*6.5)=48 years-old).

Results

It is well-established that adult speakers respect locality constraints on *wh*-movement. The purpose of this experiment is therefore not to discover evidence of locality constraints on *wh*-movement in the adult grammar, but to establish two important facts about the experiment itself. First, that the videos induce a preference for the object held by the bear in ambiguous sentence contexts (i.e., the control condition). This fact can be said to be established if adult participants prefer to answer with the object held by the bear in the control condition. Second, that the critical (*wh*- question) condition can reveal knowledge of locality constraints on *wh*-movement in mature speaker of English. As previously discussed, the locality constraint tested in these experiments is an unusual case given the commonly discussed examples in the generative linguistics literature. This fact – again, a fact about the experiment itself, and not the knowledge of speakers – is established if adult participants prefer to answer with the object used as the instrument in the critical (*wh*- question) condition. Because this is a grammatical constraint, adults should show a strong preference to answer with the object used as the instrument.

The results show that (a) adults prefer to answer with the object held by the bear in the control condition, and (b) adults prefer to answer with the object used as the instrument in the critical condition (Figure X). A Welch two sample t-test reveals a significant effect of syntactic condition ($t(95) = -13.162$, $p < .001$). In the critical condition, adults preferred the object used as the instrument 100% (96/96) of trials, indicating a strong preference for the adverbial interpretation of the critical question *Which flower did she touch the bear with?* Adults never arrived at the adjectival interpretation of the critical question, which would violate the locality constraint. This indicates that the test question is an effective test for knowledge of this locality constraint in mature English speakers.

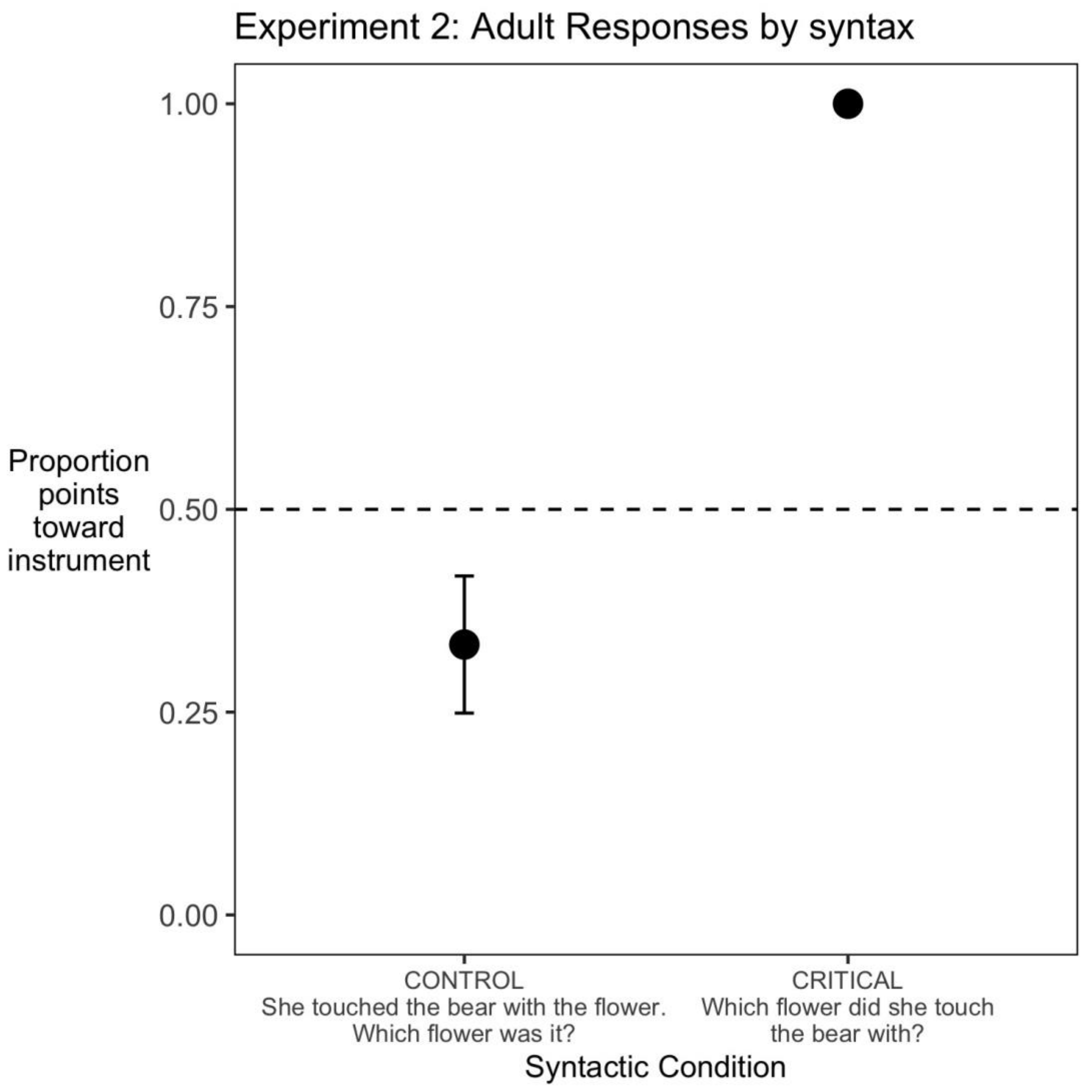


Figure 13. Experiment 3 results showing adult preference for the instrument in each syntactic condition.

In the control condition, adults preferred to answer with the object held by the bear, choosing it 65% (62/96) of trials (Figure X reports adult behavior as a preference for the object used as the instrument: 35% (34/96 trials). Most participants (n=11) had mixed response profiles (choosing the blanket held by the bear for some, but not all trials), and the remaining participants (n=5) chose the object held by the bear 100% of trials. Only 1 participant never chose the object held by the bear.

Discussion

These findings confirm that this experiment is an effective measure of locality constraints on *wh*-movement by showing that (a) in ambiguous contexts, adult speakers show a preference for the object held by the bear, indicating that they've assigned an adjectival interpretation to the phrase *with the blanket* in sentences like *She hugged the bear with the blanket.* Furthermore, (b) when prompted with a *wh*-question (which blocks the adjectival interpretation in the adult grammar) adults show a unanimous preference for the object used as the instrument.

Adult behavior on this task aligns with the behavior of 3 year-olds from Experiment 1. The hypothesis that the 3 year-olds' behavior is driven by knowledge of locality constraints on *wh*-movement predicts that if 4 and 5 year-olds were tested on this experiment, the overall result would remain stable, but the variation would decrease with age. If the behavior of older children deviated from this trajectory, this might invite a closer look at learning theories of island constraints.

The point raised in the discussion of Experiment 1 about the prosodic differences between the critical (*wh*-question) condition and the control condition still stands. Although differences between the two are unavoidable and natural, this non-syntactic explanation for the results must be addressed to fortify the claim that children's (and adult) behavior is driven by knowledge of island constraints. Experiment 3 addresses this concern by testing children on relative clauses in the critical condition instead of *wh*-questions.

4.5 Experiment 4: Relative clauses (children)

Experiment 4 was designed to address concerns with prosodic differences between the critical condition and the control condition in Experiment 2. *The bear with* appears sentence-finally in the critical condition, but sentence-medially in the control condition. As a result, these words have slightly longer durations in the critical condition than the control condition. It's possible that this low-level difference was driving children's behavior in Experiment 2. To eliminate this concern, Experiment 4 introduces a new critical condition (63) to compare with the control condition (64), featuring a relative clause in the subject position. Putting the relative clause in the subject position allows for *the bear with* to be sentence-medial in both conditions, and therefore pronounced as similarly as possible.

63. The flower [op_1 she touched the bear with $\_{}_1$] is so pretty! Which flower is it?

64. She touched the bear with the flower. Which flower was it?

Exploring other construction types (in addition to *wh*-questions) has the additional benefit of broadening our understanding of the development of syntax in the 3 year-old age range. Research on relative clause islands in 3 and 4 year-olds is suggestive that young children might have this knowledge, but differences in the control condition make these results difficult to interpret (Roeper and de Villiers, 1995). Testing more than one construction contributes to a more robust study of the acquisition of abstract syntactic knowledge.

Due to testing limitations imposed by COVID-19, child participants had to be tested online, instead of in person, which introduced minor changes to the procedure, although the materials were the same as those used in Experiment 1 (with the exception of the audio for the critical (relative clause) condition).

Materials

The materials used in this study are identical to those used for testing children online. The new critical condition was recorded by the same speaker who recorded the original stimuli (the author). Instead of presenting the trials as one long video, each familiarization-test pair was embedded into a single slide as part of a Slides deck.

Procedure

Two major difficulties arose from the virtual presentation: (1) it was no longer possible to demonstrate to children how they need to actually point at the objects on the screen. It was essential to train children to point because this is the dependent measure. Since many children have experience using the computer mouse to point at things on the screen, they were tempted to do so during the experiment. This measure is unreliable because it takes too long for children to maneuver the mouse to what they'd like to select. (2) because children were tested in their homes, they were using whatever devices they had access too, which caused variation that is typically controlled for in laboratory settings (e.g., the size of the screen, the angle of the camera, and the position of the child relative to the screen). This variation introduced difficulties for coders trying to ascertain what children were pointing at, which ultimately caused significant data attrition to the sample.

To address (1) and (2), a brief practice session asking children to point at familiar objects on the screen was added to the beginning of the experiment. For example, children saw an image of shoe and a ball, and a pre-recorded prompt asked them to point toward the shoe. Children found this exercise intuitive, and it was effective at getting children to point using their hands at the objects on the screen.

Participants

35 children ages 3-4 participated in this online study (mean=3;5;20, range=3;0;2 to 3;9;27). 5 children were excluded due to technical difficulties, unwillingness to participate, or coder uncertainty, leaving a final sample of 30 children. 14 children participated in the critical condition, and 16 participated in the control condition.

Results

The response rate was low compared to Experiment 2: children responded to 78/96 (81%) of trials in the control condition, and 70/84 (83%) of trials in the critical condition. Most of this attrition wasn't because children were unresponsive, but rather was due to coding difficulties with identifying where children were pointing on the screen from the recordings.

In the critical condition, children pointed toward the object used as the instrument 59% (41/70) of trials. In the control condition, children pointed toward the object used as the instrument 63% (49/78) of trials. A Welch two sample t-test revealed no effect of syntactic condition $t(143.62) = 0.52502$, $p = 0.6$. The experiment provides little insight into children's linguistic knowledge because it's difficult to interpret children's behavior in the critical (relative clause) condition without a baseline preference for the object held by the bear in the control condition. Because children prefer the object used as the instrument in ambiguous contexts (i.e., the control condition), it's not clear whether knowledge of locality constraints on *wh*-movement influences their behavior in the critical (relative clause) condition or not.

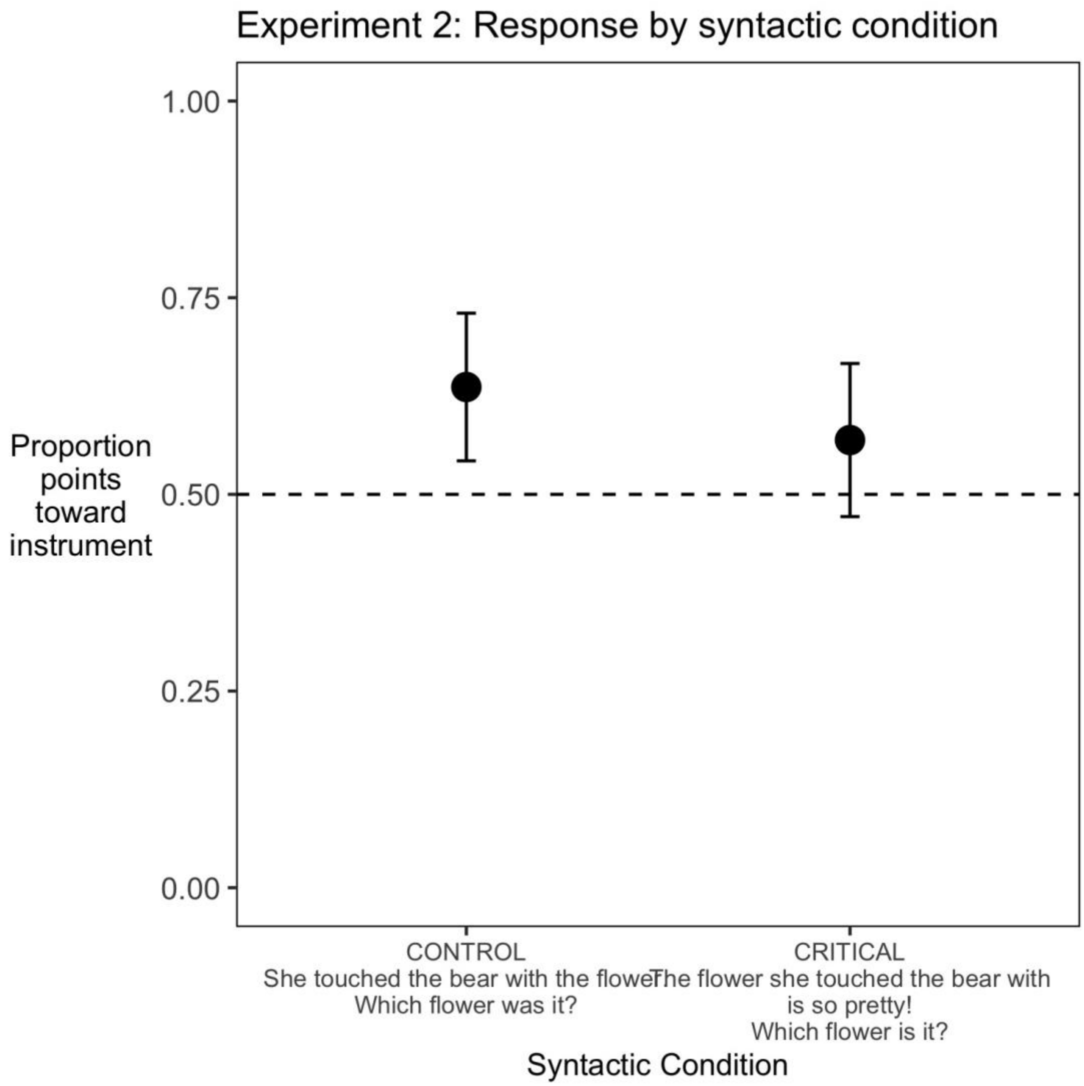


Figure 14. Experiment 4 results, showing children's preference for the instrument in each syntactic condition.

Discussion

Poor internet connection was prevalent throughout testing, and contributed to poor recording quality that hindered experimenter coding – it's possible that children had difficulty watching the videos and hearing the questions on their end as well. The context which supports the adjectival interpretation is delicate, and speakers tend to prefer the adverbial interpretation of *with* PPs when utterances are presented without the right contextual support. It's possible that children weren't able to attend to the familiarization videos in the same way as they did during in person testing, and they reverted to the default adverbial interpretation of the PP in the ambiguous context (e.g., *She hugged the bear with the blanket. Which blanket was it?*). Critically, this experiment failed because the necessary baseline was not established in the control condition – and not because of children's behavior in the relative clause condition.

Without confirming that adult participants behave as expected in this experiment, it's not possible to be certain what's driving children's behavior. Experiment 5 tests adult English speakers using the same materials and design, adapted as a survey administered online (like Experiment 3).

4.6 Experiment 5: Relative clauses (adults)

Experiment 5 tests adult speakers on the same linguistic stimuli and materials as Experiment 4. The same procedure used to test adults in Experiment 3 was used in Experiment 5.

Materials

The same materials used to test adults in Experiment 3 were used in Experiment 5, except for the questions in the critical (relative clause) condition. Like in Experiment 4, relative clauses were used to probe for knowledge of locality constraints (65) and were compared to syntactic controls like (66).

65. The flower [op_1 she touched the bear with $\_\_{1}$] is so pretty! Which flower is it?

66. She touched the bear with the flower. Which flower was it?

Procedure

The same procedure used to test adults in Experiment 3 was used in Experiment 5.

Participants

Adult participants (n=30, mean age=25, females=14) included in the final sample were native speakers of American English. 4 participants were excluded from analysis because they took over 30 minutes to complete the experiment (n=2), failed a filler question (n=1), or were an age outlier (n=1).

Results

This experiment provides an adult baseline to measure the child results against from Experiment 3. The results of Experiment 3 were inconclusive, but results from the adult experiment will be valuable for future work investigating what went wrong.

Adult participants preferred to choose the object used as the instrument in the relative clause condition, but in the control condition their preference flips, and they prefer to choose the object held by the bear. A Welch two sample t-test reveals a significant effect of syntactic condition ($t(98.288)=-13.825$, $p < .001$). In the relative clause condition, adults chose the object used as the instrument 99% (89/90) of trials, showing a strong preference for the VP adjunct interpretation with almost no variance. As discussed in Experiment 2, this experiment is not designed to establish whether adult speakers have knowledge of locality constraints; this is a precondition of testing for the emergence of this knowledge in children. The single incorrect response observed in the relative clause condition may indicate that the relative clause condition in the current experiment is a little more difficult than the *wh*-question condition in Experiment 2.

In the control condition, adults chose the object held by the bear 70% (63/90) of trials (adults chose the object used as an instrument 30% (27/90) of trials). Because syntax is manipulated as a between-subjects factor, this results replicates findings from Experiment 2 where adults (n=16) chose the object held by the bear 65% (62/96) of trials in the control condition. The control conditions in both Experiment 2 and Experiment 4 are identical; the only difference between the two Experiments is the utterance tested in the critical condition (*wh*-question vs. relative clause).

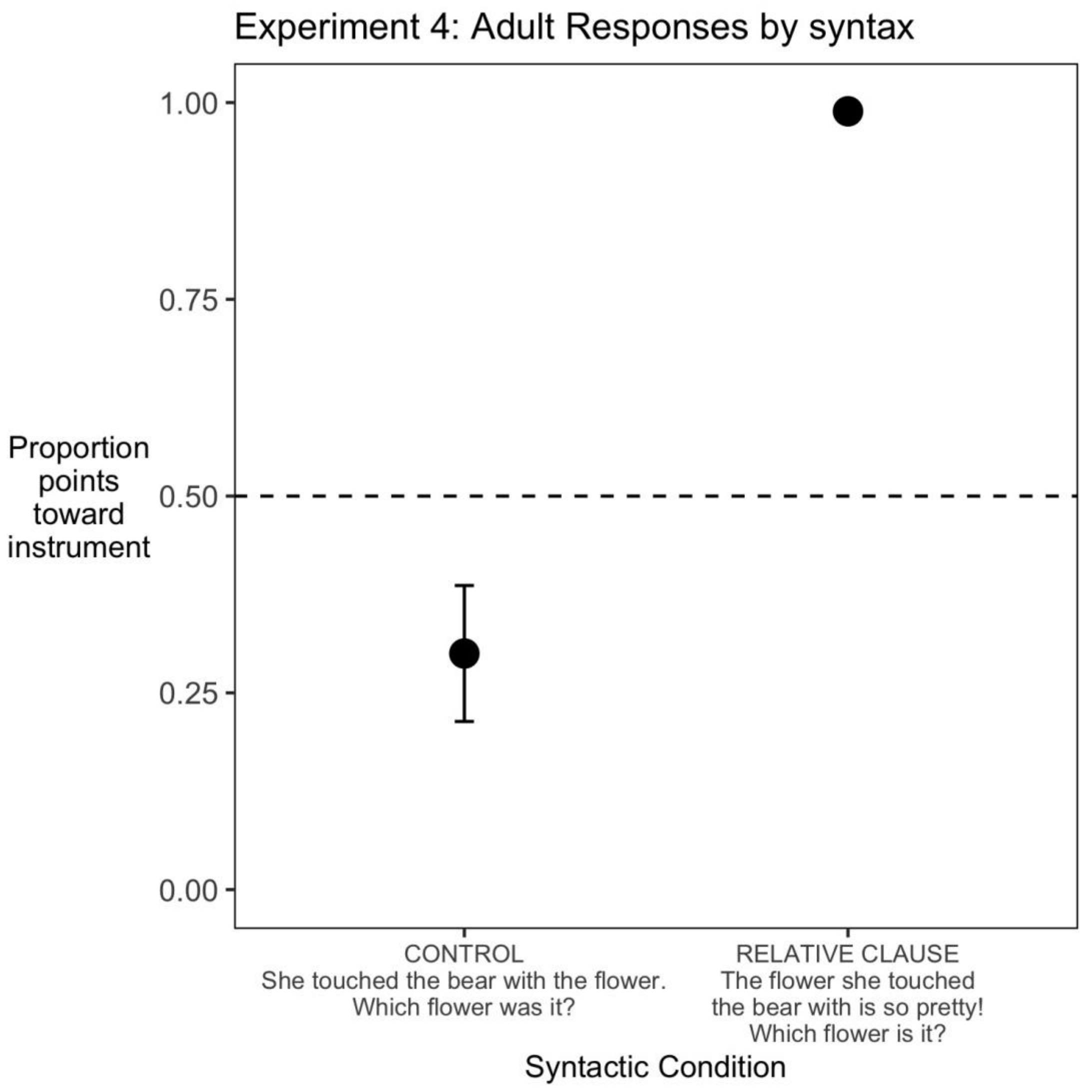


Figure 15. Experiment 5 results showing adult response preferences for the instrument (proportion) by syntactic condition.

Discussion

Adult behavior in this experiment indicates that like *wh-* questions, relative clauses are an effective way to probe knowledge of locality constraints on *wh*-movement in this experimental paradigm. This experiment essentially reproduces findings from Experiment 2. Because adult speakers behaved as expected in the experiment, it's unlikely that something different about the experiment is responsible for children's behavior in Experiment 4. It's possible that difficulties introduced by testing young children on this experimental paradigm online (during a global pandemic) introduced too much noise to the experiment presentation, sabotaging the delicate context required for accessing the adjectival interpretation in the control condition. Future work following up on Experiment 4 is needed to establish whether children's knowledge of locality constraints on *wh*-movement is general enough to apply to constructions other than *wh*-questions.

4.7 General Discussion

The behavior of children in Experiment 2 is consistent with the hypothesis that children respect locality constraints on *wh*-movement, but difficulties with the control condition in Experiment 4 makes it difficult to interpret. Adult results in Experiment 5 suggest that nothing about the experimental design/materials developed for Experiment 4 caused any difficulties. The main difference between Experiment 2 and Experiment 4 was that the former was conducted in person, while the latter was conducted online via Zoom. It's possible that online presentation of these experiments is not an effective measure of grammatical knowledge in children, given the delicate nature of the contextual support required for the adjectival interpretation of the PP in the control condition. Further work is needed to understand what went wrong.

Although adults behaved as expected on this task (Experiments 3 and 5), I would like to note that participants in the control condition expressed a lot of doubt (personal communication) as to whether they were "doing the task correctly". I take their concerns to indicate that both interpretations came to mind during the experiment, and they weren't sure which answer was the "right" one. This was by design, but it raises concerns for the child participants. If children experienced the same tension during the experiment, then this undoubtably added noise to the control condition.

Returning to the main result, one experiment is insufficient to adjudicate between innate and learning proposals for the acquisition of islands. Current research adapts this experiment for infants, to probe for knowledge of island constraints in even younger learners. Learning proposals for islands in which children acquire knowledge of islands from the distribution of *wh*-movement in their language predict that there will be some age (younger than 3 years-old) at which children have immature knowledge of islands. Innate proposals predict that there will not exist an age at which children have knowledge of *wh*-movement, without knowledge of island constraints. It's worth noting that learning proposals in which children jointly infer islands and *wh*-movement make the same empirical prediction as innate proposals, in this case. If the empirical gap is eliminated entirely, further steps will have to be taken distinguish between these two very different explanations for knowledge of island constraints in very young learners.

If locality constraints are induced through experience, then previous estimates of the amount of data children induce this knowledge from are inflated compared to reality (Pearl & Sprouse, 2013). Learning models of islands must not assume linguistic knowledge or cognitive resources which aren't available to children younger than 3 years-old, and must estimate the amount and complexity of input based on the experience of children younger than 3 years-old. If it becomes impossible for learning models to succeed under these conditions, then the plausibility of learning proposals for island constraints must be re-evaluated.

4.8 Conclusion

Locality constraints on movement impose structure-based limitations on abstract syntactic relations. The complexity of the constraints relative to the input make them an excellent candidate for innate linguistic knowledge (Chomsky, 1977), and has inspired exploration on the role of the input in the acquisition process (Pearl & Sprouse, 2013). The current study demonstrates that 3 year-olds already have knowledge of these constraints (Experiment 1), a younger age than previously attested in the literature (Fetters & Lidz, 2016, de Villiers et al., 1990, Goodluck et al., 1990, de Villiers & Roeper, 1995; 1995). It's still unknown whether children younger than 3 years-old have knowledge of island constraints or not. I argue that determining the youngest age at which children have knowledge of island constraints will illuminate the role of the input in models of the acquisition of locality constraints on movement. If very young learners have knowledge of locality constraints, and there exists no age at which knowledge of these constraints is immature, then the role of the input is likely to be quite small, making learning proposals for islands less plausible explanations than innate proposals. This research takes a small step toward an empirically supported theory of the acquisition of island constraints.

Chapter 5: Discussion of Pearl & Sprouse (2013)

5.1 Summary of Pearl & Sprouse (2013)

Pearl & Sprouse (2013) proposes and tests an'empirically grounded' learning model for how children learn (i.e., induce) island constraints from experience (i.e., linguistic input). The Pearl & Sprouse (2013) learner relies on knowledge of *wh*-movement to induce knowledge of island constraints. The model is built on the assumption that knowledge of island constraints and *wh*-movement are separable not just in theory, but *in practice*. This is at odds with how generative theories define *wh*-movement, which is often diagnosed by its island sensitivity (Chomsky, 1977). In the model, children first acquire knowledge of *wh*-movement, which means that they identify a syntactic relation between the filler and the gap positions. Then, children use the syntactic distribution of this relation to discover the syntactic environments where it does not hold, thus inducing island constraints.

In the Pearl & Sprouse (2013) learning model, the learner tracks the extraction paths of *wh*- argument questions to learn that *wh*- movement is subject to island constraints. Adult acceptability judgements are used as a proxy for target knowledge – in particular, the super-additive effect of island constraints on acceptability (Wagers, Sprouse & Phillips, 2013). For example, adult speakers consistently rate adjunct clause island violations, *whether* island violations, complex NP island violations, and sentential subject island violations worse than syntactic baselines, controlling for other factors known to influence acceptability, like syntactic complexity and length of the extraction path. The additional hit to acceptability scores that can only be attributed to island violations is called the super-additive effect (Wagers et al., 2013). In the learning model, the statistical probability of the syntactic sequence is used as a proxy for these acceptability judgment scores. The purpose of the learning model is to test the hypothesis that children induce island constraints from their linguistic experience by showing that this is possible in principle.

For inductive approaches to language learning, the structure of the linguistic input has a great influence on the learning outcome (compared to learning as inference), because what is learned is largely determined by the experience. The advantage of using this model for island constraints is that these utterances are guaranteed to be absent from child-directed speech – not just because they're ungrammatical, but because they are multiclausal *and* have at least one instance of *wh*- movement. For an inductive learner, the battle is over before its begun: the learner won't get exposure to island violations, and so will assign them a probability of zero. A study of child-directed speech reveals that there is however one inductive learning problem that the model will still need to overcome. In child-directed speech, grammatical, long-distance extraction from complement clauses is just as rare as ungrammatical long-distance extraction (i.e., island violations) (see also Omaki 2010 for similar findings). If children inferred that all unobserved structures are ungrammatical, then they would acquire a grammar that does not generate long-distance extraction from complement clauses (which would be a subset of the grammar of English). Pearl & Sprouse (2013) refers to this as the induction problem.

When each *wh*- argument question in child-directed speech is tallied up, then the utterances with grammatical long-distance extractions are just as rare as island violations. By breaking the utterances up into smaller syntactic units, and tallying those units separately, it's possible for the learner to overcome this induction problem. The model breaks the extraction path into smaller, structured units comprised of three subjacent Phrasal nodes, called trigrams. Each trigram is assigned a probability. Because clauses have highly repetitive phrasal structure, the most frequent trigrams will have high counts, creating a frequency distribution with a long tail. This is the key to solving the induction problem: the extraction paths out of complement clauses *always* contain high probability trigrams. The would-be extraction paths out of islands, on the other hand, always contain at least one unobserved trigram, drawn from the tail of the distribution. The offending trigram brings the probability of the extraction path to floor. Because the learner interprets low probability as a signal of ungrammaticality (e.g., Clark & Lappin 2011), the learner can use the primary linguistic data to discover that *wh*-movement is subject to certain locality constraints – without incorrectly concluding that all complement clauses block *wh*-movement. The next sections describe in detail how the model defines the learning input, the child's linguistic knowledge, the learning algorithm, and the target knowledge.

Learning Input

The learning input (training set) is comprised of 200,000 *wh*- argument questions (subject and object *wh*-questions). The training set was created by annotating 6 corpora of child-directed speech from the CHILDES database (MacWhinney, 2000) with phrase structure: the Adam, Eve and Sarah corpora from the Brown data set (Brown, 1973), the Valian data set (Valian, 1991), and the Suppes data set (Suppes, 1974).

Utterances were parsed automatically using a freely available syntactic parser, hand-checked by trained annotators and then vetted again for remaining errors. The final data set included child-directed speech to 25 children (ages 1 to 5 years-old) and 31,247 utterances which contained *wh*-words and verbs (making these utterances likely to be *wh*-questions).

Only 295 utterances (0.9% of the 31,247[6] likely *wh*-questions) contained grammatical long-distance *wh*- movement (i.e., extraction of the direct object from a tensed VP complement; for example, *What did she think* [$_{CP}$ *that Clara* [$_{VP}$ *bought* _] *at the farmer's market*] ?). Pearl & Sprouse (2013) determines that there is an induction problem for acquiring mature knowledge of *wh*-movement because the best evidence for long-distance *wh*-movement is extremely scarce in the child's linguistic experience. Following results from Hart & Risley (1995), Pearl & Sprouse (2013) estimates that children are exposed to about 333,000 utterances per year (on average), and *wh*-argument questions make up about 20% of these utterances (see Chapter 2 for a review of studies which suggest that this might be an overestimate). If it takes 3 years to acquire knowledge of island constraints (approximately from ages 2 to 5 years-old), then children must induce knowledge of island constraints from approximately 200,000 utterances (20% of 1,000,000 utterances), hence the training set of 200,000 *wh*-argument questions.

Learning Algorithm

To induce knowledge of island constraints, the model tracks syntactic phrases in the

[6] A second tally estimates 20,923 *wh*-questions (Pearl & Sprouse, 2013: Table 2); 295/20,923=1% of *wh*-questions are long-distance object extractions from tensed VP complements.

extraction path (the structure between the filler and the gap of the *wh*-dependency, not including the CP which contains the filler). The entire set of phrases which contain the gap but exclude the filler are referred to as the *trigram sequence*. Trigram sequences are broken down into trigrams, which are subjacent sets of three phrasal nodes in the extraction path, assigned by a sliding window. Lexical information is omitted, except for CP which is lexically subcategorized (e.g., $CP_{\textit{that}}$, CP_{null}, $CP_{\textit{whether}}$, and $CP_{\textit{if}}$). Without this subcategorization information, the model would have no way to distinguish grammatical long-distance extraction from *whether* and *if* island violations (and would fail to overcome the induction problem). Each trigram is assigned a probability, determined by its relative frequency. Trigrams that are never observed have a probability of 0 (or an extremely low number, for convenience).

Examples (38-39) illustrate the trigram breakdown for an example of grammatical *wh*- extraction (38) and an example of an ungrammatical *wh*-extraction (39), holding the string constant (example (37) is the declarative baseline). Example (38) shows grammatical extraction from the object of the VP complement clause, *the vegan empadinha*. Example (39) shows ungrammatical extraction from within the subject of the VP complement clause, a violation of the subject condition. Note that because extraction of the whole subject from the complement clause is also degraded (i.e., *that*-trace effects), subject condition violations are commonly illustrated by showing extraction from the matrix subject. The purpose of using this atypical example (39) is to show that it's the extraction path that determines the trigram sequence (compare (39) to (38)). Examples (38) and (39) are string identical, but have different probabilities associated with them because they have different sequences and different trigrams, showing that the probabilities are determined by the structure of the extraction path.

67. Jéssica thought that the cat with one white foot stole the vegan empadinha from the platter

68. What did Jéssica think that the cat with one white foot stole _ from the platter?

    What did [$_{TP}$ Jéssica [$_{VP}$ think [$_{CP}$ that [$_{TP}$ [the cat with one white foot] [$_{VP}$ stole _] [from the platter]]]]]]

    Sequence: start-TP-VP-$CP_{\textit{that}}$-TP-VP-end

    Trigrams: start-TP-VP

    TP-VP-$CP_{\textit{that}}$

    VP-$CP_{\textit{that}}$-TP

    $CP_{\textit{that}}$-TP-VP

    TP-VP-end

69. What did Jéssica think that the cat with _ stole the vegan empadinha from the platter?

    What did [$_{TP}$ Jéssica [$_{VP}$ think [$_{CP}$ that [$_{TP}$ [$_{NP}$ the cat [$_{PP}$ with _]]] [stole the vegan empadinha from the platter]]]]

    Sequence: start-TP-VP-$CP_{\textit{that}}$-TP-NP-PP-end

    Trigrams: start-TP-VP

    TP-VP-$CP_{\textit{that}}$

    VP-$CP_{\textit{that}}$-TP

    $CP_{\textit{that}}$-TP-NP

NP-PP-end

Importantly, island constraints themselves are not built into the model – instead, the learner leverages syntactic biases which Pearl & Sprouse (2013) argues are necessary for learning other syntactic phenomena. Minimally, the learner must have sufficient syntactic knowledge and sentence processing abilities to parse sentences. Pearl & Sprouse (2013) remains neutral as to the source of this knowledge (innate vs. learned). The learner must also represent *wh*- extraction paths as trigrams (and not tetragrams, bigrams, etc.), track the frequencies of trigrams and phrasal projections, and use these frequencies to calculation the probabilities of trigram sequences and *wh*-extraction paths. The learner presupposes that *wh*- argument questions are essential for inducing island constraints, and the input is restricted only to these constructions. Finally, the model's success at overcoming the induction problem is driven by the lexical subcategorization of CP. These assumptions are revisited in Section X for discussion.

Target Knowledge

During the acquisition process, the learner parses the extraction path of the *wh*-question (one of 200,000 *wh*- questions randomly selected from the parsed corpus) as a trigram sequence and assigns a probability to it. After this training period, the success of the model is evaluated by comparing the probabilities that the learner generates for each *wh*- question to the adult acceptability judgment scores reported in Sprouse, Wagers and Phillips (2012). The learner has successfully learned island constraints if it assigns a lower probability to sentences which contain island violations compared to the syntactic controls (i.e., sentences which contain island structures (but no extraction) and sentences which contain long-distance *wh*-movement (but no island violations)). The learner is tested on conditional adjuncts, complex noun phrases, sentential subjects and *whether* islands. To ensure that the model has also learned that grammatical long-distance questions are probable and thus overcome the induction problem, the model is also tested on long-distance questions.

This section introduces the results of the adult acceptability judgment experiment from Wagers, Sprouse & Phillips (2012) that Pearl & Sprouse (2013) uses to evaluate the success of their learning model. Wagers, Sprouse, and Phillips (2012) elicited acceptability judgments from adult speakers for grammatical *wh*- questions and *wh*- questions which contained island violations. In their experiments, the location of the gap (MATRIX vs. EMBEDDED) is crossed with the syntactic complexity (NON-ISLAND vs. ISLAND) in a 2x2 design. *Island* is used here to refer to the structure which blocks *wh*-movement (regardless of whether the sentence contains an island violation). Sentences in the EMBEDDED | ISLAND condition contain island violations. For example, (40a) features *wh*- movement from a complex NP *the claim that Lily forgot*. Sentences in the MATRIX | ISLAND condition control for the effect of syntactic complexity on acceptability with sentences that have both *wh*- movement and a complex NP island, but the *wh*- movement does not cross the island (40b). Sentences in the NON-ISLAND | EMBEDDED condition contain long-distance *wh*-movement from VP complements (40c) (controlling for the length of the *wh*- extraction path), and finally sentences in the MATRIX | NON-ISLAND condition contain short *wh*-movement and an embedded clause.

70. Complex NP Islands (Pearl & Sprouse 2013:30)

a.* What did the teacher make the claim that Lily forgot _? EMBEDDED |

ISLAND

b. Who _ made the claim that Lily forgot the necklace? MATRIX | ISLAND

c. What did the teacher claim that Lily forgot _? EMBEDDED | NON-ISLAND

d. Who _ claimed that Lily forgot the necklace? MATRIX | NON-ISLAND

The average acceptability score is calculated for each condition, and then normalized and plotted in an acceptability-rating space for each island (conditional adjuncts, complex noun phrases, sentential subjects and *whether* islands) (Figure X). The negative slope of the line connecting the two non-island structure conditions reflects a preference for shorter extraction paths: the EMBEDDED | NON-ISLAND condition is rated lower than the MATRIX | NON-ISLAND condition. This preference is also apparent in when comparing the MATRIX | ISLAND and the EMBEDDED | ISLAND conditions (the dotted line). The figure reveals that participants give higher ratings to sentences without islands than sentences with islands: the MATRIX | NON-ISLAND score is consistently higher than the MATRIX | ISLAND score. These syntactic controls indicate that the length of the *wh*- extraction affects acceptability as well as the structural complexity of the sentence. If there is an effect of island violations on acceptability, then sentences in the EMBEDDED | ISLAND condition will be rated even lower than those in the MATRIX | ISLAND condition and the EMBEDDED | NON-ISLAND condition. This is visualized as a steeper negative slope in the dotted line connecting the two island conditions compared to the solid line connecting the two non-island conditions.

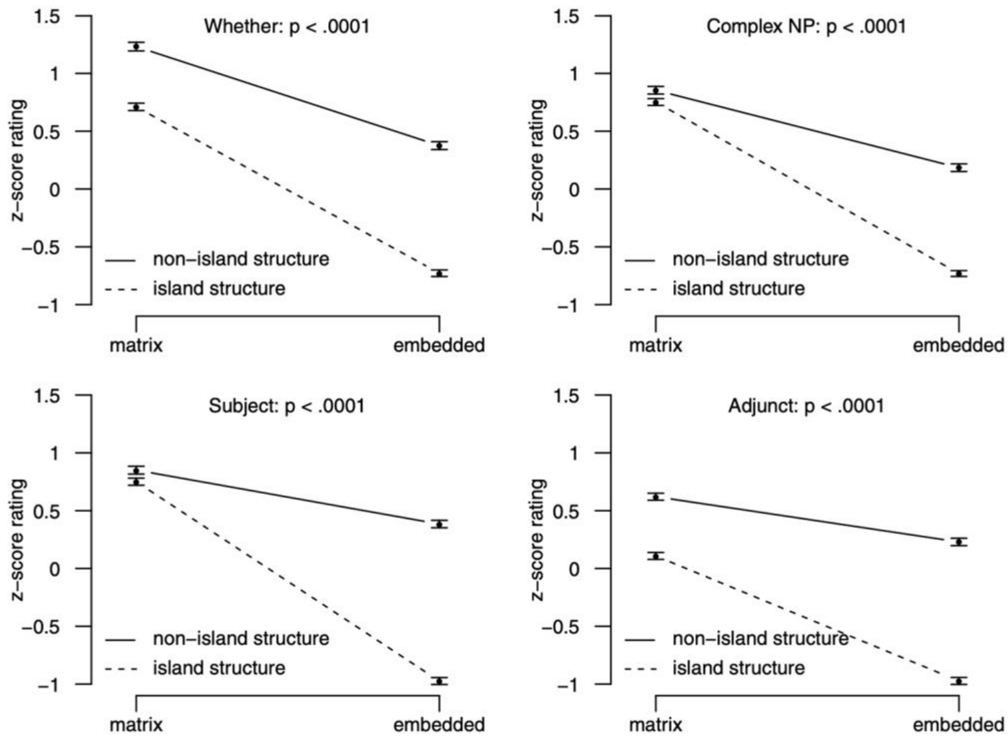


Figure 16. Interaction plots of adult acceptability judgments (Sprouse, Wagers & Phillips 2012). Island violations are visualized as interaction plots of acceptability judgment scores.

Results

Pearl & Sprouse (2013) concludes that the learner succeeds in discovering that conditional adjuncts, complex NPs, *whether* islands and sentential subjects block *wh*-movement. The probabilities are transformed and reported as log probabilities to make them easier to compare to one another (the untransformed probabilities are extremely small numbers). Log probabilities closer to zero (i.e., more positive) represent higher acceptability ratings, and log probabilities further from zero (i.e., more negative) represent lower acceptability ratings. Importantly, the learner overcomes the induction problem and learns that *wh-* movement from tensed clausal complements is grammatical (has a non-zero probability). Figure X (Pearl & Sprouse 2013: Figure 5) shows pairs of non-parallel lines, reproducing the acceptability judgment patterns elicited from adult speakers (Wagers, Sprouse and Phillips 2006), grouped by island constraint. The slope of the dotted line in each plot is steeper than the slope of the solid line, which indicates that the utterances with island violations had lower probabilities than the probability predicted by extraction length alone.

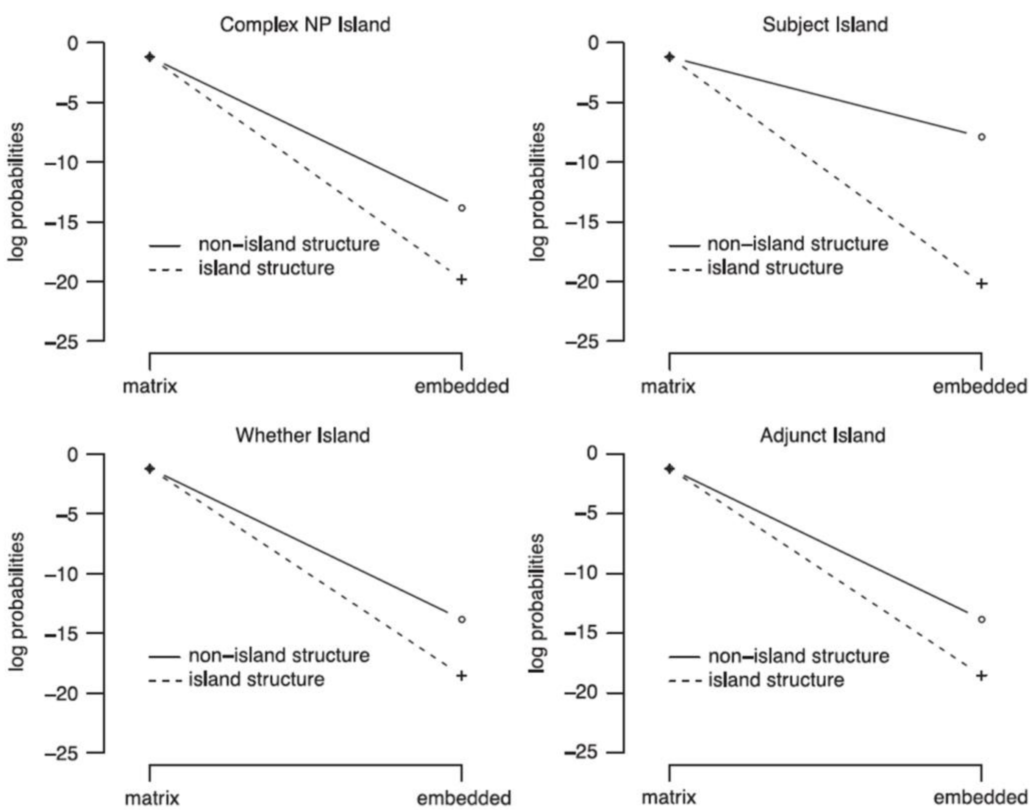


Figure 17. Pearl & Sprouse (2013: Figure 5): Log probabilities derived from a learner using child-directed/ambient speech.

The model's behavior with respect to island violations mirrors the behavior of adult speakers (Table 6). For example, *whether* island violations (-18.54) are 4.7 less than the log probability of an embedded object dependency (-13.84). This indicates that the model considers utterances containing *whether* island violations as nearly 5 times less acceptable than long-distance object *wh*- questions.

Table 6. Log probabilities of each extraction type in child-directed speech (Pearl & Sprouse 2013: Table 3). The first 3 rows are grammatical dependencies, the last 4 rows contain island violations.

Extraction Type	Sequence	Example	Child-Directed Speech
Matrix Subject	TP	Who _ ate the peanutes	-1.21
Embedded Subject	TP-VP-CP_{null}-TP	Who did Paulina suspect _ ate the peanuts	-7.89
Embedded Object	TP-VP-CP_{that}-TP-VP	What did Paulina suspect that Jad ate _	-13.84
Complex NP	TP-VP-DP-CP_{that}-TP-VP	*What did Paulina repeat [the claim that Jad ate _]	-19.81
Subject	TP-VP-CP_{null}-TP-DP-PP	*What did Paulina suspect [the man with _] ate the peanuts	-20.17
Whether	TP-VP-$CP_{whether}$-TP-VP	*What did Paulina wonder [whether Jad ate _]	-18.54
Adjunct	TP-VP-CP_{if}-TP-VP	*What did Paulina notice [if Jad ate _]	-18.54

Table 7. Trigram sequence probabilities (Pearl & Sprouse (2013): 46).

Line	Trigram Sequence	Probability (child-directed speech)
1	start-TP-VP	.42
2	TP-VP-end	.40
3	TP-VP-TP	.031
4	TP-VP-CP_{null}	.0073
5	TP-VP-NP	.0015
6	CP_{that}-TP-VP	.000044
7	VP-NP-CP_{that}	.0000012
8	DP-CP_{that}-TP	.0000012
9	CP_{null}-TP-NP	.0000012
10	TP-NP-PP	.0000012
11	TP-VP-$CP_{whether}$	.0000012
12	TP-VP-CP_{if}	.0000012

5.2 Review of Pearl & Sprouse (2013)

One weakness of Pearl & Sprouse (2013) is that it fails to commit to being either (a) a learning model of acceptability judgments or (b) a learning model of grammatical knowledge. The model aspires to inform theories of how island constraints are learned, and since these are grammatical constraints, the model therefore aspires to be option (b). Recall that the experiments in Wagers, Sprouse and Phillips (2012) successfully isolate (and quantify) the contribution that grammatical knowledge makes to overall acceptability judgments by showing that sentences which contain island violations are rated lower beyond what is predicted by other factors, like the sentence complexity and sentence length (controlling for lexical information whenever possible). This work successfully untangles the effects that grammaticality – namely, island constraints – has on acceptability from other effects which are not of interest to syntactic theorists. Although the Pearl & Sprouse (2013) model aspires to be a learning model of grammatical knowledge, it uses probability scores to explain both acceptability and grammaticality without addressing the question of how children ultimately acquire the representations which have prompted theorists to define grammaticality separately from acceptability. At the same time, the model fails to capture all the same effects on acceptability as the Wagers, Sprouse and Phillips (2012) experiments do. It captures the effect of the length of *wh*- extraction path, but fails to capture the effect of syntactic complexity (e.g., the presence of an island structure in a sentence). The end result is a partial reproduction of the adult acceptability scores presented in Wagers, Sprouse and Phillips, (2012), which raises questions about whether the learning model would be successful if it considered the probability of the entire sentence.

The learning model uses probability as a proxy for both grammaticality and acceptability. The probabilities assigned to syntactic trigram sequences represent grammaticality: unattested trigrams have a zero probability, and low probability trigram sequences which contain a zero probability trigram are ungrammatical. The probabilities assigned to syntactic trigram sequences also represent acceptability: the longer the sequences, the lower the probability; the product of n numbers between 0 and 1 will always be smaller than the product of n+1 numbers between 0 and 1. It would be a mistake to attribute the low probabilities of longer utterances to grammaticality because sentence length is not a grammatical factor. But, is it a mistake to attribute the low probability trigrams to acceptability, instead of grammaticality? Recall that the results in Wagers, Sprouse and Phillips show that syntactic complexity – the presence of an island in a grammatical *wh*- question – has the effect of lowering acceptability judgment scores, even though the sentence is grammatical. People tend to rate low frequency sentence types with lower scores than high frequency sentence types. The more syntactic elements loaded into a single sentence, the less frequent the syntactic frame will be. As a consequence, it will be rated lower than simpler sentences, but not because it's ungrammatical. The Pearl & Sprouse (2013) learner assigns a probability to the *wh*- extraction path – not the entire sentence – so syntactic information outside the *wh*- extraction path is lost. As a result, the model assigns the same probability to matrix questions with islands (e.g., *Who made the claim that Lily forgot the necklace?*) and without islands (e.g., *Who claimed that Lily forgot the necklace?*) (Figure X) even though adults show a clear preference for matrix questions without islands (Wagers, Sprouse, and Phillips 2012). In the plots, the plus and the open circle overlap completely in the matrix conditions (Figure X).

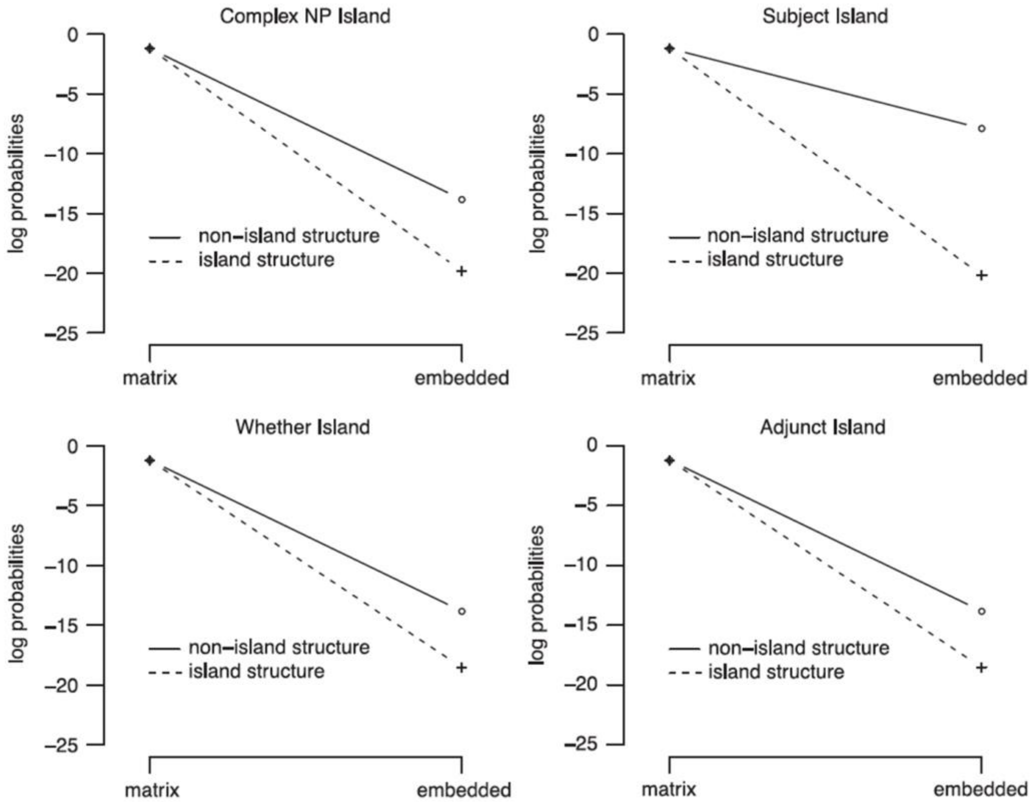


Figure 18. Pearl & Sprouse (2013: Figure 5): Log probabilities derived from a learner using child-directed/ambient speech.

Pearl & Sprouse (2013) observes something similar in a footnote on page 41, but uses sentence length (not syntactic complexity) as an example:

> "This shows that actual process of generating acceptability judgments is likely more nuanced than the basic implementation in the current algorithm. One clear difference is that the current algorithm does not factor in the portion of the utterance beyond the gap position, whereas the actual process in humans likely does. For example, *Who saw it?* is not judged as equivalent to *Who thought that Jack said that Lily saw it?*, even though both are IP dependencies. Similarly, the current algorithm does not factor lexical or semantic properties into the

judgments, whereas the actual process in humans likely does. This is why experimental studies have to balance the lexical, structural, and semantic properties of the experimental materials, as Sprouse, Wagers & Phillips (2012a) did" (Pearl & Sprouse, 2013: footnote on page 41).

The same point is made in the following passage, which also clearly indicates that this model is intended as a model of grammatical knowledge, and not acceptability:

"Although these results demonstrate that our modeled learner can acquire the general super-additive interaction pattern observed in the actual acceptability judgment experiments, it should be noted that there are noticeable differences between the observed acceptability judgments and the inferred grammaticality preferences learned by this model. The reason for this is that actual acceptability judgments are based on dozens of factors that are not included in this model. For example, lexical items, semantic probability, and processing difficulty have all been demonstrated to impact acceptability judgments (Schütze 1996; Cowart 1997; Keller 2000; Sprouse 2009). The inferred grammaticality of this particular model would constitute only one (relatively large) factor among many that affect acceptability. In other words, the grammaticality preferences of this model are themselves limited to the dependency alone—they ignore all of the other properties of the sentence." (Pearl & Sprouse, 2013: page 44)

Nonetheless, the blurred line between acceptability and grammaticality is positively acknowledged:

"The results also suggest that two desirable properties of acceptability judgments fall out of this algorithm: (i) a general preference for shorter dependencies, and (ii) a qualitative distinction between long dependencies and ungrammatical dependencies (at least in principle)" (Pearl & Sprouse, 2013: page 47).

As a brief aside, the current algorithm *does* factor lexical properties into judgments: the model uses lexical information to determine that CP_{null} and CP_{that} allow extraction, but CP_{if} and $CP_{whether}$ do not. Returning to the main thread, the Pearl & Sprouse (2013) results are missing a baseline. The model does not consider material outside the *wh*- extraction path to be relevant to its learning, but it's a reasonable assumption that a learner which assigns probabilities to syntactic trigrams might do so more generally, to acquire other grammatical properties of their language. If this is the case, then this kind of learning model has another induction problem to solve: the zero probability trigrams are a source of ungrammaticality in *wh*- argument questions, but these trigrams are rare in the linguistic input regardless of whether they're used in *wh*-argument questions or not. Why wouldn't a learning model of this variety (i.e., an inductive learning model) decide that these structures are ungrammatical across the board? Pearl & Sprouse (2013) argues that the non-parallel lines in their plot demonstrate that their model succeeded at learning island constraints. If the matrix-island condition reflected the fact that these sentences contain extremely rare syntactic trigrams, would the slopes of the lines connecting the matrix and embedded conditions still be non-parallel in these plots, or would the baseline probability in the matrix-island condition be equal to the embedded-island condition? If so, then the effect of island constraints on grammaticality would disappear, and the model would not appear to succeed. I argue that this missing baseline is critical to determine whether the model has succeeded in acquiring island constraints.

As touched on briefly in the paragraphs above, another weakness of the model is that the learner overcomes the induction problem by tracking lexical information on CP. Pearl & Sprouse (2013) motivate this choice by observing that lexical subcategorization must be learned anyway. While this is true, it's not obvious why lexical information on CP exclusive of the other phrasal projections in the *wh*-extraction path should be considered by the learner.

It's reasonable to expect that the learner would use some syntactic information to infer syntactic constraints, but the training dataset for the learning model is a perfectly parsed set of utterances. Obviously, children who are in the process of acquiring the grammar of their first language will not have access to this kind of linguistic data. At 2 years-old (the age at which Pearl & Sprouse (2013) estimates that children begin inducing island constraints), infant syntactic representations are imperfect because their grammatical knowledge is incomplete, their sentence processing mechanisms are still developing, their lexicons are still growing, and for many other reasons. It's clear that strategically, a complex model which takes into account all these factors is the next step after implementing a simpler model. But, it's interesting to explore how much the learning outcome depends on the learner correctly identifying the position of the gap. For example, in the adult grammar, sentence (41) is unambiguously a question about the instrument of eating – *a fork* would be a fitting answer to question (41).

71. What did she eat the cupcake with?

As discussed in Chapter 3, sentence (41) is ambiguous for any child who does not have knowledge of island constraints – it could be a question about the decoration on the cupcake, in which case *sprinkles* would be a fitting answer. In the Pearl & Sprouse (2013) model, if sentence (41) were included in the training set, it would not be ambiguous – it would have only one parse: the grammatical one. The trigram sequence would be start-TP-VP-PP-end. This means that the Pearl & Sprouse (2013) learner induces knowledge of island constraints from sentences which have been parsed to respect island constraints. Island constraints are supposed to be the target of learning, but they are already built into the training dataset. While it's true that adult speakers will only produce sentences which are unambiguous when it comes to potential island violations, children can't know that (unless knowledge of island constraints is innate), so the perfectly parsed trees contribute to the solution of the problem, in this case.

The Pearl & Sprouse (2013) model is committed to the hypothesis that children acquire *wh*- movement in a series of steps, advancing from one knowledge state to the next until they reach a stable knowledge state. This theoretical commitment creates the opportunity for variation in the ways that *wh*- movement can be represented in early grammars, including the hypothesis that *wh*- movement proceeds in *one fell swoop*, and not as a series of smaller movements. It's not clear whether the Pearl & Sprouse (2013) approach is committed to cyclic *wh*- movement or not. But, treating the 'extraction path' as the meaningful unit, as opposed to breaking it into the smaller, cyclically-determined paths, is suggestive of a *one fell swoop* conceptualization of *wh*- movement. Although modern syntactic theory has made serious attempts to link cyclicity with island constraints, this learning model is ambivalent about whether the learner treats long-distance movement as one extraction or two. A major challenge for the Pearl & Sprouse (2013) learner will be to explain how knowledge of cyclic movement is acquired. If Pearl & Sprouse (2013) assumes that this is innate knowledge, then the *wh*-extraction paths that the learner represents should reflect this knowledge.

Pearl & Sprouse (2013) discusses at length the consequences of several learning biases, including the fact that the model focuses only on the syntax *wh*- extractions in argument questions, and that tracking trigrams (but not bigrams) allows the model to succeed. The learner must already know that the syntax of these constructions is information which should be tracked, which indicates that the learner knows to attend to the syntax of the *wh*- extraction path (but not the syntax of polar questions, for example). Further discussion, see Pearl & Sprouse (2013) and Phillips (2012).

5.3 Conclusion

This chapter summarizes the Pearl & Sprouse (2013) learning model, which explores the possibility that learners induce island constraints from properties in their linguistic input as a second step after acquiring knowledge of *wh*- movement. By tracking the syntax of the *wh*- extraction paths, the model uses probability as an index of grammaticality to determine that subjects, complex NPs, *whether* complements and *if* clauses block *wh*- movement. One weakness of this learning model is that probability is used to explain both grammaticality and acceptability without an explanation of how the child is supposed to identify when low probability is a signal of grammaticality versus when it's a signal of acceptability. Unlike the Wagers, Sprouse and Phillips (2012) experiments, this model is missing the baseline which controls for the effect of syntactic complexity, making it unclear whether or not the model has actually succeeded at acquiring island constraints. Furthermore, the learner has the benefit of inducing island constraints from unambiguous syntactic representations (which have been parsed by speakers with knowledge of island constraints). It would be interesting to consider how ambiguous data might affect learning outcomes. Finally, the model fails to address how the learner acquires cyclic representations of *wh*- movement, even though cyclicity is the standard explanation for one of the island constraints which the model is said to learn. In the concluding chapter, I discuss how the behavioral findings in this dissertation bear on learning models of island constraints.

Chapter 6: Conclusion

This dissertation improves our understanding of the acquisition of syntax by introducing two novel studies which target early syntactic representations of *wh*-movement and island constraints. Chapter 3 presents an experiment which leverages the effects of immature sentence processing mechanisms to target knowledge of *wh*-movement at 19 months-old. Infants parse *what* as the direct object in *wh*- questions (e.g., *What is she pulling with the gop?*) and learn that *the gop* refers to the puller. In the control condition, infants incorrectly parse *the gop* as the pullee in intransitive declaratives (e.g., *She's pulling with the gop!*), reproducing the parsing error reported in Lidz et al., (2017). Infants succeed at representing the *wh*- question accurately because they use *what* to satisfy the transitivity preferences of *pull*, a strategy which allowed 19 month-olds to succeed at parsing transitive declaratives (e.g., *She's pulling that thing with the gop!*) in previous experiments (Lidz et al., 2017). These results are consistent with recent work which shows that infants represent the co-occurrence of fronted *wh*- phrases and a gap at 18 months-old (Perkins & Lidz, 2020), and corroborate claims that 20 month-old infants respond appropriately to *wh*- questions during language comprehension tasks because they have acquired knowledge of syntax (as opposed to relying heavily on the experimental context).

This result shows that 19 month-olds recognize that *wh*- object questions are transitive frames, even though the direct object has been displaced from its canonical argument position (see Perkins 2019 for similar findings). Chapter 3 argues that 19 month-olds succeed at this task by predictively resolving the *wh*- dependency to avoid making the parsing error characteristic of 19 month-olds tested in this paradigm (Lidz et al., 2017). This interpretation of the results suggests that early sentence processing strategies are influenced by syntactic knowledge. Upon acquiring knowledge of *wh*-movement, infants deploy that knowledge during sentence processing. This experiment provides insight into how emergent syntactic representations guide sentence processing mechanisms. The study presented in Chapter 3 indicates that infants represent *what* as the direct object in *wh*- questions, but cannot provide finer-grained insights into the nature of these early representations. It remains to be seen whether early representations of *wh*- movement are cyclic or island sensitive, but establishing that infants represent the dependency between the filler and the gap is a first step toward building a complete model of the acquisition of syntax.

This dissertation contributes to our model of language acquisition by adding to the foundation of behavioral research on the acquisition of syntax. If 19 month-olds have acquired knowledge of *wh*- movement and the ability to apply that knowledge during sentence processing, then there is potential to observe behavioral evidence for island constraints during infancy. This potential is contingent on many other factors, such as syntactic knowledge of island structures. For example, if an infant does not represent a clause as an adjunct, then even if the infant represents the clausal adjunct constraint, their behavior would indicate that they've violated the island constraint when in fact they've mis-represented the structure of the sentence. In order to observe island sensitivity during infancy, infants must have acquired the relevant syntactic knowledge, and even then, the difficulties with testing infants on knowledge of syntax might sabotage our ability to detect any signal of knowledge of island constraints (even if the knowledge is there).

Chapter 4 shifts the focus from early representations of *wh-* movement to early representations of island constraints. Previous studies have found evidence for knowledge of island constraints around 4 to 5 years-old (de Villiers et al., 1990; de Villiers & Roeper, 1995a, 1995b; Fetters & Lidz, 2016; Goodluck et al., 1992), but it's not clear when over the course of development children acquire this knowledge. For generative theorists, the results showing that 19 month-olds represent *wh-* movement invite the following question: if knowledge of island constraints is innate, and 19 month-olds represent *wh-* movement, then why isn't there any behavioral evidence for knowledge of island constraints until 3 years later in the developmental timeline? Chapter 3 presents a novel task designed to target knowledge of island constraints at 3 years-old. Part of making the task appropriate for this younger age group involved moving away from traditional island constraints, which tend to depict extraction from a clause. Instead, the island tested in this study is smaller (adjunct to NP), which reduces the complexity of the task considerably. This experiment asks children a *wh-* question which is unambiguous to adult speakers, but is ambiguous for any speaker who does not have knowledge of island constraints. For example, adult speakers could only answer a question like *Which blanket did she hug the bear with?* with the instrument of hugging, and not a blanket that the bear is holding. Results show that 3 year-olds consistently give adult-like responses to these questions, while giving mixed responses to an ambiguous syntactic control (e.g., *She hugged the bear with the blanket! Which blanket was it?*). Experiment X tests adult speakers on the same stimuli and yields the same results. Experiments X and X test for island sensitivity in children and adults using relative clauses (e.g., *The blanket she hugged the bear with is so pretty! Which* *blanket is it?*), but child results were inconclusive, likely due to difficulties with adjusting the experiment to online testing. This result corroborates the behavioral findings that 4 to 5 year-olds represent island constraints, but suggests that this knowledge emerges earlier than has been previously assumed.

There are two well-known approaches for how knowledge of island constraints is acquired. The generative approach proposes that knowledge of island constraints is not learned, and cannot be learned because the linguistic input lacks the information necessary to learn that *wh*- movement is constrained (e.g., Chomsky 1965; Hornstein & Lightfoot 1985). Knowledge of island constraints is a consequence of the nature of linguistic representations, and is therefore inseparable from knowledge of *wh*-movement. Non-generative approaches propose that island constraints can be induced from the linguistic input (e.g., Pearl & Sprouse, 2013). In the Pearl & Sprouse (2013) learning model, children learn *wh*- movement, and then they induce island constraints from the primary linguistic data. The Pearl & Sprouse learner tracks the syntax of extraction paths in *wh*- argument questions to determine the probability of the syntactic structure in each path. Results show that paths which contain islands are assigned low probabilities compared to paths which do not contain islands (where probability is a proxy for grammaticality/acceptability). Chapter 4 summarizes the learning model and raises questions about the linking assumptions between probability, acceptability and grammaticality and the move to restrict the input to the model to include only the extraction path.

For the Pearl & Sprouse (2013) model, *when* knowledge of island constraints is acquired is linked to *how* it's acquired because the 'learning' proceeds induction.

Induction is driven largely by experience, so children must have sufficient exposure to *wh-* object questions to induce island constraints. This learning model estimates that children induce island constraints between 2 and 5 years old, over a 3 year period. It's not clear whether this entire period is necessary for the model to converge. Because the model treats unobserved trigrams as unacceptable/ungrammatical, sufficient experience must be accumulated such that the model doesn't reject every *wh-* question. The results from Chapters 3 and 4 suggest that the age range estimated in Pearl & Sprouse (2013) is both too long and is shifted toward older children. If children have acquired knowledge of *wh-* movement at 19 months, and island constraints by 3 years-old, and they learn island constraints from their linguistic experience with *wh-* object questions, then they must do so using half of the data estimated in the Pearl & Sprouse (2013) learning model. If the model fails to converge with half the training data, then it will be interesting to see how it would need to be revised to succeed. Shifting the age range to younger children also has consequences for the quality of the training data. The syntactic complexity of child-directed speech is correlated with the child's age, so long-distance *wh-* questions are even less frequent in speech to infants than they are in speech to children. These data points are crucial to the success of the learning model because they allow it to learn that long-distance movement is acceptable/grammatical, but not across certain clause boundaries.

Future work is needed to build a complete model of the acquisition of *wh-* movement and island constraints. More behavioral studies on the early representations of long distance *wh-* movement are needed to bridge the infant research on *wh-* movement to the research on older children and island constraints. Do these early syntactic representations treat long distance movement as *one fell swoop*, or does movement proceed cyclically, stopping at each clause boundary? For generative approaches, evidence that children treat long-distance movement as shorter steps would indicate that they have correctly identified the syntactic dependency as *wh*- movement. The work in this dissertation inches scientific progress toward this goal.

Bibliography

Ambridge, B., & Goldberg, A. E. (2008). *The island status of clausal complements: Evidence in favor of an information structure explanation*.

Atkinson, E. E. (2016). *ACTIVE DEPENDENCY COMPLETION IN ADULTS AND CHILDREN: REPRESENTATIONS AND ADAPTATION*.

Bach, E., & Horn, G. M. (1976). Remarks on" conditions on transformations". *Linguistic Inquiry*, *7*(2), 265–299.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." Journal of Statistical Software, 67(1), 1–48. doi:10.18637/jss.v067.i01.

Berwick, R. C., Pietroski, P., Yankama, B., & Chomsky, N. (2011). Poverty of the stimulus revisited. *Cognitive Science*, *35*(7), 1207–1242.

Boeckx, C. (2008). The fine structure of intervention in syntax. In *Aspects of the Syntax of Agreement* (pp. 174–194). Routledge.

Boeckx, C. (2012). *Syntactic Islands*. Cambridge University Press.

Broen, P. A. (1972). *The Verbal Environment of the Language-Learning Child. ASHA Monographs, No. 17.*

Cairns, H. S., & Hsu, J. R. (1978). Who, why, when, and how: A development study. *Journal of Child Language*, *5*(3), 477–488.

Cheng, L. L.-S., & Corver, N. (2006). Lines of inquiry into wh-movement. In *Wh-movement: Moving on* (pp. 1–18). MIT Press.

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Chomsky, N. (1973). Conditions on Transformations. In S. Anderson & P. Kiparsky (Eds.), *A Festschrift for Morris Halle*. Holt, Rinehart and Winston.

Chomsky, N. (1976). On the nature of language. *Annals of the New York Acadamy of Sciences*, 46–57. https://doi.org/10.1111/j.1749-6632.1976.tb25470.x

Chomsky, N. (1977). On wh-movement. In P. Culicover, A. Akmajian, & T. Wasow (Eds.), *Formal Syntax* (pp. 71–133). Academic Press.

Chomsky, N. (1980). On cognitive structures and their development: A reply to Piaget. In M. Piattelli-Palmarini (Ed.), *Language and Learning: The debate between Jean Piaget and Noam Chomksky* (p. 409). Routledge & Kegan Paul.

Chomsky, N. (1981). Principles and parameters in syntactic theory. In N. Hornstein & D. Lightfoot (Eds.), *The Logical Problem of Language Acquisition* (pp. 32–75). Longman.

Chomsky, N. (1986). *Barriers*. MIT Press.

Cinque, G. (1990). Types of A-bar Dependencies. *MIT Press*.

Citko, B. (2014). *Phase theory: An introduction*. Cambridge University Press.

Davies, W. D., & Dubinsky, S. (2003). On extraction from NPs. *Natural Language & Linguistic Theory*, *21*(1), 1–37.

de Villiers, J., & Roeper, T. (1995a). Barriers, Binding, and Acquisition of the DP-NP Distinction. *Language Acquisition*, *4*(1–2), 73–104. https://doi.org/10.1080/10489223.1995.9671660

de Villiers, J., & Roeper, T. (1995b). Relative clauses are barriers to wh-movement for young children. *Journal of Child Language*, *22*(2), 389–404. https://doi.org/10.1017/S0305000900009843 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). Springer Netherlands. https://doi.org/10.1007/978-94-011-3808-6_10

Deane, P. (1991). *Limits to attention: A cognitive theory of island phenomena*.

Erteschik-Shir, N. (1973). *On the nature of island constraints.*

Fetters, M., & Lidz, J. (2016). Early Knowledge of Relative Clause Islands and Island Repair. *Proceedings of the 41st Annual Boston University Conference on Language Development*, *1*, 236–249.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2006). MacArthur-Bates Communicative Development Inventories, Second Edition (CDIs) [Database record].

Fisher, C., Jin, K., & Scott, R. M. (2020). The Developmental Origins of Syntactic Bootstrapping. *Topics in Cognitive Science*, *12*(1), 48–77. https://doi.org/10.1111/tops.12447

Fox, D., & Lasnik, H. (2003). Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry*, *34*(1), 143–154.

Furrow, D., Nelson, K., & Benedict, H. (1979). Mothers' speech to children and syntactic development: Some simple relationships. *Journal of Child Language*, *6*(3), 423–442.

Gagliardi, A., Mease, T. M., & Lidz, J. (2016). Discontinuous development in the acquisition of filler-gap dependencies: Evidence from 15- and 20-month-olds.

Language Acquisition: A Journal of Developmental Linguistics, *23*(3), 234–260. https://doi.org/10.1080/10489223.2015.1115048

Givón, T. (1979). From discourse to syntax: Grammar as a processing strategy. In *Discourse and syntax* (pp. 81–112). Brill.

Goldberg, A. E. (2013). 10 Backgrounded constituents cannot be "extracted." *Experimental Syntax and Island Effects*, *221*.

Goodluck, H., Foley, M., & Sedivy, J. (1992). Adjunct Islands and Acquisition. In H. Goodluck & M. Rochemont (Eds.), *Island Constraints: Theory, Acquisition and Processing* (pp. 181–194). Springer Netherlands. https://doi.org/10.1007/978-94-017-1980-3_6

Goodluck, H., Sedivy, J., & Foley, M. (1989). WH-Questions and Extraction from Temporal Adjuncts: A Case for Movement. *Wh-Questions and Extraction from Temporal Adjuncts: A Case for Movement*, *28*, 123–130.

Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, *86*(2), 366.

Hollich, G., Rocroi, C., Hirsh-Pasek, K., & Golinkoff, R. (1999). *Testing language comprehension in infants: Introducing the split-screen preferential-looking paradigm*. 15–18.

Hornstein, N., & Lightfoot, D. (1981). Introduction. In N. Hornstein & D. Lightfoot (Eds.), *The Logical Problem of Language Acquisition* (pp. 9–31). Longman.

Hornstein, N., & Lightfoot, D. (1985). *The logical problem of Language Acquisition*.

Huang, J. (1982). *Logical relations in Chinese and the theory of grammar*. Massachusetts Institute of Technology.

Huttenlocher, J., Levine, S., & Vevea, J. (1998). Environmental Input and Cognitive Growth: A Study Using Time-Period Comparisons. *Child Development*, *69*(4), 1012–1029. https://doi.org/10.1111/j.1467-8624.1998.tb06158.x

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*(3), 337–374. https://doi.org/10.1016/S0010-0285(02)00500-5

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*(4), 573–633.

Lidz, J. (2018). The explanatory power of linguistic theory. In N. Hornstein, H. Lasnik, P. Patel-Grosz, & C. Yang (Eds.), *Syntactic Structures after 60 Years: The Impact of the Chomskyan Revolution in Linguistics* (Vol. 129, pp. 225–242). Walter de Gruyter GmbH & Co KG.

Lidz, J., & Lasnik, H. (2017). The Argument from the Poverty of the Stimulus. In I. Roberts (Ed.), *The Oxford Handbook of Universal Grammar*.

Lidz, J., White, A. S., & Baier, R. (2017). The role of incremental parsing in syntactically conditioned word learning. *Cognitive Psychology*, *97*, 62–78.

Liu, Y., Ryskin, R., Futrell, R., & Gibson, E. (2019). *Verb Frequency Explains the Unacceptability of Factive and Manner-of-speaking Islands in English.* 685–691.

Merchant, J. (2008). Variable island repair under ellipsis. *Topics in Ellipsis*, *1174*(2008), 132–153.

Newport, E., Gleitman, L., & Gleitman, H. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. Snow & C.

Ferguson (Eds.), *Talking to children: Language input and acquisition*. C. U. P.

Noveck, I. A. (2001). When children are more logical than adults: Experimental investigations of scalar implicature. *Cognition*, *78*(2), 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Obenauer, H.-G. (1985). *On the identification of empty categories*.

Omaki, A. (2010). *Commitment and flexibility in the developing parser* [Ph.D., University of Maryland]. https://www.proquest.com/docview/762232044/abstract/A08C11996623401EPQ/1

Omaki, A., & Lidz, J. (2015). Linking Parser Development to Acquisition of Syntactic Knowledge. *Language Acquisition*, *22*(2), 158–192. https://doi.org/10.1080/10489223.2014.943903

Otsu, Y. (1981). *Universal grammar and syntactic development in children: Toward a theory of syntactic development*. Massachusetts Institute of Technology.

Pearl, L. (2021). Poverty of the Stimulus Without Tears. *Language Learning and Development*, 1–40. https://doi.org/10.1080/15475441.2021.1981908

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. https://doi.org/10.1080/10489223.2012.738742

Perkins, L., & Lidz, J. (2021). Eighteen-month-old infants represent nonlocal syntactic dependencies. *Proceedings of the National Academy of Sciences*, *118*(41). https://doi.org/10.1073/pnas.2026469118

Phillips, C. (2013). On the nature of island constraints II: Learning language and innateness. In J. Sprouse & N. Hornstein (Eds.), *Experimental Syntax and Island Effects* (pp. 132–157). Cambridge University Press.

Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Bradford Books.

Radford, A. (1994). The syntax of questions in child English [*]. *Journal of Child Language*, *21*(1), 211–236.

Richards, N. (2014). A-bar movement. *The Routledge Handbook of Syntax*, 167–191.

Rizzi, L. (1990). *Relativized minimality.* The MIT Press.

Roeper, T., & Villiers, J. de. (2011). The acquisition path for wh-questions. In *Handbook of generative approaches to language acquisition* (pp. 189–246). Springer.

Ross, J. R. (1967). *Constraints on variables in syntax*. Massachusetts Institute of Technology.

Ross, J. R. (1969). Guess who? In R. Binnick, A. Davison, G. Green, & J. Morgan (Eds.), *Papers from the 5th Regional Meeting of the Chicago Linguistic Society* (pp. 252–286). Chicago Linguistic Society.

Rottman, I., & Yoshida, M. (2013). Sluicing, Idioms, and Island Repair. *Linguistic Inquiry*, *44*(4), 651–668. https://doi.org/10.1162/LING_a_00142

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object wh-questions. *Infancy*, *4*, 423–436. https://doi.org/10.1207/S15327078IN0403_06

Snedeker, J. (2013). Children's Sentence Processing. In R. P. G. van Gompel (Ed.), *Sentence Processing*.

Spelke, E. (1976). Infants' intermodal perception of events. *Cognitive Psychology*, *8*(4), 553–560.

Sprouse, J., & Hornstein, N. (2013). Experimental syntax and island effects: Toward a comprehensive theory of islands. In J. Sprouse & N. Hornstein (Eds.), *Experimental Syntax and Island Effects* (pp. 1–17). Cambridge University Press.

Szabolcsi, A. (2006). Weak and strong islands. *The Syntax Companion, Eds M. Everaert and H. van Riemsdijk (Oxford: Blackwell)*, 479–531.

Szabolcsi, A., & Den Dikken, M. (1999). Islands. *Glot International*, *4*(6), 3–8.

Tabor, W., Juliano, C., & Tanenhaus, M. K. (1997). Parsing in a Dynamical System: An Attractor-based Account of the Interaction of Lexical and Structural Constraints in Sentence Processing. *Language and Cognitive Processes*, *12*(2–3), 211–271. https://doi.org/10.1080/016909697386853

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Miller & P. D. Eimas (Eds.), Speech, language, and communication (pp. 217–262). Academic Press. https://doi.org/10.1016/B978-012497770-9.50009-1

Trueswell, Sekerina, I. A., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*(2), 89–134. https://doi.org/10.1016/S0010-0277(99)00032-3