ABSTRACT

Title of dissertation: COMPARATIVE PSYCHOSYNTAX

Dustin Alfonso Chacón, Doctor of Philosophy, 2015

Dissertation directed by: Professor Colin Phillips
Department of Linguistics

Every difference between languages is a "choice point" for the syntactician, psycholinguist, and language learner. The syntactician must describe the differences in representations that the grammars of different languages can assign. The psycholinguist must describe how the comprehension mechanisms search the space of the representations permitted by a grammar to quickly and effortlessly understand sentences in real time. The language learner must determine which representations are permitted in her grammar on the basis of her primary linguistic evidence. These investigations are largely pursued independently, and on the basis of qualitatively different data. In this dissertation, I show that these investigations can be pursued in a way that is mutually informative. Specifically, I show how learnability concerns and sentence processing data can constrain the space of possible analyses of language differences.

In Chapter 2, I argue that "indirect learning", or abstract, cross-contruction syntactic inference, is necessary in order to explain how the learner determines which complementizers can co-occur with subjects gaps in her target grammar. I

show that adult speakers largely converge in the robustness of the *that*-trace effect, a constraint on complementation complementizers and subject gaps observed in languages like English, but unobserved in languages like Spanish or Italian. I show that realistic child-directed speech has very few long-distance subject extractions in English, Spanish, and Italian, implying that learners must be able to distinguish these different hypotheses on the basis of other data. This is more consistent with more conservative approaches to these phenomena (Rizzi, 1982), which do not rely on abstract complementizer agreement like later analyses (Rizzi, 2006; Rizzi & Shlonsky, 2007).

In Chapter 3, I show that resumptive pronoun dependencies inside islands in English are constructed in a non-active fashion, which contrasts with recent findings in Hebrew (Keshev & Meltzer-Asscher, ms). I propose that an expedient explanation of these facts is to suppose that resumptive pronouns in English are ungrammatical repair devices (Sells, 1984), whereas resumptive pronouns in island contexts are grammatical in Hebrew. This implies that learners must infer which analysis is appropriate for their grammars on the basis of some evidence in linguistic environment. However, a corpus study reveals that resumptive pronouns in islands are exceedingly rare in both languages, implying that this difference must be indirectly learned. I argue that theories of resumptive dependencies which analyze resumptive pronouns as incidences of the same abstract construction (e.g., Hayon 1973; Chomsky 1977) license this indirect learning, as long as resumptive dependencies in English are treated as ungrammatical repair mechanisms.

In Chapter 4, I compare active dependency formation processes in Japanese

and Bangla. These findings suggest that filler-gap dependencies are preferentially resolved with the first position available. In Japanese, this is the most deeply embedded clause, since embedded clauses always precede the embedding verb(Aoshima *et al.*, 2004; Yoshida, 2006; Omaki *et al.*, 2014). Bangla allows a within-language comparison of the relationship between active dependency formation processes and word order, since embedded clauses may precede or follow the embedding verb (Bayer, 1996). However, the results from three experiments in Bangla are mixed, suggesting a weaker preference for a lineary local resolution of filler-gap dependencies, unlike in Japanese. I propose a number of possible explanations for these facts, and discuss how differences in processing profiles may be accounted for in a variety of ways.

In Chapter 5, I conclude the dissertation.

COMPARATIVE PSYCHOSYNTAX

by

Dustin Alfonso Chacón

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
2015

Advisory Committee:
Professor Colin Phillips, Chair/Advisor
Professor Howard Lasnik, Advisor
Professor Norbert Hornstein
Professor Ellen Lau
Professor Manel Lacorte, Dean's Representative

© Copyright by
Dustin Alfonso Chacón
2015

Preface

Chapter 2 reports research that was jointly conducted with Michael Fetters, Margarent Kandel, Eric Pelzl, and Colin Phillips. Margaret Kandel is identified as MK. Chapter 3 reports research that was jointly conducted with Rebecca Kraut and Colin Phillips. Rebecca Kraut is identified as RK. Finally, Chapter 4 reports on research conducted with Mashrur Imtiaz, Sikder Monoare Murshed, Shirsho Dasgupta, Mina Dan, and Colin Phillips. Mashrur Imtiaz and Sikder Monoare Murshed are referred to as MI and SMM in Chapter 4. Parts of all three of these chapters are submitted with these collaborators listed as co-authors.

Dedication

In memory of Gina.

Acknowledgments

Putting my name on this dissertation seems almost fraudulent. Almost every idea embedded in this thesis has been inspired and refined by many people, both inside and outside of University of Maryland Language Science community. Apart from the ideas themselves, the actual writing process was only feasible with the endless support of many friends and family members.

First, it goes without saying that I owe much of my success in the past 5 years to my advisor Colin Phillips. Colin is an amazing and challenging person to work with, and I know that I would not have achieved nearly as much as I have without his support and, perhaps more importantly, his criticisms. He's been there to answer e-mails at any hour of the day and from any corner of the globe (my inbox currently has over 1,400 threads with him) or calming me down during freakouts in our morning meetings. The level of dedication he has given to both his students and the Language Science community is truly astounding. I can only hope to be as a thoughtful and committed of a mentor to others in the future.

My co-advisor, Howard Lasnik, has been an invaluable resource and important foil to Colin. Howard's enthusiasm and apparently endless knowledge of all things in linguistics and outside of the field was always energizing. Howard's willingness to consider every odd idea, analysis, and speculation I've offered over the last 5 years is probably why I'm still in this business. Additionally, Howard has been great at keeping me grounded, reminding me that I'm still at the beginning of my career.

I also owe an enormous debt to Jeffrey Lidz, Ellen Lau, and Norbert Hornstein.

Although I did not work nearly as closely with these three as I did with Howard and Colin, all three have independently contributed significantly to the way that I approach problems in language. My biggest regret in grad school is not spending more time collaborating with these three, and I am more than grateful for the conversations I did have with them. I also owe gratitude to Alexander Williams, Valentine Hacquard, Omer Preminger, Naomi Feldman, Phil Resnik, Hal Daumé, and Bill Idsardi for many thought-provoking conversations the past 5 years.

My colleagues at University of Maryland are some of the most colorful people that I've had the opportunity to meet. Aaron Steven White is one of the smartest and most frustrating people ever, but I love him all the same anyway, especially since he is the only person I know who can stop mid-way through a conversation at a party to explain what a hypercube is. His music and beer tastes are also pretty choice. Kate Harrigan feels like a friend that I've known my whole life. She and her husband Jeff Vitols have in many ways been a two-person support network over the last 5 years. I also would like to think that I played some role in her becoming an animal person. Mike Fetters has been an endless source of interesting conversation and amusing anecdotes, and has been a great video game and comic book buddy. Shota Momma has given me a run for my money as a psycholinguist one year my junior and twice my calibre. He was also a pleasure to literally run with, as well as to spend caffeinated weekends at Marie Mount Hall with. Megan Sutton and Alexis Wellwood both feel like big sisters to me, and have been role models to me in both my academic and personal life. Thank you for your support, and for the many late nights talking with me about God-knows-what.

I would also be remiss to not acknowledge Brad Larson, Shevaun Lewis, Dave Kush, Sol Lago, Dan Parker, Alix Kowalski, Eric Pelzl, Julie Gerard, Morgan Moyer, Rachel Dudley, Darryl McAdams, Anton Malko, Zach Stone, Zoe Schlueter, and many other people I've had the privilege of meeting and working with over the years. I also wish to acknowledge the many collaborators that I've met at CASL, especially Anne David and Tom Conners, who gave me an excuse to think about Bangla for the past 5 years. I also owe thanks to my two research assistants – Margarent Kandel and Rebecca Kraut – who likely taught me more than I taught them.

Many of my projects that I pursued in graduate school are directly or indirectly inspired by my love of Bangla. I've had the privilege of knowing a great number of ever-patient Bangla speakers and linguists for the past 13 years. The most obvious debt I owe is to the Hasan family. In addition to being a surrogate family to me over many years, they've provided endless judgments and encouragement in my odd Bangla obsession. Thank you Aniqua, Shaheer, Nayyara, and Rafiq. Also, I would be remiss to not mention Samira Choudhury, Nandita Rahman, and Fareesa Dastagir for their support over the years. For their assistance on various projects, I owe thanks to Awalin Sopan, Nabila Ahsan, Titir Santra, Rashad Ullah, and my transcontinental Bangla syntax cohort – Arunima Choudhury, Priyanka Biswas, Diti Bhadra, and Naira Khan. I also owe a great debt to my collaborators in Dhaka and Kolkata, Sikder Monoare Murshed, Mina Dan, Mashrur Imtiaz, and Shirsho Dasgupta. ধন্যবাদ সবাইকে!

There is no chance that I would have completed this dissertation without the

support and love of my friends outside othe walls of Marie Mount Hall. Despite all of my experiences over the last 5 years, the memories that are doubtlessly most deeply etched in my memory are: moshing to Reptar and philosophizing about David Lynch with Daniel Serrano, many ill-advised trips to Cornerstone or dim-sum with Cameron Mozafari and Edward Phillips, smashing the state with Ryan Johnson, misadventures with cats and less desirable fauna with Anna Petrone, and watching Andrew Bockover and Aniqua Hasan become real adults with spouses and families (despite my protestations). I love you guys.

This acknowledgements section would be radically incomplete without mentioning my District Karaoke family. Some of the weirdest and most embarrassing moments of my life during grad school were at Irish Channel, and that's saying a great deal. Thank you for the the Minday nights – Alek Fazlipour, Anamaria Bentea, Morgan Moyer, Harris Walker, Zach Zagar, Brette Fishman, Erin Cutler, Paul Burgess, Dylan Echter, Ariela Fleisig, Jen Liberman, Keisha Vaughan, Tom Caiazza, Andrew Kim, Michael Braun, Sarah Fathallah, Anish Arwindekar, Jillian Rubino, Ellie Cossa, Jon Jones, Lisa Gregory, Shelly Sullivan Nuessle, Lauren Haman, and Lindsey Thomas. And, a special thanks to Mindi Westhoff and Kevin Machnik for making Minday Nights possible in the first place. We love rules!

Thanks are also due to my pals and the staff at the Board and Brew, who supplied the fuel with which this dissertation was written – endless coffee, Sprecher's hard root beer, and Qwarriors

I also want to thank my large, chaotic family for their attempts to understand what I do over the years. Thank you for your love and support: my mom Jodie

Bunker, my grandparents Pam and Ken Edel, and my step-family Don and Zach Bunker. Also, thank you to my dad Tony, and the whole extended Chacón tribe past and present for your love and support – Diana, Gina, Erin, Josh, Amber, Landon, Alix, Kira, Dominic, Stephen, Cameron, Jordan, Hayden, Gabriel, Skylar, Noelle, Damien, Drina, Connie, Marty, Bea, Martha, Charlotte, Bill, Nora, Pauline, Becky, Fred, and of course, Grandma Mary.

Lastly, I want to thank my funders for the past five years. This work was partially funded by the National Science Foundation Graduate Research Fellowship Program, the Sperry Foundation Beinecke Fellowship Program, and the University of Maryland Flagship Fellowship Program.

Contents

List of Tables xii

List of Figures xiv

1 Introduction 1
1.1 Grammatical Variation and Language Acquisition 9
1.2 Syntax and Psycholinguistics of Filler-Gap Dependencies 14
1.2.1 Active Dependency Formation 15
1.2.2 Locality 17
1.3 Roadmap 25

2 Testing Learnability Commitments:
The *That*-Trace Effect in English, Spanish, and Italian 27
2.1 Introduction 27
2.2 Dimension of Variation: Complementizer-Gap Interactions 29
2.3 Experiments 1–2: Acceptability of *That*-Trace Configurations 46
2.3.1 Rationale 46
2.3.2 Experiment 1 48
2.3.2.1 Methods 49
2.3.2.2 Materials 50
2.3.2.3 Results 53
2.3.2.4 Discussion 54
2.3.3 Experiment 2 56
2.3.3.1 Methods 57
2.3.3.2 Materials 57
2.3.3.3 Results 58
2.3.3.4 Discussion 59
2.3.4 Systematicity Within English Speakers 60
2.4 Primary Linguistic Data 68
2.4.1 P&S's Model 68
2.4.2 Spanish and Italian Child-Direct Speech Corpus Study 73
2.4.2.1 Methods 74
2.4.2.2 Results 75
2.5 Wider Theoretical Implications 79
2.6 Conclusion 85

3 Finding Parses:
Active Dependency Formation and Resumptive Pronouns in English and Hebrew 87
3.1 Introduction . . . 87
3.2 Dimension of Variation: Distribution of Resumptive Pronouns . . . 90
3.2.1 Syntax/Semantics of Resumptive Pronouns . . . 90
3.2.1.1 Resumptive Pronouns Are Different Between Languages . . . 93
3.2.1.2 Resumptive Pronouns are Different Across Constructions . . . 94
3.2.1.3 Resumptive Pronouns are Universally Similar . . . 96
3.2.2 Psycholinguistics of Resumptive Pronouns . . . 98
3.2.2.1 Psycholinguistics of English Resumptive Pronouns . 98
3.2.2.2 Psycholinguistics of Hebrew Resumptive Pronouns . 102
3.2.2.3 Mechanisms of Resumptive Dependency Formation . 108
3.3 Experiments 3–4: Resumptive Dependency Construction . . . 113
3.3.1 Experiment 3a . . . 116
3.3.1.1 Methods . . . 116
3.3.1.2 Materials . . . 117
3.3.1.3 Results . . . 119
3.3.1.4 Discussion . . . 120
3.3.2 Experiment 3b . . . 122
3.3.2.1 Methods . . . 122
3.3.2.2 Materials . . . 122
3.3.2.3 Results . . . 123
3.3.2.4 Discussion . . . 124
3.3.3 Experiment 3c . . . 125
3.3.3.1 Methods . . . 125
3.3.3.2 Materials . . . 126
3.3.3.3 Results . . . 127
3.3.3.4 Discussion . . . 127
3.3.4 Experiment 4 . . . 130
3.3.4.1 Methods . . . 131
3.3.4.2 Materials . . . 132
3.3.4.3 Results . . . 136
3.3.4.4 Discussion . . . 139
3.4 Primary Linguistic Data . . . 141
3.4.1 Methods . . . 143
3.4.2 Results . . . 145
3.4.3 Discussion . . . 147
3.5 Wider Theoretical Implications . . . 150
3.6 Conclusion . . . 152

4 Locality and Word Order in Bangla Active Dependency Formation 153
4.1 Introduction . . . 153
4.2 Dimension of Variation: Word Order and Active Dependency Formation . . . 155
4.3 Experiments 5–7 . . . 161
4.3.1 Experiment 5 . . . 161
4.3.1.1 Methods . . . 162
4.3.1.2 Materials . . . 163
4.3.1.3 Results . . . 166
4.3.1.4 Discussion . . . 167
4.3.2 Experiment 6 . . . 168
4.3.2.1 Methods . . . 169
4.3.2.2 Materials . . . 170
4.3.2.3 Results . . . 174
4.3.2.4 Discussion . . . 175
4.3.3 Experiment 7 . . . 177
4.3.3.1 Methods . . . 179
4.3.3.2 Materials . . . 179
4.3.3.3 Results . . . 182
4.3.3.4 Discussion . . . 182
4.4 Wider Theoretical Implications . . . 184
4.4.1 Mechanics of the Filled-Gap Effect . . . 186
4.4.2 Structured Search vs. Linear Search . . . 191
4.4.3 Representational Differences . . . 195
4.5 Conclusion . . . 199

5 Conclusion 203

List of Tables

2.1 Patterns of complementizer-gap interactions in English, Spanish, and French. 36
2.2 Results of Experiment 1. Mean z-scored ratings by language and condition with standard error of the mean and differences-in-differences of the means by language. 54
2.3 Results of Experiment 2. Mean z-scored ratings by language and condition with standard error of the mean and differences-in-differences of the means by language. 59
2.4 Results of Experiment 1. Mean z-scored ratings by condition and dialect group with standard error of the mean and differences-in-differences of the means by dialect group. 63
2.5 Counts of *wh*-distance dependencies in Pearl & Sprouse's (2013) English child-directed corpus. Relativization dependencies counts collected from DAC. 69
2.6 Details of the corpora used in the Spanish and Italian corpus study. . 74
2.7 Occurrences of *wh*-distance dependencies in the Spanish and Italian child-directed speech corpora from CHILDES. 76
2.8 Results from Gilligan's (1987) typological survey. Each row corresponds to a pair of grammatical properties – referential null subjects (NS), non-referential null subjects, post-verbal subjects (Post-VP Subj), and insensitivity to the *that*-trace effect (no *that*-t). Each column corresponds to the number of languages found that feature both properties (yes-yes), only the first (yes-no), only the second (no-no), or neither (no-no). The row corresponding to the *that*-trace strategy is bolded, with the cell corresponding to counterexamples to the *that*-trace learning strategy boxed. 80
2.9 Findings from Nicolis's (2008) study on creole languages. Each row corresponds to a language, and each column corresponds to a grammatical property – non-referential null subjects (non-referential null subj), postverbal subjects (postverbal subj), and insensitivity to the *that*-trace constraint (no *that*-t). 84

3.1 Proportion of gap-less responses by condition in Experiment 3a . . . 120
3.2 Proportion of gap-less responses by condition in Experiment 3a . . . 124
3.3 Proportion of gap-less responses by condition in Experiment 3c . . . 127

3.4 Results from Experiment 4. Units are in milliseconds. Numbers in parentheses correspond to one standard error of the mean. Bolded values highlight a region with significant or marginally significant differences. 137
3.5 Results from Hebrew corpus study, coded by RK 146

List of Figures

2.1 Plot of DDs in Experiment 1 by latitude and longitude of participant. Color of each point corresponds to that participant's DD, or sensitivity to the *that*-trace effect. Black corresponds to strong *that*-trace effects, grey corresponds to insensitivity to *that*-trace effects, and red corresponds to preference for subject extraction over complementizers. 64

2.2 Plot of DDs in Experiment 2 by latitude and longitude of participant. Color of each point corresponds to that participant's DD, or sensitivity to the *that*-trace effect. Black corresponds to strong *that*-trace effects, grey corresponds to insensitivity to *that*-trace effects, and red corresponds to preference for subject extraction over complementizers. 65

2.3 Histogram of participants by DDs in Experiment 1, with density plot overlayed in red. Bin width is 0.1. Mean is 1.11 66

2.4 Differences in z-scored ratings in subject extraction conditions and object extraction conditions by participant in Experiment 1. 67

4.1 Division of participants in Experiment 5 into lists. 163

4.2 Sample image of a vignette with visual traces left in each critical location in Experiment 1. 166

4.3 Proportion main clause response by word order in Experiment 5. Error bars correspond to two standard errors of the mean. Proportions are collapsed across list types and locations. 167

4.4 Mean reading times by region in milliseconds for the main verb first conditions in Experiment 6. Line type corresponds to extraction type. Error bars represent two standard errors from the mean. 175

4.5 Mean reading times by region in milliseconds for the embedded verb first conditions in Experiment 6. Line type corresponds to extraction type. Error bars represent two standard errors from the mean. 176

4.6 Mean ratings of target items by condition in Experiment 7. Error bars represent two standard errors from the mean. 183

Chapter 1: Introduction

The primary goal of syntactic theory is to explain the properties of grammars, or the cognitive systems that assign structured representations to sentences that map sound to meaning. One component of this program is identifying which properties are invariant across languages, typically stated as constraints on possible structures. Another major component is identifying the range of variation between grammars. Traditionally, research in syntactic theory is conducted by examining patterns of acceptability judgments taken from native speakers of the language or languages under study, often native-speaking linguists. On the basis of these judgment patterns, syntacticians propose analyses of the representations that are built by the grammar.

The representations posited by the syntactician on the basis of these judgments are also the representations that children must learn to build in language acquisition, partially as a function of her input. Additionally, these representations must be constructed in real-time comprehension by the adult language user. However, data from language acquisition or sentence comprehension rarely inform syntacticians' theories of the syntactic representations themselves. This is because it is generally difficult to make falsifiable time-locked predictions solely on the basis of a

grammatical theory (Phillips & Wagers, 2007; Phillips & Parker, 2014). Behavioral evidence is typically only useful for arbitrating between different candidate analyses of syntactic representations with grammar-parser linking hypotheses in tow, which themselves make presumptions about the properties of grammatical representations. Similarly, it has been difficult to use language acquisition data to verify hypotheses about specific grammatical analyses, apart from determining whether a particular syntactic construction or constraint is likely innate on the basis of the age of onset (Lukyanenko *et al.*, 2014; Sutton, 2015). Additionally, interpreting children's behavioral data requires careful understanding of the learner's parsing capacities, which raises a number of independent complications (Omaki & Lidz, 2014). Thus, the relationship between syntactic theorizing and language acquisition and sentence processing has been largely unidirectional, in the sense that syntacticians typically define the range of possible syntactic representations that must be assumed in language acquisition or psycholinguistic research.

In this dissertation, I argue that learnability concerns and sentence processing data can be useful for delimiting the range of syntactic theories. I take an explicitly comparative approach to this problem. Dimensions of grammatical variation are particularly informative, because variables properties of languages must necessarily be learned. In other words, differences between individual languages cannot be "hard-wired" into the learner's initial state. Instead, the learner must infer these properties from her linguistic input. In this dissertation, I argue that comparative syntactic analyses can be interpreted as licensing conditions on inferences across constructions. If two constructions are abstractly "the same", i.e., instantiations

of the same abstract structure, then the learner can infer one construction on the basis of observing the other. This then means that specific analyses of language differences can be evaluated against realistic child-directed speech. Analyses that link phenomena for which learners have ample positive evidence to phenomena for which they have little or no positive evidence provide an elegant explanation for learners grow into adults that know both constructions.

In Chapter 2, I show that learnability considerations place strict boundaries on possible analyses of the differences of the *that*-trace effect and other complementizer-gap interactions observed in English, Spanish, and Italian. In Chapter 3, I make a similar argument for the syntactic characterization of resumptive pronouns inside islands in English and Hebrew. In Chapter 3, I also show that differences in behavior in real-time comprehension can similarly constrain analyses of syntactic differences. Since processing behaviors are not observable, differences in processing behaviors must derive from grammatical properties that are observable to the learner. In Chapter 3, I argue that differences in processing behavior in resumptive pronoun constructions in English and Hebrew can be explained by appealing to syntactic differences between the two languages, which I argue is more compatible with certain analyses of resumption phenomena than others. In Chapter 4, I show differences in the processing of filler-gap dependencies into embedded clauses in Japanese and Bangla. Unlike English and Hebrew, I argue that these findings do not necessarily imply appeals to different syntactic representations between the two languages. Instead, I suggest that these findings are due to differences in the time course of critical cues.

I discuss two well-known cases in which language acquisition data have been argued to inform syntactic theory here. The first case is the delay of Principle B effects. Chien & Wexler (1990) argue that some children deploy adult-like grammatical knowledge of the contrast in sentences like (1) and (2) at different ages. For adults, *him* may not corefer with the local subject *Dale* in (1), or be bound by the local subject *every cop* in (2). Chien & Wexler's (1990) studies show that some children accept the local binding relation in sentences like (1), but not in sentences like (2). Grodzinsky & Reinhart (1993) propose that this difference in onset times in language acquisition arises because these interpretations are ruled out by different constraints (Reinhart, 1983, 1986). This contrasts with analyses in which both constructions are ruled out by the same grammatical constraint, i.e., Chomsky's (1981) Principle B. Although these findings may not be reliable for methodological reasons (Elbourne, 2005; Conroy *et al.*, 2009), Grodzinsky & Reinhart's (1993) reasoning is particularly relevant to the argument in this dissertation. In Grodzinsky and Reinhart's argument, syntactic theories are valuable for identifying which grammatical properties "go together" as instances of the same abstract construction or constraint, and which grammatical properties are separable. Discovering dissociations in psycholinguistic data can then be used to argue for analyses that make the same dissociations.

(1) Dale$_j$ trusts him$_{i/*j}$

(2) Every cop$_j$ trusts him$_{i/*j}$]

Snyder (2001) argues that associations in the behavior of learners can motivate

associations in syntactic analyses. He shows that a number of compounding constructions and secondary predication constructions cross-linguistically correlate, and also have virtually identical onsets in language acquisition trajectories. He proposes that this is because they are ultimately instances of the same abstract construction. Once one of these constructions is learned, the other follows. This then explains why these properties correlate tightly crosslinguistically and are learned at the same time.

In this dissertation, I will largely abstract away from the behavior of actual learners. Instead, arguments about language acquisition will focus at a higher level of analysis. Instead of focussing on when children demonstrate knowledge of grammatical properties, which also depends on children's sentence processing capabilities in experimental settings, I instead focus on the informativity of the learner's linguistic environment. If comparative syntactic analyses are understood as statements of which properties can be learned simultaneously, they can be compared against realistic estimations of the learner's input. Analyses that more easily explain how learners infer linguistic properties for which they have no positive evidence should then be favored to analyses that do not.

In sentence comprehension, there are very few cases of truly divergent processing profiles for otherwise similar constructions. Instead, processing profiles are largely uniform across languages. However, one well-known cross-language difference is the difference in attachment preferences for ambiguous relative clauses. In both English and Spanish, relative clauses can be ambiguous between modifying the closer, embedded noun phrase, or a more distant, higher noun phrase. For instance,

in (3), the relative clause *that was standing on the balcony* can modify either *the maid* or *the actress*, referred to as high attachment resolution and low attachment resolution respectively. The same ambiguity holds in Spanish. However, Cuetos & Mitchell (1988) find that English-speakers prefer the low attachment resolution, i.e., *actress*, whereas Spanish speakers prefer the high attachment resolution, i.e., *criada* 'the maid'.

(3) Someone shot the maid of the actress [that was standing on the balcony]

(4) Alguien disparó contra la criada de la actriz [que estaba en el balcón]
Someone shot against the maid of the actress that was on the balcony
'Someone shot the maid of the actress that was standing on the balcony'

Languages seem to largely split in whether their preferences pattern like English or like Spanish (see Augurzky 2005), and preferences diverge even within the same language across constructions (Gibson *et al.*, 1996; de Vicenzi & Job, 1995; Fernández, 2003). Arguably, the most commonly accepted explanation for these cross-linguistic differences is to posit that, although the structures underlying these strings are largely the same, high attachment parses are more frequent in languages with high attachment preferences, and *vice versa* for low attachment languages. On this account, the parser then "fine tunes" its preferences in accordance with the differences in probabilities (Mitchell & Cuetos, 1991; Cuetos *et al.*, 1996). Other solutions have been proposed. For instance, Fodor (1998a,b) proposes that there are default intonation contours assigned to sentences in reading experiments, which in turn biases the preferred parse in ambiguity resolution.

More recently, Grillo & Costa (2014) argue that there is an ambiguity in the Spanish sentence in (4) compared to the English sentence in (3). Spanish has a construction called a "pseudo-relative", in which a clause headed by *que* 'that' may serve as a secondary predicate for an object of a sentence (Cinque, 1992). For instance, the sentence in (5) may be assigned the structure in (6), which is similar to a small clause interpretation in English. Importantly, Grillo & Costa (2014) argue that this is necessarily a "high attachment" for the clause. Grillo & Costa (2014) argue that universal attachment ambiguity resolution preferences will select the pseudorelative interpretation in languages where that parse is grammatically sanctioned, as in Spanish. When this representation is not available, as in English, then the relative clause is preferentially attached to the closer noun, as in English (Frazier, 1978).

(5) Vi al hijo del médico que corría
I saw to the son of the doctor that was running
'I saw the son of the doctor that was running'

(6) Vi al [$_{DP}$ [$_{NP}$ hijo del médico] [$_{CP}$ que corría]]

Regardless of the source of the difference in these preferences, each account attributes the parsing difference to some independently-motivated grammatical property of the language, while keeping properties of the comprehension mechanisms largely constant. Grillo & Costa's (2014) analysis presumes that comprehension mechanisms universally obey the same attachment ambiguity resolution preferences. The difference between English and Spanish on this account is that Spanish affords the ambiguous string another analysis, which is preferred on independent, universal

grounds. This is desirable, since it locates cross-language differences to the differences in the space of representations generated by the grammars of the two languages. Cross-language differences in possible grammatical constructions has some reflection on the learner's linguistic input, but cross-language differences in processing profiles do not. I use a similar form of argument for resumptive dependencies in Chapter 3, where I posit that resumptive dependencies are actively pursued in Hebrew because they are grammatical (Keshev & Meltzer-Asscher, ms), and therefore in the set of parses that the comprehender can construct actively, and that this is not the case in English.

These cases of the delay in Principle B and relative clause attachment ambiguity show that differences in grammars can explain differences in psycholinguistic data. Syntactic analyses which link specific phenomena together can explain how learners infer constructions for which they have no positive evidence. Additionally, differences in processing profiles can be used to argue that certain representations are not in the set of representations that the comprehension systems can construct in real time.

In the rest of this chapter, I discuss the relationship between comparative syntax and language acquisition in more detail, specifically focussing on the debate over Chomsky's (1981) Principles and Parameters model and its relation to indirect learning. Afterwards, I discuss the properties of filler-gap/movement dependencies, which much of this thesis focuses on.

1.1 Grammatical Variation and Language Acquisition

Regardless of the richness of the learner's initial state, properties of all languages cannot be innately prespecified, given that there are differences between languages. Comparative syntax can help define the limits of what grammatical properties absolutely *must* be learned. Properties that are observed to vary between languages, even superficial or trivial properties, reflect some kind of choice point that the learner must resolve on the basis of some data in her linguistic input. Importantly, not all of these choices are independent. Theories of comparative syntax can be useful in delineating what the relationships between these differences are, and how across properties given primary linguistic input might work, and what data are relevant for these inferences.

To illustrate, consider basic word order. Famously, languages are largely either head-initial or head-final (e.g., Greenberg 1963). Suppose that all languages were uniformly and transparently head-initial or head-final, i.e., this was the only necessary decision for determining a language's word order. If so, then the learner can detect her language's base word order by observing simple declarative transitive clauses, among many other possible constructions. For instance, an English-learner can detect that English is head-initial by observing that verbs precede objects in sentences like (7-a), and Bangla speakers can infer that Bangla is head-final by observing sentences like (7-b).

(7) a. Dale loves Annie.
b. Dale Annie-ke bhalobaše
Dale Annie-Acc loves

'Dale loves Annie'

However, as more variation in word order is uncovered, the space of possible grammars that the learner must consider increases. For the word order typology example, the set of (superficially) head-initial and head-final languages are separable into finer categories. Different analyses of these finer differences may link together different constructions, changing how the learner might leverage her linguistic input.

For instance, German is traditionally analyzed as underlyingly SOV, with obligatory V-to-C raising and raising of an additional phrase to the left edge on the clause. Thus, the sentence in (8-a) receives a substantially different analysis than its English counterpart in (7-a). For the linguist and for the learner, the cue that this representation for German is appropriate is clause-final positioning of untensed verbs (10), and "blocking" of V-to-C when the complementizer is overt in embedded contexts, revealing the base order, shown in (11) (den Besten, 1983; Travis, 1984, *inter alia*).

(8) a. Dale liebt Annie.
Dale loves Annie.
'Dale loves Annie'
b. [$_{\text{CP}}$ Dale$_i$ liebt$_j$ [$_{\text{TP}}$ t_i [$_{\text{VP}}$ liebt t_j]]]

(9) a. Dale loves Annie.
b. [$_{\text{CP}}$ [$_{\text{TP}}$ Dale [$_{\text{VP}}$ loves Annie]]]

(10) a. Dale wird Annie lieben
Dale will Annie love
'Dale will love Annie'
b. [$_{\text{CP}}$ Dale$_j$ wird$_i$ [TP t_j [$_{\text{VP}}$ t_i [$_{\text{VP}}$ Annie lieben]]]]

(11) a. Harry denkt, dass Dale Annie liebt.
Harry thinks that Dale Annie loves
'Harry thinks that Dale loves Annie'
b. Harry denkt, [$_{\text{CP}}$ dass [$_{\text{TP}}$ Dale [$_{\text{VP}}$ Annie liebt]]]

This comparison between English and German demonstrates two things. First, it shows that as the hypothesis space that the learner must navigate becomes more fine-grained, so too do the particular observations that the learner must make in order to select the right analysis. The learner can no longer simply observe that verbs precede objects to infer that the base word order is SVO, for instance, since this word order is also compatible with an underlying SOV word order with obligatory V-raising, as in German. Instead, the learner must rely on other evidence to distinguish these hypothesis. Secondly, this example demonstrates that multiple constructions may count as "successes" for selecting one grammar over the other. In German, head-finality for untensed and embedded verbs are both cues that the appropriate analysis for apparent head-initiality in simple transitive clauses like (8-a) is due to V-raising. Thus, if the two properties in (10) and (11) are both cues for V-raising, then the learner needs only to observe one cue to select a V-raising grammar, and the related constructions should follow[1]. If the analysis of German is incorrect, then other data may be crucial for the learner. Thus, comparative syntactic analyses make commitments to the underlying representational differences between languages, which in turn make different commitments to inferences across related grammatical properties.

[1]See Gibson & Wexler (1992); Fodor (1998c); Yang (2002); Sakas & Fodor (2012) for discussion on how particular learning mechanisms fare on learning basic word orders with a realistic hypothesis space.

To put this more concretely, if two or more properties are linked by an abstract representation, then the learner can observe one of these properties, infer the more abstract representation, and thereby infer the rest of the linked properties. I call this "indirect learning", and indirect learning is the focus of Chapters 2 and 3.

This way of conceptualizing language variation as reflecting the space of hypotheses the learner considers has a long history (i.e., Jakobson 1941). However, it enjoyed the most attention in Chomsky's (1981) "Principles and Parameters" model, in which the dimensions of grammatical variation were explicitly modeled as parameters, or binary options prespecified in Universal Grammar, that the learner must set. Parameter-setting approaches to language have recently fallen out of favor for a variety of reasons, due to the scarcity of robust typological generalizations across languages (Newmeyer, 2004, 2005), the popularity of "micro-comparative" approaches to comparative syntax (Kayne, 2005), theory-internal considerations in Minimalist syntax (Uriagereka, 2007; Boeckx, 2010), and the shift towards functionalist approaches to typology (Hawkins, 1999, 2004; Haspelmath, 2008). Meanwhile, more sophisticated statistical approaches to language acquisition are argued to reduce the reliance on prespecified domain-specific knowledge (Chater & Manning 2006; Perfors *et al.* 2011, but see Yang 2004; Gagliardi & Lidz 2015). Thus, there appears to be a tacit consensus that language variation is less structured than initially thought, and that a link between language variation and learning mechanisms is less crucial for explaining the properties of either.

However, it is surely premature to dismiss parametric accounts, given that there are few successful, concrete demonstrations of how the learner acquires difficult-

to-observe phenomena, i.e., phenomena that have few or particularly subtle cues in the primary linguistic input. Additionally, many difficult-to-observe phenomena vary across languages, and therefore necessarily must be inferred from the learner's primary linguistic input. Indirect learning is the only viable explanation for how difficult-to-observe phenomena are learned, as envisioned in parametric models[2]. Additionally, evaluating comparative syntactic analyses against realistic language input is informative for the theorist to determine which properties of an analysis are most relevant for achieving explanatory adequacy.

Although learning across constructions has been addressed in language acquisition (Hyams, 1989; Snyder, 2001; Culbertson, 2010), these studies do not explicitly investigate links between difficult-to-observe and easy-to-observe properties, which is arguably the kind of indirect learning that carries the most explanatory power. For instance, Snyder's (2001) analysis of predication and noun-noun compounding may very well be correct. However, it is unclear whether either of the relevant constructions are particularly rare in the learner's input. Thus, the learner may not need these constructions to be abstractly linked. Conversely, focussing on difficult-to-observe properties is arguably a more useful strategy for discovering what kinds of inferences must be available to the learner.

[2]It is important to note that non-parametric frameworks can model this kind of learning as well, as acknowledged by Newmeyer (2004). In this chapter, my focus is on the necessity of indirect learning as a learning strategy as a consequence of a deductive learning structure, not the precise formal distinctions between parameters and rule lists. These issues are independent.

1.2 Syntax and Psycholinguistics of Filler-Gap Dependencies

This dissertation largely focusses on the syntax and psycholinguistics of filler-gap dependencies. This is because their syntactic properties and processing profiles are well-described, and because many of these properties have been argued to be innate. A filler-gap dependency is the relation between a phrase that occurs in a position distinct from where it is interpreted. In order to correctly assign an interpretation to (12), for instance, *who* must be understood as the object of the verb *saw*, even though *who* appears in a different clause, and there is no surface object for the verb *saw*, indicated by the gap (__). In syntactic theory, these dependencies are often analyzed as movement dependencies or A′-dependencies, and the gaps are sometimes formalized traces or copies (Fiengo, 1977; Chomsky, 1981, 1995), but I will use the theoretically neutral language of filler-gap dependencies.

(12) **Who** did Dale say [$_{\text{CP}}$ that Sarah saw __] ?

One of the more striking facts about filler-gap dependencies is that they are subject to locality biases. That is, filler-gap dependencies are preferentially resolved in positions that are close to the filler. Additionally, there are locality constraints on possible and impossible filler-gap dependencies, such that filler-gap dependencies that are not locally resolved are perceived as unacceptable. Defining these different locality biases on filler-gap dependencies and potentially unifying them has received a great deal of attention in the psycholinguistic literature. In this section, I describe the relevant notions of locality, and how they relate. First, I describe active depen-

dency formation, or the psycholinguistic processes by which filler-gap dependencies are anticipatorily resolved in real-time sentence comprehension. Afterwards, I describe the notions of locality that constrain active dependency formation and also influence the acceptability of non-local filler-gap dependencies.

1.2.1 Active Dependency Formation

Many studies reveal that the parser actively constructs filler-gap dependencies. This is called active dependency formation. (Crain & Fodor, 1985; Stowe, 1986; Frazier & Clifton Jr., 1989; Garnsey *et al.*, 1989; Nicol & Swinney, 1989; Bourdages, 1992; Nicol *et al.*, 1994; Traxler & Pickering, 1996; Kaan *et al.*, 2000; Felser *et al.*, 2003; Sussman & Sedivy, 2003; Phillips *et al.*, 2005). The evidence for active dependency formation comes from a variety of experimental paradigms performed in many languages and constructions. The most well-known illustration is the filled-gap effect. The filled-gap effect is an increase in processing difficulty when a predicted resolution site is disconfirmed by later linguistic material (Crain & Fodor, 1985; Stowe, 1986). In a self-paced reading task, Stowe (1986) found increased reading times in sentences like (13-a) at the pronoun *us* compared to control conditions, like (13-b). She argues that this indicates that the parser initially attempts to interpret *who* as the object of the verb *bring*. Upon encountering *us*, this analysis must be rejected, because this initial gap commitment is no longer compatible with the sentence. Thus, there is a penalty for revision, and the parser must continue searching for a later resolution site. Crucially, this indicates that the parser does not wait for

unambiguous resolution sites, but instead attempts resolution before encountering a possible resolution site.

(13) a. My brother wanted to know **who** Ruth would bring us home to __ at Christmas.
b. My brother wanted to know if Ruth would bring us home to Mom at Christmas.

Another demonstration of active dependency formation is the plausibility mismatch effect, or an increase in processing difficulty when the early gap commitment yields an implausible interpretation. For instance, in an eye-tracking study, Traxler & Pickering (1996) found increased fixation times at the verb *wrote* in (14-a) compared to the control in (14-b). These results show that the parser first attempts to resolve the relativization dependency with the verb *write*, which yields an implausible interpretation in (14-a), since cities are not typically the object of writing events. At this point, the parser can revise its prediction, again incurring a cost.

(14) a. We like **the city that** the author wrote unceasingly and with great dedication about __ while waiting for a contract.
b. We like **the book that** the author wrote unceasingly and with great dedication about __ while waiting for a contract.

Stowe's (1986) and Traxler & Pickering's (1996) findings establish that the parser privileges resolving a filler-gap dependency early over waiting for an unambiguous resolution site. These results do not suggest the mechanism by which filler-gap dependencies are preferentially resolved early, however. For instance, active dependency formation processes may be a special instance of a set of general strategies that attempt to satisfy the greatest number of grammatical constraints

moment-by-moment (de Vincenzi, 1991; Pritchett, 1991; Aoshima *et al.*, 2004). On this view, the parser attempts to resolve the filler as the object of the verb in (13) and (14) because this is optimal with respect to the goals of the parser. On this view, the position in which the comprehension system attempts to resolve the filler-gap dependency will vary, depending on which position satisfies the most relevant criteria. Another hypothesis is that the parser must actively maintain features of the filler in working memory (Wanner & Maratsos, 1978; Wagers & Phillips, 2009), which may be costly (Just & Carpenter, 1992; Gibson, 1998). If so, then the parser is predicted to preferentially resolve the dependency with the first position temporally available, regardless of structural position. Thus, understanding the processes underlying active dependency formation partially depends on defining the relevant notions of locality, which I discuss in the next section.

1.2.2 Locality

Filler-gap dependencies have a number of properties that distinguish them from other syntactic dependencies. One such property is their sensitivity to locality constraints. Filler-gap dependencies are perceived as acceptable over an arbitrarily long distance, as demonstrated in (15).

(15) **Who** did Lucy say that Andy knew …that Dale suspected __ ?

However, there are certain configurations that a filler-gap dependency cannot cross. These configurations are called "islands" in the theoretical syntax literature (Ross,

1967; Chomsky, 1977). A number of island phenomena are exemplified in (16).

(16) a. *__Who__ did Dale comfort the woman [$_{CP}$ that saw __]?
b. *__Who__ did Dale hear the rumor that [$_{CP}$ Leo scared __]?
c. *__Who__ did Dale doubt [$_{DP}$ Lucy's rumor about __]?
d. *__Who__ did Dale wonder [$_{CP}$ whether Bob frightened __]?
e. *__Who__ does Dale think [$_{CP}$ who saw __ behind Laura's bed]?
f. *__Who__ did the fact that [$_{CP}$ Sarah saw __] surprise Dale?
g. *__What__ did Dale ruminate [$_{CP}$ while Harry examined __]?
h. *__Who__ did [Dale suspect __ and Harry interrogate Leland]?
i. *__Why__ did Dale remember that Ben was suspicious __ ?
j. *__Why__ did Dale say that nobody was innocent __ ?

The precise formulation of island constructions has been a major focus of research in generative syntax (e.g., Chomsky 1977, 1981, 1986; Huang 1982; Lasnik & Saito 1992). The inventory of island constructions is largely uniform across languages. Additionally, the constructions into which filler-gap dependencies may not resolve are complex, typically multiclausal, and thus likely rare in the learner's input. It is unclear how the learner might learn these restrictions on the basis of their input. For these reasons, island constraints have been argued to be evidence for innately prespecified linguistic knowledge (see Pearl & Sprouse 2013 and Phillips 2013 for discussion).

In real-time processing, many studies show that the comprehension system does not actively construct filler-gap dependencies in island configurations. For instance, Stowe (1986) found no increased reading times at *Greg's* in sentences like (17-a), compared to a control like (17). Importantly, *Greg's* occurs in an island configuration, since extraction out of this noun phrase is unacceptable, as shown in (18). This suggests that the comprehender does not attempt to construct the

filler-gap dependency headed by *what* in this position, otherwise there would be a filled-gap effect. Thus, it appears that the comprehender does not actively construct the filler-gap dependency here. Similarly, Traxler & Pickering (1996) find no difference in reading times at *wrote* regardless of whether the filler was a semantically/pragmatically plausible object of the verb. Again, this verb occurs inside an island, and thus the lack of a plausibility mismatch effect is consistent with the claim that filler-gap dependencies are not actively constructed inside islands.

(17) a. The teacher asked **what** [$_{\text{island}}$ the silly story about Greg's older brother] was supposed to mean __ .

We like the city that [$_{\text{island}}$ the author who wrote unceasingly and with great dedication] saw __ while waiting for the contract.

(18) *The teacher asked **what** [$_{\text{island}}$ the silly story about __] was supposed to mean something.

One interpretation of these findings is that dependencies in islands are blocked by grammatical constraints, and that the comprehender can rapidly use this information to filter out ungrammatical representations (Phillips, 2006; Wagers & Phillips, 2009). In other words, the comprehender's preference to resolve a filler-gap dependency early is suppressed in island contexts. On this view, the locality constraints that define where filler-gap dependencies can resolve grammatically are not necessarily the same as the locality constraints that guide where the comprehender attempts to construct filler-gap dependencies in real time. In other words, there are two distinct notions of locality – one relevant to defining island configurations, and one relevant to defining how active dependency formation processes work outside

island contexts. The alternative interpretation of these data holds that filler-gap dependencies inside island contexts may be blocked because they are particularly costly or difficult to construct (Pritchett, 1991; Kluender & Kutas, 1993; Kluender, 1998, 2005; Hawkins, 1999; Hofmeister & Sag, 2010). I call these the resource-based accounts of island phenomena. On this account, the locality constraints relevant for describing the island phenomena as in (16) are the same constraints that guide the parser's preferred resolution sites more generally. Thus, this account requires that the notions of locality make reference to the same properties of the representations.

There is evidence that suggests the resource-based accounts of island phenomena are insufficient for explaining these data. For instance, Phillips (2006) and Wagers & Phillips (2009) show that the comprehension systems can construct filler-gap dependencies in syntactic contexts that are ordinarily islands, if there is a possible continuation of the sentence that can extraordinarily license the gap, i.e., parasitic gap and across-the-board constructions respectively. Similarly, Yoshida *et al.* (2014) show that cataphoric dependencies, which are actively constructed much like filler-gap dependencies (Kazanina *et al.*, 2007), are not suppressed in island contexts. Thus, it is unlikely that active dependency formation processes are suppressed because the comprehension system is incapable of constructing dependencies in these configurations. Lastly, Sprouse *et al.* (2012a,b) shows that working memory capacity does not correlate with the strength of island violations, which they argue is predicted on resource based accounts.

Additionally, I argue that the notions of locality that are relevant in defining island constraints and active dependency formation are sensitive to different

aspects of the linguistic representation. Importantly, islands are defined in terms of "structural locality", e.g., island constraints are defined over hierarchical properties of the syntactic representation, rather than linear or temporal distance (Rizzi, 2013). This can be easily seen in comparing (19) to (20). The sentence in (19-a) is perceived as ungrammatical, even though a synonymous sentence with a longer filler-gap dependency is perceived as acceptable, shown in (20-a).

(19) a. *__Who__ did [the fact that Sarah saw __] surprise Dale?
b. __Who__ did [the fact that Sarah saw Bob] surprise __ ?

(20) a. __Who__ did it surprise Dale [that Sarah saw __] ?
b. __Who__ did it surprise __ [that Sarah saw Bob]?

Although island constraints are defined in terms of hierarchical structure, the locality biases that guide where the comprehender attempts to resolve filler-gap dependencies outside islands are sensitive to linear order. I call this "linear locality". This can be seen most clearly in Japanese, a head-final language in which embedded clauses always precede the embedding verb, as shown in (21-a). Japanese allows filler-gap dependencies via a scrambling operation in which phrases are moved from their canonical position for emphasis, focus, or stylistic reasons (Saito, 1985). This is illustrated with a clause-internal scrambling in (21-b) and a multiclausal scrambling (21-c).

(21) a. Dale-ga Annie-ni [$_{CP}$ Theresa-ga Laura-ni ringu-o
Dale-NOM Annie-DAT Theresa-NOM Laura-DAT ring-ACC
ageta-to] itta
gave-that said
'Dale told Annie that Theresa gave Laura the ring'

b. **Annie-ni** Dale-ga ___ [$_{CP}$ Theresa-ga Laura-ni ringu-o
Annie-Dat Dale-NOM ___ Theresa-NOM Laura-DAT ring-ACC
ageta-to] itta
gave-that said
'Dale told Annie that Theresa gave Laura the ring'

c. **Laura-ni** Dale-ga Annie-ni [$_{CP}$ Theresa-ga ___ ringu-o
Laura-Dat Dale-NOM Annie-DAT Theresa-NOM ___ ring-ACC
ageta-to] itta
gave-that said
'Dale told Annie that Theresa gave Laura the ring'

If readers preferentially resolve filler-gap dependencies with the first verb linearly available, then there should be an embedded-clause preference in Japanese. This is because the first verb in multiclausal sentences is the most deeply embedded verb. Conversely, if filler-gap dependencies preferentially resolve with the closest structural position, then there should be a main clause resolution preference, even though the main clause is the linearly most distant verb.

Results from studies in Japanese reveal that a linearly locality bias guides filler-gap dependency resolution preferences. Aoshima *et al.* (2004) found increased reading times at the embedded dative NP *titioya-ni* 'father-DAT' in (22-a) compared to (22-b). They interpreted this as a filled gap effect – i.e., the parser first attempts to construe *dono kodomo-ni* 'which child-DAT' as an argument of the embedded verb, and then is forced to reanalyze upon encountering *titioya-ni* 'father-DAT'[3]. This implies that the comprehension system attempts resolution with the first position linearly available, regardless of depth of embedding, which is most relevant

[3]Aoshima *et al.* (2004) also argue that this filled-gap effect is evidence that the parser resolves filler-gap dependencies with the first *verb* available, and not the first gap position available. This is because the canonical position for dative arguments precedes sentential complements, as indicated in (21-b).

to structural locality. Thus, the notion of locality that drives active dependency formation appears to be distinct from the notions of locality that are relevant for describing island constructions. It therefore appears difficult to reduce island constraints to those notions of locality that guide where the comprehender preferentially resolves filler-gap dependencies.

(22) a. **Dono kodomo-ni** hahaoya-wa ___ [$_{CP}$ otetudaisan-ga
which child-Dat mother-Top ___ housekeeper-Nom
daidokoro-de titioya-ni obento-o watasita-to] iimasita-ka?
kitchen-at father-Dat lunch-Acc handed-that said-Q?
'Which child did the mother tell ___ [$_{CP}$ that the housekeeper handed over lunch to father in the kitchen?]'

b. Dono kodomo-ga hahaoya-ni [$_{CP}$ otetudaisan-ga daidokoro-de
which child-Dat mother-Top housekeeper-Nom kitchen-at
titioya-ni obento-o watasita-to] iimasita-ka?
father-Dat lunch-Acc handed-that said-Q?
'Which child told mother [$_{CP}$ that the housekeeper handed over lunch to father in the kitchen?]'

This implies an a cross-language difference in filler-gap dependency resolution sites, as pointed out by Omaki *et al.* (2014). Head-initial languages should show a global bias for main clause resolution in multiclausal sentences, since the main verb is also the first verb in the string. Omaki *et al.* (2014) confirmed this cross-language prediction using a Question After Story task (de Villiers *et al.*, 1990). In their studies, English- and Japanese-speaking participants were exposed to globally ambiguous filler-gap dependencies, as in (23).

(23) a. **Where** did Emily tell someone [$_{CP}$ that she hurt herself] ?
b. **Doko-de** Emily-chan-wa [$_{CP}$ asi-o kegasita-to] itteta-no.
Where-at Emily-Dim-Top foot-Acc hurt-that was saying-Q
'Where was Emily saying that she hurt her foot?'

In this sentence, both *where* and the scrambled locative *doko-de* 'where' may be interpreted as a modifier of either the main or embedded verb. Omaki *et al* show that both adults and children answer with the location of the event denoted by the first verb, regardless of depth of embedding or language. Importantly, this shows that the parser's bias towards resolution with the first verb is robust across languages, and that this bias persists even when both potential resolutions are revealed to be available. Additionally, these data provide further support for the claim that the comprehender prefers actively constructing filler-gap dependencies in the first position linearly available, in accordance with a linear locality bias. This contrasts with the structural notion of locality which is relevant for defining island constraints.

In Chapter 4, I replicate Omaki *et al.*'s (2014) findings in Bangla, a language in which embedded clauses may appear on either side of the embedding verb. This permits a within-language replication of Omaki *et al*'s studies. In Bangla, filler-gap dependencies preferentially resolve with the first verb, regardless of its depth of embedding. In follow-up studies, however, I show that the filled-gap effect does not necessarily index these preferences. I suggest that the reason for this difference is rooted in the mechanisms that underlie detecting a filled gap. Thus, I argue that there is a universal bias to resolve filler-gap dependencies with the first position linearly available, although the specific processing profiles used to probe for this bias may differ.

1.3 Roadmap

In this chapter, I laid out the way that learnability concerns in language acquisition and real-time adult comprehension data interact with syntactic analyses of language differences. Specifically, I claimed that analyses of cross-language differences can be understood as defining which variable properties are linked, and which are dissociated. I argued that this provides a metric for using learnability concerns to help decide between analyses. Namely, analyses that link phenomena that are easily observable in the learner's input to phenomena that are unobservable should be preferred. Additionally, I argued that active dependency formation processes are sensitive to multiple locality constraints. Specifically, filler-gap dependencies are preferentially resolved in the first position linearly available, but only if this is a grammatical resolution site. I later argue that this can be used to diagnose what positions are grammatical resolution sites for filler-gap dependencies.

In Chapter 2, I discuss the relationship between syntactic analyses of complementizer-gap interactions in Spanish, Italian, and English and realistic linguistic input. I show that the learner's input does not support a learning strategy in which she must learn on a complementizer-by-complementizer basis whether that complementizer permits subject extraction. Instead, I argue for older analyses which link complementizer-gap interactions to other constructions, such as Rizzi's (1982).

In Chapter 3, I discuss the relationship between syntactic analyses of resumptive pronouns in English and Hebrew and the mechanisms by which resumptive dependencies are formed in these languages. I show that resumptive dependencies

in English are not actively constructed inside islands, consistent with the observation that filler-gap dependency resolution is suppressed inside islands. This contrasts with findings from Hebrew, in which resumptive dependencies are actively constructed (Keshev & Meltzer-Asscher, ms). I explain this difference by positing that resumptive pronouns inside islands are grammatical in Hebrew, and thus actively constructed in real-time. Resumptive dependencies are ungrammatical in English, and therefore not actively constructed. I argue that this contrast is likely not directly learnable. Instead, learners must rely on indirect learning to determine whether resumptive dependencies are grammatical, and that particular analyses are more suitable for this indirect learning than others.

In Chapter 4, I present the results from experiments conducted in Bangla, a language with particularly flexible word order. This property allows for a within-language investigation of the role that linear and structural locality play in the resolution preferences of filler-gap dependencies. The findings from these studies show that filler-gap dependencies are preferentially resolved with the first position linearly available. However, I show that the filled-gap effect is not observed in preverbal clauses in Bangla, unlike Japanese. I suggest that this is more indicative of the mechanisms used to detect filled gaps, rather than any representational difference between the two languages.

Finally, I conclude the dissertation in Chapter 5.

Chapter 2: Testing Learnability Commitments: The *That*-Trace Effect in English, Spanish, and Italian

2.1 Introduction

In this chapter, I show how learnability concerns and analyses of child-directed speech can be used to constrain the space of possible comparative syntactic analyses. As described in Chapter 1, comparative syntactic analyses can be interpreted as statements about which phenomena are "connected" by abstract structure. If two constructions α and β are linked via the same abstract structure S, then by virtue of observing α the learner can infer S, thereby inferring β. This is what I called indirect learning in Chapter 1. Importantly, I argue that this is the most plausibly way by which the learner can infer constructions for which she lacks positive evidence.

Viewed this way, different syntactic analyses can be interpreted as different licensing conditions on indirect learning. This can then be evaluated against realistic linguistic input, to determine which analyses better explain how the learner infers linguistic properties for which there is no positive evidence in the linguistic environment. In this chapter, I focus on the *that*-trace constraint. I argue that indirect learning is necessary for learners exposed to Spanish and Italian linguistic input to

infer the (superficial) absence of the *that*-trace constraint in these languages. This is more in-line with traditional analyses of the variability of the *that*-trace effect (Perlmutter, 1971; Rizzi, 1982) than later formulations (Rizzi & Shlonsky, 2007; Pesetsky & Torrego, 2001).

First, in section 2.2, I describe the range of variation that the learner must navigate. Specifically, I compare the interactions between complementizers and gap positions in English, Spanish/Italian, and French in complementation and relativization. I also briefly survey the range of analyses proposed to capture these data. I argue that a modern analyses, such as Rizzi & Shlonsky's (2007), are compatible with a view in which the learner must decide for each complementizer in her language whether it admits subject extraction. In section 2.3, I present data from two acceptability judgment tasks performed in English and Spanish which demonstrate that the cross-language differences attested in section 2.2 are robust. After establishing that the cross-language variation is robust, I present the results from corpus studies in Spanish and Italian in section 2.4. In conjunction with Pearl & Sprouse's (2013) English corpus study, I argue that the distribution of complementizers and subject gaps domain-specific not provide a basis for the learner to infer the cross-language contrast confirmed in 2.2. I therefore conclude that some complementizer-gap interactions, specifically the superficial inapplicability of the *that*-trace effect in Spanish/Italian, must be learned indirectly, in the sense described in Chapter 1.

2.2 Dimension of Variation: Complementizer-Gap Interactions

In this section, I present the three contrasts that I focus on in this chapter. I argue that the accounts that have been proposed either are overly restrictive and therefore cannot capture the full range of data, or are too permissive and therefore cannot explain how the learner select her target grammar given her primary linguistic input, as described in section 2.4.

Descriptively, research in comparative syntax reveals that different languages exhibit different constraints on complementizer-gap combinations across constructions and complementizers. These constraints are sensitive to construction type (i.e., relativization complementizers vs. sentential complementation complementizers), are sensitive to morphophonological properties of the complementizer, and gap position (subject vs. object extractions). Traditional analyses on these phenomena have typically focussed on variation in the acceptability of subject extraction over a sentential complementation complementizer, i.e. the *that*-trace effect (Perlmutter, 1971; Chomsky & Lasnik, 1977; Chomsky, 1981; Rizzi, 1982, 1986). However, other evidence, specifically the so-called "anti-*that*-trace" effect (Pesetsky, ms) and *que/qui* alternation in French (Perlmutter, 1971; Kayne, 1981; Rizzi, 2006; Rizzi & Shlonsky, 2007) suggest that these earlier analyses must be amended to permit specific complementizers in specific constructions to extraordinarily allow subject extraction. I argue that once this move is made, whatever mechanisms are used to inhibit or permit subject extractions in these cases can be extended to cover account for the full range of data. If so, this then obviates the need for traditional analyses

of the variability of the *that*-trace constraint. I walk through the relevant data, then discus how they have been analyzed. Afterwards, I discuss how these analyses relate to the broader question of how the learner draws the right inferences for her language on the basis of her linguistic input.

The sentences in (1) show that English permits long-distance *wh*-movement of subjects or objects over a null (or absent) complementizer, represented here as $\varnothing_{comp}$. When the complementizer is overt (e.g., *that*), subject extraction is not permitted, shown in (1-d), whereas object extraction over an overt complementizer is acceptable, shown in (1-c), as well as subject extraction over a null/absent complementizer, shown in (1-b). The acceptability of subject extraction over the overt complementizer *that* is called the *that*-trace effect (Chomsky & Lasnik, 1977; Chomsky, 1981; Rizzi, 1982).

(1) a. **Who** did Dale say $\varnothing_{comp}$ Sarah saw __ ?
b. **Who** did Dale say $\varnothing_{comp}$ __ saw Bob?
c. **Who** did Dale say that$_{comp}$ Sarah saw __ ?
d. ***Who** did Dale say that$_{comp}$ __ saw Bob?

This constraint applies to a variety of complementizers. For this reason, the *that*-trace constraint is sometimes called the "comp-trace effect", since it is general to complementizers. For instance, object extraction over the complementizer *for* is acceptable, but not subject extraction (Bresnan, 1977). Thus, the constraint that applies in (1) may not be specific to the complementizer *that*.

(2) a. **Who** would Dale prefer for Mike to find __ ?

b. ***Who** would Dale prefer for __ to find Bob?

There are also constraints on complementizer-gap collocations in relativizations. Local subject and object relativization over an overt relativization complementizer (represented here as $that_{rel}$) are grammatical, whereas local subject relativization over a null relativization complementizer (represented as $\varnothing_{rel}$) is not. I call this the anti-*that*-trace effect, following Pesetsky (ms). Importantly, this shows that the constraint underlying the *that*-trace effect in (1) either must be relativized to complementation alone, or must be somehow evaded in relativization constructions. Additionally, this establishes that it is possible for phonologically identical complementizers ($that_{comp}/that_{rel}$ and $\varnothing_{comp}/\varnothing_{rel}$) to obey different constraints. This implies that the learner cannot generalize from complementation complementizers to relativization complementizers, nor generalize between morphophonologically identical complementizers.

(3) a. **The ring** that$_{rel}$ Annie wore __ has gone missing
b. **The ring** that$_{rel}$ __ frightened Dale has gone missing
c. **The ring** $\varnothing_{rel}$ Annie wore __ has gone missing
d. *__The ring__ $\varnothing_{rel}$ frightened Dale has gone missing

Importantly, this is not a distinction between dependency types (i.e., relativization vs. *wh*-movement), but rather complementizer type (relativization vs. complementation). For instance, long-distance subject relativization over a sentential complementation complementizer is ungrammatical, shown in (4). This is an instance of the *that*-trace constraint we see above with long-distance *wh*-movement in (1), and unlike the pattern with local relativization in (3).

(4) a. **The ring** that everyone knows $\varnothing_{comp}$ __ frightened Dale has gone missing.

b. *__The ring__ that everyone knows that __ frightened Dale has gone missing.

To sufficiently describe these data, syntactician's analytic tools must be able to state constraints that apply to extraction site, complementizer choice, and complementizer type. This has not been trivial to do satisfactorily (see Kayne 1981; Rizzi 2006; Rizzi & Shlonsky 2007). Additionally, syntactic theory must be capable of explaining why these constraints occur at all, which constraints are universal, and what the relation between these constraints are. Insofar as these constraints are universal, they can plausibly be "hard-wired" into the learner. If so, then learning is not necessary, meaning that the learner's input may not be relevant. However, findings from Romance suggest that both the *that*-trace effect and the anti-*that*-trace effect are not universal. Thus, both the *that*-trace effect and the anti-*that*-trace effect must be learned.

First, I describe facts in Spanish. These facts are most well-studied in Italian (Rizzi, 1982, 1986), but Spanish is the same in the relevant respects (Torrego, 1984). In Spanish, complementizers are obligatorily overt in both complementation and relativization constructions:

(5) a. Dale dijo que Sarah vio a Bob
Dale said that Sarah saw to Bob
b. *Dale dijo $\varnothing_{comp}$ Sarah vio a Bob
Dale said Sarah saw to Bob
'Dale said that Sarah saw Bob'

(6) a. El anillo que llevaba Annie está perdido.
the ring that wore Annie is lost.
b. *El anillo $\varnothing_{rel}$ llevaba Annie está perdido.
the ring wore Annie is lost
'**The ring** that Annie wore is lost'

Extraction of subjects is licensed over both complementation complementizers, shown in (7), and relativization complementizers, shown in (8). These findings imply that the *that*-trace constraint and the anti-*that*-trace constraint either must be language-specific, or must be somehow avoided in Spanish. Either way, the learner must be capable of inferring the right complementizer-gap interactions for her language from the relevant child-directed speech.

(7) a. ¿A quién dijo Dale que vio Sarah?
to whom said Dale that saw Sarah
'**Who** did Dale say that Sarah saw __ '
b. ¿Quién dijo Dale que vio a Bob?
who said Dale that saw to Bob
'**Who** did Dale say (that) __ saw Bob?'

(8) a. El anillo que llevaba Annie se ha perdido
the ring that wore Annie self has lost
'**The ring** that Annie wore __ is lost'
b. El anillo que le asustó a Dale se ha perdido
the ring that him scared to Dale self has lost
'**The ring** that __ scared Dale is lost'

Next, I describe the constraints on complementizer-subject gap extractions in French. Although I do not focus on French in this chapter, these data are relevant because they motivate a popular amendment to analyses of the *that*-trace effect. Namely, they motivate the appeal to complementizer-specific rules, which potentially changes the analytic strategies for the English and Spanish facts. First, like in Spanish, French complementizers are obligatorily overt, as shown in (9) and (10).

(9) a. Dale a dit que Sarah a vu Bob
Dale has said that Sarah has seen Bob
b. *Dale a dit Sarah a vu Bob
Dale has said Sarah has seen Bob

'Dale said that Sarah saw Bob'

(10) a. La bague qu' Annie portait
the ring that Annie wore
b. *La bague Annie portait
the ring Annie wore
'**The** ring that Annie was wearing __ '

However, unlike Spanish or English, the form of the complementizer changes for subject extractions, with either complementation complementizers or relativization complementizers. The complementizer *que* may introduce a complement clause that contains a gap site in any position apart from the subject position. For instance, the object *wh*-extraction in (11-a) is acceptable over the complementizer *que*, but not the subject extraction in (11-b):

(11) a. Qui Dale a dit que Sarah a vu?
who Dale has said that Sarah has seen?
'**Who** did Dale say that Sarah saw __ ?'

b. *Qui Dale a dit que/qu' a vu Bob?
who Dale has said that has seen Bob?
'**Who** did Dale say that __ saw Bob?'

Unlike *que*, the complementizer *qui* only permits subject extraction:

(12) a. *Qui Dale a dit qui Sarah a vu?
who Dale has said that Sarah has seen?
'**Who** did Dale say that Sarah has seen __ ?'
b. Qui Dale a dit qui a vu Bob?
Who Dale has said that has seen Bob?
'**Who** did Dale say (that) __ saw Bob?'

Thus, the learner exposed to French input must infer that *qui* alone licenses

subject extraction, across both complementation complementizers and relativization complementizers. Conversely, *que* is reserved for non-subject extractions as an "elsewhere" case, as shown in (13).

(13) a. La bague qu' Annie portait est perdue.
the ring that Annie wore is lost
'**The ring** that Annie wore __ is lost'

b. *La bague que/qu' a fait peur à Dale est perdue
the ring that has made fear to Dale is lost
'**The ring** that __ scared Dale is lost'

c. *La bague qui Annie portait est perdue
the ring that Annie wore is lost
'**The ring** that Annie wore __ is lost'

d. La bague qui a fait peur à Dale est perdue
the ring that has made fear to Dale is lost
'**The ring** that scared Dale is lost'

The patterns described above are summarized in Table 2.1. As mentioned earlier, a theory of grammar must minimally be able to describe this array of data, and ideally must also show how a learner can infer the correct array of complementizer-gap interactions when exposed to realistic primary linguistic data.

In the rest of this section, I turn to specific analyses of the patterns above. To a large extent, these analyses have focussed on the *that*-trace effect and its apparent inapplicability in languages like Spanish/Italian[1]. Most accounts attribute the lack of surface realization of the *that*-trace effect to some independent property of those languages, i.e., this inapplicability of the *that*-trace effect in Spanish is indirectly

[1]Spanish and Italian are identical in the relevant respects (Torrego, 1984). Most of the analyses that I discuss in this section use Italian data, but I use Spanish data in this chapter.

English

	that		∅	
	Subject	Non-Subject	Subject	Non-Subject
WH	*	✓	✓	✓
RC	✓	✓	*	✓

French

	que		*qui*	
	Subject	Non-Subject	Subject	Non-Subject
WH	*	✓	✓	*
RC	*	✓	✓	*

Spanish/Italian

	que/che	
	Subject	Non-Subject
WH	✓	✓
RC	✓	✓

Table 2.1: Patterns of complementizer-gap interactions in English, Spanish, and French.

learned. However, I argue that it is unclear how these analyses can extend to account for the *que/qui* contrast or the relativization/complementation construction contrast without committing the learner to track complementizer-gap interactions (Rizzi & Shlonsky, 2007). In section 2.4, I argue that this is insufficient

The first detailed analysis of these facts was given by Perlmutter (1971). Perlmutter proposes that there is a constraint that requires all clauses (S nodes) to contain a surface subject. On this view, the *that*-trace constraint is a subcase of the ban on subject-less S nodes. That is, the representation in (14) is ungrammatical because it contains a clause without a surface subject.

(14) *Who did Dale [$_{\text{VP}}$ say [$_{\text{S}}$ that [$_{\text{VP}}$ saw Bob]]]

This predicts that all sentences with subject extractions are ungrammatical,

including subject extractions over a null/absent complementizer, as in (1-b). To account for grammatical subject extractions, Perlmutter proposes an "S-pruning" rule that deletes S nodes that only dominate a VP. Thus, the filter on subject-less S nodes is vacuously satisfied:

(15) *Who did Dale [$_{VP}$ say [$_{S}$ [$_{VP}$ saw Bob]]] ⇒ Who did Dale [$_{VP}$ say [$_{VP}$ saw Bob]]

For Perlmutter, the requirement that S nodes contain a surface subject is variable. Perlmutter proposes that this constraint does not apply in languages like Spanish, due to the availability of sentences like (16). Since the *that*-trace constraint is a subcase of this constraint, subject extraction over an overt complementizer is trivially available, as in (17).

(16) a. Llegó.
arrived
'(He) arrived.'
b. Llueve.
rains
'It rains.'

(17) ¿Quién dijo Dale [$_{S}$ que [$_{VP}$ vio a Bob]]

At a larger grain of analysis, Perlmutter's analysis of *wh*-extraction patterns in English and Spanish require that the learner determine whether her language requires surface subjects. If surface subjects are optional, as in Spanish, there is no violation with subject extraction. If surface subjects are obligatory, then the learner infers that subject extraction over an overt complementizer is ungrammatical. In Perlmutter's analysis, it is unclear whether the S-deletion rule that permits sub-

ject extraction over absent complementizers is universal. If it is universal, then by virtue of learning that subjects are obligatory, the learner can also infer that subject extraction over null/absent complementizers is acceptable[2].

Perlmutter extends this analysis to the French data. Unlike Spanish, French requires overt subjects, demonstrated in (18). Thus, French appears to contradict Perlmutter's (1971) account, since subject extractions are permitted over *qui*. To account for this, Perlmutter proposes that the complementizer *qui* satisfies the subject requirement in French, because *qui* is homophonous with the *wh*-phrase *qui* 'who'. Thus, the morphological properties of *qui* serve as a cue that this complementizer exceptionally permits subject extraction.

(18) a. *(Il) est arrivé
He is arrived
'He arrived.'
b. *(Il) pleut
It rains
'It rains.'

Importantly, Perlmutter's (1971) analysis does not require the learner to track complementizer-gap distributions in the linguistic environment to make any of these inferences. Additionally, it accounts for the so-called *que/qui* contrast without However, his account predicts that every language in which the complementizer is homophonous with *who* permits subject extraction over it. If this is not the case, then this reintroduces the need to learn that *qui* has some special property which allows subject extraction. Additionally, the anti-*that*-trace effect in English rela-

[2]Presumably, learning that subjects are obligatory can be accomplished through indirect negative evidence, namely, a conspicuous lack of subject-less sentences, see Yang (2002)

tivization is largely unaccounted for in Perlmutter's (1971) analysis. Without further elaboration, his analysis should predict the same properties for relativization as complementation.

Perhaps the most famous account of the *that*-trace effect is the treatment given by Chomsky (1981) and Rizzi (1982, 1986, 1990). Chomsky attributes the *that*-trace effect to the Empty Category Principle (ECP), a presumably universal principle of universal grammar that captures a wide range of restrictions on movement dependencies. The ECP requires that the trace, or the base position of the movement dependency, be in a local relationship with its antecedent, the moved phrase, or a number of other privileged syntactic elements. Although the ECP accounts for a wide range of facts, the ECP account of the *that*-trace constraint specifically largely rests on stipulations of specific complementizers by positing that specific complementizers block the licensing relation.

Unlike Perlmutter's (1971) analysis of the overt/null complementizer asymmetry, the ECP analysis relies on attributing specific properties to complementizers that determine whether they permit subject extraction. On this analysis, there is nothing obviously inherent to null complementizers that requires that they permit antecedent government, and conversely there is nothing inherent about overt complementizers that requires that they block antecedent government. If each complementizer must be analyzed as either allowing or blocking subject extraction, then whatever idiosyncratic features of *that*that block subject extraction can be attributed to the null relativization complementizer and French *que*, and similarly the idiosyncratic properties of ∅ that permit subject extraction can be attributed

to the relativization complementizer *that*and French *qui*. However, this increase in analytic power has consequences for the learner. In effect, she must determine for each complementizer whether it permits subject extraction, which I later argue is unachievable given realistic linguistic input. On the other hand, if the properties of *that*that block subject extraction and ∅ that permit subject extraction are universal, then this partially simplifies the learning problem. If so, then an alternative explanation for the anti-*that*-trace effect and the *que/qui* contrast must be advanced, since these appear to be counterexamples to the generalization that overt complementizers like *that*block extraction, but not null complementizers.

Rizzi (1982, 1986, 1990) extends Chomsky's (1981) ECP analysis of the *that*-trace effect to Italian, which patterns like Spanish in the relevant respects (Torrego, 1984). Following Perlmutter (1971); Taraldsen (1980), Rizzi argues that the presence of "rich agreement" permits the canonical, pre-verbal subject position to be vacant, or satisfied by a null pronominal element, as in (19). Rizzi (1982, 1986) proposes that this frees the overt lexical subject to occur in a postverbal position, through a right adjunction operation. This explains why languages that permit null subjects also permit post-verbal subjects. This is the so-called null subject parameter.

(19) a. *pro* llueve.
pro rains
'It rains.'
b. *pro* llegó.
pro arrived
'(He) arrived.'

(20) *pro* vio a Bob Sarah.
pro saw to Bob Sarah
'Sarah saw Bob.'

Rizzi assumes that the constraint that blocks extraction over the overt complementizer *that*also applies to overt complementizers in Italian. In other words, the *that*-trace constraint is universal. To explain why languages like Italian and Spanish permit subject extraction, Rizzi proposes that subject extraction through the post-verbal subject position, as in (20), circumvents the *that*-trace constraint. Put differently, by virtue of allowing post-verbal subjects, languages like Italian and Spanish permit an parse in which the subject gap position is in a position that satisfies the ECP. Since English does not permit post-verbal subjects, this parse is not available. This is illustrated in (21). In later work, Rizzi calls this the "skipping strategy", since the lexical subject need not move to the canonical, preverbal subject position (Rizzi, 2006; Rizzi & Shlonsky, 2007). Instead, it may "skip" into the next clause.

(21) a. ¿Quién dijo Dale que vio a Bob ?
who said Dale that saw to Bob
'Who did Dale say (that) saw Bob?'
b. *¿Quién$_i$ dijo Dale [$_{\mathrm{CP}}$ t'_i [$_{\mathrm{C'}}$ que [$_{\mathrm{TP}}$ t_i [$_{\mathrm{T'}}$ vio a Bob]]]]
c. ¿Quién$_i$ dijo Dale [$_{\mathrm{CP}}$ t'_i [$_{\mathrm{C'}}$ que [$_{\mathrm{TP}}$ [$_{\mathrm{T'}}$ vio a Bob t_i]]]]

At a larger grain of analysis, Rizzi's (1982) analysis provides multiple "cues" for indirectly learning the absence of the *that*-trace effect in Spanish and Italian. Besides agreement facts, the learner could infer that there is no surface *that*-trace effect by detecting that subjects are not obligatory or that subjects may appear in a post-VP position. In fact, Safir (1985) and Safir & Jaeggli (1989) propose that absence of the *that*-trace effect is only related to post-verbal subjects, whereas the other properties logically can vary independently. However, versions of the "skipping strategy" that

do not rely on a relationship between post-verbal subjects and absence of the *that*-trace effect must rely on a relationship between rich agreement and absence of the *that*-trace effect instead.

Later, Rizzi (2006) and Rizzi & Shlonsky (2007) rejects the need for extraction to precede via this post-VP subject position. This is partially motivated by the fact that languages like Brazilian Portuguese do not appear to allow post-verbal subjects like Spanish or Italian, although they permit subject extraction over an overt complementizer subjects (Chao 1981, but see Menuzzi 2000). Instead, subjects may be extracted from their initial, VP-internal position (Koopman & Sportiche, 1991). If so, then post-verbal subjects are not necessarily the cue by which the learner can infer that subject extractions over overt complementizers are grammatical. On this view, the relation between rich agreement and subject extraction over an overt complementizer is critical, unlike Safir (1985) and Safir & Jaeggli (1989).

ECP-based frameworks have also been extended to account for the *que/qui* alternation and the $\varnothing_{comp}/\varnothing_{rel}$ contrast as well. Kayne (1976) and Pesetsky (1982) propose that *qui* is a complementizer that undergoes abstract agreement with the subject, which in effect exceptionally allows subject extraction over the complementizer. Rizzi (2006) and Rizzi & Shlonsky (2007) extend this reasoning to the anti-*that*-trace effect and similar phenomena in other languages. Crucially, absent an obvious morphological agreement cue, this agreement can presumably only be learned by tracking subject-gap distributions.

Pesetsky & Torrego (2001) propose an analysis of the *that*-trace effect that is quite different than Rizzi's analyses and its descendants. On this view, the English

complementizer *that* is the result of raising an auxiliary to the complementizer position. They propose that the constraint that blocks auxiliary raising with subject questions, as in (22), also applies in embedded contexts, to prevent simultaneous subject raising and auxiliary raising (i.e., overt complementizer *that*).

(22) a. **Who** __ left?
b. ***Who** did __ leave?

Although this analysis relies on substantially different technology than the ECP accounts, it raises many of the same concerns. It is unclear to what degree the complementizer is universally a raised auxiliary, and whether this varies on a complementizer-by-complementizer basis. The Pesetsky & Torrego analysis raises another analytic and learning question. Crucially, their analysis relies on the ability for an auxiliary to syntactically raise, yet still appear in its base position, since auxiliaries in embedded contexts surface in their base position, as shown in (23).

(23) a. **Who** did Dale say that Harry will investigate __ ?
b. ***Who** did Dale say (that) will Harry investigate __ ?

On Pesetsky & Torrego's analysis, it is unclear how the learner should infer that the complementizer *that* is a raised auxiliary, since the auxiliary surfaces in its base position. If complementizers are universally raised auxiliaries, then their analysis only still requires some mechanism such as the skipping strategy to explain why the *that*-trace effect seems to not apply in languages like Spanish. However, if there are languages with overt complementizers and no covert auxiliary raising, then it is unclear how the learner exposed to English data should rule out an analysis in

which *that* is a complementizer, and not a raised auxiliary. In either case, it is unclear whether this analysis requires specific complementizers to serve as "loopholes" to an otherwise robust generalization.

Analyses in which complementizers undergo abstract subject agreement or through some other special mechanism that extraordinarily permit subject extraction provides the syntactician the technical language to describe the array of complementizer-gap interactions described here. Importantly, the relevant notion of agreement or its equivalent must be abstract in order to distinguish between morphologically identical complementizers, such as $that_{comp}/that_{rel}$ and $\varnothing_{comp}/\varnothing_{rel}$. Once this theoretical move is made, however, it is unclear to what extent the "skipping strategy" is needed. For instance, a logical extension of analyses which attribute specific properties that permit or block subject extraction to specific complementizers is to explain all subject-gap interactions through this mechanism. In other words, if the English relativization complementizer $that_{rel}$ permits subject extraction through abstract agreement, then so too might be the Spanish and Italian complementation complementizers. If so, then there is no need to link the absence of the *that*-trace effect in Spanish and Italian to any other grammatical property. If so, then apparent counterexamples like Brazilian Portuguese, which do not exhibit post-verbal subjects but permit subject extraction over an overt complementizer (Chao, 1981; Menuzzi, 2000), are no longer problematic.

In fact, there are some reasons for rejecting the skipping strategy. Newmeyer (2004, 2005), following Gilligan (1987), argues that many of the properties that should correlate on Rizzi's analysis have many counterexamples. These counterex-

amples are especially problematic from a learning perspective, since it implies that the link between the relevant constructions is not absolute, and in effect reintroduces the learning problem that these links could resolve. Additionally, Sobin (1987, 2002, 2009) claims that the *that*-trace effect is not robustly observed across all varieties of English, which are otherwise similar in the relevant grammatical properties. If these findings are reliable, then a complementizer-by-complementizer learning strategy can better accommodate the variation between English varieties. This is strong motivation for abandoning locating the *that*-trace effect in universal grammar, and linking its absence to other grammatical features.

However, I argue that indirect learning is in fact necessary for learning the absence of the *that*-trace effect in Spanish/Italian. In section 2.4, I argue that there is likely insufficient evidence in the English, Spanish, or Italian learner's input to directly infer whether overt subject extraction over overt complementizers is permitted on a complementizer-by-complementizer basis. Instead, I argue that the learner must rely on some indirect cue, such as post-verbal subjects, as in Rizzi's (1982) original proposal. Before turning to the distributions of subject-gap alternations that the learner are exposed to, however, I first demonstrate that the *that*-trace effect is robust in English speakers, and that the contrast between English- and Spanish-speaking adults is robust.

2.3 Experiments 1–2: Acceptability of *That*-Trace Configurations

In this section, I investigate whether the contrast between English and Spanish is robust, and whether the *that*-trace constraint is reliably observed between English speakers. If either of these conditions fail, it is unclear that the learner is learning anything about the *that*-trace constraint, i.e. drawing inferences from data to select a hypothesis. If the *that*-trace constraint is not reliably observed in English speakers, then it is not obvious that there is any difference between English and languages like Spanish. In other words, there may no be reliable difference between languages, meaning that the learner does not need to select between a grammar that observes the *that*-trace constraint and one that does not. Similarly, if there is systematic variation between dialects of English, as claimed by Sobin (1987, 2002, 2009), then the absence of the *that*-trace constraint cannot be attributed to any grammatical property that is shared between varieties of English that obey the constraint and those that do not. This severely complicates the learning problem. In this section, I report on the results of a cross-linguistic acceptability judgment task that uses approximately lexically-matched materials between Spanish and English.

2.3.1 Rationale

To my knowledge, the first experimental studies on the *that*-trace effect were conducted by Sobin (1987, 2002, 2009). In a 3-point judgment task, Sobin (1987) showed that English-speaking participants rated *that*-trace violations lower than grammatical conditions, but higher than strong island violations. Sobin claimed that

this demonstrates that the *that*-trace constraint is therefore not reliably observed between individuals or between dialects. This is surprising, since his data in fact shows clear sensitivity to the *that*-trace constraint. Additionally, he claims that the *that*-trace constraint is absent in Midwestern American English. However, there are no control groups in any of Sobin's studies. Thus, the cross-dialect claim does not follow from these findings. His results are equally compatible with the claim that the *that*-trace effect is simply a less severe violation, but uniform across dialects and individuals.

Cowart (1996, 2003) investigated sensitivity to the *that*-trace effect in a design that corrected for some methodological concerns in Sobin's original design. Cowart (1996) manipulated extraction site (subject or object) and complementizer (present or absent) in an acceptability judgment task, in which he found a penalty for subject extraction over an overt *that*. This confirms the claim that English-speakers largely find *that*-trace configurations to be unacceptable. Additionally, Cowart (2003) conducted a large-scale acceptability judgment study in 5 different dialect locations, and found no differences between dialect groups with respect to the *that*-trace constraint. This suggests that the *that*-trace constraint is robust across English varieties. He also found that, across multiple testing sessions, the variability between participants was smaller than the variability within subjects. Cowart therefore concludes that participants are largely uniform in their judgments within groups, as well.

To our knowledge, the contrast between English- and Spanish-type languages has not been investigated systematically. In an experiment similar to Cowart's (1996), Featherston (2005) showed that German speakers show sensitivity to the

that-trace constraint, which was also replicated in a cross-language experiment by Keller & Alexopoulou (2005). However, this is not so informative, since German patterns more like English than Spanish/Italian (Bayer & Salzmann, 2012). Thus, this does not illustrate whether the relevant cross-language difference is robust.

There were two goals in Experiments 1 and 2. The first goal was to determine whether the reported contrast between English and Spanish is robust. To this end, I conducted two acceptability judgment studies in English and Spanish that were approximately lexically- and structurally-matched. This was intended to control for other potential grammatical differences between the two languages as much as possible. The second goal was to probe whether there were any detectable systematic differences between English varieties. For this reason, I ran an additional sample of English-speakers, and examined the responses by participant and geographic location.

2.3.2 Experiment 1

Experiment 1 was a replication of Cowart's (1996) study conducted in English and Spanish with approximately similar materials. Like Cowart, I hypothesized that the there should be a penalty for subject extraction over an overt complementizer in English. In contrast, I predict no such effect in Spanish. I also included an additional larger group of English participants to probe for dialect differences and individual differences. If Midwestern American English speakers are not sensitive to the *that*-trace constraint, then I predict that this population's responses should track

the responses of the Spanish-speaking participants more than the other English-speaking participants. Finally, if speakers within a linguistic community converge on the same grammar, the effect should be roughly the same across individuals. I address these last two concerns in section 2.3.4.

2.3.2.1 Methods

Experiment 1 was an acceptability judgment task administered on Ibex[3]. Participants were asked to rate the acceptability of sentences on a 1–7 scale. The experiment was explained to the participants on screen, with several example sentences illustrating how to use the scale. The items were displayed onscreen one by one, and participants either assigned a score to the sentence by clicking on buttons corresponding to the values 1–7, or by pressing the corresponding number key on the keyboard. The target items were distributed across four lists in a 2×2 Latin Square design, and the order of the items were randomized by participant. The materials for Experiments 1 and 2 were administered in the same session, but we report the results separately.

Participants were recruited from Amazon Mechanical Turk[4] and were compensated 3 USD. There were 120 total English-speaking participants (3 excluded), and 24 Spanish-speaking participants. English speakers had to have completed a minimum of 1,000 HITs (Amazon Mechanical Turk tasks) with a 95% approval rating, and must have been from the USA. A lower threshold was used for Spanish-speaking

[3] `http://spellout.net/ibexfarm`
[4] `http://www.mturk.com`

participants, due to the smaller number of available participants on Mechanical Turk. Spanish speakers had to have completed a minimum of 100 HITs with a 75% or greater approval rating. I requested demographic information for each participant, including location, income, and years of education. Income ranges were in increments of 20,000 USD (0-20,000 USD, 20,000-40,000 USD, 40,000-60,000 USD, 60,000-80,000 USD, 80,000-100,000 USD, 100,000+ USD). This demographic information was collected to test for possible effects of socioeconomic status and effect of location in probing for variability between individuals in the English sample. For the between-language analyses, I restricted analysis to the first 24 English-speaker participants to keep equal sample sizes between the two languages, and included the rest of the participants in analysis of dialect influence and individual differences. There were no participants in the highest two income bins in either language in the initial sample.

2.3.2.2 Materials

There were 16 sets of target items and 20 complexity-matched fillers. I manipulated the presence of the complementizer and the position of the gap (object or subject) to replicate Cowart's (1996) design. The materials are illustrated in (24) and (25).

(24) **Spanish**:

a. *Object gap, {complementizer/no complementizer}*:

¿**Con quién** sugirió Álvaro **{*∅/que}** horneará Lucía un pastel ___
with who suggested Alvaro **{∅/that}** will bake Lucia a cake ___

para la fiesta?
for the party?

'**Who** did Allen suggest {∅/**that**} Lucy will bake a pie with __ for the party?'

b. *Subject gap, {complementizer/no complementizer}*:

¿**Quién** sugirió Álvaro **{*∅/que}** horneará __ un pastel con
who suggested Alvaro **{∅/that}** will bake __ a cake with
Lucía para la fiesta?
Lucia for the party?

'**Who** did Allen suggest {∅/***that**} __ will bake a pie with Lucy for the party?'

(25) **English**:

a. *Object gap, {complementizer/no complementizer}*:
'**Who** did Allen suggest {∅/**that**} Lucy will bake a pie with __ for the party?'

b. *Subject gap, {complementizer/no complementizer}*:
'**Who** did Allen suggest {∅/***that**} __ will bake a pie with Lucy for the party?'

Materials were approximately lexically matched in both English and Spanish. However, due to divergent grammatical properties between the two languages, there were some systematic differences between the English and Spanish materials. In the English materials, the grammatical role of the *wh*-filler was ambiguous at the beginning of the sentence, because all four conditions began with the question word *who*. Subject extractions were later disambiguated by the lack of a subject in the embedded clause, and object extractions were disambiguated by a transitive verb or a preposition missing an object. In the Spanish materials, the grammatical role of the *wh*-filler is unambiguous. This is because Spanish does not permit preposition stranding, and marks animate direct objects with the preposition *a* 'to' ("*a*-marking", Fernández Ramírez 1986). Thus, *quién* 'who' is unambiguously marked

as a subject, whereas object extractions are marked as such by the presence of a preposition. Additionally, in the English materials, the subject (*Allen*) and embedding verb (*suggested*) immediately followed the filler. In the Spanish materials, the embedding verb (*sugirió* 'suggested') followed the *wh*-filler, because Spanish requires subject-verb inversion with long-distance *wh*-extraction (Torrego, 1984). The complementizer followed the embedding verb in English and the main clause subject in Spanish. In the embedded clause, the English materials contained the embedded subject (*Lucy*), followed by the predicate (*will bake a pie with (Lucy)*), followed by a sentence-final adjunct (*for the party*). The Spanish materials contained the same elements, except the embedded subject (*Lucía*) followed the embedded verb, again because of obligatory subject-verb inversion with subject extraction.

For the English materials, I expected significantly lower ratings for the subject extraction, present complementizer condition, corresponding to the *that*-trace effect. This is because the *that*-trace constraint is specifically a constraint on subject extraction over an overt complementizer, and thus only the subject extraction, present complementizer condition violates a grammatical constraint. Thus, it should have significantly reduced ratings compared to the other three conditions. For Spanish, I expected lower ratings for the no complementizer conditions, because Spanish requires a complementizer with embedded clauses. I predicted no interaction between subject extraction and complementizer if there is no surface effect of the *that*-trace constraint in Spanish. The conditions with predicted lower ratings are marked with * in (24) and (25).

2.3.2.3 Results

For analysis, I z-scored the ratings by participant. For each participant, I divided the difference between each rating and that participant's mean rating by that participant's standard deviation. This controls for some variability in participants' use of the scale and transforms the data into a continuous measure across participants, although participant's individual ratings are still discrete. I used the `lmerTest` package in R (Kuznetsova *et al.*, 2014) to fit mixed effects models for each language with z-scored response as dependent variable. Main effect of extraction site, complementizer, income, education, and age were included in the analysis. The interaction between extraction site and complementizer was also included. Participants and items were included as random effects. In the English data, there was a main effect extraction site ($\hat{\beta} = 0.58 \pm 0.11, t(70) = 5.47, p < 0.0001$) and an interaction of extraction site and complementizer ($\hat{\beta} = -1.52 \pm 0.15, t(44) = -9.93, p < 0.0001$). For the Spanish model, there was a significant effect of complementizer ($\hat{\beta} = 0.90 \pm 0.11, t(370) = 7.83, p < 0.0001$), and a significant interaction of extraction site and complementizer ($\hat{\beta} = -0.60 \pm 0.16, t(370) = -3.73, p = 0.0002$).

Means and differences-in-differences of the means (DDs) are reported in Table 2.2. DDs are a measure of the superadditive interaction of extraction site and complementizer (Maxwell & Delaney, 2003; Sprouse *et al.*, 2012a, 2013). These are calculated by subtracting the ratings of that *that*-trace conditions from the subject extraction controls (here, the no complementizer condition), and subtracting the difference between object conditions from this value. Thus, this gives a measure

	English	*Spanish*
No complementizer, subject extraction	0.89 ± 0.08	-0.40 ± 0.07
No complementizer, object extraction	0.30 ± 0.07	-0.54 ± 0.08
Overt complementizer, subject extraction	-0.50 ± 0.06	-0.13 ± 0.09
Overt complementizer, object extraction	0.46 ± 0.07	0.35 ± 0.09
DD	1.54	0.62

Table 2.2: Results of Experiment 1. Mean z-scored ratings by language and condition with standard error of the mean and differences-in-differences of the means by language.

of the "extra" unacceptability of subject extraction with an overt complementizer. When evaluating DDs, Sprouse *et al.* (2013) suggest drawing a distinction between statistical significance and practical significance, or a significance level that meets some further criterion beyond statistical significance. They suggest that DDs above 0.5 standard deviations from the mean meet this criterion, because this is the size of an effect that should be detectable without statistical tests (Cohen, 1988, 1992). For this reason, I follow Sprouse *et al* in treating this as the minimal criterion for practical significance.

2.3.2.4 Discussion

The English results are largely consistent with our predictions, but the Spanish results are surprising. In English, there was an interaction of extraction site and complementizer, reducing ratings in sentences with subjects extracted over an overt complementizer. This is consistent with the hypothesis that English-speakers are sensitive to the *that*-trace constraint. The reduction in ratings with subject extraction was unexpected and was not observed in Cowart's studies, however. For the Spanish results, I predicted that the conditions without complementizers would be

rated lower than the overt complementizer conditions, because complementizers are obligatory in Spanish. The data support this prediction. However, there was also an interaction of extraction site and complementizer, lowering ratings with subject extraction over an overt complementizer. Taken at face value, this suggests that there is no difference between English and Spanish with respect to the *that*-trace constraint, contrary to the widely reported facts in the comparative syntax literature. If so, this finding undermines our claim that sensitivity to the *that*-trace constraint must be learned indirectly, because properties that do not vary between languages may be innate, and therefore not learned.

However, it is unclear whether it is reasonable to assign the same interpretation to the interaction effect in the English and Spanish results. In English, the interaction effect is expected because both subject extractions and overt complementizers are rated highly, and thus the expected value for subject extraction overt an overt complementizer is high, unless the *that*-trace constraint applies. In Spanish, however, complementizers are obligatory, and thus the baselines are different than in the English study. Thus, it may be the case that Spanish has an independent bias against subject extraction that is revealed in the complementizer conditions, but is overshadowed in the no complementizer conditions by the degradedness of missing complementizers. In other words, the complementizer factor has a different status in the Spanish study than it did in the English study.

Experiment 2 resolved this issue by probing for sensitivity to the *that*-trace constraint in English and Spanish by manipulating the adjacency of the complementizer and the preverbal subject gap, which is shown to ameliorate the *that*-trace

constraint (Bresnan, 1977; Culicover, 1993; Sobin, 2002). This manipulation is possible in both English and Spanish, and thus the baseline conditions are less likely to introduce confounds in the between-language interpretation of the results.

2.3.3 Experiment 2

Like Experiment 1, the goal of Experiment 2 was to establish the contrast in sensitivity to the *that*-trace constraint in English and Spanish to determine to what degree the *that*-trace constraint must be learned across languages. Instead of manipulating complementizer presence as in Experiment 1, I manipulated whether an adverbial intervened between the complementizer and subject gap position. This has been shown to mitigate the *that*-trace violation in English, shown in (26).

(26) a. *__Who__ did Dale say that __ saw Bob?
b. __Who__ did Dale say that, to everyone's surprise, __ saw Bob?

This manipulation permits probing for sensitivity to the *that*-trace constraint in a way that does not depend on manipulating complementizer type. This allows for more stable baselines across languages. If English speakers are sensitive to the the *that*-trace constraint, then they should rate sentences with subject extractions over gap-adjacent complementizers lower than sentences with subject extraction with an adverbial intervening between the complementizer and subject gap. If Spanish speakers are sensitive to the *that*-trace constraint, there should be a similar reduction. However, if Spanish speakers are not sensitive to the *that*-trace constraint, there should be no reduction in ratings.

2.3.3.1 Methods

Experiment 2 was conducted simultaneously with Experiment 1.

2.3.3.2 Materials

The materials in Experiment 2 were largely similar to the materials in Experiment 1. Instead of manipulating complementizer, I manipulated adverbial position. If the adverbial modified the main clause, it preceded the complementizer. In this case, the complementizer and the preverbal subject gap in the embedded clause were adjacent. This should yield a *that*-trace constraint violation. If the adverbial modified the embedded clause, then it followed the complementizer. This broke up adjacency between the complementizer and preverbal subject position. This should ameliorate the *that*-trace violation (Bresnan, 1977; Culicover, 1993; Sobin, 2002). For English, there should be an interaction of adverbial position and extraction site, because the main clause adverbial, subject extraction condition. If Spanish speakers are not sensitive to the *that*-trace constraint, then there should be no interaction of these factors. As before, * indicates predicted lower ratings. The materials are exemplified in (27) and (28).

(27) **Spanish**:

a. *Object gap, {main clause adverbial/embedded clause adverbial}*:

¿**Con quién** remarca Juana **{cada año que / que cada año}**
with who remarks Joanna **{every year that / that every year}**
va Carmen ___ de pesca?
goes Carmen ___ of fishing

'**Who** does Joanna remark **{every year that / that every year}**

Carmen goes fishing with __ ?'

b. *Subject gap, {main clause adverbial/embedded clause adverbial}*:

¿**Quién** remarca Juana **{cada año que / que cada año}**
who remarks Joanna **{every year that / that every year }**
va __ de pesca con Carmen?
goes __ of fishing with Carmen

'**Who** does Joanna remark **{every year that / that every year}** __ goes fishing with Carmen?'

(28) **English**:

a. *Object gap, {main clause adverbial/embedded clause adverbial}*:
'**Who** does Joanna remark **{every year that / that every year}** Carmen goes fishing with __ ?'

b. *Subject gap, {main clause adverbial/embedded clause adverbial}*:
'**Who** does Joanna remark **{*every year that / that every year}** __ goes fishing with Carmen?'

2.3.3.3 Results

The results for Experiment 2 were analyzed in a similar fashion as Experiment 1. I used the `lmerTest` package in R to fit mixed effects models for each language with z-scored response as dependent measure. Main effects of extraction site, adverb position, education, income, and age were included in the analysis. Interaction effects of extraction site and adverb position were also included. Participants and items were included as random effects. For the English model, there was a significant effect for low income ($\hat{\beta} = 0.40 \pm 0.14, t(24) = 2.89, p = 0.008$). There was also an interaction of extraction site and adverbial ($\hat{\beta} = -0.41 \pm 0.18, t(100) = -2.30, p = 0.02$). In the Spanish model, there was a main effect of extraction site ($\hat{\beta} = -0.27 \pm 0.11, t(369) = -2.55, p = 0.01$).

Means and differences-in-differences (DDs) are reported in Table 2.3. If a

	English	*Spanish*
Main clause adverbial, subject extraction	-0.69 ± 0.06	-0.35 ± 0.08
Main clause adverbial, object extraction	-0.17 ± 0.08	0.01 ± 0.08
Embedded clause adverbial, subject extraction	-0.17 ± 0.07	-0.29 ± 0.08
Embedded clause adverbial, object extraction	-0.11 ± 0.07	-0.02 ± 0.08
DD	0.45	0.09

Table 2.3: Results of Experiment 2. Mean z-scored ratings by language and condition with standard error of the mean and differences-in-differences of the means by language.

DD of 0.5 is a reasonable approximation for "practical significance", then the English results again suggest sensitivity to the *that*-trace constraint, as this DD value approaches practical significance. This is not the case for the Spanish results.

2.3.3.4 Discussion

The results of Experiments 1 and 2 strongly suggest that English speakers are sensitive to the *that*-trace constraint, but Spanish speakers are not. In Experiments 1 and 2, English-speaking participants gave lower ratings to sentences with an overt complementizer immediately adjacent to a subject gap, as predicted. For Spanish-speakers, there was an apparent *that*-trace effect in Experiment 1. I suggested that this was an artifact of the experimental design. The results from Experiment 2 were consistent with this, since there was no evidence for sensitivity to the *that*-trace constraint in Spanish.

These data serve to establish that this cross-language contrast is robust. This in turn implies that the learner must distinguish whether her language is sensitive to the *that*-trace constraint on the basis of her linguistic input, either by directly tracking the distribution of complementizers and gaps or through some other indirect

means. This therefore requires that there be sufficient evidence to base this inference on. As I argue in section 2.4, learners of Spanish must indirectly learn the absence of the *that*-trace effect, either by observing post-verbal subjects or null subjects. This in turn limits the space of possible syntactic analyses available to the theorist, namely, insensitivity to the *that*-trace constraint must be conditioned on some other linguistic property.

However, as discussed in section 2.3.1, there are claims that the *that*-trace effect is not robust between speakers of English. Even if English-speakers as a group robustly demonstrate sensitivity to the *that*-trace constraint, there may be a subset of speakers that do not. If speakers in fact are not uniform in their sensitivity to the *that*-trace constraint, this potentially complicates the relation between the learner's linguistic experience and the grammar she develops (Han *et al.*, 2007). Before addressing the distribution of complementizers and subject gaps in child-directed speech, I first address the systematicity between speakers of English in section 2.3.4.

2.3.4 Systematicity Within English Speakers

The results from Experiments 1 and 2 are consistent with the claim that English-speakers are sensitive to the *that*-trace constraint. However, as discussed in 2.3.1, the *that*-trace constraint is claimed to vary across dialects of English and potentially between individuals (Pesetsky, 1982; Sobin, 1987). Specifically, Sobin (1987) claims that speakers of Midwestern American English are not sensitive to

the *that*-trace constraint, and Sobin (2009) suggests that the *that*-trace constraint is variable even between individuals. If there are significant dialect differences, this requires pinpointing on a dialect-by-dialect basis which cues the learner uses to determine whether the *that*-trace constraint applies. If there is significant individual variation, this might imply that the learner does not use their linguistic experience to determine whether their target grammar is sensitive to the *that*-trace constraint. Thus, in-group variation may severely complicate my primary claim in this chapter.

As reported in section 2.3.1, Cowart (1996, 2003) has investigated the claim that there is systematic dialect variation between varieties of English as made by Sobin (1987, 2002, 2009). Cowart (2003) found no effect of geographic location on a large-scale study with the same manipulations as my Experiment 1. Additionally, he showed that the variability in the *that*-trace conditions within participants between testing sessions was greater than the variability between participants within a testing session. He concludes that this suggests the variation within English-speakers' responses in Sobin's studies is likely not due to systematic idiolectal differences, but rather experimental noise.

I argue that my results are largely in line with (Cowart, 2003). In this section, I explore the responses from all 120 English-speaking participants (3 excluded). I only use the data from Experiment 1 because the effect was clearer in Experiment 1 than in Experiment 2, and thus any potential dialect or idiolectal differences are more likely to be distinguishable from noise. First, I investigate whether there is a significant difference between the responses from Midwestern and non-Midwestern participants. Secondly, I test whether the attested variability in the performance of

our participants is determined by location of participant.

First, I investigate the proposed differences between speakers of Midwestern American English and other varieties of American English. Each participant was asked to provide his or her zip code in Experiment 1. Participants that reported zip codes in North Dakota, South Dakota, Nebraska, Kansas, Minnesota, Wisconsin, Iowa, Missouri, Michigan, Indiana, Illinois or Ohio were coded as Midwestern. Our sample consisted of 97 non-Midwestern participants, and 20 Midwesterners. The means and DDs of the means by dialect group are displayed in Table 2.4. Again, both dialect groups are well above the threshold of "practical significance", with a DD value above 0.5. Thus, both Midwesterns and non-Midwesterners as a group appear to obey the *that*-trace constraint.

After coding participants for Midwestern/non-Midwestern status, I constructed a mixed effects model using the `lmerTest` package in R. The dependent measure was the z-scored responses from all the participants from Experiment 1. Main effects of extraction, complementizer, Midwestern status, age, income, and education were included in the analysis. Interaction effects for complementizer and extraction, and complementizer, extraction site, and Midwestern status were included. Participants and items were included as random effects. There were main effects of extraction ($\hat{\beta} = 0.24 \pm 0.05, t(1758) = 4.65, p < 0.0001$) and complementizer ($\hat{\beta} = 0.12 \pm 0.05, t(1754) = 2.38, p = 0.02$). There was also a significant interaction of extraction and complementizer ($\hat{\beta} = -1.12 \pm 0.07, t(1753) = -15.48, p < 0.0001$). There were no effects of Midwestern status, no interaction of complementizer, extraction site, and Midwestern status. These results also show that with all 120 (3

excluded) participants included in the analysis, the *that*-trace constraint persisted, with no difference between Midwestern and non-Midwestern participants.

	Midwesterners	*Non-Midwesterners*
No complementizer, subject extraction	0.78 ± 0.08	0.60 ± 0.04
No complementizer, object extraction	0.50 ± 0.09	0.37 ± 0.04
Overt complementizer, subject extraction	-0.42 ± 0.07	-0.37 ± 0.03
Overt complementizer, object extraction	0.44 ± 0.08	0.49 ± 0.03
DD	1.15	1.10

Table 2.4: Results of Experiment 1. Mean z-scored ratings by condition and dialect group with standard error of the mean and differences-in-differences of the means by dialect group.

Apart from Sobin's (1987) claim about Midwestern dialects, there may be other dialect differences hidden in the data. Thus, it's worth checking whether there is any relationship between location and DDs. To this end, I plotted the DD scores of each participant in Experiments 1 and 2 against their reported geographic location, shown in Figures 2.1 and 2.2. If there is a group of participants with low sensitivity to the *that*-trace constraint grouped in one geographic region, then this is evidence for a dialect without the *that*-trace constraint. Conversely, if there are low-sensitivity participants are intermixed with high-sensitivity participants, or if there are no low-sensitivity participants, then there is no evidence for systematic dialect differences in this data.

In Figure 2.1, participants are overwhelmingly sensitive to the *that*-trace constraint in Experiment 1, as the majority of participants have high DD scores for Experiment 1. The few who have lower DD scores form no obvious geographic cluster. The results for Experiment 2 plotted in Figure 2.2 are more diverse, with a greater number of participants falling in the lower DD ranges. There is an ap-

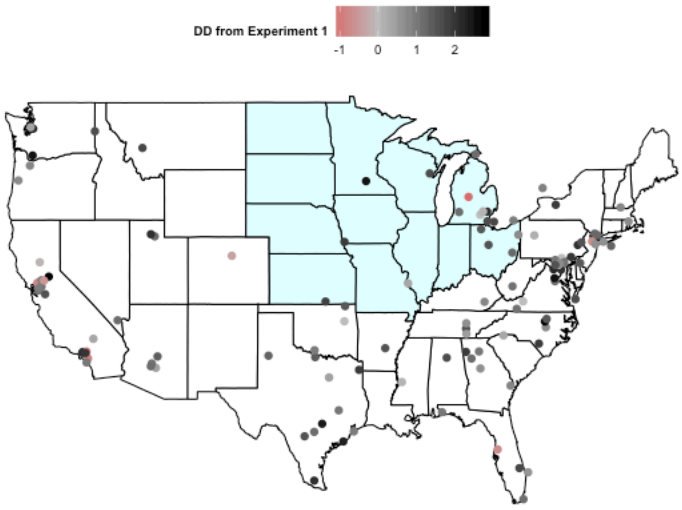


Figure 2.1: Plot of DDs in Experiment 1 by latitude and longitude of participant. Color of each point corresponds to that participant's DD, or sensitivity to the *that*-trace effect. Black corresponds to strong *that*-trace effects, grey corresponds to insensitivity to *that*-trace effects, and red corresponds to preference for subject extraction over complementizers.

parent clustering of low-sensitivity participants in Michigan. However, since these same participants exhibited high DDs in Experiment 1, it is unlikely that these participants do not exhibit the *that*-trace constraint. Thus, these data provide no evidence for the claim that there is systematic dialect variation, consistent with Cowart's (2003) argument.

However, there is some variability between individuals. If this variability is due arbitrary idiolect differences, then this might suggest that presumably similar linguistic input across individuals does not uniformly lead the learner to select the same grammar. On the other hand, if this variability is simply due to experimental noise, then this does not invalidate the primary claim in this chapter. Distinguishing these alternatives is difficult without using a test-retest paradigm, as performed by

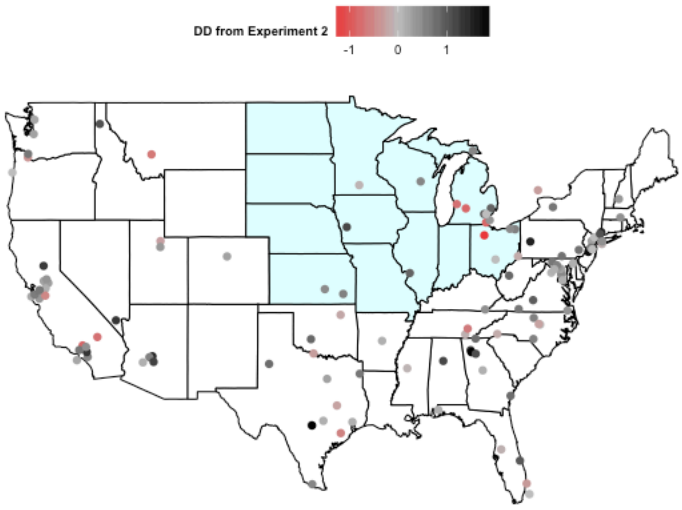


Figure 2.2: Plot of DDs in Experiment 2 by latitude and longitude of participant. Color of each point corresponds to that participant's DD, or sensitivity to the *that*-trace effect. Black corresponds to strong *that*-trace effects, grey corresponds to insensitivity to *that*-trace effects, and red corresponds to preference for subject extraction over complementizers.

Cowart (2003). Additionally, drawing the connection between individual variation in a judgment task and divergent *grammars* is non-trivial. For instance, there may be systematic differences between individuals, but this may be due to variation in facility in giving acceptability judgments or parsing complex materials.

The extreme case can be defined as follows: if the child's input completely underdetermines the correct grammar, then the learner arbitrarily selects one of the two "options" with equal probability (Han *et al.*, 2007). If so, then approximately half of the learners will select a grammar that is sensitive to the *that*-trace constraint, and the other half will not. If so, then participants' responses in an acceptability judgment task are predicted to be bimodally distributed, with approximately half showing sensitivity to the *that*-trace constraint (i.e., exhibiting high DD scores in

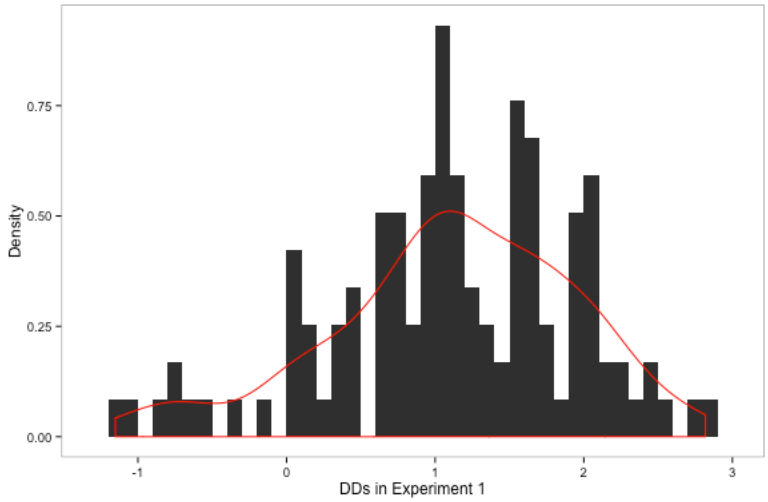


Figure 2.3: Histogram of participants by DDs in Experiment 1, with density plot overlayed in red. Bin width is 0.1. Mean is 1.11

Experiment 1), and the other half not showing sensitivity (i.e., exhibiting a DD score of 0). Plotting the histogram of participants' DD scores in Experiment reveals that there is a largely unimodal distribution. Importantly, there is no spike of DD scores around 0, corresponding to a a significant population that is insensitive to the *that*-trace constraint.

Similarly, plotting the differences in subject extraction conditions against the differences in the object extraction conditions in Experiment 1 reveals that participants' responses largely centered on 0 for object extraction, but around -1.02 for subject extraction conditions, as shown in 2.4. If some participants' grammars did not penalize subject extraction over overt complementizers, i.e. there was no difference for subject extraction over an overt complementizer compared to a null complementizer, then the differences between the subject conditions should be centered around 0 for those participants. This shows that participants' responses largely

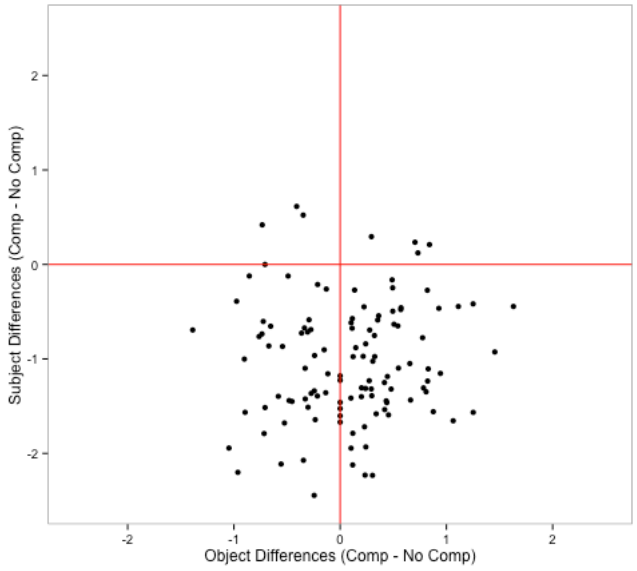


Figure 2.4: Differences in z-scored ratings in subject extraction conditions and object extraction conditions by participant in Experiment 1.

indicate a penalty for subject extraction over overt complementizers by virtue of their skew. This too supports the claim that our sample is largely drawn from the a population that is uniformly sensitive to the *that*-trace constraint.

In this section, I established that English-speakers are uniformly sensitive to the *that*-trace constraint, and that Spanish speakers are not sensitive to the *that*-trace constraint. This then implies that the learner must be capable of inferring whether her target languages exhibits this constraint on the basis of her linguistic input. This in turn implies that the learner must have access to sufficient data to inform this decision. One analysis, suggested in 2.2, implies that the learner determines for each complementizer whether subject extraction is licensed over it. An alternative analysis, initially suggested by Rizzi (1982), implies that learners of languages like Spanish and Italian infer that the *that*-trace constraint does not apply in their language by virtue of observing some other, superficially unrelated

construction. I argue that the learner's linguistic exposure is unlikely to support direct, complementizer-by-complementizer learning. Thus, an analysis like Rizzi's (1982) is motivated on learnability grounds.

2.4 Primary Linguistic Data

The previous sections establish that the sensitivity to the *that*-trace constraint must be learned, since there is reliable variation between linguistic communities and reliable uniformity within linguistic communities. This suggests that the primary linguistic input that learners are exposed to must be informative and distinct enough for the learners to select the right grammar. In this section, I investigate the distributions of complementizers and gaps in English, Spanish, and Italian to determine what data is accessible to the learner. In section 2.4.1, I describe Pearl & Sprouse's (2013) (P&S) findings on learnability of island constraints on filler-gap dependencies. In section 2.4.2, I describe the results of a study on Spanish and Italian child-directed speech. I argue that the distributions of complementizers and gaps in each language are underinformative for determining whether the *that*-trace constraint applies. The learner must therefore depend on an indirect cue, such as the availability of postverbal subjects, as suggested in 2.2.

2.4.1 P&S's Model

P&S describe a model that they argue is capable of learning a number of constraints on filler-gap dependencies. This model is notable because it makes simple,

Total utterances:	101,383	
Total *wh*-dependencies:	20,923	
Total relativizations:	774	
	that	∅
Subject *wh*-dependencies:	0	13
Object *wh*-dependencies:	2	159
Subject relativizations:	349	0
Non-subject relativizations:	320	29

Table 2.5: Counts of *wh*-distance dependencies in Pearl & Sprouse's (2013) English child-directed corpus. Relativization dependencies counts collected from DAC.

explicit assumptions about the learner's prior biases and inference mechanisms, and succeeds at learning a variety of island effects given realistic primary linguistic input. The model learns constraints on dependencies by tracking the probability of the path of nodes that the dependency crosses. For more detail, see Pearl & Sprouse (2013).

For their experiments, P&S trained their model on five corpora of child-directed speech drawn from the CHILDES databank (MacWhinney, 2000). This corpus contained 101,838 total utterances, and a total of 20,923 long-distance dependencies. P&S take this to be an estimate of 10% of the learner's input, which they take to be approximately 1 million utterances, or the number of sentences that learners hear from birth to age three according to Hart & Risley (1995). This is a somewhat conservative estimate, since children are still learning by age 3. I follow P&S in using these estimates for comparison's sake.

Although their model is successful in learning some island constraints, their model fails to capture the *that*-trace effect. Phillips (2013) points out that this is because embeddings with complementizers are quite rare compared to embeddings

without complementizers in English. Specifically, P&S's corpus contained only 2 instances of object extractions over an overt complementizer. There were also 0 instances of subject extractions over an overt complementizer, i.e., there were no *that*-trace violations.

If the English learner must infer an interaction between subject extraction and complementizer type, she could feasibly do this by comparing the expected number of subject extractions over an overt complementizer with her input. However, since the ratio of extractions over an overt complementizer to extractions over null complementizers is low (2:172), and the ratio of multiclausal subject extractions to multiclausal object extractions is low (13:161), the learner's expected number of subject extractions over an overt complementizer would also be low. Thus, on the hypothesis that subject extraction over an overt complementizer is grammatical, the learner's expected value for observed subject extractions over an overt complementizer may very well be 0. Put differently, the fact that the learner observes no subject extractions over an overt complementizer is consistent with either such extractions being ungrammatical or being grammatical but proportionately rare.

The relevant data are also too rare in absolute terms, as well. Scaling P&S's findings to more realistic proportions, one can estimate that if the learner hears approximately one million sentences in three years (Hart & Risley, 1995), then the learner hears approximately 20 long-distance object extractions over an overt complementizer, and a marginal number of long-distance subject extractions over an overt complementizer. Abstracting over differences in learners' input as a function of age, the learner hears approximately 333,000 sentences per year, of which 6 or

7 are long-distance object extractions. This amounts to approximately one long-distance object extraction every two months. Thus, the learner that must infer the *that*-trace constraint from the distribution of gaps and complementizers must reliably detect, parse, and distinguish from noise these very rare object extractions, and be capable of inferring the subject-object asymmetry on this basis. In other words, the learner must reject the hypothesis that all extractions over an overt complementizer are rare by detecting the very occasional object extractions over an overt complementizer, and contrasting this with the number of subject extractions over an overt complementizer that she is exposed to. This is made even less plausible by the observation that English-learners all converge on the same grammar – i.e., it must be the case that every learner is sensitive enough to pick up on these marginal distributional facts, and that the input is uniform enough across different individuals' experience. Thus, I argue that it is unlikely that the learner can uniformly infer that her grammar obeys the *that*-trace constraint by directly inspecting the distribution of complementizers and gaps in English child-directed speech.

Interestingly, P&S's corpus reveals that the learner's linguistic input might permit direct learning of the anti-*that*-trace effect. If so, then these data may not need to derive from more abstract principles or otherwise be "hard-wired" into the learner's initial state. I extracted all NP nodes dominating an S node from P&S's corpus. This yielded all NPs containing a relative clause, but also a number of other constructions, including noun phrases with complement clauses. I then hand-coded each line for whether it contained a relativization dependency, and excluded all lines that did not contain a relativization dependency. Each relative clause was

then coded for whether the dependency resolution site was a gap or a resumptive pronoun (see Chapter 3 for details), whether the gap/resumptive pronoun was inside an island, whether the antecedent of the gap/resumptive pronoun was marked definite or indefinite, and whether the gap/resumptive pronoun occurred in a subject or non-subject position. The results of this corpus study are given in Table 2.5. Recall that the "anti-*that*-trace effect" is the ban of local subject relativization over a null/absent complementizer, i.e., an interaction between subject extraction, construction type, and complementizer type. This corpus search found an abundance of local relativizations over an overt complementizer for both subject and non-subject arguments. However, relativizations over a null/absent complementizer are substantially rarer than relativizations over an overt complementizer, in contrast with complementation complementizers. Supposing that these findings are representative of what the learner hears in 3 years, then she is likely exposed to approximately 290 object extractions, or about 97 object extractions per year, or approximately 8–9 per month. Unlike the *that*-trace effect, this may be sufficiently many observations for the learner to infer the anti-*that*-trace effect in relativization dependencies, and thus this may be directly learnable. In other words, analyses that do not directly extend to the anti-*that*-trace effect may not be problematic, since these effects are presumably directly observable in the learner's input.

Thus, although the anti-*that*-trace effect is likely directly learnable on the basis of the English learner's input, the *that*-trace constraint is not likely to be directly learnable. However, this may not be problematic, if the learner's initial state is sufficiently biased towards grammars that obey the *that*-trace constraint.

In other words, if learners of languages like Spanish or Italian must infer that the *that*-trace constraint does not apply, then the English learner could simply "default" into a grammar with the appropriate properties. This is consistent with conservative learning strategies, such as the "subset principle" (Berwick, 1985; Wexler & Manzini, 1987), which posits that learners opt for more restrictive grammars until confronted with positive evidence for a more permissive grammar. Put differently, if English learners have nothing to learn, then the misleading distributions of sentences in English are actually irrelevant, as long as the learners are not otherwise predisposed to track these data.

2.4.2 Spanish and Italian Child-Direct Speech Corpus Study

The goal of this corpus study was to determine whether insensitivity to the *that*-trace constraint is detectable in the distribution of gaps in Spanish and Italian child-directed speech. Namely, I investigate whether there are a greater number of subject extractions over a complementizer in these languages compared to English. This would be a sufficient cue to learners of these languages to infer that subject extractions are grammatical. If Spanish and Italian child-directed speech does not contain a robust number of subject extractions over a complementizer, then the Spanish and Italian learners must infer insensitivity to the *that*-trace constraint indirectly, i.e., on the basis of post-verbal subjects.

	Corpus	*N*	*Age*	*Total Lines*
Spanish:	Aguirre	1	1;7–2;10	24,866
	BecaCESNo	17	2;0–5;0	14,280
	Marrero-Albalá	3	2;3–4;11	8,963
Italian:	Antelmi	1	2;2–3;4	3,228
	D'Odorico	6	1;04–2;06	11,142
	Roma	1	1;4–4;0	2,957
	Tonelli	3	1;05–2;05	29,519

Table 2.6: Details of the corpora used in the Spanish and Italian corpus study.

2.4.2.1 Methods

We annotated utterances in three corpora of Spanish child-directed speech and four corpora of Italian child-directed speech. DAC annotated the Spanish corpus, and MK annotated the Italian corpus. Our corpora were drawn from the Romance section of the CHILDES database (MacWhinney, 2000). The Spanish corpus consisted of the Aguirre, BecaCESNo, and Marrero-Albalá corpora (Aguirre, 2004; Benedet *et al.*, 2004; Albalá & Marrero, 2004), and the Italian corpus consisted of the Roma, Tonelli, D'Odorico, and Antelmi corpora (Antinucci & Parisi, 1973; Tonelli *et al.*, 1995; D'Odorico & Carubbi, 2003; Antelmi, 2004). Both sets of corpora consisted of naturalistic interactions between parents and children. The details of the corpora are described in Table 2.6.

We first extracted all sentences in the corpora that contained a question word or the relative clause complementizer (*que* in Spanish or *che* in Italian). We excluded all sentences that contained an embedded clause but no long-distance dependency, which for our purposes consisted of *wh*-dependencies and relativization dependencies. For each dependency, we annotated its type (*wh*-dependency or rel-

ativization). We also annotated the grammatical role of the filler, marking each dependency headed as a subject, non-subject argument, adverbial, argument of an existential verb, or argument of a copula. We distinguished adverbials as a separate class, because adverbs are not sensitive to the *that*-trace effect (Lasnik & Saito, 1992), and we distinguished existential and copular verb arguments because the syntactic position of these arguments may be different than canonical subjects. Our analysis in this section focuses on the subject/non-subject argument distinction, and we leave aside analysis of adverb dependencies and copular/existential construction argument dependencies. The resulting corpus and documentation is available on DAC's website.

2.4.2.2 Results

The raw counts for results for the Spanish and Italian corpora are presented in Table 2.7. The Spanish corpus contained 48,109 total utterances, and the Italian corpus contained 46,846 sentences. Each corpus is slightly less than half the size of P&S's corpus. In the Spanish corpus, there were 6,124 total long-distance dependencies, and in the Italian corpus there were 7,235 total long-distance dependencies.

Interestingly, long-distance subject extractions are rare in both Spanish and Italian child-directed speech, similar to English. Assuming that these sentences are representative of the typical learner's experience and that she hears approximately 1 million sentences in three years (Hart & Risley, 1995), Spanish learners hear approximately 120 long-distance subject extractions and Italian learners approximately 20

	Spanish		*Italian*	
Total Utterances:	48,109		46,846	
Filler-Gap Dependencies:	6,124		7,235	
Subject relativization:	296		357	
Argument relativization:	426		227	
	Long	**/ Total**	**Long**	**/ Total**
Argument *wh*-filler:	49	/ 1,415	4	/ 1,490
Subject *wh*-filler:	6	/ 665	1	/ 308

Table 2.7: Occurrences of *wh*-distance dependencies in the Spanish and Italian child-directed speech corpora from CHILDES.

in three years. For Spanish learners, this is equivalent to 40 long-distance subject extractions per year, or 3 to 4 per month, or about one per week. For Italian learners, this is equivalent to 6 to 7 per year, or approximately one every two months. Thus, English learners apparently rarely – if ever – hear sentences that violate the *that*-trace effect, whereas Spanish and Italian learners hear some tokens during the course of language acquisition.

However, it is unclear whether these 120/1 million and 20/1 million observations are sufficient for Spanish or Italian learners to determine that their grammar is distinct from one that generates 0/1 million. In section 2.4.1, I argued that 20/1 million observations of long-distance object extractions over a complementizer are likely insufficient for an English learner to infer the subject-object asymmetry in multiclausal extractions. This extends to the Italian learner as well – the 20/1 million estimated observations of a *that*-trace configuration is likely insufficient for every Italian learner to reliably infer that such extractions are permissible. The facts are somewhat murkier for Spanish, since we have a higher rate of long-distance subject extractions compared to Italian. However, this number is likely artificially inflated,

since the Marrero-Albalá corpus contained five of the six total long-distance subject extractions. Moreover, of these five, three were directed to one child. It is possible that this child's experience was exceptional, and not representative of the typical learner's actual input.

These findings suggest that the distributions of gaps and complementizers in English, Spanish, and Italian are not sufficient for the learner to make the correct inference for each language when exposed to realistic child-directed speech from that language. This strongly implies that learner must indirectly learn the inapplicability of the *that*-trace constraint. Syntactic analyses in which the inapplicability of the *that*-trace constraint is conditioned on some other property is one mechanism by which this can be accomplished.

In section 2.2, I described several accounts that link the inapplicability of the *that*-trace constraint to independently observed properties of Spanish and Italian, namely post-verbal subjects and null subjects. Luckily, these are robustly represented in child-directed speech. To demonstrate this, I sampled 1,000 random sentences without replacement from the Spanish corpus, and annotated each sentence for whether it contained a postverbal subject or a null subject. I excluded copular and existential constructions from this analysis. In this sample of 1,000 sentences, 40 contained postverbal subjects, and 244 contained null subjects. Scaling this to 1 million sentences, this means that the typical learner is exposed to approximately 40,000 postverbal subjects and 244,000 null subjects over the course of three years. Both of these indirect cues are substantially more frequent than the relevant direct cues. Thus, insensitivity to the *that*-trace constraint is likely easily learnable on

the basis of Spanish and Italian learner's input, if the learner's prior grammatical knowledge is such that the insensitivity to the *that*-trace constraint is linked to these phenomena.

The commonality between the English, Spanish, and Italian child-directed speech is that (1) multiclausal extractions are rare, and (2) subject long-distance dependencies are rarer than object long-distance dependencies. Thus, any learning strategy that depends on detecting the presence or absence of multiclausal subject extractions is potentially risky. Importantly, this does not apply to the anti-*that*-trace effect – the input to the typical learner is likely sufficient for her to infer whether relativization complementizers permit subject extraction on the basis of local subject relativizations. Importantly, this requires that the learner does not not infer across constructions – i.e., English learners should not infer that since local subject relativizations over $that_{rel}$ is grammatical, so too is subject extraction over $that_{comp}$.

This is complicated by the *que/qui* contrast in French as described in section 2.2, i.e. subject extraction is licensed over a complementizer if and only if the complementizer is *qui*, in both complementation and relativization constructions. We did not perform a corpus study in French, but suppose that the distributions in French are similar to those in English, Spanish, and Italian, i.e. there are very few subject extractions over an overt complementizer, but plenty of local subject relativizations. If inference from relativization complementizers to complementation complementizers is not permitted, and learning from long-distance *wh*-dependencies over a complementizer is risky, it is unclear how French learners ought to infer that

subject extraction over *qui* is grammatical.

I discuss how these results relate to the specific grammatical accounts, and broader theoretical discussions regarding the *that*-trace constraint and syntactic variation in the next section.

2.5 Wider Theoretical Implications

As described in Chapter 1, the precise characterization of syntactic variation in generative syntax has received renewed interest in recent years (Baker, 2001; Uriagereka, 2007; Newmeyer, 2004, 2005; Roberts & Holmberg, 2005; Bieberauer *et al.*, 2010), as has the appropriate analysis of the *that*-trace effect and the null subject parameter (Pesetsky & Torrego, 2001; Rizzi & Shlonsky, 2007; Chomsky, 2013). The findings in this chapter have some consequences for these issues. In this section, I turn to the ways our approach interfaces with research in syntactic theory.

Parametric theory is a clear framework for describing how the learner might learn variable properties of her grammar. However, as discussed in Chapter 1 and in this chapter, the empirical foundation of parametric analyses, including Rizzi's (1982) null subject parameter. In this chapter, I argued that indirect learning is the only realistic strategy by which the learner can detect difficult-to-observe properties that vary across languages, by conditioning these difficult-to-observe properties on abstractly related easy-to-observe properties. If these connections are not reliable, then this reintroduces the learnability problem that indirect learning is intended to solve. Thus, these abstract relationships must be robust cross-linguistically. For

	yes-yes	*yes-no*	*no-yes*	*no-no*
Referential NS - Nonref. NS	24	0	15	2
Referential NS - Post-VP Subj	22	49	11	15
Referential NS - No *that*-t	5	3	2	1
Nonreferential NS - Post-VP Subj	14	25	1	1
Nonreferential NS - No *that*-t	7	2	0	1
Post-VP Subj - No *that*-t	**4**	**0**	**3**	**4**

Table 2.8: Results from Gilligan's (1987) typological survey. Each row corresponds to a pair of grammatical properties – referential null subjects (NS), non-referential null subjects, post-verbal subjects (Post-VP Subj), and insensitivity to the *that*-trace effect (no *that*-t). Each column corresponds to the number of languages found that feature both properties (yes-yes), only the first (yes-no), only the second (no-no), or neither (no-no). The row corresponding to the *that*-trace strategy is bolded, with the cell corresponding to counterexamples to the *that*-trace learning strategy boxed.

instance, if detecting null subjects or post-verbal subjects does not deterministically imply insensitivity to the *that*-trace constraint, then the learner must still determine whether her language is sensitive to the *that*-trace constraint somehow. In a typological perspective, this means that there should be no language that has post-verbal subjects but still exhibits the *that*-trace effect.

The null subject parameter's typological predictions have been tested in Gilligan's (1987) survey of 100 languages. Gilligan found that many of the properties that are predicted to covary by Rizzi's (1982) null subject parameter are systematically correlated. However, Gilligan also finds that these correlations are rarely without exception. Newmeyer (2004, 2005) cites these counterexamples as evidence against parametric theories. Conversely, Roberts & Holmberg (2005) argue that the presence of these correlations are strong evidence for parametric theories of variation, and the exceptions are explained through other means. The relation between typological generalizations and abstractly related constructions is unclear.

Although Gilligan's study is quite broad and detailed, very few of Gilligan's 100 languages have data on the *that*-trace effect. This is important, because this is arguably the only difficult-to-observe property of the phenomena that are subsumed under Rizzi's (1982) null subject parameter. Interestingly, there are no counterexamples to the generalization that languages with post-verbal subjects also lack sensitivity to the *that*-trace constraint, which is the source of insensitivity to the *that*-trace constraint (Rizzi, 1982; Safir, 1985; Safir & Jaeggli, 1989). Gilligan (1987) finds eight languages that are consistent with this prediction (i.e., four languages exhibit both post-verbal subjects and *that*-trace constraint insensitivity, and four languages lack post-verbal subjects but have *that*-trace constraint sensitivity). Crucially, he finds no languages with post-verbal subjects that are sensitive to the *that*-trace constraint, which would be counterexamples to the generalization that post-verbal subjects deterministically imply *that*-trace constraint insensitivity.

Put differently, all the languages for which Gilligan has data either conform to the generalization that post-verbal subjects imply insensitivity to the *that*-trace constraint, or are insensitive to the *that*-trace constraint through some other means. As far as Gilligan's data indicates, exposure to post-verbal subjects necessarily implies insensitivity to the *that*-trace constraint.

The three languages that Gilligan (1987) finds that are sensitive to the *that*-trace constraint without post-verbal subjects are Basque, Yoruba, and Papiamentu (see also Chao 1981 and Menuzzi (2000) on Brazilian Portuguese). Examining each of these languages is illustrative for clarifying why these cases are not problematic for the *that*-trace learning strategy.

First, the *that*-trace constraint plausibly has no surface realization in Basque because Basque is a head-final language with clitic complementizers. This is illustrated in (29), adapted from Uriagereka (1992). In (29-a), the subject *Jon-ek* 'John-ERG' appears in its canonical clause-initial position. In (29-b), the subject gap is presumably not adjacent to the complementizer, which surfaces as the clitic *-ela* on the embedded auxiliary *du*. If the *that*-trace constraint only applies to adjacent subject gaps and complementizers (e.g. Bresnan 1977), then Basque speakers can easily find a parse that does not violate the *that*-trace constraint.

(29) a. Jon-ek lana egin du
Jon-ERG work-the.ABS make 3.SG.have.3SG
'John did the work'

b. Nor-k esan duzu ___ lan-a egin du-**ela**?
Who-ERG say 3.SG.have.2.SG work-the.ABS make 3.SG.have.3SG-**that**
'Who did you say that did the work?'

Secondly, Yoruba subject extraction avoids violating the *that*-trace constraint by virtue of not permitting subject gaps at all, according to Carstens (1987). In Yoruba, *wh*-movement typically leaves a gap, as observed in English, (30). However, subject extractions in Yoruba are always mediated by a pronoun, *ó*, which may be a "resumptive pronoun" restricted to the subject position (e.g., Engdahl 1985, see Chapter 3). Thus, Yoruba speakers can assign a parse to sentences with subject extractions over a complementizer that does not violate the *that*-trace constraint if the *that*-trace constraint is stated in terms of gaps, because subject gaps are disallowed. Additionally, this resumptive pronoun solution is unlikely to raise the learnability

problems associated with the *que/qui* contrast. This is because monoclausal subject extractions also require a pronominal form in place of a gap, as shown in (31-a). Thus, the learner can infer this feature of Yoruba by observing monoclausal subject questions, and without needing to make any precarious inferences across constructions. In other words, the Yoruba learner does not need to rely on long-distance extractions in sentences with multiple levels of embedding to learn that subject gaps are not permitted, which ought to be directly learnable.

(30) a. Àìná ka ìwé yìí lánǎ
Aina read book this yesterday
'Aina read this book yesterday'
b. Kí ni Àìná kà __ ?
what FOC Aina read
'What did Aina read __ ?'

(31) a. Tani$_i$ *(ó$_i$) ń ko̩rin?
who 3.SG ASP sing?
'Who is singing?'
b. Tani$_i$ Fúnmi sò [$_{CP}$ pé *(ó$_i$) lo̩]?
Who Fumi said that 3.SG go
'Who did Funmi say that left?'

The third language of interest in Gilligan's (1987) study is Papiamentu. The insensitivity to the *that*-trace constraint in Papiamentu is somewhat more difficult to explain. However, Nicolis (2008) shows that Papiamentu is not an isolated case. In a study of nine creole languages, Nicolis shows that six of creole languages are insensitive to the *that*-trace constraint, even though all nine lack post-verbal subjects.[5]

[5]Nicolis notes Haitian Creole as a language with nonreferential null subjects, but with *that*-trace constraint sensitivity. This seems to contradict his generalization that nonreferential null subjects are an exceptionless cue to *that*-trace constraint insensitivity. However, this seems to be an analytic choice. Déprez (1994) shows that the complementizer *ki* is only permitted with subject extraction, (32). Nicolis assumes that *ki* is an "agreement-complementizer" akin to French *qui*, and it is in this sense that this language is sensitive to the *that*-trace constraint. In either case, it appears that

	Non-Referential Null Subj	*Postverbal Subj*	No *that*-t	*Lexifier*
Berbice Dutch Creole	+	−	+	Dutch
Cape Verdean Creole	+	−	+	Portug.
Haitian Creole	+	−	−	French
Jamaican Creole (basilect)	+	−	−(?)	English
Jamaican Creole (mesolect)	?	−	−	English
Kriyol	+	−	+	Portug.
Jamaican Creole Creole	+	−?	+	French
Papiamentu	+	−	+	Portug. Spanish Dutch
Saramaccan	+	−(?)	+	Portug. Spanish

Table 2.9: Findings from Nicolis's (2008) study on creole languages. Each row corresponds to a language, and each column corresponds to a grammatical property – non-referential null subjects (non-referential null subj), postverbal subjects (postverbal subj), and insensitivity to the *that*-trace constraint (no *that*-t).

These results are summarized in Table 2.9.

Nicolis (2008) claims that null expletives are a more reliable predictor of *that*-trace insensitivity instead of post-verbal subjects. For him, this is because null expletives alone can fill the canonical subject position in case of subject extraction, as in Rizzi's (1982) "skipping strategy" analysis. He also suggests that this proposal

subject extraction over a complementizer is permitted, which is consistent with Nicolis' prediction that null expletive languages uniformly are insensitive to the *that*-trace constraint.

(32) a. Kimown ou kwè (*ki) Mariz te wè a?
Who 2.Sg think that Mariz Past see
'Who do you think that Mariz saw?
b. Kimown ou kwè (ki) pati?
Who 2.Sg think (that) left?
'Who do you think that left?'

captures much of Gilligan's (1987) original data, once the distinction between null expletives vs. null pronouns are carefully distinguished. He agrees that there are likely several ways to avoid violating the *that*-trace constraint, including extraction from a postverbal subject position. Whether his solution for these creole languages generalizes without raising additional learnability problems requires further research, however.

2.6 Conclusion

In this chapter, I argued that the learner must indirectly learn whether her language is sensitive to the *that*-trace constraint, because the distributions of gaps and complementizers in realistic linguistic input across languages are underinformative. This implies that learners must have access to an inferential mechanism by which they can learn whether their language is sensitive to the *that*-trace constraint on the basis of some other linguistic property that is detectable in the input. Certain analyses of the *that*-trace constraint insensitivity in Spanish and Italian transparently provide the learner with such an inferential mechanism, e.g. Rizzi (1982) proposes that post-verbal subjects imply insensitivity to the *that*-trace constraint. This contrasts with an analysis in which the learner must determine for each complementizer whether it permits subject extraction, which I argue is a reasonable extension of analyses that allow individual complementizers to permit subject extraction.

Additionally, this chapter provides a new perspective to debates on grammatical variation and parameter theory, as described in Chapter 1. Much of this debate

has focussed on the formal character of linguistic variation, to what extent the dimensions of grammatical variation are part of the learner's innate linguistic endowment, and how reliably these approaches describe linguistic typology (Newmeyer, 2004, 2005; Roberts & Holmberg, 2005). However, both sides of this debate have remained relatively silent on what role competing theories play in explaining successful language acquisition. In my view, this is perhaps the domain in which approaches to grammatical variation make the most concrete predictions, and also have the most at stake.

In the next chapter, I apply this reasoning to the distributional properties of resumptive pronouns inside islands in English and Hebrew. Combined with novel results in the psycholinguistics of resumptive pronouns, I argue that the learner must infer that resumptive pronouns inside island contexts are grammatical when exposed to Hebrew data, which I argue requires indirect learning.

Chapter 3: Finding Parses: Active Dependency Formation and Resumptive Pronouns in English and Hebrew

3.1 Introduction

In Chapter 2, I demonstrated how syntactic analyses can be constrained by learnability concerns and the distributional properties of realistic child-directed speech. I argued that comparative syntactic analyses can be interpreted as licensing conditions on inferences across constructions, and that certain cross-construction inferences were necessary to explain how the learner determines the grammatical properties of her language on the basis of her linguistic experience. In this chapter, I extend this analytic framework to the phenomenon of resumptive pronouns in English and Hebrew. Here, I show that the mechanisms used for constructing resumptive dependencies in English and Hebrew are different. I argue that resumptive dependencies are grammatical in island contexts in Hebrew, but not in English. I also show that this likely only learnable through indirect learning, in the sense of Chapter 2.

Resumptive pronouns are pronouns that appear as the tail end of a filler-

gap dependency under certain conditions, as demonstrated in Hebrew in (1) and in English in (2). In (1-a), the relativization dependency headed by *ha-iš* resolves with a gap, whereas in (1-b) it resolves with the pronoun *oto*. In the English example, the relativization dependency headed by *the things* similarly resolves with the pronoun *they*.

(1) a. **ha-iš** še ra'iti ___
the-man that I saw ___
b. **ha-iš** še ra'iti oto
the-man that I saw him

(2) a. ? These are **the things** that we don't know what they are
b. * These are **the things** that we don't know what ___ are

Like the *that*-trace effect, the properties of resumptive pronouns vary across languages (Sells, 1984; McCloskey, 2006; Asudeh, 2012). In this chapter, I focus on the contrast between resumptive pronouns in Hebrew and English. In Hebrew, resumptive pronouns freely vary with gaps outside island constructions in Hebrew, as demonstrated in (1). Conversely, resumptive pronouns are typically reserved for island constructions in languages like English[1]. Hebrew is often described as a language with "grammaticized" resumptive pronouns, unlike English.

In this chapter, I argue that resumptive dependencies in island constructions in English and Hebrew display distinct processing profiles in real-time comprehension. Specifically, I argue that resumptive pronoun dependencies in English are constructed in a pronoun-driven fashion, as defined in section 3.2.2. This contrasts

[1]However, this is not always the case, see Ariel (1990). As dependencies increase in length and complexity, the availability of resumptive pronouns in English increase (Dickey, 1996; Alexopoulou & Keller, 2003). I will largely focus on the availability of resumptive pronouns in island contexts in English, since this is the environment in which their properties have been studied most carefully.

with recent findings in Hebrew (Keshev & Meltzer-Asscher, ms), which suggest that resumptive dependencies are constructed actively. Critically, I contend that real-time processing profiles are unobservable. Thus, the Hebrew learner must infer that she should actively construct resumptive dependencies on the basis of some property of Hebrew linguistic input, and the English learner must infer the right strategy for constructing resumptive dependencies on the basis of English linguistic input. These processing strategies must be indirectly learned, due to the unobservability of processing profiles.

I propose that this difference arises because resumptive pronouns in island contexts are grammatical in Hebrew, but are ungrammatical "repair-strategies" in English, to be described in section 3.2.2. If resumptive pronouns are grammatical in island contexts in Hebrew, but not English, then the learner must infer this on the basis of her linguistic experience. I argue that the learner of Hebrew must indirectly learn this, as in Chapter 2. I argue that this must be indirectly learned in Hebrew, by observation of resumptive pronouns outside island contexts. I argue that not all theories of resumptive pronouns in Hebrew necessarily license this inference. Thus, the combination of the learnability concerns and real-time comprehension data limit the space of syntactic analyses. Specifically, I argue that this data favors analyses in which resumptive pronoun dependencies in Hebrew, inside and outside of island constructions, must be a "natural class", to the exclusion of resumptive pronoun dependencies in English.

The structure of this chapter is largely the same as in Chapter 2. In section 3.2, I briefly describe the properties and analyses of resumptive pronouns. In section

3.3, I describe a series of experiments on the processing of resumptive pronouns, and describe how it contrasts with recent findings in the psycholinguistics of resumptive pronouns in Hebrew (Keshev & Meltzer-Asscher, ms). Afterwards, I describe the results of corpus studies conducted in English and Hebrew in section 3.4. Finally, in section 3.5 I discuss the implications of these findings for syntactic theory, and then conclude.

3.2 Dimension of Variation: Distribution of Resumptive Pronouns

The characterization of resumptive dependencies has been of interest in theoretical linguistics since Ross (1967), and more recently in the psycholinguistic literature. Traditionally, in the syntax literature, resumptive pronouns are described as amnestying movement violations across languages. However, the nature of this island amnestying and the precise description of the cross-language differences has remained controversial. In psycholinguistics, the focus has been on whether resumptive pronouns are grammatical constructions, or whether they ungrammatical strategies used to facilitate comprehension or production. In this section, I briefly survey the properties of resumptive pronouns, and a number of syntactic and psycholinguistic analyses.

3.2.1 Syntax/Semantics of Resumptive Pronouns

The syntactic and semantic properties of resumptive pronouns vary across languages and across constructions within the same language (McCloskey, 2006;

Asudeh, 2012). Additionally, the full range of variation across all languages and constructions has not been systematically mapped out. For this reason, I restrict my attention to distributional properties of resumptive pronouns, leaving aside many important syntactic and semantic properties. I discuss some of those properties in this section, but will largely set them aside in the rest of the chapter.

The data that I focus on are given in (3)–(6). The sentences in (3), copied from (1) above, show that in relativizations in Hebrew, resumptive pronouns are in "free variation" with gaps outside islands. Conversely, resumptive dependencies are not acceptable tails for relativization dependencies outside islands in English, shown in (4).

(3) a. **ha-'iš**-še ra'iti ___
the-man that-I.saw ___
b. **ha-'iš** še-ra'iti 'oto
the-man that-I.saw him

(4) a. **The man** that I saw ___
b. *__The man__ that I saw him

By definition, gaps are not available in island contexts in either English or Hebrew. However, relativization dependencies resolving with a resumptive pronoun are described as acceptable in Hebrew, shown in (5). In English, relativization dependencies may resolve with resumptive pronouns in island constructions with some improved status in English, shown in (6).

(5) a. *ra'iti 'et **ha-yeled** še-Dalya makira 'et ha-iša še-ohevet
I.saw ACC **the-boy** that-Dalya knows ACC the-woman that-loves
___
___

b. ra'iti 'et **ha-yeled** še-Dalya makira 'et ha-iša še-ohevet
I.saw Acc **the-boy** that-Dalya knows Acc the-woman that-loves
oto
him

(6) a. *These are **the things** that we don't know what __ are
b. (?) These are **the things** that we don't know what they are

The question that I address in this section is whether the resumptive dependencies in (3)–(6) are formally similar – i.e., do they have the same grammaticality status and the same representational properties. I argue that resumptive pronouns in Hebrew should be analyzed as fundamentally the same grammatical construction, to the exclusion of resumptive pronouns in English, which are ungrammatical "repair strategies".

Surprisingly, there are few few syntactic treatments of resumptive pronouns that make specific commitments to the full range of data in (3)–(6), with a few exceptions (Asudeh, 2012). However, I argue that there are roughly three categories of analyses. One category treats resumptive pronouns in Hebrew as fundamentally "the same" at a coarse level of analysis, to the exclusion of English resumptive pronouns, regardless of whether they occur in island contexts. Another category treats resumptive pronouns in island contexts as fundamentally "the same", to the exclusion of resumptive pronouns outside of islands. Finally, the third category treats all resumptive pronouns as fundamentally the same. The specific analyses in each of these categories can be motivated by a wide array of theoretical and empirical concerns, but are similar at the relevant level of abstraction. In this section, I sketch a few notable accounts in each category before turning to psycholinguistic accounts

of resumptive pronouns.

3.2.1.1 Resumptive Pronouns Are Different Between Languages

Chomsky (1977) provides an account of resumptive pronouns in Hebrew that is largely compatible with an analysis in which resumptive pronouns are representationally the same. Chomsky (1977) proposes that only filler-gap dependencies are movement dependencies, or island-sensitive relations between fillers ("operators") and traces/gaps. On this view, resumptive dependencies inside islands in Hebrew, as in (5), necessarily cannot be movement dependencies. Instead, he proposes that resumptive dependencies in Hebrew are a binding relation (Hayon, 1973), a fundamentally different syntactic relation than movement.

Chomsky (1977) suggests that resumptive dependencies in all languages and across all constructions must necessarily be operator-pronoun binding relations. It is unclear what the between-language difference is on his analysis. However, it is possible to extend Chomsky's (1977) framework to account for why resumptive pronouns are ungrammatical outside islands in English, but are grammatical in Hebrew. If relativization dependencies that resolve in gaps alone are movement dependencies, then it follows that resumptive dependencies outside islands in Hebrew are not movement dependencies. Thus, Hebrew permits island-insensitive filler-pronoun binding. However, generalized operator-pronoun is disallowed in English, as evidenced by the absence of resumptive pronouns outside islands in English. Thus, resumptive dependencies in islands in English must have a different representational status,

either as contextually-defined operator-pronoun binding, or as a formative that is not grammatically sanctioned, as described in section 3.2.2. Put differently, on this extension of Chomsky's (1977) analysis, the absence of general operator-pronoun binding in English outside island contexts is an indication that operator-pronoun bounding inside islands are not grammatically sanctioned in the same way that generalized operator-pronoun binding is in Hebrew.

3.2.1.2 Resumptive Pronouns are Different Across Constructions

The most common analyses posit that resumptive pronouns are formally similar inside islands across languages, but formally different inside and outside islands in Hebrew. Resumptive pronouns in islands in English and Hebrew have the same representational character, but resumptive pronouns outside islands in Hebrew receive an alternative treatment.

Ross (1967) provides the first analysis of resumptive pronoun dependencies in English. Ross (1967) argues that movement dependencies are all relations between a filler and a pronoun at an abstract level of representation. In English, there is an obligatory "chopping" rule that deletes the pronoun, yielding a gap. For him, island constraints are not constraints on movement dependencies at all, but on chopping. On this view, resumptive pronouns in islands in English and Hebrew can be analyzed as typical filler-gap dependencies with suppressed chopping. Ross's (1967) analysis makes no commitment to the nature of resumptive pronouns outside islands in Hebrew, but some language-specific grammatical construct must be responsible for

these.

Later work on resumptive pronouns relies on Chomsky's (1995) notion of "Last Resort", or an economy condition that prioritizes universally available grammatical operations (i.e., movement) over language-specific grammatical operations (i.e., resumption). For instance, Aoun *et al.* (2001) propose that resumptive pronouns outside islands in Lebanese Arabic display a number of properties diagnostic of movement dependencies outside islands, but resumptive pronouns inside islands do not display these properties. On their analysis, resumptive pronouns are essentially markers of the gap site outside islands. True resumption, i.e., operator-pronoun binding, is reserved to island contexts. This implies that in grammaticized resumptive languages, there is a formal distinction between pronouns inside and outside islands, which may in principle be independent. A similar proposal on the basis of very different syntactic and semantic evidence is made by (Demirdache & Percus, 2011).

Erteschik-Shir (1992) also proposes that resumptives inside islands have a universal representation, whereas resumptive pronouns outside islands in Hebrew have a different analysis. She follows Doron (1982) in suggesting that resumptive pronouns have specific semantic/pragmatic restrictions that do not apply to gaps in Hebrew. However, Ariel (1990) shows that many of these semantic/pragmatic restrictions disappear in particularly complicated or long filler-resumptive dependencies. Thus, Erteschik-Shir (1992) argues that Hebrew has two resumptive pronouns – a grammaticized resumptive pronoun that has specific semantic/pragmatic restrictions, and an ungrammatical resumptive pronoun that is used in island contexts and

other taxing constructions (Sells's (1984) "intrusive pronoun", see 3.2.2).

I argue against these analyses, because I argue in section 3.3 that resumptive pronouns in islands in English and Hebrew display divergent processing profiles. These differences can be expediently explained by positing that resumptive pronouns in islands are grammatical in Hebrew, but not in English. Furthermore, I argue that resumptive pronouns inside and outside islands in Hebrew must be representationally linked. This is because, as I argue in section 3.4, the only feasible way a learner is likely to infer that resumptive pronouns are grammatical inside islands is to observe them outside islands. This inference is licensed in analyses where resumptive dependencies are representationally the same across constructions, e.g. generalized operator-pronoun binding.

3.2.1.3 Resumptive Pronouns are Universally Similar

Shlonsky (1992) offers an analysis in which resumptive pronouns inside and outside islands have the same representational properties across constructions. Like Aoun *et al.*'s (2001), he extends Chomsky's (1995) "Last Resort" analysis of resumptive pronouns, i.e., language-specific strategies like resumptive pronouns are only permitted when universal properties like movement are disallowed. He notes that in Palestinian Arabic, another grammaticized resumptive language, relativization dependencies obligatorily resolve with a resumptive pronoun. He proposes that this is because the relativization complementizer in Palestinian Arabic indirectly blocks movement, requiring resumptive pronouns to be used in constructions that

otherwise would not be islands. He extends this analysis to apparent optionality in languages like Hebrew, and proposes that Hebrew has two phonologically identical complementizers, one which always blocks movement like in Palestinian Arabic, and one which permits movement like in English. In effect, Shlonsky argues that all resumptive pronouns in Hebrew amnesty movement violations, but the number of constructions that block movement is greater in Hebrew than in English.

This subsection only provides a cursory survey of the theoretical and empirical concerns that analyses of resumptive pronouns have addressed. Some analyses stress the importance of the superficial form of resumptive dependencies, which treat resumptive dependencies as either uniform across languages and constructions, or uniform in "grammaticized resumptive pronoun" languages like Hebrew. Conversely, other analyses attempt to explain the differences in syntactic and semantic properties between constructions within the same language by appealing to representational differences within the same language. In this chapter, I argue that learnability and processing concerns suggest that resumptive dependencies in Hebrew should be analyzed in the same way, but resumptive dependencies in English should not be analyzed as grammatical. Instead, English resumptive dependencies should be analyzed as ungrammatical repair mechanisms, as proposed in recent work in psycholinguistics. In the next section, I briefly describe the extant psycholinguistic data on resumptive pronouns in both English and Hebrew, and the different analyses of resumption as a repair mechanism.

3.2.2 Psycholinguistics of Resumptive Pronouns

The syntactic analyses described above presume that resumptive dependencies are well-formed grammatical representations in both English and Hebrew. However, there is a parallel discussion in the psycholinguistic literature investigating whether resumptive dependencies in English are grammatically well-formed or ungrammatical repair mechanisms that the comprehension or production mechanisms use in exceptional circumstances. These views largely stem from earlier work that argues that resumptive pronouns are not grammatical in English, because they have unusual formal properties compared to grammaticized resumptive languages (Kroch, 1981; Sells, 1984; Cresswell, 2002). First, I survey the psycholinguistic literature for English resumptive pronouns, and then for Hebrew resumptive pronouns. Afterwards, I outline three hypotheses for how the comprehension systems might construct resumptive dependencies in real time.

3.2.2.1 Psycholinguistics of English Resumptive Pronouns

Chao & Sells (1983) and Sells (1984), following Kroch (1981), argue that resumptive dependencies in English are ungrammatical repair strategies that the production mechanisms to use to repair a plan in which a filler-gap dependency resolves in an island (Zukowski & Larsen, 2004; Ferreira & Swets, 2005; Asudeh, 2012). Chao & Sells (1983) argue that resumptive dependencies are not grammatical in English because they display a number of syntactic and semantic properties that are not observed in grammaticized resumptive pronoun languages. For instance, they note

that resumptive pronouns in English cannot take a quantificational antecedent in English, (7), but may in Hebrew, (8).

(7) a. I met **the linguist** that Kate forgot if Thora had seen him before
b. *I met **no linguist / every linguist** that Kate forgot if Thora had seen him before.

(8) **Kol gever** še-Dina xoševet še-hu ohev et Rina.
Every man that-Dina thinks that-he loves ACC Rina
'Every man that Dina thinks that he loves Rina'

According to Chao & Sells (1983), this demonstrates that resumptive pronouns in Hebrew are mediated by a true syntactic binding relation, whereas the relation between the filler and the resumptive in English is not syntactic. Instead, the relation between the filler and the resumptive pronoun is constructed via the discourse model, i.e., outside of the syntactic representation. They argue that this explains the contrast in (7)–(8) because there is no discourse representation for quantificational antecedents like *no linguist* or *every linguist*.

Later work supports the hypothesis that resumptive pronouns in English are used to repair early production plans in which filler-gap dependencies resolve in islands (Zukowski & Larsen, 2004; Ferreira & Swets, 2005). For instance, Ferreira & Swets (2005) argue that when a speaker produces a resumptive pronoun, they initially commit to a plan in which a filler-gap dependency resolves in an island construction, and then revise the plan in real-time to include a resumptive pronoun. They show that participants can be primed into producing sentences with a filler-gap dependency resolving into an island context, in which they overwhelmingly produced resumptive pronouns in these environments. Importantly, there was an increase in

onset times earlier in the sentence compared to control conditions. These findings are consistent with the claim that resumptive dependencies are repair strategies used by production mechanisms to avoid producing utterances with gaps in islands.

Additionally, controlled acceptability studies suggest that the extent to which resumptive dependencies remedy an island violation in English has been overstated, consistent with the claim that they are ungrammatical constructions in English. For instance, Alexopoulou & Keller (2007) show that resumptive pronouns inside islands are rated as low as gaps, and even lower in many constructions. This has been widely replicated (McDaniel & Cowart, 1999; Omaki & Nakao, 2010; Heestand *et al.*, 2011). These findings also imply that resumptive dependencies are not grammatical constructions, consistent with the analysis in which English resumptive dependencies are ungrammatical repair strategies.

However, the precise interpretation of these results is unclear. Ackerman *et al.* (2014) find that resumptive dependencies are accepted at a greater rate than their gapped alternatives in a forced-choice task, which contradicts the claim that resumptive dependencies are less acceptable than their gapped counterparts. Similarly, Beltrama & Xiang (2013) find that resumptive pronouns were rated higher than their gapped counterparts in island contexts when participants were asked to rate the comprehensibility of a sentence, in contrast to the acceptability or grammaticality of the sentence. Keffala (2013) suggests that the decreased acceptability of resumptive dependencies inside islands is partly due to the reanalysis processes involved, which increase costs and therefore lower acceptability (Sprouse, 2008). Since the comprehension system typically avoids constructing filler-gap dependen-

cies in island contexts, as described in Chapter 1, a resumptive dependency requires suppressing an initial preference to resolve the dependency outside of the island. Thus, the findings from controlled acceptability judgment tasks do not obviously reveal that resumptive dependencies are ungrammatical in English.

Interestingly, there is one systematic exception to the generalization that resumptive dependencies are rated lower than their gapped counterparts. McDaniel & Cowart (1999) show an improvement for resumptive pronouns in subject position in an island immediately adjacent to a complementizer compared to gaps. They argue that this is because resumptive pronouns not only repair an island violation in this position, but also a *that*-trace constraint violation, as discussed in Chapter 2 (Alexopoulou & Keller, 2007; Omaki & Nakao, ms). However, if resumptive pronouns are ungrammatical in English, it is not clear why one ungrammatical structure should be preferred to another (but see Omaki & Nakao ms).

Even if resumptive dependencies are ungrammatical, English speakers can clearly assign an interpretation to them. However, they has been very little work on the real-time comprehension of resumptive pronouns in English. In a self-paced reading task, Hofmeister & Norcliffe (2013) show that participants read resumptive pronouns more quickly in longer filler-gap dependencies (9) compared to shorter filler-gap dependencies, shown in (10). Additionally, they find that subsequent regions are read more quickly than in their gapped alternatives immediately after the resumptive pronoun.

(9) Mary confirmed that there was **a prisoner** who the prison officials had acknowledged that the guard helped { him / __ } to make a daring escape.

(10) The prison officials had acknowledged that there was **a prisoner** that the guard helped { him / __ } to make a daring escape.

Hofmeister & Norcliffe (2013) argue that this demonstrates that the comprehension mechanisms can rapidly use resumptive pronouns. However, these data say do not clarify the mechanisms by which the comprehension mechanisms link a filler with a resumptive pronoun, only that they are constructable in real-time and can affect reaction times accordingly. The finding that resumptive dependencies can be constructed in real-time is not incompatible with the claim that they are ungrammatical repair strategies. I describe three hypothetical mechanisms for constructing resumptive dependencies in section 3.2.2.3, that I test in section 3.3. However, before outlining these hypotheses, I describe the psycholinguistic work on resumptive dependencies in Hebrew.

3.2.2.2 Psycholinguistics of Hebrew Resumptive Pronouns

Like in English, recent work suggests that the acceptability of resumptive dependencies in Hebrew has also been overstated. Farby *et al.* (2010) showed that resumptive pronouns are rated worse than their gapped counterparts outside island contexts. Additionally, they find that resumptive dependencies are rated more highly than their gapped counterparts inside islands, unlike the English findings. Interestingly, this improvement is very slight. However, there are complications with these studies. For instance, it is unclear whether Farby *et al.*'s (2010) ungrammatical fillers were appropriately complexity-matched. This could change how

participants used the scale, artificially skewing the results towards the low end of the scale.

Keshev & Meltzer-Asscher (ms) correct for these oversights in an acceptability judgment task. They compare the ratings of resumptive pronouns and islands in two different island constructions, the coordinate structure islands, shown in (11), and complex noun phrase islands, shown in (12). In both sentences, there is a relativization dependency headed by *ha-iša* 'the woman', which resolves with either a resumptive pronoun (*ota*) or a gap. The islands are demarcated with square brackets. They find that ratings improve for sentences containing resumptive pronouns in complex noun phrase constructions compared to their gapped alternatives as in (12), but not for resumptive dependencies in coordinate structure constructions as in (11). This suggests that resumptive dependencies can amnesty complex noun phrase island violations, but not coordinate structure island violations.

(11) a. Ha-šotrim hekiru et **ha-iša** še-ha-xašudim [daxafu et
the-cops knew Acc **the-woman** that-the-suspects [pushed Acc
ha-melcar ve takfu ota be-mis'ada yukratit]
the-waiter and attacked her in-restaurant upscale]
'The cops knew **the woman** that the suspects [pushed the waiter and attacked her in an upscale restaurant]'

b. Ha-šotrim hekiru et **ha-iša** še-ha-xašudim [daxafu et
the-cops knew Acc **the-woman** that-the-suspects [pushed Acc
ha-melcar ve takfu __ be-mis'ada yukratit]
the-waiter and attacked __ in-restaurant upscale]
'The cops knew **the woman** that the suspects [pushed the waiter and attacked __ in an upscale restaurant]'

(12) a. Ha-šotrim hekiru et **ha-iša** še-ha-xašudim [še-takfu
The-cops knew Acc **the-woman** that-the-suspects [that-attacked
ota] daxafu et ha-melcar be-mis'ada yukratit
her] pushed Acc the-waiter in-restaurant upscale
'The cops knew **the woman** that the suspects [that attacked her] pushed

the waiter in an upscale restaurant'

b. Ha-šotrim hekiru et **ha-iša** še-ha-xašudim [še-takfu
The-cops knew Acc **the-woman** that-the-suspects [that-attacked
___] daxafu et ha-melcar be-mis'ada yukratit
___] pushed Acc the-waiter in-restaurant upscale
'The cops knew **the woman** that the suspects [that attacked ___] pushed the waiter in an upscale restaurant'

Keshev & Meltzer-Asscher also examine the time course of resumptive dependency formation in these sentences. They performed a self-paced reading task with a filled-gap effect paradigm with sentences like (13) and (14). In these sentences, the critical conditions contained a relativization dependency that ultimately resolves with a resumptive pronoun. In between the relativization head and the resolution site, there are two potential resolution sites that filled by a noun phrase, one embedded in an island and one outside the island. In (13-a), the relativization headed by *ha-iša* 'the woman' resolves with the resumptive pronoun *ota*. The two regions of interest are the filled gap inside the complex noun phrase island, *ha-melcar* 'the waiter', and the filled-gap after the complex noun phrase island, *ha-tabax* 'the cook'. The control, given in (13-b), lacks a relativization dependency. In (14-a), the relativization headed by *ha-iša* 'the woman' again resolves with the resumptive pronoun *ota* 'her'. The first region of interest is the filled-gap *ha-melcar* 'the water', which the object of the verb in the first conjunct verb phrase, and the next filled-gap *ha-tabax* 'the cook' is the object of the verb in the second conjunct verb phrase. Although both of these filled gaps are embedded in a coordinate structure island, the comprehension system has no evidence that the first filled-gap effect is embedded in the island until encountering the conjunct *ve* 'and'. Keshev & Asscher-Meltzer find

that there is an increased reading time at the filled gap *ha-tabax* 'the cook' after the complex noun phrase in (13-a), and at the filled gap *ha-melcar* 'the waiter' before the evidence of the island in (14-a). This indicates that the comprehension system attempts to resolve the relativization dependency with these positions, which would be grammatically well-formed resolution sites. Interestingly, there is also increased reading times at the filled gap region *ha-melcar* 'the waiter' in the complex noun phrase island in (13-a), but not in the filled gap *ha-tabax* 'the cook' embedded in the coordinate structure island in (14-a). In other words, Keshev & Asscher-Meltzer find increased reading times in positions in which filler-gap dependencies can grammatically resolve, and also in positions where resumptive pronouns are rated more highly than their gapped counterparts in their acceptability judgment tasks. Keshev & Asscher-Meltzer interpret this as indicating that the comprehension systems actively attempt resolving relativization dependencies with positions in which resumptive pronouns are considered acceptable. They interpret these findings as suggesting that resumptive dependencies are actively pursued when acceptable, specifically in islands like complex noun phrase island constructions.

(13) a. Ha-šotrim mekirim et **ha-iša** še-ha-xašud [še-takaf
the-cops know ACC **the-woman** that-the-suspect [that-attacked
et ha-melcar] kilel et ha-tabax axrey še-hu daxaf
ACC the-waiter] cursed ACC the-cook that-he pushed her
ota be-misada yukratit be-Tel Aviv
in-restaurant upscale in-Tel Aviv
'The cops know the woman that the suspects that attacked the waiter cursed the cook that he pushed (her) in an upscale restaurant in Tel Aviv'

b. Ha-šotrim imtu et ha-divuax še-ha-xašud [še-takaf
the-cops verified ACC the-report that-the-suspect [that-attacked

et ha-melcar] kilel et ha-tabax axrey še-hu daxaf et
ACC the-waiter] cursed ACC the-cook before that-he pushed ACC
ha-iša be-misada yukratit be-Tel Aviv
the-woman in-restaurant upscale in-Tel Aviv
'The cops' verified the report that the suspect that attacked the waiter cursed the cook before he pushed the woman in an upscale restaurant in Tel Aviv'

(14) a. Ha-šotrim mekirim et **ha-iša** še-ha-xašud [takaf et
the-cops know ACC **the-woman** that-the-suspect [attacked ACC
ha-melcar ve-kilel et ha-tabax] axrey še-hu daxaf ota
the-waiter and-cursed ACC the-cook] before that-he attacked her
be-misada yukratit be-Tel Aviv
in-restaurant upscale in-Tel Aviv
'The cops know the woman that the suspect attacked the waiter and cursed the cook before he attacked (her) in an upscale restaurant in Tel Aviv'

b. Ha-šotrim imtu et ha-divuax še-ha-xašud [takaf et
the-cops verified ACC the-report that-the-suspect [attacked ACC
ha-melcar ve-kilel et ha-tabax] axrey še-hu daxaf et
the-waiter and-cursed ACC the-cook] before that-he attacked ACC
ha-iša be-misada yukratit be-Tel Aviv
the-woman in-restaurant upscale in-Tel Aviv
'The cops verified the report that the suspect attacked the waiter and cursed the cook before he attacked the woman in an upscale restaurant in Tel Aviv'

This suggests an important contrast between the English and Hebrew parsers. In Hebrew, the parser actively constructs relativization dependencies in positions in which gaps are grammatically licensed, and also in islands in which resumptive pronouns are considered acceptable. However, the mechanisms of filler-gap dependency formation in English appears to be different. Resumptive dependencies are largely restricted to island constructions in English, as discussed earlier. However, there is ample evidence that shows that English speakers do not attempt filler-gap dependency resolution in island contexts, as discussed in Chapter 1. Thus, there appears

to be a contrast in processing profiles between English and Hebrew. I contend that this is a non-trivial acquisition problem, since processing profiles are not observable to the learner. Thus, the English learner must learn to not actively resolve filler-gap dependencies with resumptive pronouns on the basis of English input, and the Hebrew learner must learn to actively resolve filler-gap dependencies with resumptive pronouns on the basis of Hebrew input. I argue that these processing data follow if resumptive dependencies inside islands are ungrammatical in English, but are grammatical in Hebrew. If the set of representations that the comprehension system can actively construct are a subset of the set of grammatically well-formed representations, then the English comprehension system will not actively construct resumptive dependencies in island contexts, whereas the Hebrew comprehension system can.

This requires a more careful investigation into how resumptive dependencies are formed in English, if they are not constructed actively. In the next section, I outline three possible hypotheses for how the English parser might achieve this without actively constructing resumptive dependencies. In Experiments 3 and 4, I show that resumptive dependencies are constructed in a "pronoun-driven" fashion, as I define in the next section. Afterwards, in section 3.4, I describe how the learner might indirectly learn that resumptive pronouns in islands are grammatical when exposed to Hebrew data, but not when exposed to English data.

3.2.2.3 Mechanisms of Resumptive Dependency Formation

In this section, I articulate three possible hypotheses for how the comprehension system might construct a resumptive dependency in real time. The first hypothesis that I consider contends that resumptive dependencies are formed just when the comprehension system detects that there are no possible grammatical gap positions available, most likely at the end of the sentence. I call this the end of sentence hypothesis. The second hypothesis contends that the comprehension system constructs a resumptive dependency immediately upon encountering a pronoun in an island when there is an open filler-gap dependency. On this view, the comprehension system ranks resolving an open dependency with a pronoun over waiting for a later grammatically licensed gap site. I call this the filler-driven hypothesis. Lastly, the third hypothesis contends that the resumptive dependency is only created when the pronoun selects the open filler-gap dependency head as its antecedent. This attributes resumptive dependency formation processes to typical pronoun-antecedent dependency formation mechanisms. I call this the pronoun-driven hypothesis. In section 3.3, I argue for the pronoun-driven hypothesis.

To demonstrate these three hypotheses, consider the resumptive pronoun *they* which resumes the dependency headed by *the things* in (15), repeated from (4). For each of the three hypotheses under consideration, I describe the steps by which the comprehension mechanisms construct the resumptive dependency in (15), highlighting the differences in time course predictions that these hypotheses make.

(15) These are **the things** that we don't know what they are.

First, consider the error-driven hypothesis. On this hypothesis, the resumptive dependency is only considered when the comprehension system detects no gap for the filler-gap dependency. Upon encountering the cue to build the relativization dependency (*the things that*), the comprehender constructs an abstract filler-gap dependency in advance of any confirming bottom-up evidence, indicated here by (__). Upon encountering the island boundary *what*, active dependency formation is suppressed. By the end of the sentence, the filler-gap dependency is still open, and thus the comprehender must reanalyze the sentence so that the dependency resolves with the pronoun. This is illustrated in (16). This hypothesis predicts that resumptive dependencies are only constructed when forced, and not necessarily at the resumptive pronoun site.

(16) **End of Sentence Hypothesis**:

a. *Open filler-gap dependency:*
These are **the things** that … (__)

b. *Encounter island; suppress active dependency formation:*
These are **the things** that we don't know what …

c. *Encounter end of sentence; detect unresolved filler-gap dependency:*
These are **the things** that we don't know what they are. (__ ?)

d. *Reanalyze pronoun as resumptive:*
These are **the things** that we don't know what they are.

In contrast, the filler-driven hypothesis holds that the comprehender attempts resolution with the pronoun immediately. In (15), the filler-gap dependency is opened upon encountering the cue to build the relativization dependency (*the things that*), and then again suppressed upon encountering the island boundary (*what*).

However, upon encountering the pronoun, the gap prediction is immediately reactivated, and the filler-gap dependency resolves with the pronoun. This is illustrated in (17).

(17) **Filler-driven Hypothesis**:

a. *Open filler-gap dependency:*
These are **the things** that … (___)

b. *Encounter island; suppress active dependency formation:*
These are **the things** that we don't know what …

c. *Encounter pronoun, construct resumptive dependency:*
These are **the things** that we don't know what they …

d. *Finish sentence:*
These are **the things** that we don't know what they are.

On this hypothesis, resumptive dependencies are ranked above filler-gap dependencies. In other words, the comprehender immediately discharges the dependency upon encountering a potentially resumptive pronoun, instead of anticipating a grammatically licensed gap outside of the island. Importantly, this is different from active dependency formation, since the resumptive dependency is not actively pursued until encountering bottom-up evidence, namely, the resumptive pronoun.

Finally, I describe the pronoun-driven hypothesis. Variants of the pronoun-driven hypothesis have been proposed by several authors (Erteschik-Shir, 1992; Alexopoulou & Keller, 2007), however there has been very little empirical investigation to verify it. On this hypothesis, the comprehender progresses through the sentence, and actively constructs the abstract filler-gap dependency outside the island. Upon encountering the pronoun, however, the pronoun instigates a retrospective search for an antecedent. When the pronoun selects the open filler-gap dependency head

as its antecedent, the pronoun is extraordinarily treated as the resolution site of the dependency. This is distinct from the filler-driven hypothesis since the resumptive dependency formation is not necessarily constructed, since it relies on the pronoun's retrospective search for an antecedent and not the filler's prospective search for a resolution site.

(18) **Pronoun-driven Hypothesis**:

a. *Open filler-gap dependency:*
These are **the things** that … (__)

b. *Encounter island; suppress active dependency formation:*
These are **the things** that we don't know what …

c. *Encounter pronoun, instigate search for antecedent:*
These are **the things** that we don't know what they … (__)

d. *…if the pronoun's antecedent = filler, construct resumptive dependency:*
These are **the things** that we don't know what they …

e. *…otherwise, continue searching for gap:*
These are **the things** that we don't know what they … (__)

It is worth commenting on the way that the pronoun might be "substituted" for the resolution site of the open dependency. Resumptive dependencies involve two distinct relations – the filler-gap dependency is extraordinarily discharged with the pronoun, and the pronoun selects the filler as its antecedent. The latter process is taken to be a result of typical pronoun-antecedent relation construction on the pronoun-driven hypothesis. However, it is unclear how this should then force abandonment for search for a later gap site. One possible analysis is that the active dependency formation processes make an early commitment to a gap. Gaps instigate pronoun-like retrieval events (Gordon *et al.*, 2001, 2002, 2004; Wagers, 2008). Resumptive pronouns might "substitute" in for a gap just in case the parser erroneously

treats the pronoun's retrieval event as satisfying the expectation for a gap retrieval event. Alternatively, resumptives may be formed as a way to integrate a filler-gap dependency in particularly taxing environments, in which the parser sacrifices close grammatical fidelity in favor of discharging dependencies earlier than later. This may partially explain why resumptive dependencies are rated more highly in complicated structures or over long distances (Ariel, 1990; Dickey, 1996; Alexopoulou & Keller, 2007). On both accounts, resumptive dependencies are less a repair strategy from the perspective of the comprehender, but rather an error or a perhaps partially stochastic artifact of noisy encoding.

The pronoun-driven hypothesis makes two predictions that are not made by the filler-driven hypothesis. Namely, the preference for resumptive dependencies will be proportional to the accessibility of the *wh*-filler as an antecedent for the pronoun. In other words, if the *wh*-filler is easily accessible, then resumptive dependencies should be more easily constructed compared to constructions in which the *wh*-filler is less accessible. Additionally, if resumptive dependencies are created by noisy encoding, then pronoun-driven hypothesis also predicts small or non-categorical preferences for resumptive dependencies. In other words, if the parser truly avoids constructing ungrammatical representations, then there should still be a general bias against constructing resumptive dependencies.

In section 3.3, I report on a series of offline sentence fragment completion tasks and an eye-tracking study to investigate which of these three hypotheses better describes the mechanisms used to construct resumptive dependencies. I argue that the evidence is most consistent with the pronoun-driven hypothesis. This further

highlights the contrast between resumptive dependency formation in English and Hebrew. I propose that resumptive dependencies are actively constructed in Hebrew because they are grammatical, but not in English. This is most consistent with analyses in which resumptive dependencies inside islands are grammatical in Hebrew, but not in English. Afterwards, I address the distributional properties of child-directed speech in English and Hebrew in section 3.4.

3.3 Experiments 3–4: Resumptive Dependency Construction

The goal of Experiment 3 was to investigate the mechanisms of resumptive dependency formation in English. Experiment 3 consists of three subexperiments. Each subexperiment largely uses the same design. For this reason, I first describe the general design for all three experiments before describing each individual subexperiment.

Experiment 3 is a series of offline sentence completion tasks in which participants are instructed to complete a sentence fragment that contains an open filler-gap dependency and a potentially resumptive pronoun. This is illustrated in (19).

(19) The bridesmaid speculated **which groomsman** [$_{NP}$ the speech [$_{CP}$ that <u>he</u> prepared ___]] could offend …

a. …the bridal party because of the crude language
b. …___ at the reception because of the crude language

The materials were constructed to manipulate the relative accessibility of the *wh*-filler as its antecedent, and to maximize the availability of the resumptive dependency. Each experiment had slightly different materials, which were intended

to manipulate the availability of the *wh*-filler as an antecedent for the pronoun. However, each target contained a main clause with a verb (*speculated*) that takes a complement clause with a fronted *wh*-phrase (*which groomsman*). The subject of the complement clause was an inanimate noun phrase (*the speech*) that was modified by a relative clause. This noun phrase was inanimate to ensure that it was not interpreted as a potential antecedent for the pronoun, which was always an animate pronoun. The subject of the relative clause modifying the embedded clause subject was the potentially resumptive pronoun immediately adjacent to the complementizer. This is the privileged position in which resumptive pronouns are rated more highly than their gapped counterparts (McDaniel & Cowart, 1999; Omaki & Nakao, ms). Afterwards, there was an auxiliary, followed by a verb (*offend*). These verbs were chosen to be strongly transitive, and the *wh*-filler was always a semantically suitable object for the verb.

Participants were instructed to provide a completion so that the sentence was syntactically and semantically well-formed. For the example prompt in (19), they could provide a response which contains a direct object for the last verb in the prompt (*offend*), as in (19-a), or a response that does not, as in (19-b). A response that does not contain a direct object, as in (19-b), can be understood as a parse in which the *wh*-dependency headed by *which groomsman* resolves as the object of the last verb in the prompt (*offend*). This is because this last verb was strongly biased to be transitive.

In these experiments, the dependent measure is the proportion of responses that lack gaps, as in (19-a). Participants are always capable of providing assigning

an interpretation to the sentence in which the *wh*-dependency resolves with the last transitive verb in the prompt. Thus, resumptive dependencies are never forced, and there is always a grammatical response in which the open filler-gap dependency can resolve with a grammatical gap position. However, if the parser prefers a resumptive interpretation, in which the filler-gap dependency discharges with the pronoun, then there should be an increase in gap-less responses. This is because when the parser confidently commits to a resumptive dependency, it should abandon search for a later gap site. If so, then an increase rate of gap-less responses is an indirect measure of the rate of confident commitments to a representation in which the filler-gap dependency discharges with the resumptive pronoun.

This experimental paradigm can be used to distinguish the hypotheses described in section 3.2.2.3. The end of sentence hypothesis predicts a low rate of gap-less responses. This is because prompts never force the comprehender to resolve the filler-gap dependency with the pronoun. Thus, participants should be free to resolve the filler-gap dependency with the transitive verb missing an argument. Conversely, on the filler-driven hypothesis, I predict a high proportion of gap-less responses in all conditions containing a potentially resumptive pronoun compared to controls. This is because resumptive pronouns should be preferentially interpreted as the tail of the *wh*-dependency as soon as the reader encounters the pronoun. Finally, the pronoun-driven hypothesis predicts that the proportion of gap-less responses should vary as a function of the accessibility of the *wh*-filler as an antecedent for the pronoun. The three subexperiments manipulate the syntactic prominence and availability of alternative, non-filler antecedents in the sentence. If the rate

of gap-less responses correlates with the relative accessibility of the *wh*-filler as antecedent compared to alternative antecedents in the sentence. Experiments 3a–3c test these hypotheses by varying the relative accessibility of the *wh*-filler as an antecedent for the resumptive pronoun.

3.3.1 Experiment 3a

The goal of Experiment 3a was to determine whether a resumptive pronoun could complete an open filler-gap dependency when the only antecedent available to the pronoun is the *wh*-filler.

3.3.1.1 Methods

Experiment 3a was a sentence fragment completion task. Participants were instructed to read a sentence fragment carefully, and then provide a response. Participants were instructed to ensure that their response made the sentence syntactically and semantically well-formed. Before performing the experiment, participants were given examples of "complete" sentences in which all syntactic and semantic roles were discharged, and "incomplete" sentences in which verbs were either missing arguments or a filler-gap dependency was not resolved. In case a fragment could be interpreted as complete, participants were instructed to provide an adjunct or other information. Examples of these completions were also provided. Twenty participants were drawn from Amazon's Mechanical Turk platform[2] participated in Experiment 3a. They received 2 USD for participation. They also also had to have

[2] `http://www.mturk.com`

completed a native English speaker qualification task, have had completed 500 tasks on Mechanical Turk with a 90% or higher approval rating, and be from the United States. The study took approximately 10 minutes.

3.3.1.2 Materials

There were 16 sets of target items presented in a 2 × 2 Latin Square design with 24 complexity-matched fillers. The presence of a *wh*-dependency (+WH, −WH) and the form of the subject embedded in an island (−Pro,+Pro) were manipulated. Materials are exemplified in (20).

(20) a. **+WH,+Pro**:
The bridesmaid speculated **which groomsman** the speech [that he prepared] could offend …

b. **+WH,−Pro**:
The bridesmaid speculated **which groomsman** the speech [that Andrew prepared] could offend …

c. **−WH,+Pro**:
The bridesmaid speculated to the groomsman that the speech [that he prepared] could offend …

d. **−WH,−Pro**:
The bridesmaid speculated to the groomsman that the speech [that Andrew prepared] could offend …

Each target item in Experiment 3a began with a noun phrase that either has a strong gender bias or is definitionally gendered (e.g., *the bridesmaid* is necessarily feminine). In the +WH conditions, there was a *wh* complement clause with a dislocated *wh*-noun phrase (*which groomsman*), and in the −WH conditions there was an indirect object (*to the groomsman*) followed by a non-*wh* complement clause. The +WH conditions therefore have a filler-gap dependency that must be resolved,

whereas the −WH conditions do not. In Experiment 3a, the head of the *wh*-noun phrase / indirect object noun phrase had the opposite gender bias than the main clause subject. These gender biases were counterbalanced, such that half of the target items had a masculine-biased main clause subject and a feminine-biased *wh*-noun phrase / indirect object noun phrase, and half of the target items had a feminine-biased main clause subject and a masculine-biased *wh*-noun phrase / indirect object noun phrase. The embedded clauses all contained a subject that was inanimate (*the speech*). This noun phrase was inanimate to ensure that it was unlikely to be selected as the antecedent for the animate, (potentially) resumptive pronoun. This subject noun phrase is modified by a relative clause, which is the island that hosts the potentially resumptive pronoun in the +Pro conditions. In Experiment 3a, the pronoun always matched the gender feature of the *wh*-noun phrase in the +WH conditions (*which groomsman*), and the indirect object in the −WH conditions (*to the groomsman*). In the −Pro conditions, this subject is a proper name (*Andrew*). Afterwards, there is an auxiliary followed by a verb that is biased to a transitive interpretation (*could offend*). This is to draw out the contrast between gap-less and gapped responses.

In the −WH conditions, I predict a high rate of gap-less completions. This is because there is no filler-gap dependency to license a gap, and thus participants should always provide an argument for the transitive verb (*offend*). Similarly in the +WH,−Pro condition, I predict a low rate of gap-less completions. This is because the *wh*-dependency must resolve as the object of the last verb in the prompt, otherwise it does not receive a thematic role. The critical condition is the +WH,+Pro

condition, since this condition contains both the open filler-gap dependency and the (potentially) resumptive pronoun. On the end of sentence hypothesis, I expect a high rate of gap-less completions, since on this hypothesis the resumptive dependency is only constructed when forced to. On the filler-driven and pronoun-driven hypothesis, I predict an increase in gap-less reductions in the +WH,+Pro condition compared to the +WH,−Pro condition. This is because on either hypothesis, the comprehender should construct the resumptive dependency in some proportion of trials. On the filler-driven hypothesis, this is because the comprehender always prefers constructing a filler-resumptive dependency when available. On the pronoun-driven hypothesis, this is because the only antecedent available to the pronoun is the *wh*-filler, and thus the open filler-gap dependency head is readily available.

3.3.1.3 Results

For analysis, I coded each participant's response for two criteria: (1) semantic and syntactic suitability of the production given the fragment, and (2) the presence of a gap in the completion. I coded for semantic/syntactic suitability because participants in pilot versions of this study often misinterpreted the fragment, and thus provided syntactically or semantically ill-formed fragments. Semantic and syntactic suitability was determined on a local coherence basis, and did not take into consideration whether the filler-gap dependency had a resolution site. I assumed that any completions that were ambiguous as containing a gap. This may partially inflate the number of gapped completions. This was a conservative coding decision that was

intended to avoid inflating the number of gap-less responses in the critical condition, which may have yielded a spurious resumptive effect. In Experiment 3a, 97.2% of the responses were semantically and syntactically suitable given the prompt. The unsuitable responses were excluded from analysis.

The proportions of syntactically/semantically suitable, gap-less responses are given in Table 3.1.

	−Pro	+Pro
+WH	20%	34%
	15/60	26/50
−WH	100%	100%
	80/80	80/80

Table 3.1: Proportion of gap-less responses by condition in Experiment 3a

For analysis, I fit a logit mixed effects models using the `lme4` and `lmerTest` packages in R (Bates *et al.*, 2014; Kuznetsova *et al.*, 2014) treating gapped responses as the dependent variable. Main effect of subject type was included in the analysis. Because there were no gapped responses in the −WH conditions, I did not include a factor for dependency. Participants and items were included as random effects. There was a main effect of subject type ($\hat{\beta} = -0.77 \pm 0.38, z = -2.03, p = 0.05$).

3.3.1.4 Discussion

There were several findings from Experiment 3a. First, there was a significant increase in gap-less responses when there is a potential resumptive pronoun in the subject position of an island. I take this increase in gap-less responses to index the availability of the resumptive dependency, i.e., by virtue of resolving the

open filler-gap dependency with the resumptive pronoun, participants abandoned their preference for resolution outside of the island. This demonstrates that this methodology is useful for probing resumptive dependency formation mechanisms.

Interestingly, the resumptive effect is not categorical. Although there was an increase in gap-less responses in the +WH,+Pro condition compared to the +WH,−Pro condition, the rate of gap-less responses in the +WH,+Pro condition did not approach the rate of rate of gap-less responses in the −WH conditions, which served as the baseline for sentences without an open filler-gap dependency. This may be interpreted as evidence that the comprehender only entertains resumptive pronouns in exceptional cases, e.g. as an error in encoding, as discussed in section 3.2.2.3.

Additionally, the results from Experiment 3a show that this sentence fragment completion task is a very noisy measure. The predicted rate of gap-less responses for the +WH,−Pro condition was 0%, since the filler-gap dependency needed a gap to discharge grammatically. However, 20% of the responses were gap-less in this condition. This could be for a variety of reasons. These sentences are relatively complex, and thus participants may have simply abandoned the *wh*-dependency. Alternatively, participants may have been rushing through the task, since this experiment was administered on Mechanical Turk, and participants were finishing the experiment very quickly.

Finally, the difference between the +WH,−Pro and +WH,+Pro condition does not support the end of sentence hypothesis, since there was an effect of the resumptive pronoun. However, this does not distinguish between the filler-driven and

the pronoun-driven hypotheses. Experiment 3b attempts to distinguish these hypotheses by modulating the number of antecedents available for the (potentially) resumptive pronoun.

3.3.2 Experiment 3b

Experiment 3b was another sentence fragment completion task using similar materials to the ones used in Experiment 3a. The goal of Experiment 3b was to determine whether modulating the number of antecedents available to the pronoun affected the proportion of gap-less responses. In this experiment, there was an additional potential antecedent available to the pronoun, besides the *wh*-filler. On the pronoun-driven hypothesis, the alternative antecedent should reduce the availability of the resumptive dependency. On the filler-driven hypothesis, there should be no effect of the alternative antecedent.

3.3.2.1 Methods

Experiment 3b was performed in the same way as Experiment 3a.

3.3.2.2 Materials

The materials in Experiment 3b were the similar as the materials in Experiment 3a. The crucial difference is that the main clause subjects were manipulated to have the same gender bias as the *wh*-filler / indirect object and the pronoun. An example set of stimuli for Experiment 1b is presented in (21). Note that, unlike in

Experiment 3a, the main clause subject (*the priest*) and the *wh*-filler / indirect object noun phrase (*the groomsman*) match in gender bias, and similarly match in gender feature to the potentially resumptive pronoun. Gender biases were counterbalanced across conditions.

(21) a. **+WH,+Pro**:
The priest speculated **which groomsman** the speech [that he prepared] could offend …

b. **+WH,−Pro**:
The priest speculated **which groomsman** the speech [that Andrew prepared] could offend …

c. **−WH,+Pro**:
The priest speculated to the groomsman that the speech [that he prepared] could offend …

d. **−WH,−Pro**:
The priest speculated to the groomsman that the speech [that Andrew prepared] could offend …

3.3.2.3 Results

Analysis for Experiment 3b was conducted in the same way as Experiment 3a. Participants' responses were 97.5% semantically/syntactically suitable for the prompt, again showing that participants typically constructed appropriate representations for the prompt. The proportion of gap-less responses is given in Table 3.2. A logit mixed effects model was constructed with the same structure as in Experiment 3a. There were no main effect of subject type, unlike Experiment 3a ($\hat{\beta} = 0.07 \pm 0.49, z = 0.15, p = 0.88$).

	−Pro	+Pro
+WH	26%	25%
	19/75	21/79
−WH	100%	100%
	79/79	79/79

Table 3.2: Proportion of gap-less responses by condition in Experiment 3a

3.3.2.4 Discussion

In Experiment 3b, there was no difference in the +WH,−Pro and +WH,+Pro conditions. This indicates that the potentially resumptive pronoun had no effect on the rate of gap-less productions. The contrast between Experiment 3a and Experiment 3b suggests that resumptive dependencies are more easily constructable when the only antecedent available to the pronoun is the open filler-gap dependency head. These findings are most compatible with the pronoun-driven hypothesis, because this hypothesis predicts that resumptive dependencies are constructed by the pronoun's search for an antecedent. On this hypothesis, the lack of resumptive effect in Experiment 3b can be interpreted as an effect of the main clause subject, a particularly accessible alternative antecedent. In other words, the pronoun preferably selects the main clause subject as its antecedent, since it is a syntactically prominent referent (Foraker & McElree, 2007), forcing the filler-gap dependency to search for a gap after the pronoun, yielding similar gap rates in the +WH,−Pro and +WH,+Pro conditions.

There is an alternative interpretation for the lack of the resumptive effect in Experiment 3b, however. Suppose that the comprehender ranks resumptive dependencies higher than accommodating an extrasentential antecedent, but both are

ranked as undesirable constructions. On this account, the resumptive effect observed in Experiment 3a arises because selecting the *wh*-filler is the only intrasententially available antecedent. Conversely, in Experiment 3b, there is no resumptive effect because there is an alternative antecedent available, and thus the pronoun is not forced to select the *wh*-filler as its antecedent. This analysis is very similar to the pronoun-driven view as described in section 3.2.2.3. The crucial difference is that whereas the pronoun-driven hypothesis predicts that the availability of the resumptive dependency will inversely track the availability of alternative intrasentential antecedents, the last resort hypothesis predicts no resumptive effect if there are *any* alternative intrasentential antecedents. Experiment 3c investigates this prediction.

3.3.3 Experiment 3c

There were two main goals for Experiment 3c. The first was to determine whether the resumptive effect present in Experiment 3a and absent in Experiment 3b is observed when there is an alternative antecedent in a position less prominent than the main clause subject position. This helps distinguish the pronoun-driven hypothesis from the low priority hypothesis, described in section 3.3.2.4. The second goal was to replicate the manipulation between Experiments 3a and 3b within the same study.

3.3.3.1 Methods

The methods for Experiment 1c were the same as in Experiments 1a and 1b.

3.3.3.2 Materials

The materials in Experiment 3c were similar to the materials used in Experiments 3a and 3b, except with a few differences in the main clause to permit the alternative antecedent for the pronoun to be a less prominent position. The alternative antecedents were an oblique argument to a verb with an expletive subject. The ±WH manipulation in Experiment 3c was different than in Experiment 3a and 3b. The +WH conditions had an open *wh*-dependency as in Experiments 3a and 3b. The −WH conditions also had a *wh*-dependency that resolved in an intermediate clause (__ *said that*). Lastly, there was no ±Pro manipulation, since all target items had a pronoun in the subject position in the island. Instead, the gender bias of the alternative antecedent was manipulated to either match or mismatch with the gender of the pronoun (*the priest* or *the bridesmaid*). In Experiment 3c, there were two resumptive conditions – the +WH,−Match condition forced the pronoun to select the *wh*-filler as its antecedent, like in Experiment 3a, and the +WH,+Match condition allowed the pronoun to select the oblique argument as its antecedent. This allows a within-experiment test of the resumptive effect found in Experiments 3a and 3b. The materials are exemplified in (22).

(22) a. **+WH,+Match**:
It was explained to the priest **which groomsman** the speech [that he prepared] could offend …

b. **+WH,−Match**:
It was explained to the bridesmaid **which groomsman** the speech [that he prepared] could offend …

c. **−WH,+Match**:
It was explained to the priest **which groomsman** __ said that the speech [that he prepared] could offend …

	+Match	−Match
+WH	41%	59%
	32/79	44/75
−WH	99%	97%
	79/80	77/80

Table 3.3: Proportion of gap-less responses by condition in Experiment 3c

d. **−WH,−Match**:
It was explained to the bridesmaid **which groomsman** __ said that the speech [that he prepared] could offend …

3.3.3.3 Results

Data analysis was the same as in Experiments 1a and 1b. 98.4% of participants' responses were syntactically and semantically suitable given the prompt. The proportion of gap-less responses is given in Table 3.3. I fit a logit mixed effect model like in Experiments 3a and 3b, treating responses with gaps as the dependent variable. Main effects of alternative antecedent gender (±Match) were included in the analysis. Participants and items were included as random effects. There was a main effect of alternative antecedent gender ($\hat{\beta} = -1.36 \pm 0.47, z = -2.92, p = 0.004$).

3.3.3.4 Discussion

There were two main findings in Experiment 3c. First, both +WH conditions had a gap-less response rate substantially greater than expected if the comprehender avoided construing the pronoun as a resumptive pronoun. In other words, the high rate of gap-less responses in both +WH conditions compared to the rate of gap-less responses in Experiments 3a and 3b suggests that in both conditions the

pronoun resolves the filler-gap dependency some proportion of the time, triggering abandonment of gap search.

The second crucial finding is that there was a significant difference in the response rate between the +Match and −Match conditions. If the interpretation that the pronoun in both the +WH conditions in Experiment 3c were interpreted as a resumptive pronoun, then this difference between the +Match and −Match conditions suggest that non-prominent alternative antecedents modulate the accessibility of a resumptive dependency. Put differently, resumptive dependencies are constructed, albeit at a lesser rate, even when the pronoun has an accessible alternative antecedent. This is consistent with the pronoun-driven view, in which the pronoun accesses an antecedent in proportion to its prominence, and is interpreted as a resumptive just in case it is coreferential with the *wh*-filler. These findings are not predicted on the low priority hypothesis, in which the comprehender systematically avoids constructing *wh*-filler-pronoun dependencies if there is any other intrasentential antecedent available for the pronoun.

Taken together, the results from Experiments 3a–3c largely support the pronoun-driven hypothesis. To summarize, upon reaching the pronoun, a retrospective search of an antecedent is executed. If there is only a *wh*-filler available, then this is selected as the antecedent, and a resumptive dependency can be constructed. When there are antecedents apart from the *wh*-filler available, then the pronoun selects these as its antecedent proportional to their accessibility. If the alternative antecedent is in a particularly prominent position, e.g. the main clause subject, then the pronoun overwhelmingly selects this referent as its antecedent. If this alternative antecedent

is in a less prominent position, then the pronoun selects it only some proportion of the time. This yields a marginal and diminished rate of resumptive dependency formation, respectively.

Experiments 3a–3c focussed on the interactions between a *wh*-dependency and a potentially resumptive pronoun, and the results imply that resumptive dependencies in English are constructed in a pronoun-driven fashion. This contrasts with Keshev & Meltzer-Asscher's (ms) findings in Hebrew, which suggest an active search for a resumptive pronoun in island contexts. I argue that this is because resumptive pronouns in island contexts are ungrammatical repair mechanisms in English, but are grammatical in Hebrew. However, there are two differences between Experiments 3a–3c and Keshev & Meltzer-Asscher's (ms) experiments. First, Experiments 3a–3c investigated *wh*-dedendencies, whereas Keshev & Meltzer-Asscher's (ms) studies investigated relativization dependencies. This may impact the results, since resumptive pronouns are canonically found in relativization constructions, and only marginally in *wh*-dependencies (e.g., Ariel 1990; Prince 1990). Thus, although English resumptive dependencies are likely not actively constructed, resumptive relativization dependencies may have a different profile than resumptive *wh*-dependencies. Additionally, Experiments 3a–3c were offline tasks, whereas the Hebrew results are from a self-paced reading task. An online task on resumptive dependency formation in English might reveal that readers consider resumptive dependencies to a greater degree in real-time processing than suggested by these results, and then ultimately reject them. For these reasons, we investigate the time course of resumptive dependency formation in an online task in Experiment 4.

3.3.4 Experiment 4

The results from Experiments 3a–3c suggest that resumptive dependencies are likely constructed in a pronoun-driven fashion in English, in contrast with the active resumptive dependency formation process claimed for Hebrew resumptive dependencies (Keshev & Meltzer-Asscher, ms). As argued earlier, this is explained if resumptive dependencies in islands are grammatically licensed in Hebrew but not English, as suggested in the psycholinguistic literature. However, these differences in processing profiles between the two languages might partially be due to the different nature of the tasks or to the differences in construction type investigated. The goal of Experiment 4 was to probe for evidence of a pronoun-driven resumptive dependency formation in an online task.

Specifically, in Experiment 4, I test whether the comprehender considers the an open filler-gap dependency head as a possible antecedent for a (potentially) resumptive pronoun. The critical manipulation is the number of gender-matching antecedents available to the pronoun, as in Experiments 3a and 3b. Previous work reveals that gazes on pronouns with no grammatically accessible antecedents are substantially longer than gazes on sentences with grammatically accessible antecedents (Badecker & Straub, 2002; Sturt, 2003). Additionally, some research suggests that pronouns with multiple antecedents also increase gaze duration, the "multiple mismatch effect" (Badecker & Straub 2002, but see Chow *et al.* 2014. In Experiment 4, we test to see whether open filler-gap dependency heads impact the gaze duration of a potentially resumptive pronoun. Crucially, if gaze durations are shorter at

the pronoun in sentences in which the filler-gap dependency head is a possible antecedent compared to conditions in which the pronoun has no available antecedent, this is strong evidence that the parser (momentarily) considers the pronoun-filler dependency, as needed on the pronoun-driven view. Additionally, if there is a multiple mismatch effect in sentences in which the pronoun matches in morphological features with both the open filler-gap dependency head and another potential antecedent compared to sentences in which it only has one potential antecedent, this too is evidence that the open filler-gap dependency head is momentarily considered an antecedent for the pronoun.

3.3.4.1 Methods

Experiment 4 was an eye-tracking study. Forty-eight participants were drawn from the University of Maryland community and received credit for completing the experiment. All participants were native speakers of English and had normal or corrected vision. Participants were tested individually in a quiet room in one session lasting 20–30 minutes, and then participated in an unrelated experiment afterwards. Eye movements were recorded using an EyeLink 1000 eye-tracker (SR Research, Toronto, Ontario, Canada), which interfaced with a PC computer. Participants were seated with their chin and forehead stabilized by the eye-tracker apparatus. They were seated 32 inches from an LCD monitor, which displayed the stimuli at 4.6 characters per degree of visual arc. Stimulus presentation and interface with the eyetracker was implemented with the EyeTrack software suite (University of

Massachusetts, Amherst). The eye-tracker has an angular resolution of $0.25^{\circ}-0.5^{\circ}$. Viewing was binocular, but only the right eye was recorded. Eye-movements were recorded with a sampling rate of 1000Hz.

Sentences were presented in 12-point fixed-width Courier font in two lines. The line break was located after the head noun, before the relative pronoun. This location for the line break ensured that the relative pronoun, the potentially resumptive pronoun, and the entire contents of the island were all on the second line, as discussed in the next section. A calibration procedure was performed before the experiment, and re-calibration was carried out between trials if needed. Each participant was instructed to read for comprehension as naturally as possible. Each trial began with only a gray square on the left edge of the display. The participant triggered the appearance of the sentences by fixating on the square. The trial ended when participants read to the end of the line and pressed a button on a hand-held controller. After each item, there was a yes-no comprehension question, which participants answered using the hand-held controller. Participants received no feedback for correct or incorrect responses.

3.3.4.2 Materials

There were 24 sets of target items presented in a 2 × 2 Latin Square design, with 36 complexity-matched fillers. Like Experiment 3c all the target items contained a potentially resumptive pronoun (*he/she*) in the subject position of an island, with an open filler-gap dependency that grammatically resolves with a gap

after the island. Unlike Experiment 3c, the crucial filler-gap dependency is a relativization dependency. The main clause subject and the pronoun were manipulated to either match or mismatch with the gender bias of a relativization head, such that either the main clause subject, the head of a relativization dependency, both, or neither matched in gender with the potentially resumptive pronoun. The materials are exemplified in (23).

(23) a. **+Subject Match,−Filler Match**
The popular actress worked with **the artsy director** // who the fans tha/t she g/reeted /on /the set /had /expected /John /to /hire /for the next blockbuster

b. **−Subject Match,−Filler Match**
The popular actor worked with **the artsy director** // who the fans tha/t she g/reeted /on /the set /had /expected /John /to /hire /for the next blockbuster

c. **+Subject Match,+Filler Match**
The popular actor worked with **the artsy director** // who the fans tha/t he g/reeted /on /the set /had /expected /John /to /hire /for the next blockbuster

d. **−Subject Match,+Filler Match**
The popular actress worked with **the artsy director** // who the fans tha/t he g/reeted /on /the set /had /expected /John /to /hire /for the next blockbuster

In the +Subject Match, +Filler Match condition, the main clause subject (*the popular actor*), the filler (*the artsy director*), and the pronoun (*he*) all match in gender bias, counterbalanced across conditions. The −Subject Match,−Filler Match condition is identical to the +Subject Match,+Filler Match condition, except the pronoun is marked with the opposite gender bias than the main clause subject or the target filler. In the −Subject Match, +Filler Match condition, the pronoun matches in gender with the filler, but the lexical content of the main clause subject

is changed to mismatch with the pronoun. This is the "resumptive condition", since the only antecedent available for the pronoun is the filler. Finally, the +Subject Match, −Filler Match conditions has the same lexical content as the −Subject Match, +Filler condition, but with the pronoun's gender matching the main clause subject only.

The structure of these materials is largely the same as in Experiments 3a–3c, except for a few key differences. First, the filler-gap dependency of interest is a relativization dependency (in (23), headed by *the artsy director*), not a *wh*-dependency. After the relativization head, there is a relative pronoun (*who*), followed by the subject of the relative clause (*the fans*). Here, there is a plural animate subject, instead of an inanimate singular subject, as in Experiments 3a–3c. The island that hosts the potentially resumptive pronoun modifies this noun phrase, as in Experiments 3a–3c. After this island, there is an auxiliary followed by an exceptionally case marking (ECM) verb, i.e., a verb that takes an object and a clause as arguments. This object serves as a filled-gap, in the sense described in Chapter 1. Increased processing difficulty at this region implies that the reader is attempting to resolve the relativization dependency (headed by *the artsy director*) with the verb *expected*, and then revising. This allows us to determine in which conditions the comprehender abandons search for a later gap site after the pronoun, like in Experiments 3a–3c.

There were 9 regions measured, starting with the potentially resumptive pronoun until the end of the sentence. These are demarcated by slashes in (23). Note that the pronoun region includes an additional two characters on either side, since

pronouns and other short function words have a high chance of being skipped (Rayner *et al.*, 2011). The double slash (//) indicates the line break. The three regions of interest are the pronoun region (*/t (s)he g/*), the verb immediately after (*reeted/*), and the filled-gap region (*John*).

The pronoun-driven hypothesis predicts that the ±Match manipulation should have an effect on gaze duration on the pronoun region. Gaze durations on the pronoun in the +Subject Match conditions are likely to be shorter in general compared to the −Subject Match conditions if there is a bias to interpret the pronoun as coreferential with the main clause subject. However, if there is a multiple mismatch effect, or a cost associated with multiple antecedents available for a pronoun, then the pronoun-driven hypothesis predicts an increase in gaze duration for the +Subject Match,+Filler Match condition compared to the +Subject Match,−Filler Match condition. The pronoun-driven hypothesis predicts shorter gaze durations at the pronoun region for the −Subject Match,+Filler Match condition compared to the −Subject Match,−Filler Match condition as well. This is because there is no antecedent available for the −Subject Match,−Filler Match condition for the pronoun, which should cause increased processing difficulty in this region. If the open filler-gap dependency head is a possible antecedent for the pronoun, however, there should be no increase in processing difficulty in this region.

3.3.4.3 Results

Before performing data analysis, I visually inspected each trial using EyeDoctor[3] to correct for small vertical drifts. Fixations of less than 80ms in duration and within one character of the previous or following fixation were incorporated into the neighboring fixation. All remaining fixations shorter than 80ms were excluded, since readers are not likely to extract any information during short fixations (Rayner & Pollatsek, 1989). We also excluded fixations longer than 800ms.

Standard eye-tracking measures (Rayner, 1998) were calculated for each region. We report three eye-tracking measures, first pass times, regression path times, and total reading times. First-pass time is the sum of all fixation times starting with the first fixation in a region until the first fixation outside the region (either to the left or right) provided that the reader has not fixated on subsequent text (Rayner & Duffy, 1986). This measurement is an "early measures", which index of early processing steps. We also measured regression-path times, the sum of all fixation times starting with the first fixation in the region until the first fixation to the right of the region, provided that the reader has not fixated on subsequent text (Brysbaert & Mitchell, 1996). We also measured total reading times, or the sum of all fixations in a region. These latter two measures are "late measures", which index later processing steps.

88% of the comprehension questions were answered correctly. If participants responded incorrectly to a comprehension question, that trial was removed from

[3] `http://www.psych.umass.edu/eyelab/software/`

	t he g	reeted	on	the set	had	expected	John	to	hire
First pass reading times									
+Subject,+Filler	242 (6)	259 (7)	327 (11)	242 (7)	298 (10)	**254 (6)**	249 (7)	296 (8)	252 (8)
−Subject,−Filler	246 (7)	248 (7)	301 (10)	258 (7)	303 (9)	**244 (7)**	230 (7)	317 (9)	267 (8)
−Subject,+Filler	259 (7)	253 (7)	330 (13)	260 (7)	312 (11)	**264 (7)**	239 (6)	317 (11)	254 (7)
+Subject,−Filler	247 (7)	263 (7)	320 (11)	246 (7)	315 (10)	**257 (6)**	241 (7)	305 (10)	268 (10)
Regression path									
+Subject,+Filler	541 (62)	425 (39)	483 (31)	315 (33)	437 (31)	399 (35)	292 (9)	479 (48)	**357 (34)**
−Subject,−Filler	392 (28)	408 (27)	482 (35)	320 (20)	436 (25)	366 (24)	333 (28)	420 (26)	**332 (21)**
−Subject,+Filler	427 (28)	346 (14)	481 (35)	392 (37)	486 (36)	413 (38)	295 (11)	501 (45)	**368 (27)**
+Subject,−Filler	460 (41)	423 (31)	492 (42)	336 (28)	446 (33)	357 (19)	325 (19)	409 (36)	**498 (48)**
Total time									
+Subject,+Filler	**443 (20)**	390 (18)	**580 (27)**	**426 (20)**	552 (24)	427 (19)	311 (12)	549 (26)	**390 (16)**
−Subject,−Filler	**450 (22)**	347 (15)	**560 (25)**	**395 (18)**	559 (26)	390 (17)	294 (12)	496 (24)	**344 (13)**
−Subject,+Filler	**432 (21)**	335 (12)	**574 (27)**	**391 (17)**	588 (26)	424 (20)	287 (10)	538 (28)	**373 (15)**
+Subject,−Filler	**402 (18)**	356 (18)	**517 (23)**	**367 (16)**	559 (25)	400 (18)	301 (11)	512 (23)	**401 (21)**

Table 3.4: Results from Experiment 4. Units are in milliseconds. Numbers in parentheses correspond to one standard error of the mean. Bolded values highlight a region with significant or marginally significant differences.

analysis. Using the `lme4` and `lmerTest` packages in R, I constructed a mixed effect model for each region and for each measure in R. Each model included main effects of filler match and subject match, and random effects of participant and item. I coded filler match as 1 and subject mismatch as 1. This is because we were most interested in finding an interaction effect of the resumptive condition, e.g. when only the filler matched the pronoun.

First pass reading times At the verb after the end of the island boundary (*expected*), there was a marginal interaction effect of subject mismatch and filler match ($\hat{\beta} = 25\ (15), t(640) = 1.65, p = 0.10$). Additionally, at the verb immediately before the filled-gap region (*expected*), there was a marginal interaction of subject mismatch and filler match in first pass reading times ($\hat{\beta} = 25\ (15), t(640) = 1.65, p = 0.10$).

Regression path At the verb in which the relativization dependency resolves, there was a main effect of subject mismatch ($\hat{\beta} = -166\ (75), t(380) = -2.2, p = 0.03$), and a marginal effect of filler match ($\hat{\beta} = -141\ (75), t(386) = -1.9, p = 0.06$), and a marginal interaction effect of subject mismatch and filler match ($\hat{\beta} = 177\ (105), t(382) = 1.68, p = 0.09$).

Total reading times At the pronoun (*(s)he*), there was a marginal main effect of subject mismatch ($\hat{\beta} = 51\ (29), t(689) = 1.79, p = 0.07$). There was also a significant main effect of filler match in the first spillover region (*on*) ($\hat{\beta} = -62\ (43), t(818) = 2.24, p = 0.03$), and also on the subsequent spillover region (*the*

set) ($\hat{\beta} = 54\ (29), t(600) = 1.89, p = 0.06$). Finally, there was also a main effect of subject mismatch at the verb in which the relativization dependency resolves (*hire*) ($\hat{\beta} = -60\ (27), t(526) = -2.25, p = 0.025$).

3.3.4.4 Discussion

The results from Experiment 4 did not confirm the hypothesis that the comprehender considers the open filler-gap dependency head as a potential antecedent for the resumptive pronoun. At the pronoun, there were no significant effects of gender match with either the subject or the open filler-gap dependency head. Additionally, there were no effects at the filled-gap region, suggesting that the comprehender did not attempt resolution of the relativization dependency in this position differentially between conditions. In other words, the evidence from this study did not reveal any effect of the pronoun accessing the open filler-gap dependency head as an antecedent, nor "canceling" of search for the later gap.

The results from Experiment 4 do not reveal that pronouns immediately access the open filler-gap dependency head as a possible antecedent, as expected by the pronoun-driven hypothesis. However, the results from Experiments 3a–3c largely point to a pronoun-driven resumptive dependency mechanism. The failure to find evidence of the pronoun-driven mechanism in Experiment 4 may ultimately be due to the ungrammatical nature of resumptive dependencies. Suppose that the comprehension system actively suppresses accessing the open filler-gap dependency head as an antecedent for the pronoun, as a measure to avoid constructing ungrammatical

resumptive dependencies. This would explain the lack of an effect of the gender of the open filler-gap dependency head on gaze duration at the pronoun. If the parser actively suppresses constructing filler-pronoun dependencies, then resumptive dependencies must be the result of a faulty or noisy encoding system, as I suggested on the basis of the small size of the resumptive effect in Experiments 3a–3c. If so, then it is unsurprising that there are no clear effects of filler gender in Experiment 4, nor any particular interaction effect of subject gender and open filler-gap dependency head gender. At the very least, these data again underscore that resumptive dependency formation is not an active process. If the parser actively constructed filler-pronoun dependencies, then there minimally should have been a gender mismatch effect, i.e., an increase in gaze duration at the pronoun when the gender of the pronoun and the filler-gap dependency head mismatched, and likely a filled-gap effect at the object position of the first verb after the island in the conditions in which the filler-gap dependency head mismatched in gender with the pronoun.

As stated earlier, a particularly expedient way to explain the difference between active resumptive dependency formation in Hebrew and the non-active processes in English is to suppose that resumptive dependencies in Hebrew are grammatical in island contexts, whereas resumptive dependencies in islands in English are ungrammatical. On this view, the active dependency formation observed in Hebrew can be explained with a fairly trivial linking hypothesis, namely, that active dependency formation searches the space of grammatical representations (see Chapter 4 for discussion). In Hebrew, this includes resumptive relativizations into islands, but not in English. Instead, resumptive dependencies in English are constructed in a

messier, bottom-up – specifically pronoun-driven – process, due to their nature as repair strategies, or plausibly an error in encoding processes.

If this is the correct analysis of these divergent processing profiles, then this implies that the learner must infer whether her grammar permits resumptive relativization dependencies in islands on the basis of some input. I have suggested that this must be done indirectly, on the basis of observing resumptive pronouns outside of islands in Hebrew. In other words, resumptive dependencies resolving in islands are ungrammatical in English but grammatical in Hebrew, and resumptive dependencies outside and inside islands in Hebrew are abstractly related. This is most compatible with analyses in which resumptive dependencies are the same dependency representationally, since these analyses provide a clear way for the Hebrew learner to infer that resumptive pronouns inside islands are grammatical on the basis of resumptive pronouns outside islands. Alternatively, resumptive dependencies may simply be more frequent in islands in Hebrew if they are grammatical. If so, then analyses do not need to assign the same representational characteristics to resumptive pronouns inside and outside islands. I address these questions in the next section.

3.4 Primary Linguistic Data

In section 3.3, I established that resumptive dependencies in islands in English are not actively constructed, unlike in Hebrew (Keshev & Meltzer-Asscher, ms). I argued that this follows if resumptive dependencies are grammatical in Hebrew,

but not in English. This in turn implies that Hebrew child-directed speech must contain some cue that signals that resumptive dependencies are grammatical inside islands, and English child-directed speech contains some cue that signals that resumptive dependencies are not grammatical inside islands. Alternatively, Hebrew speakers might actively construct resumptive dependencies inside islands because they are simply more frequent in islands compared to English, similar to the "fine tuning" hypothesis for relative clause attachment ambiguities described in Chapter 1, regardless of the representational status of resumptive dependencies in Hebrew.

In either case, examining the distribution of resumptive pronouns inside island contexts in English and Hebrew can help clarify how active resumptive construction might be learned. If resumptive dependencies in Hebrew are frequently found in islands, then the cross-language contrast in processing profiles can be explained on this basis, without invoking a representational difference between the two. This result would be compatible with either Category I, II, or III theories, since it explains the difference in processing profiles to the probabilities defined over representations. Alternatively, if there are a substantial number of resumptive dependencies resolving in islands in Hebrew, then Hebrew learners might simply directly learn (in the sense in Chapter 2) that resumptive dependencies in islands are grammatical. In other words, if there are sufficiently many resumptive dependencies in islands in Hebrew, then the processing facts in section 3.3 are not sufficient for distinguishing between the categories of analyses described in section 3.2.1.

However, if the distribution of resumptives in islands are not distinguishable in English and Hebrew child-directed speech, then the contrast in processing profiles

cannot arise from differences in "fine tuning". Additionally, if these distributions are the same between the two languages, this would suggest that resumptive dependencies in islands cannot be directly inferred as grammatical in Hebrew, but must be indirectly learned. I argue that this is the case, and that this implies a need for analyses in which resumptive dependencies inside and outside islands in Hebrew are abstractly the same construction.

3.4.1 Methods

As in Chapter 2, we conducted analyses on two corpora. The English corpus was Pearl & Sprouse's (2013) corpus described in Chapter 2. The Hebrew corpus was the Berman longitudinal corpus (Berman & Weissenborn, 1991; MacWhinney, 2000). The Berman corpus is a longitudinal corpus containing transcripts of one-hour recordings in naturalistic, home settings with four monolingual, Hebrew-learning children in Israel. The ages of the children ranged from 1;4 to 3;3, and contained 52,876 lines, or approximately half of the size of Pearl & Sprouse's (2013) corpus.

All relative clauses were extracted from the Pearl & Sprouse (2013), as described in Chapter 2. For the Hebrew corpus, all utterances that contained the complementizer *še* were extracted. This search yielded all embedded clauses, including nominal and verbal complement clauses. Hebrew does not have a null complementizer for relativizations, so all grammatical relativizations are a subset of this search.

Each utterance in both corpora was hand-coded for whether it contained a relative clause. DAC coded the English corpus, and RK coded the Hebrew corpus. All sentences not including a relative clause were excluded from analysis. Each relativization was then coded for (1) whether the resolution site was a gap or a resumptive pronoun, (2) whether the resolution site was inside an island, (3) whether the antecedent of the gap/resumptive pronoun was marked definite or indefinite, and (4) whether the gap/resumptive pronoun occurred in a subject or non-subject position. This last feature was relevant for analysis in Chapter 2. We coded for the syntactic position and definiteness of the antecedent, because earlier work suggests that the distributions of resumptive pronouns in grammaticized resumptive and non-grammaticized resumptive languages differ in these dimensions (Prince, 1990). The subject/non-subject coding was also relevant for the analysis in Chapter 2. For the English dependencies, the form of the complementizer (*that*, null complementizer, or *wh*-relative pronoun) was annotated. This too was only relevant for the analysis in Chapter 2. For the Hebrew corpus, the morphological form of the resumptive pronoun was also coded, either as a free or bound morpheme. There are some resumptive pronouns in Hebrew that are obligatory, and that are morphologically clitics on another part of speech, such as a preposition. It is unclear whether the learner will be able to immediately analyze these resumptive pronouns as instances of the same phenomenon, since they have a different morphophonological form. This contrast is demonstrated in (24) and (25), in which the direct object resumptive pronoun *oto* may freely alternate with a gap, but the resumptive pronoun *-av* is obligatory.

(24) a. **ha-iš** še ra'iti oto
the-man that I saw him
b. **ha-iš** še ra'iti ___
the-man that I saw ___
'The man that I saw ___ '

(25) a. **ha-sefer** še šamati al-av
the-book that I heard about-him
b. *ha-sefer še šamati al-___
the-book that I heard about-___
'The book that I heard about ___ '

3.4.2 Results

As described in Chapter 2, there was a total of 774 relativization dependencies in the Pearl & Sprouse corpus, which contained 101,838 lines. In all relativizations, there were two sentences that contained resumptive pronouns. These are demonstrated in (26).

(26) a. Do you see **anything else** you know what it is?
b. Um, they get all kinds of money for **those foolish idiotic** pictures that they have to telling you underneath what it is

Interestingly, in both cases, the resumptive pronoun occurs in the subject position immediately adjacent to a fronted *wh*-element. As described in section 3.2.2, the complementizer-adjacent subject position is a privileged position for resumptive pronouns. The increased judgments for resumptive pronouns in this position may follow from the fact that children hear resumptive pronouns in this position, but not in other positions. However, it is also notable that the second sentence is independently ungrammatical, due to the selectional restriction violation **they have*

Total sentences:	52,876
Total relativizations:	719
Total resumptive pronouns:	155
Total RPs in islands:	0

Table 3.5: Results from Hebrew corpus study, coded by RK

to telling you. Since there are only two instances of resumptive pronouns, one in an independently ungrammatical sentence, resumptive dependencies in islands may be indistinguishable from noise for the learner. Therefore, the learner is unlikely to infer from English child-directed speech that resumptive dependencies are grammatical in island contexts.

Additionally, there are three other relativization dependencies that are ungrammatical, given in (27). Assuming that the learner is unlikely to infer that these dependencies are grammatical, it seems reasonable to suppose that the learner will likely not treat the sentences in (26) as positive evidence that resumptive relativizations are grammatical in islands in English.

(27) a. I'm breaking the stick that the motor of I made.
b. I still think something that flies in the air and maybe you hold
c. I'm putting another girl that's going to and now she goes.

The results of the Hebrew corpus are given in Table 3.5. There were 719 relativization dependencies in Hebrew out of 52,876 lines. Importantly, of these relativizations in Hebrew, none resolve inside island constructions.

Thus, it is unlikely that the Hebrew parser actively predicts resumptive dependencies in island contexts simply because they are more frequent than in English. Additionally, these findings imply that the learner exposed Hebrew data does not

have sufficient data to infer that resumptive relativization dependencies in islands are grammatical. Thus, I argue that the difference in processing profiles observed in section 3.3 necessitates positing that resumptive dependencies in islands in Hebrew are grammatical, but not in English, and that the Hebrew learner must learn this indirectly.

In the Hebrew corpus, there were 155 resumptive dependencies outside islands. Of these 155 relativizations, 42 were free morphemes. By the reckoning used in Chapter 2, Hebrew-learning children hear approximately 840 sentences that contain an optional resumptive dependency within the first 3 years of life, and 2,260 morphologically obligatory resumptive dependencies. I submit that even if the learner attends to only one kind of resumptive pronoun, both kinds are sufficiently present in Hebrew child-directed speech for her to infer that they are grammatical. On a Category I theory, the learner can therefore infer that resumptive dependencies in islands are also grammatical.

3.4.3 Discussion

Resumptive pronouns are virtually absent inside island contexts in both English child-directed speech and Hebrew child-directed speech. This is perhaps surprising, given that informal reports suggest that resumptive pronouns are produced quite regularly in naturalistic speech (e.g., Prince 1990). This asymmetry between child-directed speech and intuitions about adult-directed productions may ultimately be due to complexity differences in child-directed speech vs. adult-directed

speech. Resumptive pronouns are more frequent in complex constructions in English, and complexity may vary as a function of age (Kunert *et al.*, 2011) and context (White, 2014). The factors may skew the distribution of resumptive pronouns in the English corpus partially. These concerns are of broader interest than just characterizing the distribution of resumptive pronouns in child-directed speech, but are important to a wider range of concerns regarding learnability of islands and other multiclausal phenomena. Alternatively, the impression that resumptive pronouns are produced at a high frequency in English may simply be overstated.

The goal of this corpus study was to determine whether the rate of resumptive dependencies inside islands were different in island contexts, and to determine what evidence the Hebrew learner could theoretically capitalize on to infer that resumptive dependencies in islands are grammatical. This is motivated by the finding that resumptive dependencies in Hebrew are constructed actively (Keshev & Meltzer-Asscher, ms), but passively in English, shown in section 3.3. However, processing profiles are not learnable, since they are not observable (Fodor, 1998a). Thus, this cross-language difference in processing must be rooted in something observable. I argued that this difference cannot arise simply from differences in frequencies of resumptive dependencies in islands, leading to differently "fine tuned" adult parsers. Additionally, it cannot be the case that the Hebrew learner can infer that resumptive dependencies in islands are grammatical directly, because there is no positive evidence in Hebrew child-directed speech. However, this conclusion may depend on the interpretation of the bound/free morpheme contrast. Borer (1984) claims that morphologically bound resumptive pronouns are obligatory because preposi-

tional phrases, in which they are typically found, are islands in Hebrew. If so, then the approximately 2,260 morphologically bound resumptive pronouns that the Hebrew learner observes in three years of life are observations of resumptive dependencies inside islands (Omer Preminger, p.c.). If the data is understood this way, then the Hebrew learner can directly learn that resumptive dependencies inside islands are grammatical. However, this only follows if the learner antecedently knows that prepositional phrases are islands in Hebrew. Additionally, it requires that the learner knows that obligatory, morphologically bound resumptive clitics are instances of the same class as morphologically free resumptive pronouns in other islands. In this chapter, I do not focus on the full extent of differences in resumptive dependencies cross-linguistically, instead only focussing on the differences between English and Hebrew. However, it is likely easy to modify the learning strategy outlined here to accommodate other classes of grammaticized resumptive languages. For instance, Zaenen *et al.* (1981) show that in Swedish, there are grammaticized resumptive pronouns reserved for the subject position like English. These resumptive pronouns appear to display the full range of semantic properties associated with gaps/traces, unlike resumptive pronouns in English or Hebrew. If the learner knows that resumptive dependencies restricted to the subject position are an instance of Swedish-type resumptives, and these are not restricted to island contexts (as they are in English), then the Hebrew learner is not likely to infer that her target grammar is like Swedish. This is because out of the 152 resumptive pronouns found in our corpus, only 29 were in the subject position, and 11 of which were in the copular constructions, which are likely not analyzed as pronouns at all, but copulas (Doron,

1986).

3.5 Wider Theoretical Implications

In this chapter, I expanded the logic used in Chapter 2 to another domain of linguistic variation, the processing profile of resumptive dependencies inside islands. Unlike the *that*-trace constraint, the proper characterization of the cross-language differences are murkier with resumptive pronouns generally, and are complicated by non-trivial interactions between the grammar and the larger comprehension/production systems, as described in 3.2. However, I argued that a simple explanation for these differences in processing profiles is to posit that resumptive relativizations in islands are grammatical in Hebrew, but not in English. This follows on work in psycholinguistics that posits that resumptive pronouns in English island contexts are ungrammatical, often with the implication that this is not the case in Hebrew. Instead, English resumptive pronouns are repair strategies, constructed via an anaphoric, extrasyntactic mechanism (Erteschik-Shir, 1992; Alexopoulou & Keller, 2007). However, the evidence for this cross-language contrast has been thin on the ground, and has focussed predominantly on judgment studies (McDaniel & Cowart, 1999; Alexopoulou & Keller, 2007; Heestand *et al.*, 2011; Omaki & Nakao, ms), which in turn have a non-obvious interpretation (Keffala, 2013). However, if my interpretation of the cross-language contrast between Keshev & Meltzer-Asscher's (ms) and the findings in 3.3 are correct, then this provides stronger evidence of the purported difference between the mechanisms underlying resumptive dependencies

in Hebrew and English.

Additionally, Keshev & Meltzer-Asscher's (ms) findings and the findings in section 3.3 have consequences for the "islands debate", discussed in Chapter 1. On resource-based accounts, filler-gap dependencies are not perceived as unacceptable due to island constraints, but because they are too costly to construct in real-time comprehension (Pritchett, 1991; Kluender & Kutas, 1993; Kluender, 1998, 2005; Hawkins, 1999; Hofmeister & Sag, 2010). On this view, any evidence that the comprehender constructs filler-gap dependencies into islands when unforced is surprising. The findings that English-speakers can construct resumptive dependencies inside islands as in Experiments 3a–3c, and that Hebrew speakers do so actively, presents serious challenges to these accounts. This is because it shows that the comprehension system can favor constructing long-distance dependencies inside island contexts as long as the dependency has the right formal characteristics. A more suitable explanation of the comprehender's suppression of filler-gap dependency resolution inside islands is to posit that the parser constructs finely-detailed representations that are sensitive to grammatical constraints. On this view, English has no active search inside islands for relativization/*wh*-dependencies because there are no representations in which they may resolve grammatically into islands. Other dependencies, however, may actively resolve into islands (Yoshida *et al.*, 2014). This contrasts with Hebrew, in which there are grammatical representations in which a relativization dependency resolves into island contexts, via resumptive pronouns. I discuss the relation between locality and active dependency formation in more detail in Chapter 4.

3.6 Conclusion

In this chapter, I presented two main findings. The first finding is that resumptive dependencies in islands in English are constructed in a bottom-up, pronoun-driven fashion. This contrasts with Hebrew, in which resumptive dependencies are constructed in an active fashion. I argued that a parsimonious explanation of this variation is to posit that in Hebrew, resumptive pronoun dependencies inside islands are grammatically licensed, whereas they are an ungrammatical construction English. If so, this implies that the learner must infer that resumptive pronouns in islands are grammatical in Hebrew. I argued that this is not possible by only relying on the distribution of resumptive dependencies in islands. Instead, I proposed that the learner must indirectly learn that resumptive dependencies inside islands are grammatical by observing resumptive dependencies outside islands. If resumptive dependencies inside and outside islands are constructed via the same grammatical mechanism in Hebrew, then this provides a parsimonious indirect learning account. However, claiming that resumptive dependencies inside islands in English and Hebrew are grammatically distinct is inconsistent with analyses that treat them as representationally the same. Additionally, the supposition that the Hebrew learner infers that resumptive dependencies inside islands are grammatical based on observing resumptive dependencies outside islands requires that the learner considers these instances of the same abstract construction, which is most compatible analyses that treat resumptive dependencies as instances of the same dependency across constructions within the same language.

Chapter 4: Locality and Word Order in Bangla Active Dependency Formation

4.1 Introduction

In the previous chapters, I argued that analyses of grammatical differences can be evaluated by spelling out their psycholinguistic commitments. In this chapter, I demonstrate a previously unattested cross-linguistic difference in real-time processing, like Chapter 2. I show that active dependency formation preferences in Bangla are sensitive to locality as defined in terms of both structure and linear order, and discussed in Chapter 1. I do this by capitalizing on word order flexibility in Bangla, which permits a within-language test of the interactions between word order and structural position. These findings contrast with findings from Japanese, in which filler-gap dependencies preferentially resolve with the first position linearly available, regardless of structural depth, as described in Chapter 1. Thus, there is a cross-language difference with regard to whether "structural locality" influences the comprehension system's preferences for filler-gap dependency resolution. Unlike Chapter 2, I argue that this cross-language difference in Japanese and Bangla does not necessarily license positing different representational analyses of filler-gap de-

pendencies in Bangla and Japanese. Instead, I suggest that these differences may largely be methodological, stemming from the relative differences in informativity at the point of the filled-gap region.

Specifically, I show that Bangla speakers preferentially resolve filler-gap dependencies with the first position that is grammatically licensed. This mirrors the findings from Japanese, described in Chapter 1. However, sensitivity to the disruption of resolution with the first position available differs across experimental paradigms. In ambiguity resolution tasks, Bangla speakers preferentially resolve a filler-gap dependency with the first position linearally available, regardless of its structural position, like Japanese speakers (Omaki *et al.*, 2014). However, Bangla speakers only exhibit the filled-gap effect when the filled-gap is in the same clause as the filler, i.e., when it is local in structural terms. This differs from findings in Japanese (Aoshima *et al.*, 2004; Yoshida, 2006).

This chapter proceeds as follows. In section 4.2, I describe the relevant syntactic properties of Bangla, highlighting how they vary from Japanese and English. In section 4.3, I present the results from three experiments – a Question After Story task, based on Omaki *et al.* (2014), a self-paced reading task, and an acceptability judgment task. The results from the question after story task show that Bangla speakers preferentially resolve filler-gap dependencies with the first position linearly available. The self-paced reading task and the acceptability judgment task suggest that disrupting dependencies from resolving in this position is not costly. I address the cross-experiment and cross-language differences in section 4.4.

4.2 Dimension of Variation: Word Order and Active Dependency Formation

In Chapters 2 and 3, I argued that the underinformativity of the learner's primary linguistic data motivated constraining syntactic analyses in specific ways to license inference across constructions. In this chapter, the dimension of grammatical variation of interest is the positioning of embedded clauses with respect to the main clause verb. This is unlikely to present the same learnability challenges as the *that*-trace constraint or resumptive pronouns, since it presumably does not require very sophisticated parsing capabilities or complicated inferential strategies to detect whether a language permits embedded clauses to precede the embedding verb as in Japanese, follow the embedding verb as in English, or both as in Bangla. However, the word order properties of Bangla are valuable for testing whether the (presumably universal) parsing principles intended to explain the behavior both the English and Japanese speakers with regard to filler-gap dependency resolution can also capture the behavior of speakers of a language like Bangla, which permits the word orders observed in both English and Japanese.

Bangla, often called Bengali, is a language spoken primarily in Bangladesh and the eastern Indian state of West Bengal. There are approximately 180 million speakers worldwide (Lewis *et al.*, 2015), with a large international diaspora. Bangla is part of the Eastern Zone of the Indo-Aryan branch of the Indo-European language family. Due to its contact with multiple language families, Bangla features many

properties typical of northern Indo-Aryan languages, Dravidian languages, and languages in the Southeast Asian language area. For more complete descriptions of the language, see Thompson (2012) and David (2015).

The variable positioning of embedded clauses is demonstrated in (1). Embedded clauses that follow the main verb position may surface with the complementizer *je*, and embedded clauses that precede the main verb may surface with the complementizer *bole*, or with *je* in a non-initial position, shown in (2). Dasgupta (2007) describes this clause-internal *je* as an "anchor", which may be a distinct category from the complementizer *bole* or the clause-initial complementizer *je*. Examples taken from Bayer (1996).

(1) a. še bollo ora ašbe
he said they will come
b. še ora ašbe bollo
he they will come said
'He said that they will come'

(2) a. chele-ṭa bollo [$_{CP}$ je tar baba ašbe]
boy-CL said that his father will come
b. chele-ṭa [$_{CP}$ tar baba ašbe bole] bollo
boy-CL his father will come that said
c. chele-ṭa [$_{CP}$ tar baba je ašbe] bollo
boy-CL his father that will come said
'The boy said that his father will come'

The precise array of syntactic and semantic properties of each embedded clause position remains unclear. Bayer (1996) shows that the complementizer *bole* cannot co-occur with certain verbs, although it is unknown what defines the class of verbs that may co-occur with *bole*. Preverbal CPs containing the non-initial anchor *je* may encode distinct discourse statuses than other embedded clauses (Bhattacharya, 2001,

2002; Bayer & Dasgupta, to appear). The derivational status and structural position of embedded clauses, both before and after embedding verbs, are contentious (Bal, 1990; Bayer, 1996, 1999, 2001; Simpson & Bhattacharya, 2000, 2003).

Next, I turn to the properties of filler-gap dependencies in Bangla. Scrambling is freely available to either the immediate preverbal position or the left-edge position, as shown in (3) and (4), (Simpson & Bhattacharya, 2000, 2003; Bayer, 2001). This results in a filler-gap dependency, like scrambling in Japanese or *wh*-movement in English. Embedded clauses that precede or follow the embedding verb both permit extraction.[1]

(3) a. jɔn bhablo [$_{CP}$ meri bollo [$_{CP}$ su hæmleṭ poṛeche]]
John thought Mary said Sue Hamlet read

b. jɔn bhablo [$_{CP}$ meri **hæmleṭ** bollo [$_{CP}$ su ___ poṛeche]]
John thought Mary **Hamlet** said Sue ___ read

c. jɔn **hæmleṭ** bhablo [$_{CP}$ meri bollo [$_{CP}$ su ___ poṛeche]]
John **Hamlet** thought Mary said Sue ___ read
'John thought that Mary said that Sue has read Hamlet'

(4) a. **tomar beṛal-ke** amra šɔbai [$_{CP}$ paš-er baṛi-r kukur ___ kamṛeche
your cat-Acc we everyone neighbor-Gen dog ___ bit
bole] šunechilam
that heard
'We had all heard that the neighbor's dog has bitten your cat'

b. **bas theke** amar didi [$_{CP}$ ɔtogulo duronto bacca ___ laphiye
bus from my sister so many uncontrollable child ___ jumping
nambe bole] bhabe ni
descend.3.Fut that think didn't
'My sister hasn't thought that so many children could jump down from a bus.

(5) **sita-ke** jɔn bhablo [$_{CP}$ je ahmed ___ dekhlo]
Sita-Acc John thought that Ahmed ___ saw

[1]Compare and contrast this with findings from Basque (Uriagereka, 1992) and Malayalam (Srikumar, 2007), in which extraction from preverbal clauses is completely unavailable and restricted by base position of the extracted phrase, respectively.

'John thought that Ahmed saw Mary'

The precise characterization of the positioning of *wh*-phrases in Bangla is controversial. For my purposes, it suffices to show that *wh*-phrases can head filler-gap dependencies, regardless of their representational properties. This suffices for motivating the experiments in this chapter. However, I describe the different perspectives on the positioning of *wh*-phrases for completeness.

In simple clauses, Bangla appears to be *wh-in-situ*, shown in (6). There is no apparent filler-gap dependency when a *wh*-phrase takes wide scope over the sentence, and the *wh*-phrase appears in its canonical position.

(6) a. raj sita-ke dekhlo
Raj Sita-ACC saw
'Raj saw Sita'

b. raj ka-ke dekhlo
Raj who-ACC saw
'**Who** did Raj see __ ?'

However, an *in-situ wh*-phrase may only take wide scope when its containing clause precedes the verb. The *wh*-phrase *ke* 'who' in the embedded clause in (7-a) may be interpreted with wide or narrow scope, i.e., this sentence may be a query about who will come, or it may be a statement that someone knows the identity of the person who will come. Conversely, when *ke* 'who' is contained in an embedded clause that follows the main verb, only this second low-scope interpretation is available, (7-b). Simpson & Bhattacharya (2000, 2003) propose that these facts follow from a CP-pied-piping rule in Bangla, i.e., the entire clause moves to the preverbal position to fix the scope of the embedded *wh*-phrase. However,

Davison (2007) shows that sentential complements containing the complementizer *bole*, which canonically appear preverbally, also allow embedded *wh*-phrases to scope over the whole sentence, as shown in (8). This seems difficult to reconcile with a CP-pied-piping analysis.

(7) a. ora [$_{CP}$ ke ašbe] šuneche
they who will come heard
'They have heard who will come'
'**Who** have they heard __ will come?'
b. ora šuneche [$_{CP}$ ke ašbe]
they heard who will come
'They have heard who will come'
*'**Who** have they heard __ will come?'

(8) ora [$_{CP}$ [$_{CP}$ dilip ka-ke khun korbe] bole] jante perechilo?
they Dilip who-Acc kill will do that to know were able
'**Who** have they come to know that Dilip killed __ ?'

A *wh*-phrase may take wide scope if it is extracted from a postverbal embedded clause. For instance, the sentences in (9) are interpreted as queries of what Sue read, who left, and which illness Ram died from, respectively. In contrast to (7-b), the *wh*-phrase does not take low scope (Simpson & Bhattacharya, 2000, 2003):

(9) a. jɔn **ki** bhablo [$_{CP}$ meri bollo [$_{CP}$ su __ poṛeche]]
John what thought Mary said Sue __ read
'**What** did John think that Mary said that Sue read __ ?'
b. jɔn **ke** bollo [$_{CP}$ __ cole gæche]
John who said __ left went
'**Who** did John say __ left?'
c. tumi **ki** **ošukh-e** bhable [$_{CP}$ ram __ mara gæche]
you which illness-Loc thought Ram __ dead go.3.Pst
'**Which illness** did you think that Ram died of __ ?'

In these examples, the *wh*-filler appears in the preverbal position, which is the

canonical focus position in South Asian languages. However, Bangla *wh*-fillers and other topicalized/focussed material may appear in several positions, including the left edge (see Manetta 2012 for discussion of the preverbal position, and Choudhury 2010 for focus strategies in Bangla).

To summarize: in Bangla, a *wh*-phrase may head a filler-gap dependency, either as a subcase of a general scrambling rule, or as part of a *wh*-movement rule to fix scope. Additionally, embedded clauses may precede or follow the main verb. A phrase may be extracted from embedded clauses on either side of the embedding verb. Schematically, this means the following four structures are permissible in Bangla:

(10) a. **Preverbal embedded clause, embedded clause resolution:**
łdots wh …[$_{\text{CP}}$ …___ …] …V …

b. **Preverbal embedded clause, main clause resolution:**
…wh …___ …[$_{\text{CP}}$ …] …V …

c. **Postverbal embedded clause, embedded clause resolution:**
…wh …V …[$_{\text{CP}}$ …___ …] …

d. **Postverbal embedded clause, main clause resolution:**
…wh …___ …V …[$_{\text{CP}}$ …] …

If the parser attempts to resolve filler-gap dependencies with the first position linearly available, as established by the cross-language findings described in Chapter 1, then the structures in (10-a) and (10-d) should be preferred to the structures in (10-b) and (10-c) respectively. This is because configurations (10-a) and (10-d) correspond to representations in which the filler-gap dependency resolves with the first verb available, whereas (10-b) and (10-c) correspond to representations in which the filler-gap dependency resolves with the distant verb. We investigate these prefer-

ences in Experiment 5.

4.3 Experiments 5–7

Experiments 5–7 investigate the locality preferences in filler-gap dependency resolution across word orders in Bangla. Experiment 5 was a Question After Story task, Experiment 6 was a self-paced reading task, and Experiment 7 was an acceptability judgment task.

4.3.1 Experiment 5

The goal of Experiment 5 was to probe for effects of word order on the preferred resolution sites of ambiguously extracted filler-gap dependencies across word orders. Experiment 5 was a within-language adaptation of Omaki *et al.*'s (2014) cross-language question after story experiments. In this task, participants watched a series of animated vignettes in which a character acted out an event in one location and reported on it in another location. Afterwards, participants were asked an ambiguous *wh*-question. Participants' responses to the ambiguous question are taken to reflect the parse that they ultimately selected. If the participant preferred to resolve the filler-gap dependency with the first position linearly available, then Bangla speakers were predicted to preferentially resolve the dependency with the main clause when the main clause verb was the first verb in the string, and with the embedded clause when the embedded verb was the first verb in the string.

4.3.1.1 Methods

Experiment 5 was adapted from Omaki *et al.*'s (2014) Question After Story task (de Villiers *et al.*, 1990). Participants were instructed in Bangla to watch a sequence of 8 vignettes, with the order of the vignettes varying between participants. At the end of each vignette, the screen displayed "Write your answer now" in Bangla (এখন আপনার উত্তরটা লিখুন।). At this point, the experimenter paused the video and instructed the participant to read the question printed on a paper questionnaire. Participants were instructed to write a brief response. This was because pilot studies revealed that participants attempted to recapitulate large portions of the story in this task. After responding, the experimenter resumed the video, which then progressed to the next vignette.

There were 96 participants recruited for Experiment 5. Forty-eight participants were collected from the student population at The University of Dhaka in Dhaka, Bangladesh, and 48 participants from the student population at Calcutta University in Kolkata, India. Bangladeshi participants were compensated 500 Bangladeshi Taka (BDT) for participating, and Indian participants were compensated 200 Indian Rupees (INR). This session took approximately 15 minutes. Experiment 5 was conducted after participants completed either Experiment 6 or after an unreported experiment. These populations were each split into two groups, a "within-subjects" and a "between-subjects" group, as discussed in section 4.3.1.2.

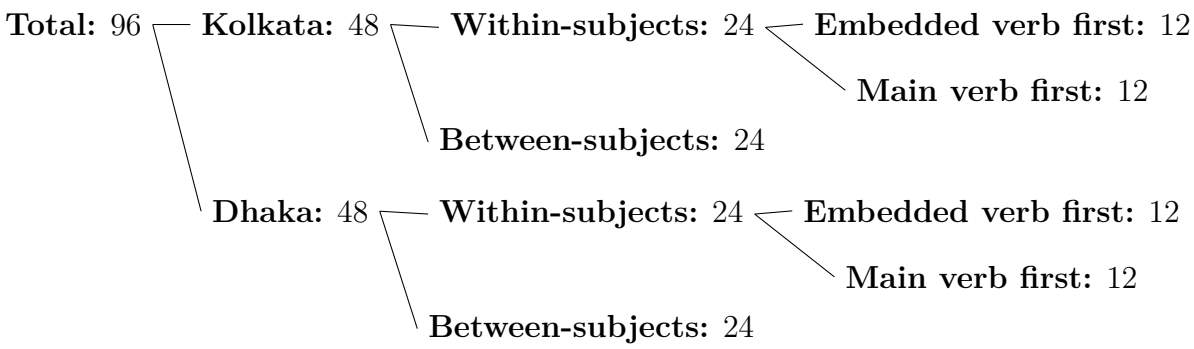


Figure 4.1: Division of participants in Experiment 5 into lists.

4.3.1.2 Materials

The materials were adapted from Omaki *et al.* (2014). The stories and audio were translated by DAC, MI, and SMM to standard colloquial Dhakaiya Bangla. Each participant saw a video with the vignettes presented in one of 8 different orders, systematically varying which video was first. Each participant received a paper questionnaire with each vignette. On the questionnaire, there was one question for each animation, 4 of which were target items and the other 4 of which were fillers. One group of 24 participants saw "within-participants" questionnaires, in which the target items alternated between main verb first and embedded verb first word orders. We counterbalanced which videos corresponded to target items and which videos corresponded to fillers. The other group of 24 was further divided into two groups for the "between-participants" group. Each group of 12 only saw main verb first word orders, and the other group of 12 only saw embedded verb first word orders. The between-subjects was intended to eliminate any potential self-priming effect within a list and replicate Omaki *et al.*'s (2014) cross-language findings within the same population. This division of participants is illustrated in Figure 4.1.

The stories were animated vignettes made from a sequence of clipart images. In each vignette, a character goes to four different locations, and performs an action in each. Afterwards, participants read an ambiguous *wh*-question, and participants were instructed to write a response to the question. A sample story from Omaki *et al.*'s (2014) English study is given in (11), and sample target item questions in Bangla are given in (12).

(11) [**Introduction**]
It was a beautiful day in spring so Lizzie decided she was going to go catch butterflies in the park.
[**1st Location**]
Her Mom and Dad weren't home, so Lizzie thought she should tell her brother or sister about going to the park, so that Mom and Dad would know where she was when they got back. She first went to her brother's room, but he was taking a nap and she couldn't tell him about catching butterflies. [**2nd Location**]
Instead, Lizzie looked for her sister. She looked all over the house but didn't see her sister anywhere! When she was about to give up, Lizzie heard her sister's voice in the basement! She went to the basement and said to her sister: "I'm gonna catch butterflies in the park!"
[**3rd Location**]
Then, on her way to the park, Lizzie passed by a parking lot and saw a butterfly near it. She walked slowly towards the butterfly, but before Lizzie could get there, another girl came along and caught the butterfly! Lizzie didn't see any more butterflies there, so she kept walking towards the park.
[**4th Location**]
There were lots and lots of butterflies in the park, and she caught one in a jar and took it home with her. She liked the one that she caught, but she wished she could have caught more butterflies.

(12) a. **Main Verb First:**

šumi kothae ækjɔn-ke boleche [$_{CP}$ je se prɔjapoti dhorbe]?
Shumi where someone-Acc told that she butterfly will catch

b. **Embedded Verb First:**

šumi kothae [$_{CP}$ se prɔjapoti dhorbe bole] ækjɔn-ke boleche?
Shumi where she butterfly will catch that one-Acc told

'Where did Shumi (=Lizzie) tell someone that she will catch butterflies'

Each vignette consisted of six phases. The first phase introduced the protagonist, and displayed him or her magnified in the center of the screen. The following four phases depicted him or her at each of the four locations. The protagonist succeeded or failed to perform some intended action as announced in the introductory phase, or succeeded or failed to report on it. In the sixth and final phase, the protagonist returns to the center of the screen, and then the story concludes.

The first two and last two locations were relevant for either the main clause event (i.e., the reporting event), or the embedded clause event (i.e., the intended action). In each location, the protagonist either successfully performed or attempted and failed to perform the intended action from the introduction phase, or successfully reported or attempted and failed to report on the intended action. In locations where the protagonist succeeded, there was a visual trace left behind (i.e., a butterfly in a bottle, or a word balloon for the reporting event). This is illustrated in (12). The contrast between successes and failures was intended to make the event-location pairings more memorable, and to ensure that the 'where' test questions were felicitous.

To avoid any potential recency bias, the ordering of the stories were counterbalanced such that the first pair of events pertained to the reporting event in half of the stories, and the embedded clause event in the other half. In each case, the story provided motivation for continuing to the next series of events. For instance, in (11), the reporting events are motivated by Lizzie's (= Shumi) need to tell her sib-

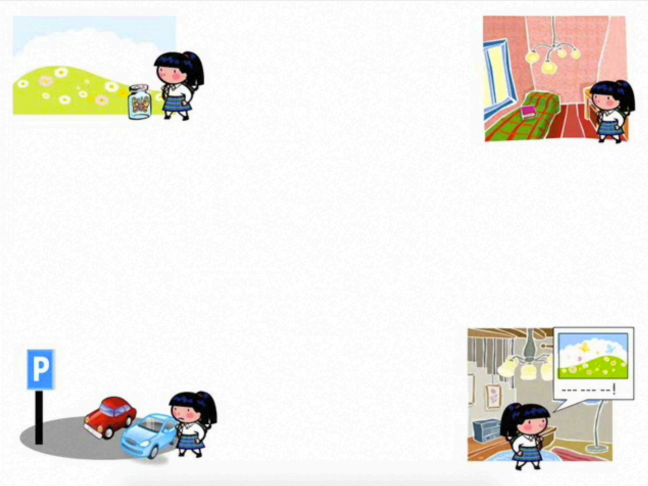


Figure 4.2: Sample image of a vignette with visual traces left in each critical location in Experiment 1.

lings where she was going. Additionally, the pairing of quadrant position and event was randomized across stories so that participants could not predict which locations would correspond to which actions. See Omaki *et al.* (2014) for more discussion.

4.3.1.3 Results

Each target item response was coded as a main clause response or an embedded clause response, depending on the named location. Responses that either failed to answer the question, or that provided both possible answers were excluded. These responses were evenly distributed between the two conditions – 24% for the main verb first condition, and 27% for the embedded verb first condition. The proportion of main clause responses by condition is plotted in Figure 4.3. I fit a logistic mixed effects model using the `lmer` and `lmerTest` packages in R (Bates *et al.*, 2014; Kuznetsova *et al.*, 2014). I fit a logit mixed effects model with main clause responses as a dependent variable. Main effects of list type (within-subjects

list or between-subjects list), word order (main verb first or embedded verb first), and location (Dhaka or Kolkata) were included in the analysis. Random effects of participant, list, were also included in the analysis. There was a main effect of word order ($\hat{\beta} = 5.47\ (1.04), z = 5.27, p < 0.0001$). The results are graphed in Figure 4.3.

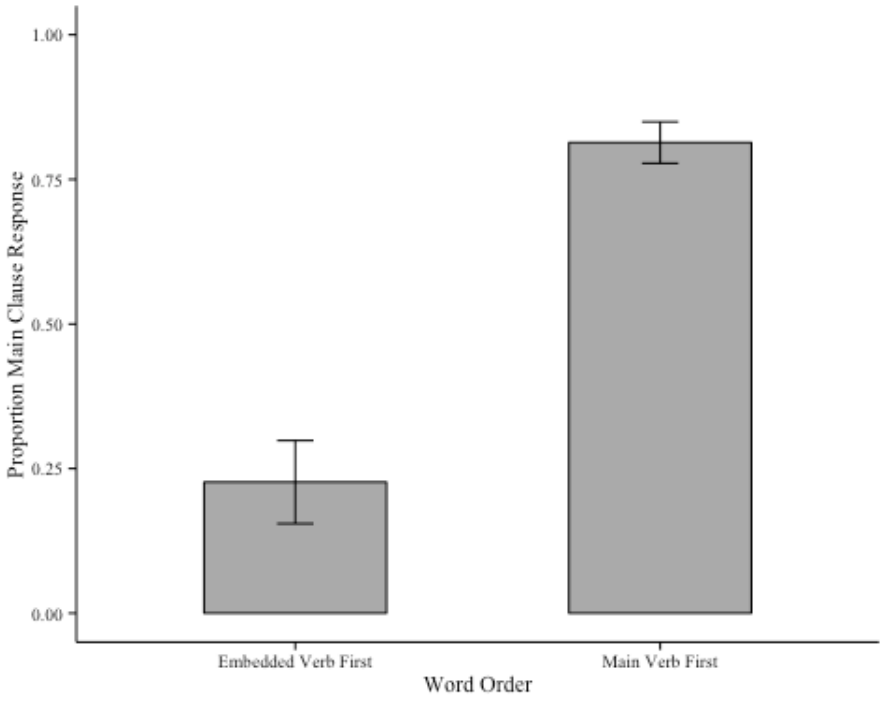


Figure 4.3: Proportion main clause response by word order in Experiment 5. Error bars correspond to two standard errors of the mean. Proportions are collapsed across list types and locations.

4.3.1.4 Discussion

These results reveal that Bangla speakers preferentially resolve a filler-gap dependency with the first position linearly available, regardless of whether this position is in the main clause or in an embedded clause, i.e., structurally distant. This replicates the cross-language findings from Omaki *et al.* (2014) within the same

language. This within-language comparison is valuable, because it "irons out" any potential cross-language confounds that may have influenced any of Omaki *et al.*'s (2014) results, for instance, the difference between *wh*-movement dependencies and scrambling dependencies (Saito, 1985). The question after story task has a few differences from classic paradigms used for probing for active dependency formation, like the filled-gap effect or plausibility mismatch effect. An advantage of this task is that it directly probes participants' preferred resolution sites instead of measuring sensitivity to disruption, as in the filled-gap paradigm. However, the question after story task is also an offline task, and thus does not directly reveal the time course of filler-gap dependency resolution. Thus, it may not necessarily index early resolution site commitments transparently. For these reasons, Experiment 6 uses the filled-gap paradigm in a self-paced reading task to probe for early commitment to resolution sites in real time.

4.3.2 Experiment 6

The goal of Experiment 6 was to probe for early commitment to an early dependency resolution site using a real-time measure. Experiment 6 uses the filled-gap paradigm, as discussed in Chapter 1. Crucially, Experiment 6 diverges from Experiment 5 by relying on detecting early resolution preferences by disrupting the linearly local resolution of the filler-gap dependency, instead of permitting participants to select their preferred parse for a globally ambiguous sentence. If the comrpehender attempts resolution with the first verb linearly available as in Experiment 5, I

predict that disrupting filler-gap resolution with the main clause in main verb first word orders will yield a filled-gap effect, and similarly that disrupting filler-gap resolution with the embedded clause in embedded verb first word orders will also yield a filled-gap effect.

4.3.2.1 Methods

Experiment 6 was conducted using Ibex software[2]. Although this platform is intended for experiments conducted on the internet, Experiment 6 was conducted in person by a native-speaker experimenter. This is because Ibex is entirely web-based, and it was possible to display Bangla characters in an easy-to-read font using a Unicode-enabled web browser. Experiment 6 was a moving window self-paced reading task (Just *et al.*, 1982). In this task, stimuli were first presented as a series of dashes. Participants pressed the spacebar to reveal the first word. Pressing the space again hid the first word, and then revealed the next word hidden under a dash. Participants proceeded word-by-word in this fashion. The experimenter asked participants to read at a comfortable but quick pace. To ensure that they attended to the stimuli, there was a yes/no comprehension question that displayed at the end of the sentence. Participants pressed the F key to respond yes, and the J key to respond no, and received on-screen reminders of these key-response pairings. They received feedback for both correct and incorrect responses. There were also a number of practice trials before the main experimental phase to familiarize the participants with the technique. All instructions, feedback, and experimental text

[2] `http://spellout.net/ibexfarm`

were in Bangla.

There were 32 participants in Experiment 6. These participants were drawn from the University of Dhaka student community. Due to a technical error, 3 of these participants' data were not recorded, and thus we report on 29 participants. Participants were compensated 500 BDT for their time.

4.3.2.2 Materials

We used a 2×2 Latin Square design for Experiment 6, crossing word order (main verb first / embedded verb first) and extraction type (argument / adjunct extraction). There were 32 target items and 48 complexity-matched fillers. A sample set of target items is given in (13).

(13) a. **Main Verb First, Argument Extraction**

rašad / jiggæša koreche / [$_{CP}$ **ka-ke** / ḍakṭar-ṭa / khubi /
Rashad ask did **who-Acc** doctor-CL very
ɔbakbhabe / rugi-ke / bolechen / [$_{CP}$ je / tini / purano /
surprisedly patient-ACC told.POL that he.POL old
haspatal-e / __ / cikitša / korechen]]
hospital-at treatment did.POL

'Rashad asked who the doctor very surprisedly told the patient that he treated __ in the old hospital'

b. **Main Verb First, Adjunct Extraction**

rašad / jiggæša koreche / [$_{CP}$ **kothae** / ḍakṭar-ṭa / khubi /
Rashad ask did **where** doctor-CL very
ɔbakbhabe / rugi-ke / bolechen / [$_{CP}$ je / tini / purano /
surprisedly patient-ACC told.POL that he.POL old
haspatal-e / ta-ke / cikitša / korechen]]
hospital-at him-ACC treatment did.POL

'Rashad asked where the doctor very surprisedly told the patient that he treated him in the old hospital'

c. **Embedded Verb First, Argument Extraction**

rašad / jiggæša koreche / [$_{CP}$ **ka-ke** / ḍakṭar-ṭa / [$_{CP}$ tini /
Rashad ask did who-Acc doctor-Cl he.Pol
purano / haspatal-e / rugi-ke / cikitša / korechen] / bole /
old hospital-at patient-Acc treatment did.Pol that
khubi / ɔbakbhabe / ___ / bolechen]
very surprisedly told.Pol

'Rashad asked who the doctor very surprisedly told that he treated the patient in the old hospital'

d. **Embedded Verb First, Adjunct Extraction**

rašad / jiggæša koreche / [$_{CP}$ **kothae** / ḍakṭar-ṭa / [$_{CP}$ tini /
Rashad ask did where doctor-Cl he.Pol
purano / haspatal-e / rugi-ke / cikitša / korechen] / bole /
old hospital-at patient-Acc treatment did.Pol that
khubi / ɔbakbhabe / ta-ke / bolechen]
very surprisedly him-Acc told.Pol

'Rashad asked where the doctor very surprisedly told that he treated the patient in the old hospital'

All target items contained three clauses, which I call the main clause, the intermediate clause, and the embedded clause for the sake of describing these materials. Each target item also contained a filler-gap dependency in the intermediate clause that is forced to resolve with the third verb due to a filled-gap blocking resolution with the second verb. All target items begin with a proper name (*rašad*) followed by an embedding verb (*jiggæša koreche*, 'asked'). This is to ensure that participants cannot anticipate the upcoming word order of the intermediate and embedded clauses. This also allows the target items to be main clause declaratives, with the *wh*-phrase taking scope in its extracted position in the intermediate clause. After the embedding verb, there is a displaced *wh*-filler in the intermediate clause. In the argument conditions, this is the accusative-case marked *wh*-phrase *ka-ke* 'who-Acc'. At this point in the argument extraction conditions, *ka-ke* could

be construed as the indirect object of a verb in the intermediate clause or as the direct object of the embedded verb in the upcoming embedded clause, for which there is no evidence at this point in these conditions. After the *wh*-region, there is a nominative-marked NP that denotes a referent of high social status (*ḍakṭar-ṭa*, 'the doctor'). This nominative marking also makes it clear that the *wh*-phrase is the head of a filler-gap dependency, because accusative-marked NPs do not canonically appear before a nominative NP. After the nominative-marked subject in the intermediate clause, the two word order conditions diverge.

In the main verb first conditions, there is a heavy adverb (*khub ɔbakbhabe*, 'very surprisedly'), followed by another accusative-marked NP (*rugi-ke*, 'patient-ACC') and the intermediate verb (*bolechen*, 'said'/'told'). The accusative-marked NP *rugi-ke* 'patient-ACC' is the filled gap for the argument extraction conditions, since it blocks resolution of the filler as the object of the verb *bolechen* 'said/told'. The adjunct extraction cases serve as a control, because the adjunct filler-gap dependency is not blocked from resolving with the intermediate clause verb. Thus, increased reading times are expected for the argument extraction conditions compared to the adjunct extraction conditions. After this intermediate clause verb, the complementizer *je* introduces the embedded clause. The embedded clause starts with a polite 3rd person nominative pronoun (*tini* 'he.POL"), an then additional adjunct (*purano haspatal-e* 'at the old hospital'.) Afterwards, there is an accusative-marked NP (*ta-ke*, 'him') in the adjunct extraction conditions followed by the embedded clause verb (*cikitša korechen*, 'treated.POL'). This verb shows politenesss agreement, matching the embedded clause subject (*tini* 'he.POL'), which is biased

to refer to the subject of the intermediate clause (*ḍakṭar-ṭa*, 'the doctor'), due to its politeness marking. The argument extraction conditions feature the same sequence of regions, except with no direct object pronoun (*ta-ke* 'him-ACC'). This permits the filler-gap dependency headed by *ka-ke* 'who-ACC' to resolve as the direct object of the embedded verb.

In the embedded verb first conditions, the politeness-marked nominative pronoun (i.e., the subject of the embedded clause, *tini* 'he.POL') immediately follows the intermediate subject (*ḍakṭar-ṭa* 'the doctor'). This sequence of two nominative-marked NPs unambiguously signals the embedded clause (see Miyamoto 2002). The embedded adverbial (*purano haspatal-e* 'at the old hospital') follows the embedded clause subject pronoun, followed by an accusative-marked NP (*rugi-ke* 'the patient-ACC') and the embedded clause verb (*cikitša korechen* 'treated.POL'). Here, the accusative-marked NP is again the filled gap in the argument extraction conditions, and the adjunct extraction condition serves as the control. After the filled gap NP, both conditions contain the embedded clause complementizer (*bole*), followed by the heavy adverbial (*khubi ɔbakbhabe* 'very surprisedly'). After the embedded clause, the adjunct extraction conditions contain an indirect object (*ta-ke* 'him-ACC') and the intermediate verb (*bolechen* 'said/told'), morphologically marked for politeness agreement with the polite subject pronoun (*tini* 'he.POL'), which again is biased to resolve with the intermediate clause subject (*ḍakṭar-ṭa* 'doctor-CL'). The argument extraction conditions again lacked the accusative-marked pronoun (*ta-ke* 'him-ACC'), which allows the filler-gap dependency to resolve with the intermediate clause verb.

4.3.2.3 Results

Analyses were conducted on comprehension task response accuracy and reading times. All reading times in sentences that received incorrect responses were removed from analysis. There were 4 participants whose accuracy fell below 75%, and were also therefore removed from analysis. The mean accuracy on the comprehension questions was 80.6% after removing these 4 participants.

For analysis, I constructed a mixed effects model for each word order condition and for each region using the `lmerTest` package in R (Kuznetsova *et al.*, 2014), with reading time as dependent measure. Main effects of extraction type were included in the analysis. Participant and item were included as random effects. For the main verb first condition, there was no effect of extraction type at the filled gap region. However, in the immediately following spillover region, there was a main effect of extraction type, raising reading times for argument extractions ($\hat{\beta} = 78\ 37, t(63.5) = 26.96, p = 0.03$). There was also a main effect of extraction type on the 2nd word of the embedded adjective, raising reading times for argument extractions ($\hat{\beta} = 74\ 30, t(258) = 2.49, p = 0.01$). In the embedded verb first conditions, there was no effect of argument extraction in any region, including the the filled-gap region or its spillover region. The reading times by region are given in Figure 4.4 for the verb first conditions, and in Figure 4.5 for the embedded verb first conditions.

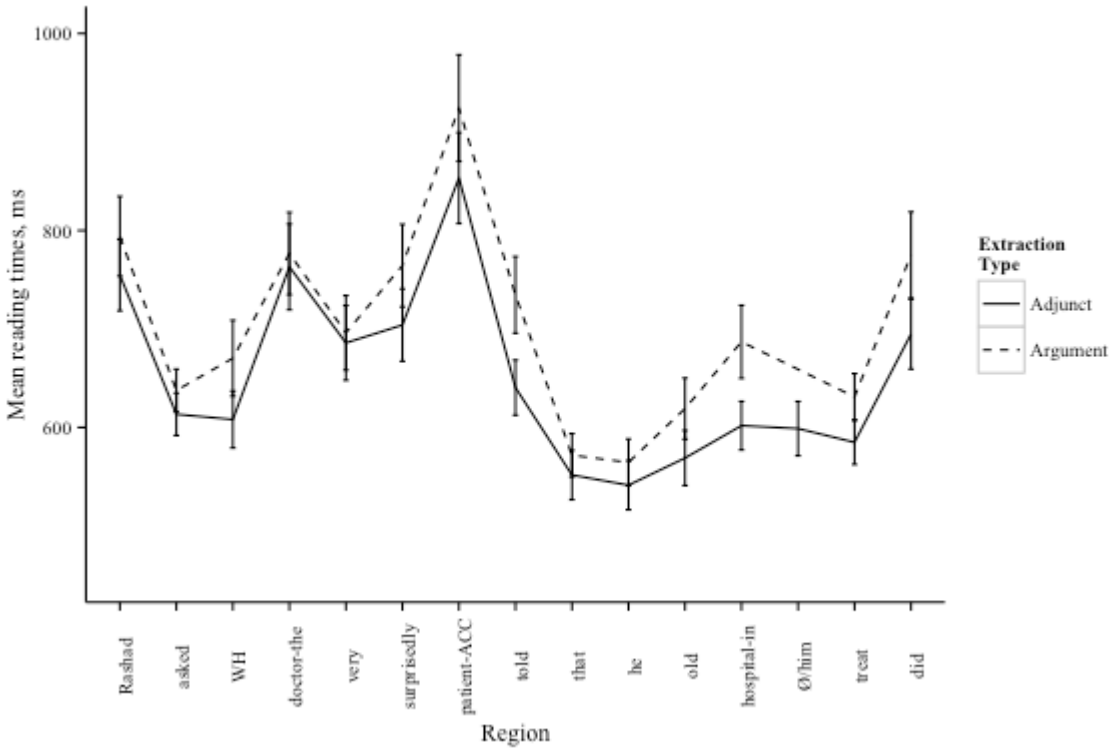


Figure 4.4: Mean reading times by region in milliseconds for the main verb first conditions in Experiment 6. Line type corresponds to extraction type. Error bars represent two standard errors from the mean.

4.3.2.4 Discussion

In Experiment 6, there was a filled gap effect with the main verb first word order conditions, but not in the embedded verb first word order conditions. *Prima facie*, this contradicts the findings from Experiment 5, in which there was a robust categorical preference to resolve the filler-gap dependency with the embedded verb in embedded verb first sentences. These tasks differed in a number of ways that might explain this difference, however. Experiment 5 probed for offline resolution preferences for globally ambiguous sentences, whereas Experiment 6 probed for early

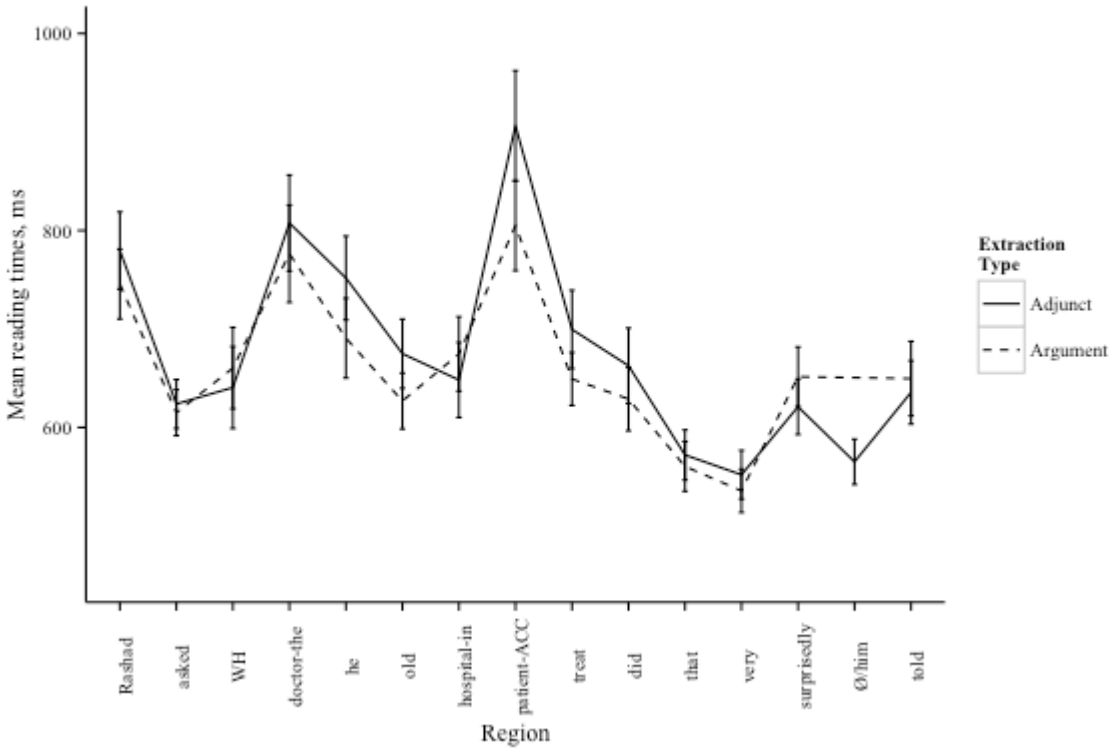


Figure 4.5: Mean reading times by region in milliseconds for the embedded verb first conditions in Experiment 6. Line type corresponds to extraction type. Error bars represent two standard errors from the mean.

commitment to early resolution by disrupting this parse in real-time.

This failure to find a filled-gap effect in preverbal embedded clauses may be an experimental artifact due to these differences between tasks. Experiment 5 was a substantially simpler task than Experiment 6. Experiment 5 was an offline task that used comparatively simple, two-clause sentences, whereas Experiment 6 was a self-paced reading task with three-clause sentences. Because of these extra complications in Experiment 6, participants may simply have been unable to maintain the filler in memory, or may have otherwise been overloaded during the task. If so, then the lack of a filled-gap effect in the preverbal embedded clause may simply reflect

abandonment of search in these conditions. For instance, suppose that the participants did not construct fully detailed representations in Experiment 6. For the main verb first conditions, participants could plausibly detect the local ungrammaticality at the filled-gap region using this strategy, because there are two accusative-marked NPs only separated by an adjunct. However, in the embedded verb first conditions, the error signal is more indirect, since it relies on matching multiple arguments to multiple upcoming verbs to detect that *wh*-filler cannot be integrated with the embedded verb. Additionally, the filler and the filled-gap are separated by a clause boundary in the main verb last word order, which may increase complexity costs, thereby reducing fidelity of the representations being built. Experiment 7 attempts to resolve some of these confounds by using a filled-gap paradigm in an offline task with substantially simpler sentences than in Experiment 6.

4.3.3 Experiment 7

The goal of Experiment 7 was to probe for sensitivity to disrupted resolution of a filler-gap dependency in various positions using a filled-gap paradigm in an acceptability judgment task. Although the filled-gap paradigm is typically used with online measures like self-paced reading, the reanalysis processes underlying the filled-gap effect are known to decrease acceptability in offline judgment tasks (Sprouse, 2008). Thus, it is possible to probe for the same effect from Experiment 6 using a task that is more natural than the self-paced reading task in Experiment 6. Additionally, the items in Experiment 7 are simpler than the items in Experiment

6, since the target items only have two levels of embedding, instead of three levels of embedding like in Experiment 7.

If there are decreased ratings when dependency resolution is blocked with the first verb across word orders, this supports the claim that filler-gap dependencies preferentially resolve with the first position available, as found in Experiment 5. This then means that the lack of a filled-gap effect in the preverbal embedded clause in Experiment 6 is likely due to some task-related difficulty. Conversely, if there is no filled-gap effect in preverbal embedded clauses, this suggests that the conflicting results from Experiment 5 and 6 reflect that these tasks are measuring different components of the filler-gap dependency resolution process.

Additionally, in Experiment 7 I investigate whether there is a filled-gap effect in the linearly distant position, i.e., the embedded clause in main verb first word orders or the main clause in the embedded verb first word orders. One possible explanation for the lack of filled gap effect with embedded verb first word orders in Experiment 6 is that the parser actively predicts a resolution site with the filler's clausemate verb, i.e., the later main clause verb. In Experiment 6, we only probed for sensitivity to disruption of resolution with in the embedded clause with embedded verb first word orders. Thus, probing for a filled-gap effect with the main verb in embedded verb first word orders may help diagnose whether the lack of filled-gap effect in embedded verb first word order in Experiment 6 was simply due to the parser attempting resolution with the later main verb.

4.3.3.1 Methods

Experiment 7 was conducted on a paper questionnaire. Participants were instructed to read the sentences carefully, and then to circle a number ranging from 1 to 7. Participants were given sample sentences with values already circled to illustrate how to use the scale, with an ungrammatical sentence marked with a score of 2 and a grammatical sentence marked with a score of 7. The experiment was conducted in 8 lists, with equal number of items per condition per list, and an equal number of items per condition across lists. The experiment lasted approximately 10–20 minutes.

Participants were drawn from the University of Dhaka and Calcutta University student population. There were 32 participants from each population. Participants in Dhaka were compensated 500 BDT for their time, and participants in Kolkata were compensated 200 INR. They either performed Experiment 4 or an unreported experiment before this task.

4.3.3.2 Materials

The materials in Experiment 7 are similar to the ones used in Experiment 6, except without the outermost embedding clause. The target items were presented in a 2×2×2 design, crossing word order (main verb first or embedded verb first), extraction type (argument extraction or adjunct extraction), and position of the filled gap (local, distant). There were 24 groups of target items, with 36 complexity-matched fillers, 18 of which were ungrammatical. The materials are exemplified in

(14).

(14) a. **Main Verb First, Local Filled Gap, Argument Extraction**

jahid **ka-ke** khubi ɔbakbhabe tar bondhu-ke boleche [$_{CP}$ je
Jahid **who-Acc** very surprisedly his friend-Acc told that
nipa parṭy-te ___ dekheche]?
Nipa party-at ___ saw

'Who did Jahid very surprisedly tell his friend that Nipa saw ___ at the party?'

b. **Main Verb First, Local Filled Gap, Adjunct Extraction**

jahid **kothae** khubi ɔbakbhabe tar bondhu-ke boleche [$_{CP}$ je nipa
Jahid **where** very surprisedly his friend-Acc told that Nipa
parṭy-te *pro* dekheche]?
party-at *pro* saw

'Where did Jahid very surprisedly tell his friend that Nipa saw (him) at the party?'

c. **Main Verb First, Distant Filled Gap, Argument Extraction**

jahid **ka-ke** khubi ɔbakbhabe ___ boleche [$_{CP}$ je nipa parṭy-te
Jahid **who-Acc** very surprisedly ___ told that Nipa party-at
tar bondhu-ke dekheche]?
his friend-Acc saw

'Who did Jahid very surprisedly tell that Nipa saw his friend at the party?'

d. **Main Verb First, Distant Filled Gap, Adjunct Extraction**

jahid **kothae** khubi ɔbakbhabe *pro* boleche [$_{CP}$ je nipa parṭy-te tar
Jahid **where** very surprisedly *pro* told that Nipa party-at his
bondhu-ke dekheche]?
friend-Acc saw

'Where did Jahid very surprisedly tell (him) that Nipa saw his friend at the party?'

e. **Embedded Verb First, Local Filled Gap, Argument Extraction**

jahid **ka-ke** [$_{CP}$ nipa parṭy-te tar bondhu-ke dekheche bole] khubi
Jahid **who-Acc** Nipa party-at his friend-Acc saw that very
ɔbakbhabe ___ boleche
surprisedly ___ told

'Who did Jahid very surprisedly tell ___ that Nipa saw his friend at the party'

f. **Embedded Verb First, Local Filled Gap, Adjunct Extraction**

jahid **kothae** [$_{CP}$ nipa parṭy-te **tar bondhu-ke** dekheche bole] khubi
Jahid **where** Nipa party-at his friend-Acc saw that very

ɔbakbhabe *pro* boleche
surprisedly *pro* told

'Where did Jahid very surprisedly tell (him) that Nipa saw his friend at the party'

g. **Embedded Verb First, Distant Filled Gap, Argument Extraction**

jahid **ka-ke** [$_{CP}$ nipa parṭy-te ___ dekheche bole] khubi
Jahid **who-Acc** Nipa party-at saw that very
ɔbakbhabe tar bondhu-ke boleche
surprisedly his friend-Acc told

'Who did Jahid very surprisedly tell ___ that Nipa saw his friend at the party'

h. **Embedded Verb First, Distant Filled Gap, Adjunct Extraction**

jahid **kothae** [$_{CP}$ nipa parṭy-te *pro* dekheche bole] khubi ɔbakbhabe
Jahid where Nipa party-at *pro* saw that very surprisedly
tar bondhu-ke boleche
his friend-Acc told

'Where did Jahid very surprisedly tell (him) that Nipa saw his friend at the party'

There were a few differences between the target items in Experiment 6 and Experiment 7. First of all, the target items in Experiment 7 contained only two clauses with *wh*-filler in the main clause, similar to the materials in Experiment 5. Additionally, the *wh*-phrase appeared in the canonical preverbal position like in Experiment 5, not the left-edge position as in Experiment 6. Lastly, the gap/indirect object pronoun variation between the argument/adjunct extraction conditions from Experiment 6 was removed. This is not problematic, because in Bangla the verb *bɔla* can either mean 'to say' or 'to tell' – i.e., the indirect object is optional. Furthermore, Bangla permits null objects in certain conditions (Simpson *et al.*, 2013). In (14), I marked the potential indirect object position for the adjunct conditions as *pro*, although this may not necessarily reflect the interpretation assigned to these

sentences.

4.3.3.3 Results

For analysis, I constructed a mixed effects model with rating as a dependent variable. Main effects for city, list, word order, filled gap position, and extraction type were included in the analysis. Interaction effects of word order, filled gap position, and extraction type were also included. Participant and item were included as random effects. There was a main effect of word order ($\hat{\beta} = 0.87\ (0.20), t(1434) = 4.27, p < 0.0001$), and a three-way interaction effect between word order, filled gap position, and extraction type, lowering ratings in the main verb first, local filled gap, argument extraction condition ($\hat{\beta} = -1.38\ (0.40), t(1434) = -3.42, p = 0.0006$). This last interaction corresponds to the filled-gap effect in Experiment 6, since ratings improve when the local resolution of a filler-gap dependency is licensed with main verb first word order. All other factors were not significant. The means are plotted in Figure 4.6.

4.3.3.4 Discussion

The results from Experiment 7 largely corroborate the results from Experiment 6. There was a decrease in acceptability judgments when the filler-gap dependency was unable to resolve with the first position linearly available only in main verb first word order conditions. Thus, the lack of a filled-gap effect with embedded verb first word orders in Experiment 6 is plausibly not due to participants assigning shallow

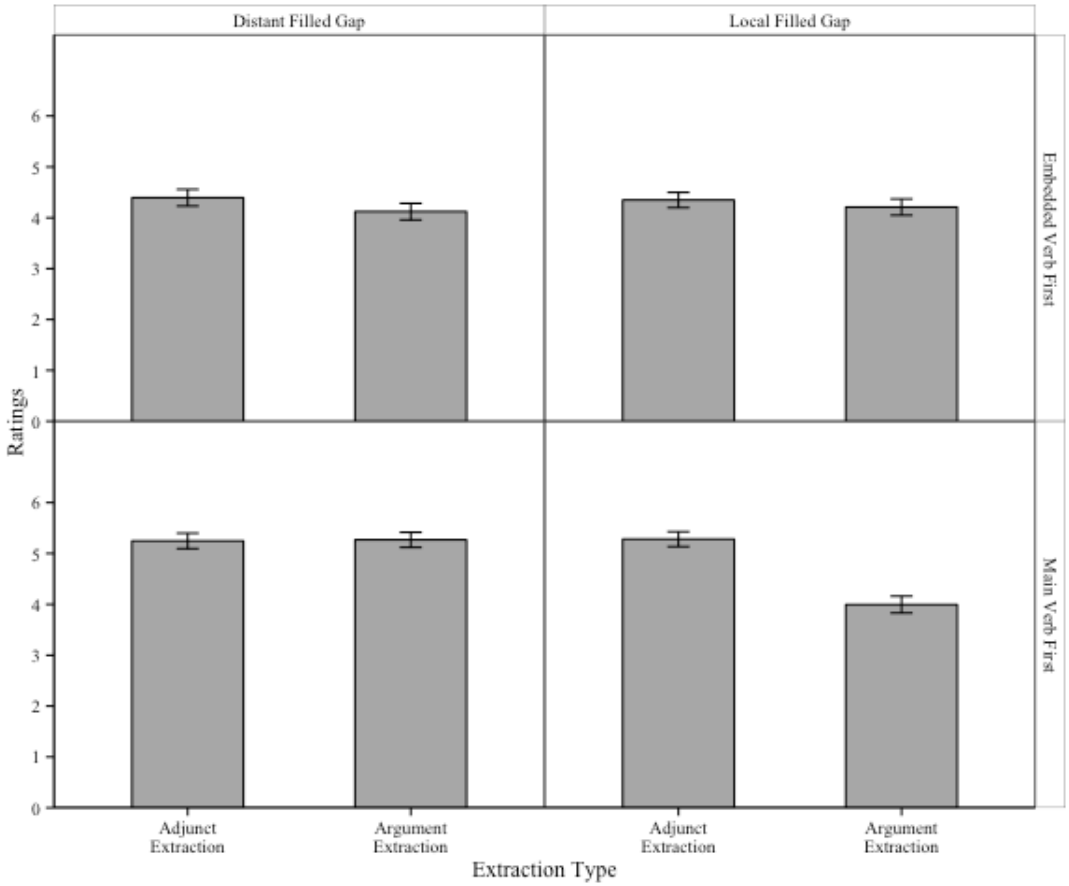


Figure 4.6: Mean ratings of target items by condition in Experiment 7. Error bars represent two standard errors from the mean.

representations to the experimental items or the unnaturalness of the task, since this was less to occur in Experiment 7. Instead, the contrasts between Experiment 5 on one hand and Experiments 6 and 7 on the other must ultimately arise from the way the Bangla parser behaves in the filled-gap paradigm and in ambiguity resolution tasks. In other words, the divergence in results from Experiment 5 on one hand and Experiments 6 and 7 on the other reveal that the processes underlying filled-gap effect detection as a means to detect locality preferences in filler-gap dependencies are different in some important way than ambiguity resolution mechanisms.

Additionally, there was no filled-gap effect in the embedded verb first word order in Experiment 7, in either the embedded clause or the main clause. This was not measured in Experiment 6. Thus, the lack of filled-gap effect with the embedded verb first word orders in Experiment 6 cannot be due to a preference for resolution with the later main verb, otherwise there would have been a filled-gap effect with the main clause verb in Experiment 7. In the next section, I speculate as to these discrepancies between Experiments 5–7 and the discrepancy between the Bangla findings and Japanese findings, and how this relates to the nature of cross-language variation. In the next section, I suggest that the differences between these results stem from the difference in sensitivity to the local unacceptability of a filled-gap with embedded verb first word orders in Bangla. This partially explains the contrast between the Bangla and Japanese findings. I argue that these results do not merit alternative analyses of Bangla and Japanese, nor positing differences in parsing strategies between these two languages.

4.4 Wider Theoretical Implications

Experiment 5 replicated the bias to resolve filler-gap dependencies with the first position available, regardless of structural position. This was a within-language replication of what had been observed between languages (Aoshima *et al.*, 2004; Omaki *et al.*, 2014). However, Experiments 6 and 7 only revealed a preference for resolution with the first verb available when the first verb was also a clausemate of the filler, in contrast with findings from Japanese. Thus, there appears to be a

contrast between the biases that guide how the parser resolves a globally ambiguous sentence in tasks like Experiment 5 and the prediction and recovery mechanisms that yield the filled-gap effect in Experiments 6 and 7. Additionally, there is a contrast between how these predictive and recovery mechanisms interact between Bangla and Japanese such that Japanese speakers do show a filled-gap effect in these constructions, whereas Bangla speakers do not. Here, I offer a number of possible explanations for these contrasts. The data at hand do not obviously support one conclusion over the others, and none capture the full range of data presented in this chapter. However, I argue that the differences within the Bangla findings and between Bangla and Japanese most likely lies in the mechanisms underlying the filled-gap effect – namely, detection of a local anomaly, and the reanalysis processes that instigate a search for an alternative representation that matches the bottom-up input. Importantly, this means that the results from Experiments 5–7 alone do not warrant proposing different representations for filler-gap dependencies in Bangla and Japanese, as I argued for resumptive dependencies in Chapter 2.

There are three possible explanations that I consider in this section. The first explanation is that there are differences in the accessibility of the relevant parses between constructions and languages. That is, the differences in sensitivity to disruption of the linearly local filler-gap dependency is due to a difference in ease of recoverability between the two languages or ease of error-detection. In other words, the reason the filled-gap effect was not observed in Experiments 6 and 7 in Bangla whereas there are filled-gap effects in similar experiments conducted in Japanese is ultimately due to the nature of mechanisms that yield the filled-gap

effect, and not the grammatical representations themselves. Another explanation is that the comprehension mechanisms use different strategies to actively construct dependencies in real-time. This explanation leverages some surface differences in filler-gap dependencies between Bangla and Japanese, and implies that superficial properties of the language can influence which processing strategy the comprehender uses. Finally, another possible explanation is that preverbal embedded clauses in Bangla have different representational status than those in Japanese which may yield distinct processing profiles once the relevant linking assumptions are carefully spelled out, similar to Chapter 2. I argue that these latter two approaches require a substantially more sophisticated linking hypothesis of the representation of a filler-gap dependency and the processes used to construct them in real-time than the data in this chapter motivates.

4.4.1 Mechanics of the Filled-Gap Effect

Traditionally, the filled-gap effect, e.g., the increased reading times at *us* in (15), is taken to index the cost of recovering from an early commitment to an erroneous parse. For instance, in (15), the parser initially commits to a representation in which *who* is the object of *bring*, then encounters *us*, detects a local ungrammaticality, rejects the current parse, and finally instigates a search for a different parse in which the filler-gap dependency resolves at a later position.

(15) My brother wanted to know **who** Ruth would bring us home to __ at Christmas.

The lack of filled-gap effect in preverbal clauses in Bangla may be attributable to a failure of any of these mechanisms to apply quickly and accurately. Consider the filled-gap effect in main verb first and embedded verb first word orders from Experiment 3, replicated in (16-a) and (16-b).

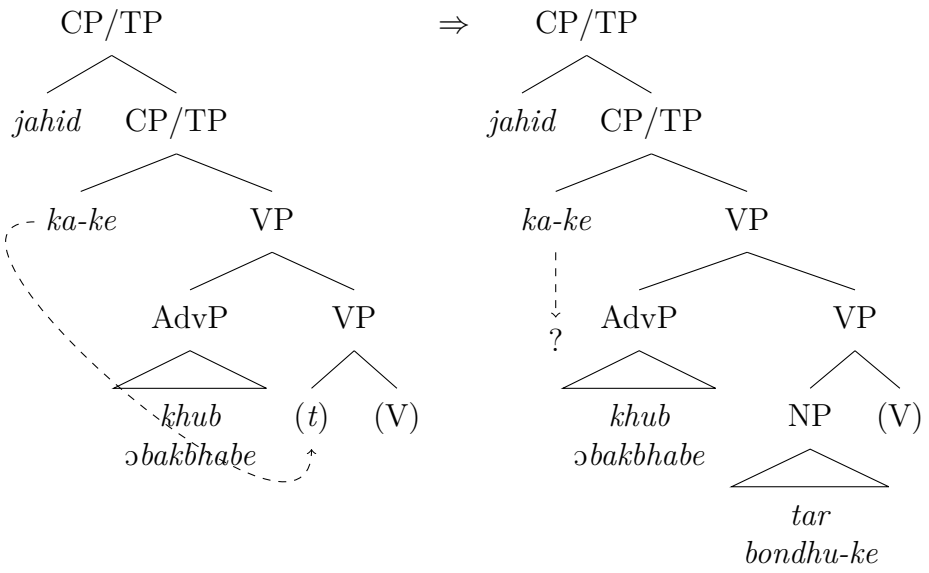
(16) a. jahid **ka-ke** khubi ɔbakbhabe tar bondhu-ke boleche [$_{CP}$ je
Jahid who-ACC very surprisedly his friend-ACC told that
nipa parṭy-te ___ dekheche]?
Nipa party-at ___ saw
'Who did Jahid very surprisedly tell his friend that Nipa saw ___ at the party?'

b. jahid **ka-ke** [$_{CP}$ nipa parṭy-te tar bondhu-ke dekheche bole] khubi
Jahid who-ACC Nipa party-at his friend-ACC saw that very
ɔbakbhabe ___ boleche
surprisedly ___ told
'Who did Jahid very surprisedly tell ___ that Nipa saw his friend at the party'

In the main verb first word order in (16-a), the filler is hypothesized to preferentially resolve with the first verb available, i.e., the main verb. Upon encountering *tar bondhu-ke* 'his friend-ACC', it is likely easy for the parser to detect the local ungrammaticality – the association of the verb *boleche* 'said' with two accusative-marked NPs. This is because both NPs are very recent and thus both plausibly in the focus of attention. Additionally, there is yet no evidence for an upcoming embedded clause at this point. Thus, the blocked resolution triggers a search for a multiclausal representation that can accommodate both NPs, which induces a cost (Miyamoto, 2002). The detectability of the error signal is therefore easily noticeable, and the alternative representation that can accommodate the open filler-gap dependency is not immediately obvious. Thus, the filled-gap sentences with main

verb first word order are similar to the more familiar filled-gap effect sentences like (15). The trees in (17) illustrate the disruption of the early commitment to a gap position (marked t) upon encountering the filled gap *tar bondhu-ke* 'his friend-ACC', and the subsequent intermediate representation in which the filler-gap dependency has no obvious resolution site.

(17)

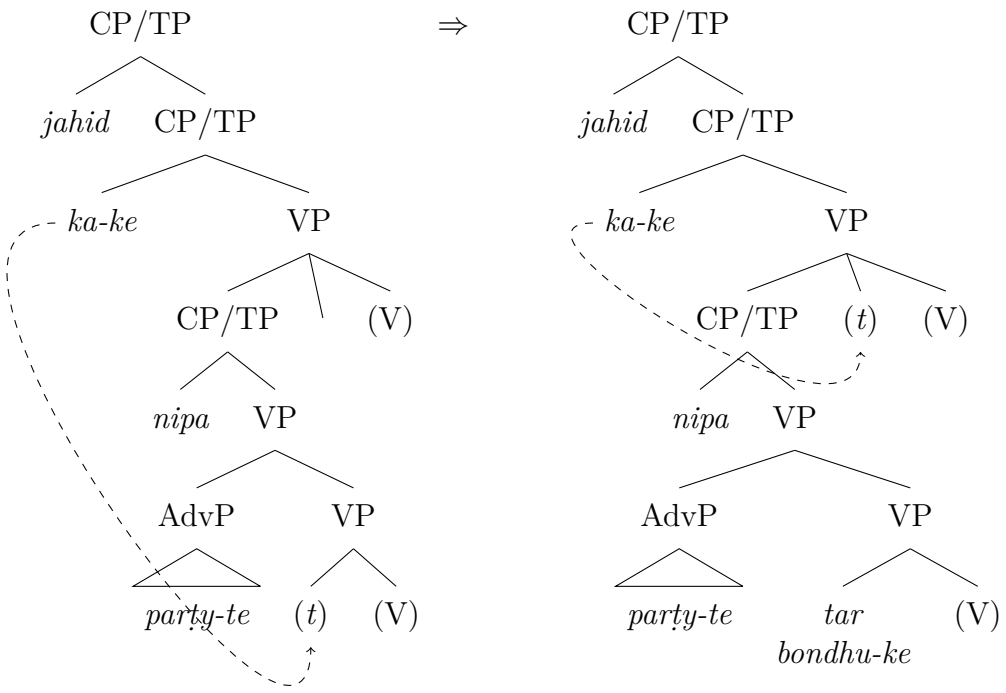


Conversely, in the embedded verb first conditions, the error signal may not be as easily detectable, because the filler is now separated by a clause boundary. Thus, the parser may not immediately detect that the verb has an extra accusative-marked NP (but see Wagers & Phillips 2014 for evidence of active dependency formation over long distances). Put differently, the error signal in this word order is only detectable by matching up arguments and verbs, and rapidly detecting that this structure requires a verb that takes two accusative-marked arguments in two different clauses. This error signal may be more difficult to detect. This difficulty may be compounded by the increased complexity of the center-embedding structure.

Thus, the local ungrammaticality in the embedded verb first word orders may be less obvious than in the main verb first word orders.

Alternatively, the availability of an alternative parse may be more obvious in the embedded verb first word orders in Bangla. Given that the parser has detected two nominative NPs, the parser may have already constructed both the embedded clause and the main clause predicates before encountering the filled gap. This may mean that the alternative parse, i.e., a parse in which the dependency resolves with the main verb, is more easily accessible than in the main verb first word order, and thus there is minimal cost in readjusting the predicted resolution site to the main VP. The shift from embedded preference to a main clause preference upon encountering the filled gap is illustrated in (18). Thus, the lack of a filled-gap effect in embedded verb first word orders may either be due to a failure to detect the local ungrammaticality, or the accessibility of the alternative, dispreferred parse.

(18)



This account essentially pins the lack of the filled-gap effect on the word order properties of the embedded verb first word orders. However, it does not immediately explain the mismatches between Bangla and Japanese, in which there are robust filled-gap effects in embedded clauses. One possible explanation is that since Japanese has more robust case-marking than Bangla, and perhaps has more cross-clausal scrambling, the detectability of the local ungrammaticality at the filled-gap site in experiments in Japanese is more easily detectable. In other words, the lack of filled-gap effects in Bangla is due to the relative undetectability of the mismatch in arguments in the embedded clause, which is more detectable in a language like Japanese. This suggests that filled-gap effects can be "turned off" or "turned on" in embedded clauses proportional to the relative detectability of the clause boundaries,

case-marking, acceptability of scrambling, and so on. Importantly, this explanation implies that filled-gap resolution preferences are universal, but the filled-gap paradigm is not always suitable for detecting these preferences. This does not imply any important differences in processing strategies across languages.

4.4.2 Structured Search vs. Linear Search

Another possible explanation of these facts relies on the observation that the scope of a *wh*-operator in Japanese is explicitly marked by a Q-morpheme, *-ka* or *-no*, but Bangla lacks an overt Q-marker[3]. The *wh*-extractions in Omaki *et al.*'s (2014) question after story tasks are only ambiguous with respect to the resolution site of the *wh*-filler, i.e., which verb it modifies. The *wh*-operator's scope is fixed over the entire sentence, as marked by the Q-morpheme *-no*, affixed to the main verb.

Yoshida (2006) explicitly compared *wh*-scrambling, which requires a Q-morpheme, to non-*wh*-scrambling dependencies, which do not bear any morphological dependency between the filler and its resolution site. Like Aoshima *et al.* (2004), he found

[3]Bayer (1996) claims that, much like Japanese, Bangla *wh*-scope can be marked by the overt scope marker *ki*, homophonous with the polar question marker:

(19) Dale ki bhablo Sarah ka-ke dekhlo
Dale Q thought Sarah who-ACC saw
'Who did Dale see that Sarah saw?'

However, I have not found a Bangla speaker that finds this construction grammatical. This discrepancy may be a result of language contact. Fixing scope by virtue of a Q-marker is allowed in Hindi (Manetta, 2012), and many Bangla speakers are competent in both Bangla and Hindi. However, even if this construction were robustly grammatical and common among Bangla speakers, the Q-morpheme *ki* precedes the *wh*-filler. Thus, the *wh*-filler would not trigger an active, prospective search for the Q-morpheme as I suggest for Japanese.

a filled-gap effect in embedded clauses for *wh*-scrambling dependencies. However, Yoshida failed to find a filled-gap effect for non-*wh* scrambling in Experiment 2 in his Chapter 3. In the relevant experiment, Yoshida (2006) compared sentences like (20-a) and (20-b) in a self-paced reading task. In (20-a), the dative-marked NP heads a scrambling dependency that ultimately resolves as an argument of the main verb *osieta* 'notified'. If the comprehender attempts to resolve the dependency with the first verb, then the embedded dative-marked NP *butyō-ni* 'department director-DAT' should block this preferred resolution, yielding a filled-gap effect. However, he found no difference in reading times at this region.

(20) a. **Zyosē-syain-ni** senmu-wa [$_{CP}$ syatyō-ga
female-worker-Dat senior director-TOP president-NOM
kaigi-de butyō-ni syōkyū-o
meeting-at department director-DAT promotion-ACC
yakusoku-sita-to] __ osieta
promise-did-that __ notified
'The senior director notified the female employee that the president promised a promotion to the department head at the meeting'

b. Zyosē-syain-wa senmu-ni [$_{CP}$ syatyō-ga
female-worker-DAT senior director-TOP president-NOM
kaigi-de butyō-ni syōkyū-o
meeting-at department director-DAT promotion-ACC
yakusoku-sita-to] osieta
promise-did-that notified
'The female employee notified the senior director that the president promised a promotion to the department head at the meeting'

Thus, there is an asymmetry in where filled-gap effects are found for long-distance scrambling even in Japanese. One possible explanation of this within-language contrast is that active dependency formation does not apply with non-*wh*-scrambling dependencies. However, it's unclear what kind of parsing architecture

would not need to resolve a non-*wh* filler-gap dependency, especially given that no other study shows such an asymmetry in dependency type[4]. A potentially more plausible explanation assigns significance to the fact that Japanese *wh*-scrambling requires a Q-morpheme, but Bangla *wh*-scrambling and Japanese non-*wh*-scrambling does not. Suppose that since *wh*-dependencies must resolve with a verb and a scope marker (which may or may not attach to the same verb as the resolution site), the Japanese parser uses a different search strategy specifically for *wh*-scrambling dependencies than other long-distance dependencies. For instance, suppose that the parser actively pursues dependency resolution with the closest structural position in the general case, i.e., Japanese and Bangla speakers typically both preferentially resolve filler-gap dependencies with a clausemate verb, regardless of linear position. However, since *wh*-scrambling dependencies must discharge with both a structural position and a specific morpheme, the parser chooses to "scan" the string linearly instead of generating a structured gap prediction. In other words, because *wh*-scrambling in Japanese is "overloaded" as a structural and morphological dependency, the parser prioritizes discharging a morpheme-to-morpheme dependency defined over the string over predicting and later confirming a gap site in a structured representation. This in effect means that the filled-gap effect in preverbal embedded clauses in Japanese does not actually reflect a universal to resolve a filler-gap dependency with the closest verb linearly available, but rather a construction-specific adaptation to resolve morphological dependencies linearly instead of structurally.

[4]The filled-gap effect in embedded clauses has been found with other filler-gap dependencies in Japanese, such as exclamative constructions (Ono *et al.*, 2006). Importantly, the exclamative construction also requires a filler-morpheme dependency. Thus, this accounts predicts that exclamative dependencies should display the same processing profile as *wh*-scrambling dependencies.

Importantly, this accounts predicts that the Bangla parser does not opt to scan the string linearly, since there is no expectation for a Q-morpheme, nor does the Japanese parser for non-*wh*-scrambling.

Although this account captures the differences between the Japanese findings and the Japanese-Bangla contrast, there are still some problems that must be resolved. For instance, recall that Aoshima *et al.* (2004) found a filled-gap effect occurred in the preverbal direct object position. Thus, if *wh*-scrambling dependencies in Japanese trigger a string-scanning operation, this supposedly linear search must be partially structured to detect that the preferred dependency is blocked before encountering the verb. In other words, scanning the string must use sufficient structural information to determine whether the presence of an object blocks resolution of the later phrase-morpheme dependency. Secondly, this account is inconsistent with findings from Experiments 5 and 7, since it cannot explain the categorical embedded verb bias with the embedded verb first word order, nor the lack of a filled-gap effect with the main verb in embedded verb first word order. In other words, the findings from these experiments suggest that the comprehender in fact preferentially resolves a filler-gap dependency with the first verb linearly available, even though there is no phrase-morpheme dependency. Thus, this account fails to explain the within-language differences in Bangla.

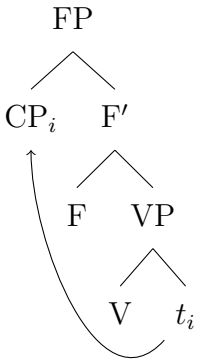
4.4.3 Representational Differences

The last explanation for the divergent findings within and between languages I consider attributes the behavioral differences to underlyingly different representations of filler-gap dependencies, which feeds into different processing strategies. Recall that in Japanese, clauses canonically surface in the preverbal position[5]. However, as mentioned in section 4.2 the status of the two clause positions in Bangla is controversial. One account holds that postverbal clauses surface in their base position, and preverbal CPs move to a higher position (Simpson & Bhattacharya, 2000, 2003; Bhattacharya, 2001, 2002; Bhattacharya & Simpson, 2012). Alternatively, CPs may base-generate preverbally, and appear after the verb via a dislocation operation (Bal, 1990; Bayer, 1996, 1999, 2001). Schematically, then, Bangla VPs headed by an embedding verb either have the representations in (21) or the representations in (22).

(21) **Head-Initial Analysis**:

Main verb first:

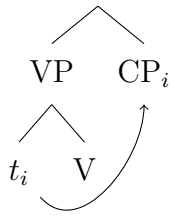
Embedded verb first:



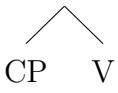
[5]Japanese does permit both preverbal and postverbal CPs, but postverbal CPs are marked in many respects, and the syntax literature virtually unanimously holds that postverbal CPs are right dislocated. See Tanaka (2001) for discussion.

(22) **Head-Final Analysis**:

Main verb first:



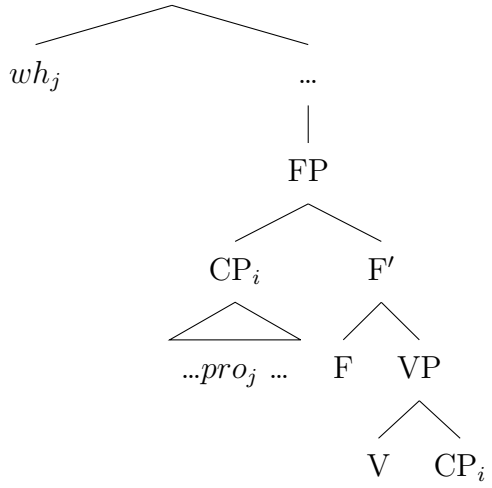
Embedded verb first:



Suppose that the head-initial analysis as in (21) is correct. If so, then preverbal CPs are moved to a higher, left-branching position in Bangla, whereas postverbal CPs surface in their base position. If so, then preverbal CPs in Bangla and Japanese are in different syntactic positions, with Bangla CPs in a derived, left-branching position, and Japanese CPs in a preverbal complement position. On this account, extraction from the preverbal CP in Bangla should be a CED violation (Huang, 1982; Chomsky, 1986; Uriagereka, 1999)[6]. Thus, movement dependencies are predicted to be unacceptable, apparently contrary to fact. However, it has also been argued that a null resumptive pronoun may resolve a filler-gap dependency into island constructions in some languages (Ishii, 1991; Nakao, 2009). If so, then we can explain apparent movement from preverbal CPs as an instance of a filler-null resumptive pronoun dependency:

[6]Jurka (2013) suggest that extraction from specifiers is improved if there is specifier-head agreement. Importantly, there is no overt morphological agreement between preverbal clauses and any detectable functional element. However, this requires more careful investigation into the verbs that license the complementizer *bole*. If there is reason to suspect that preverbal clauses in Bangla agree with some functional head, then extraction form preverbal clauses may not be degraded

(23)



If this is the correct analysis, then extraction from preverbal CPs in Bangla is a resumptive pronoun dependency, whereas extraction from preverbal CPs in Japanese is a movement dependency. In Chapter 3, I established that non-grammaticized resumptive languages like English suspend active dependency formation in island contexts, but can still construct dependencies into islands with resumptive pronouns nonetheless. If this is the correct analysis of Bangla, then the lack of active search inside of preverbal clauses follow from the fact that preverbal clauses are moved constituents. Put differently, the preverbal/postverbal alternation implies that extraction from a preverbal CP must be mediated via a null resumptive, which in turn means that filler-gap dependencies into these constructions must be constructed non-actively, as discussed in Chapter 3.

Prima facie, a hypothesis space in which preverbal clauses may permit filler-gap dependencies either through a movement operation or through a null resumptive dependency is likely to be difficult to acquire from primary linguistic data. However, this may not be problematic if the null resumptive strategy for preverbal clauses is

restricted to languages in which there is a free preverbal/postverbal alternation like Bangla. Suppose that the learner exposed to a language in which embedded clauses may freely alternate between a preverbal and postverbal position are forced to posit that the postverbal position is the base position, i.e., that this is the only grammar available which permits embedded clauses to precede or follow the main verb. In other words, the hypothesis space only permits languages in which embedded clauses surface in their base preverbal position, or languages in which embedded clauses base generate after the embedding verb, with an additional parameter that permits embedded clause movement to a higher specifier position. This would effectively ban languages that permit embedded clauses to base generate in multiple positions (motivated by Baker's (1988) UTAH), and languages that base generate with preverbal embedded clauses that move rightwardly to a specifier position. This is an effect a "soft-Antisymmetry" learner, in Kayne's (1994) sense.

If the hypothesis space is shaped this way, then the learner exposed to a language with free positioning of embedded clauses before and after the embedding verb must infer that embedded clauses uniformly base generate in a postverbal complement position. Thus, any extraction from a preverbal embedded clause cannot be movement, by deduction, preverbal embedded clauses are in specifier positions, and extraction is disallowed from specifiers. This is consistent with facts from Basque and Malayalam, languages in which from preverbal CPs are permitted, but but not postverbal CPs (Uriagereka, 1992; Srikumar, 2007).

Learning considerations aside, there are two problems with this analysis. First, this account again fails to explain the categorical embedded clause bias with embed-

ded verb first word orders in Experiment 5. It is unclear what kind of representation Bangla-speaking participants could construct with a filler-gap dependency headed by an adjunct filler in sentences like (24), as used in Experiment 5.

(24) šumi kothae [$_{CP}$ se prɔjapoti dhorbe bole] ækjɔn-ke boleche?
Shumi where she butterfly will catch that one-Acc told
'Where did Shumi tell someone that she will catch butterflies'

If resolution into preverbal embedded clauses can only be mediated via a null resumptive proform in Bangla, then Bangla speakers must have a bias to resolve with a null resumptive in sentences like (24). However, exceptionally attributing a bias to resolve with a null proform in Bangla in preverbal embedded clauses defeats the analytic strategy of pinning the apparent lack of active search on the a parser that fails to search for resumptive proforms in the first place. Additionally, resumptive pro-adverbials are not typically assumed to exist. In fact, the inability for gaps in island constructions to take an adverbial antecedent is considered evidence for the existence of null resumptive pronouns for this reason (Ishii, 1991). Thus, the null resumptive pronoun explanation is incapable of explaining how the sentence in (24) is preferentially interpreted with an embedded clause resolution, and why this parse is grammatical at all.

4.5 Conclusion

In this chapter, I presented data from three experiments in Bangla, a language in which embedded clauses may precede or follow the embedding verb. This permits

a within-language manipulation on the effect of word order in the processing of filler-gap dependencies. Results from Experiment 5, an offline ambiguity resolution task, provide further evidence for a global bias to resolve filler-gap dependencies with the first position linearly available. Moreover, Experiment 5 showed this bias by means of a within-language modulation of word order, which had been impossible in previous studies. Results from a self-paced reading task and an acceptability judgment task using a filled-gap paradigm, however, failed to show evidence of active search in preverbal embedded clauses. This contrasts with robust findings from Japanese (Aoshima *et al.*, 2004; Yoshida, 2006; Omaki *et al.*, 2014). I proposed several possible solutions for this difference between Bangla and Japanese and for the contrast between experimental paradigms. I suggested that an account which relies on a divergence between the detectability of filler-gap dependency resolution preferences in filled-gap paradigms and ambiguity resolution paradigms is likely to be the most successful explanation, although many details are left unexplored. Importantly, however, these differences do not commit the theorist to posit different representational accounts of the relevant constructions in Bangla and Japanese without a more sophisticated understanding of the independently motivated representational differences between these languages and more sophisticated linking hypotheses.

However, the bias to resolve a filler-gap dependency with the first position suggests that the parser's locality bias is ultimately sensitive to linear order across languages, as shown in Experiment 5. As discussed in Chapter 1, there are (at least) two notions of locality that are relevant for describing filler-gap dependencies. The first locality bias is whatever biases drive active dependency formation, as

discussed in this chapter and in section Chapter 3. The second is whatever notion of locality delineates acceptable representations from unacceptable representations (i.e., island constraints). Crucially, the locality bias distinguishing acceptable from unacceptable parses must be sensitive to hierarchical structure, due to contrasts like (25), in which the linearly closer resolution site is less acceptable than the linearly distant resolution site:

(25) a. *__Who__ did [$_{\text{NP}}$ the fact that Sarah saw __] surprise Dale?
b. __Who__ did [$_{\text{NP}}$ the fact that Sarah saw Bob] surprise __ ?

If the locality biases that drive active dependency formation are predominantly sensitive to linear order and not hierarchical structure, then explaining the asymmetry in (25) in terms of the biases underlying active dependency formation is non-trivial. This is a challenge to resource-based accounts of island phenomena, as discussed in Chapters 1 and 3. These accounts cannot simply rely on the biases underlying preferred filler-gap dependency resolution, since these appear to ultimately be local in nature, whereas island constraints are defined structurally. This critique persists despite the conflict between Experiment 5 and Experiments 6 and 7.

Lastly, there are some important differences between the results in Chapters 3 and 4. For the English/Hebrew contrast, the evidence in English largely implies that resumptive dependencies were not grammatical, since participants appeared to be resistant to construct them in multiple paradigms, whereas this was not the case in Hebrew. This easily lends itself to an explanation in which resumptive dependencies are ungrammatical in English, but grammatical in Hebrew. However, this

kind of explanation is unlikely to work for Bangla, since the difference ultimately appears to be in preferred representations, as opposed to possible representations. Had Experiment 5 not revealed a strong bias to resolve filler-gap dependencies with the preverbal embedded clause, then supposing that filler-gap dependency resolution into preverbal embedded clauses is ungrammatical in Bangla but grammatical in Japanese would have straightforwardly explained the differences in processing results in Experiments 6 and 7. However, this within-language contrast in Bangla (and the within-language contrast in Japanese discovered by Yoshida 2006) imply that the differences between these two languages is partially due to the way that comprehension systems construct and manage representations in real time, and not solely due to differences in the representations themselves. Thus, appealing to differences in representational inventories between the two languages is likely not sufficient for explaining the full range of data in this chapter.

Chapter 5: Conclusion

Traditionally, the goal of syntactic theory is to generate hypotheses about the range of possible representations present in a language, typically on the basis of patterns of acceptability judgments. This is a characterization of a speaker's linguistic competence. Psycholinguists hypothesize, through behavioral and theoretical studies, how these representations are built in real-time comprehension and acquired in development. However, these results rarely are then used to enrich the theories that syntacticians construct.

In this thesis, I argued that the range of analyses available can be constrained by evaluating these analyses against psycholinguistic data. Specifically, I argued that analyses of the *that*-trace effect and related phenomena and analyses of resumptive pronouns demand that the learner's input have specific distributional properties. I showed that these commitments can then be evaluated against realistic linguistic input, and moreover that certain analyses fare better than others. In Chapter 2, I argued that a possible interpretation of analyses of the *that*-trace effect imply that the learner must determine on a complementizer-by-complementizer basis whether that complementizer permits subject extractions. On this view, the learner's linguistic input largely underdetermines the range of knowledge that she demonstrates as a

linguistically competent adult. This implies that the learners must indirectly learn whether her target language is sensitive to the *that*-trace constraint, which I argue is possible on traditional analyses, such as the null subject parameter accounts.

I made a similar argument for resumptive pronouns in Hebrew and English in Chapter 3, where the relevant crosslinguistic difference was active vs. passive resumptive dependency formation, respectively. I diagnosed this difference in processing profiles as indicating that resumptive pronouns in islands are grammatical in Hebrew, but not English. This in turn implies that the learner must have sufficiently informative linguistic input to draw these differing conclusions. I argued that too likely requires indirect learning. I proposed that analyses in which resumptive pronouns in Hebrew, both inside and outside islands, form a "class" to the exclusion of resumptive pronouns in English were capable of capturing the processing differences between English and Hebrew, while licensing the inference across constructions in Hebrew. Importantly, this is more compatible with certain analyses of resumption across languages than others.

Lastly, in Chapter 4, I showed that there are apparently divergent processes underlying active dependency formation in Bangla. Although an offline task confirmed the observation that filler-gap dependencies preferentially resolve in preverbal embedded clauses, online and acceptability judgment measures failed to replicate this finding. This contrasts with findings in Japanese, where there is a robust preference to resolve filler-gap dependencies in a preverbal embedded clause. I argued that there are several ways of capturing these facts, and that a full explanation requires more serious investigation into the representational properties of Bangla

and Japanese, and a more sophisticated theory of locality in filler-gap dependency preferences across constructions and languages.

This dissertation almost uniquely focussed on filler-gap dependencies, because the dimensions of variation are well-constrained, and the syntactic and psycholinguistic properties of these dependencies are very well understood. However, the strategy used in this dissertation can extend to a number of constructions, including apparent inapplicability of Principle C in Mohawk (Baker, 1996) or Somali (Svolacchia & Puglielli, 1999), absence of quantifier-variable binding with certain pronouns in Japanese (Hoji, 1991, 1995), or subtle differences in passivization constructions in Vietnamese and Mandarin (Simpson & Ho, 2008). In fact, every finding in so-called micro-comparative syntax makes the hypothesis space the learner must navigate more nuanced, and studies in the properties of the learner's input more crucial. Additionally, well-known differences in real-time sentence comprehension, such as the differences in relative clause attachment ambiguities as described in Chapter 1, have been woefully understudied from a language acquisition perspective. Any apparent difference between two languages can be understood as a challenge to the learner, and additionally may present dimensions of variation for the comprehender. Articulating the relationships between analyses of these phenomena, realistic linguistic input, and real-time comprehension mechanisms is likely to be informative for theory-building in all three domains.

Bibliography

Ackerman, Lauren, Michael Frazier, & Masaya Yoshida, 2014. Resumptive pronouns ameliorate island violations in forced-choice tasks. Poster Presented at the Annual CUNY Sentence Processing Conference.

Aguirre, Carmen, 2004. Spanish aguirre corpus. CHILDES Data repository, oai:childes.talkbank.org:Romance-Spanish-MOR-Aguirre. Published online by TalkBank.

Albalá, José María & Victoria Marrero, 2004. Spanish marrero-albalá corpus. CHILDES Data repository, oai:childes.talkbank.org:Romance-Spanish-MOR-Marrero. Published online by TalkBank.

Alexopoulou, Theodora & Frank Keller, 2003. Linguistic complexity, locality, and resumption. In *Proceedings of the 22nd West Caost Conference on Formal Linguistics*, pp. 15–28. Somerville, MA: Cascadilla Press.

Alexopoulou, Theodora & Frank Keller, 2007. Locality, cyclicity, and resumption: At the interface between the grammar and the human sentence processor. *Language*, 88, pp. 110–160.

Antelmi, Donatella, 2004. Italian antelmi corpus. CHILDES Data respotiroy, oai:childes.talkbank.org:Romance-Italian-Antelmi.

Antinucci, Francesco & Domenico Parisi, 1973. Early language acquisition: A model and some data. In C. Ferguson & D. Slobin (eds.), *Studies of child language development*, pp. 607–619. New York: Holt.

Aoshima, Sachiko, Amy Weinberg, & Colin Phillips, 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language*, 51, pp. 23–54.

Aoun, Joseph, Lina Choueiri, & Norbert Hornstein, 2001. Resumption, movement, and derivational economy. *Linguistic Inquiry*, 32, pp. 371–403.

Ariel, Mira, 1990. *Accessing Noun-Phrase Antecedents*. London: Routledge.

Asudeh, Ash, 2012. *The logic of prominimal resumption*. Oxford: Oxford University Press.

Augurzky, Petra, 2005. *Attaching Relative Clauses in German: The Role of Implicit and Explicit Prosody in Sentence Processing*. Ph.D. thesis, University of Leipzig.

Badecker, William & Kathleen Straub, 2002. The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology*, 28(4), pp. 748–769.

Baker, Mark, 1988. *Incorporation: A theory of grammatical function changing*. University of Chicago.

Baker, Mark, 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.

Baker, Mark, 2001. *The atoms of language: The mind's hidden rules of grammar*. New York, NY: Basic Books.

Bal, Bijoy Kumar, 1990. *Comp and complementizers in Oriya and English*. Ph.D. thesis, Central Institute of English and Foreign Languages.

Bates, Douglas, Martin Maechler, Ben Bolker, & Steven Walker, 2014. lme4: Linear mixed-effects models using Eigen and s4. `http://CRAN.R-project.org/package=lme4`.

Bayer, Josef, 2001. Two grammars in one: Sentential complements and complementizers in Bengali and other South-Asian languages. In Peri Bhaskarao & Karamuri Venkata Subbarao (eds.), *The yearbook of South Asian linguistics and languages: Tokyo symposium on South-Asian languages – contact, convergence, and typology*, pp. 11–36. New Delhi: Sage Publications.

Bayer, Josef & Probal Dasgupta, to appear. Emphatic topicalization and the structure of the left periphery: Evidence from German and Bangla. *Syntax*.

Bayer, Josef & Martin Salzmann, 2012. *That*-trace effects and resumption – how improper movement can be repaired. In Patrick Brandt & Eric Fuß(eds.), *Repairs: The Added Value of Being Wrong*, pp. 275–333.

Bayer, Joseph, 1996. *Directionality and logical form*. Boston, MA: Kluwer.

Bayer, Joseph, 1999. Final complementizers in hybrid languages. *Journal of Linguistics*, 35, pp. 233–271.

Beltrama, Andrea & Ming Xiang, 2013. Resumption can rescue! islands and relative clauses in Italian and English. Poster presented at the Annual CUNY Sentence Processing Conference.

Benedet, Maira, Celis Cruz, Maria Carrasco, & Catherine Snow, 2004. Spanish becacesno corpus. CHILDES Data respository, oai:childes.talkbank.org:Romance-Spanish-MOR-BecaCESNo. Published online by TalkBank.

Berman, Ruth A. & Jurgen Weissenborn, 1991. Acquisition of word order: A crosslinguistic study. Final Report submitted to the German-Israeli Foundation for Scientific Research.

Berwick, Bob, 1985. *The acquisition of syntactic knowledge*. Cambridge, MA: MIT Press.

Bhattacharya, Tanmoy, 2001. The puzzle of Bangla comp-internal clauses. *Snippets*, 3.

Bhattacharya, Tanmoy, 2002. Peripheral and clause-internal complementizers in Bangla: A case for remnant movement. In *Proceedings of WECOL 2000*, pp. 100–112. Fresno, CA.

Bhattacharya, Tanmoy & Andrew Simpson, 2012. Sluicing in Indo-Aryan: An investigation of Bangla and Hindi. In Jason Merchant (ed.), *Sluicing in a cross-linguistic perspective*, pp. 183–218. Amsterdam: John Benjamins.

Bieberauer, Theresa, Anders Holmberg, Ian Roberts, & Michelle Sheehan (eds.), 2010. *Parametric variation: Null subjects in Minimalist theory*. Cambridge, MA: Cambridge University Press.

Boeckx, Cedric, 2010. What principles and parameters got wrong. Universitat de Autònoma de Barcelona.

Borer, Hagit, 1984. Restrictive relatives in Modern Hebrew. *Natural Language and Linguistic Theory*, 2, pp. 219–260.

Bourdages, Johanne S., 1992. Parsing complex nps in French. In Helen Goodluck & Michael S. Rochemont (eds.), *Island Constraints: Theory, acquisition, and processing*, pp. 61–87. Dordrecht: Kluwer.

Bresnan, Joan, 1977. Variables in the theory of transformations. In Peter Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal Syntax*, pp. 157–196. New York: Academic Press.

Brysbaert, Marc & Don C. Mitchell, 1996. Modifier attachment in sentence parsing: Evidence from dutch. *The Quarterly Journal of Experimental Psychology*, 49(3), pp. 664–695.

Carstens, Vicki, 1987. Extraction asymmetries in Yoruba. In David Odden (ed.), *Current Approaches to African Linguistics*. Dordrecht: Foris.

Chao, Wynn, 1981. Pro-drop languages and non-obligatory control. *University of Massachusetts Occasional Papers in Linguistics*, 7, pp. 46–74.

Chao, Wynn & Peter Sells, 1983. On the interpretation of resumptive pronouns. In Peter Sells & Charles Jones (eds.), *Proceedings of the 13th NELS Conference*, pp. 47–61. Amherst, MA: Graduate Linguistic Student Association.

Chater, Nick & Christopher D. Manning, 2006. Probabilistic models of language processing and acquisition. *TRENDS in Cognitive Science*, 10(7), pp. 335–344.

Chien, Yu-Chin & Kenneth Wexler, 1990. Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition*, 1, pp. 225–295.

Chomsky, Noam, 1977. On wh-movement. In Peter Culicover, Thomas Wasow, & Adrian Akmajian (eds.), *Formal Syntax*, pp. 71–133. New York: Academic Press.

Chomsky, Noam, 1981. *Lectures on Government and Binding: The Pisa Lectures*. The Hague: Mouton de Gruyter.

Chomsky, Noam, 1986. *Barriers*. Cambridge, MA: The MIT Press.

Chomsky, Noam, 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam, 2013. Problems of projection. *Lingua*, 130, pp. 33–49.

Chomsky, Noam & Howard Lasnik, 1977. Filters and control. *Linguistic Inquiry*, 8(3), pp. 425–504.

Choudhury, Arunima, 2010. A study of focus in Bangla.

Chow, Wing Yee, Shevaun Lewis, & Colin Phillips, 2014. Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology*, 5, p. 630.

Cinque, Guglielmo, 1992. The pseudo-relative and acc-*ing* constructions after verbs of perception. In *University of Venice Working Papers in Linguistics*. Venice: Università di Venezia.

Cohen, Jacob, 1988. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Erbaum.

Cohen, Jacob, 1992. Statistical power analysis. *Current directions in the psychological sciences*, 1(3), pp. 98–101.

Conroy, Anastasia, Eri Takahashi, Jeffrey Lidz, & Colin Phillips, 2009. Equal treatment for all antecedents: How children succeed with principle B. *Linguistic Inquiry*, 40(3), pp. 446–486.

Cowart, Wayne, 1996. *Experimental syntax: Applying objective methods to sentence judgments*. Thousand Oaks: Sage Publications.

Cowart, Wayne, 2003. Detecting syntactic dialects: The that-t phenomenon. In *39th Meeting of the Chicago Linguistic Society*.

Crain, Stephen & Janet Dean Fodor, 1985. How can grammars help parsers? In David Dowty, Lauri Karttunen, & Arnold M. Zwicky (eds.), *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives*, pp. 94–128. Cambridge, MA: Cambridge University Press.

Cresswell, Cassandre, 2002. Resumptive pronouns, wh-island violations, and sentence production. In *Proceedings of the Sixth International Workshop on Tree Adjoining Grammar and Related Frameworks (TAG+6)*, pp. 101–109. Venice: Università di Venezia.

Cuetos, Fernando & Don C. Mitchell, 1988. Cross-linguistic differences in parsing: Restrictions on the use of the late closure strategy in spanish. *Cognition*, 30(1).

Cuetos, Fernando, Don C. Mitchell, & Martin M. Corley, 1996. Parsing in different languages. In Manuel Carreiras, José García-Albea, & Núria Sebastián-Gallés (eds.), *Language processing in Spanish*, pp. 145–187. Mahwah, N.J.: Lawrence Erlbaum.

Culbertson, Jennifer, 2010. *Learning biases, regularization, and the emergence of typological universals in syntax*. Ph.D. thesis, Johns Hopkins University.

Culicover, Peter, 1993. Evidence against ECP accounts of *that*-trace effect. *Linguistic Inquiry*, 24(3), pp. 557–561.

Dasgupta, Probal, 2007. The ubiquitous complementizer. In Josef Bayer, Tanmoy Bhattacharya, & M.T. Hany Babu (eds.), *Linguistic Theory and South Asian Languages. Essays in Honor of KA Jayaseelan*, pp. 163–173. Amsterdam: John Benjamins.

David, Anne B., 2015. *Descriptive Grammar of Bangla*. New York: Walter de Gruyter.

Davison, Alive, 2007. Word order, parameters, and the extended COMP projection. In Josef Bayer, Tanmoy Bhattacharya, & M.T. Hany Babu (eds.), *Linguistic Theory and South Asian Languages. Essays in Honor of KA Jayaseelan*, pp. 175–198. Amsterdam: John Benjamins.

de Vicenzi, Marcia & Remo Job, 1995. An investigation of late closure: The role of syntax, thematic structure, and pragmatics in initial interpretation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, pp. 1303–01321.

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

de Vincenzi, Marica, 1991. *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.

Demirdache, Hamida & Orin Percus, 2011. Resumptives, movement and interpretation. In Alain Rouveret (ed.), *Resumptive pronouns at the interfaces*, pp. 367–393. Amsterdam: John Benjamins.

den Besten, Hans, 1983. On the interaction of root transformations and lexical deletive rules. In Werner Abrahams (ed.), *On the formal syntax of Westgermania*, pp. 47–131. Amsterdam: John Benjamins.

Déprez, Viviane, 1994. Haitian creole: A pro-drop language? *Journal of Pidgin and Creole Languages*, 9(1), pp. 1–24.

Dickey, Michael Walsh, 1996. Constraints on the sentence processor and the distribution of resumptive pronouns. *Linguistics in the Laboratory*, 19, pp. 157–192.

D'Odorico, Laura & Stefania Carubbi, 2003. Prosodic characteristics of early multiword utterances in Italian children. *First Language*, 23(1), pp. 97–116.

Doron, Edith, 1982. The syntax and semantics of resumptive pronouns. In *Texas linguistics forum*, volume 19.

Doron, Edith, 1986. The pronominal 'copula' as agreement clitic. In Hagit Borer (ed.), *The syntax of pronominal clitics*, pp. 313–332. Orlando: Academic Press.

Elbourne, Paul, 2005. On the acquisition of principle B. *Linguistic Inquiry*, 36(3), pp. 333–365.

Engdahl, Elisabet, 1985. Parasitic gaps, resumptive pronouns, and subject extraction. *Linguistics*, 23, pp. 3–44.

Erteschik-Shir, Nomi, 1992. Resumptive pronouns in islands. In Helen Goodluck & Michael Rochemont (eds.), *Island Constraints: Theory, Acquisition, and Processing*. Dordrecht: Kluwer.

Farby, Shira, Gabi Danon, Joel Walters, & Michal Ben-Shachar, 2010. The acceptability of resumptive pronouns in Hebrew. In *Proceedings of the Israeli Association for Theoretical Linguistics*.

Featherston, Sam, 2005. That-trace in German. *Lingua*, 115, pp. 1277–1302.

Felser, Claudia, Harald Clahsen, & Thomas F. Münte, 2003. Storage and integration in the processing of filler-gap dependencies: An erp of study of topicalization and *wh*-movement in german. *Brain and Language*, 87, pp. 348–368.

Fernández, E., 2003. Processing strategies in second language acquisition: Some preliminary results. In Elaine C. Klein & Gita Martohardjono (eds.), *The Development of Second Language Grammars: A Generative Approach*, pp. 217–239. Amsterdam: John Benjamins.

Fernández Ramírez, Salvador, 1986. *Gramática Española 4: El verbo y la oración*. Madrid: Arco/Libros.

Ferreira, Fernanda & Benjamin Swets, 2005. The production and comprehension of resumptive pronouns in relative clause "island" contexts. In Anne Cutler (ed.), *Twenty-first century psycholinguistics: Four cornerstones*, pp. 263–278. Mahway, NJ: Lawrence Erlbaum Associates.

Fiengo, Robert, 1977. On trace theory. *Linguistic Inquiry*, 8, pp. 35–61.

Fodor, Janet Dean, 1998a. Learning to parse? *Journal of Psycholinguistic Research*, 27(2), pp. 285–319.

Fodor, Janet Dean, 1998b. Parsing to learn? *Journal of Psycholinguistic Research*, 27(2), pp. 339–374.

Fodor, Janet Dean, 1998c. Unambiguous triggers. *Linguistic Inquiry*, 29(1), pp. 1–36.

Foraker, Stephanie & Brian McElree, 2007. The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, 56(3), pp. 357–383.

Frazier, Lyn, 1978. *On comprehending sentences: Syntactic parsing strategies*. Ph.D. thesis, University of Connecticut.

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

Gagliardi, Annie & Jeffrey Lidz, 2015. How nature meets nurture: Universal grammar and statistical learning. *Annual Review of Linguistics*, 1, pp. 12.1–12.21.

Garnsey, Susan M., Michael K. Tanenhaus, & Robert M. Chapman, 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research*, 18, pp. 51–60.

Gibson, Edward, Neal Perlmutter, Enriqueta Canseco-Gonzalez, & Gregory Hickok, 1996. Recency preference in the human sentence processing mechanism. *Cognition*, 59, pp. 23–59.

Gibson, Edward & Kenneth Wexler, 1992. Triggers. *Linguistic Inquiry*, 25(3), pp. 407–454.

Gibson, Edward T., 1998. Linguistic complexity: locality of syntactic dependencies. *Cognition*, 68, pp. 1–76.

Gilligan, Gary, 1987. *A Cross-Linguistic Approach to the Pro-Drop Parameter*. Ph.D. thesis, University of Southern California.

Gordon, Peter C., Randall Hendrick, & Marcus Johnson, 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), pp. 1411–1423.

Gordon, Peter C., Randall Hendrick, & Marcus Johnson, 2004. Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51, pp. 97–114.

Gordon, Peter C., Randall Hendrick, & William H. Levine, 2002. Memory-load interference in syntactic processing. *Psychological science*, 13, pp. 425–430.

Greenberg, Joseph, 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph Greenberg (ed.), *Universals of Language*, pp. 73–113. Cambridge, MA: MIT Press.

Grillo, Nino & João Costa, 2014. A novel argument for the universality of parsing principles. *Cognition*, 133(1), pp. 156–187.

Grodzinsky, Youssef & Tanya Reinhart, 1993. The innateness of binding and coreference. *Linguistic Inquiry*, 24(1), pp. 69–101.

Han, Chung-Hye, Jeffrey Lidz, & Julien Musolino, 2007. Verb-movement and grammar competition in Korean: Evidence from quantification and negation. *Linguistic Inquiry*, 38, pp. 1–47.

Hart, Betty & Todd Risley, 1995. *Meaningful differences in the everyday experience of young American children*. Baltimore: P.H. Brookes.

Haspelmath, Martin, 2008. Parametric versus functional explanations of syntactic universals. In Theresa Bieberauer (ed.), *The limits of syntactic variation*, pp. 75–107. Amsterdam: John Benjamins.

Hawkins, John, 1999. Processing complexity and filler-gap dependencies across languages. *Language*, 75, pp. 224–285.

Hawkins, John, 2004. *Efficiency and complexity in grammars*. Oxford University Press: Oxford.

Hayon, Yehiel, 1973. *Relativization in Hebrew*. The Hague: Walter de Gruyter.

Heestand, Dustin, Ming Xiang, & Maria Polinsky, 2011. Resumption still does not rescue islands. *Linguistic Inquiry*, 42(1), pp. 138–152.

Hofmeister, Philip & Elisabeth Norcliffe, 2013. Does resumption facilitate sentence comprehension? In Philip Hofmeister & Elisabeth Norcliffe (eds.), *The Core and the Periphery: Data-driven perspectives on syntax inspired by Ivan A. Sag*, pp. 225–246. Stanford, CA: CSLI Publications.

Hofmeister, Philip & Ivan A. Sag, 2010. Cognitive constraints on syntactic islands. *Language*, 86, pp. 366–415.

Hoji, Hajime, 1991. Kare. In C. Georgopoulous & R. Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, pp. 287–304. Dordrecht: Kluwer.

Hoji, Hajime, 1995. Demonstrative binding and principle b. In J. N. Beckman (ed.), *Proceedings of NELS 25*, pp. 255–271. Amherst, MA: GLSA Publications.

Huang, C.-T. James, 1982. *Logical relations in Chinese and the theory of grammar*. Ph.D. thesis, MIT.

Hyams, Nina, 1989. The null subject parameter in language acquisition. In Ken Safir & Osvaldo Jaeggli (eds.), *The null subject parameter*, pp. 215–238. Dordrecht: Kluwer.

Ishii, Yasuo, 1991. *Operators and Empty Categories in Japanese*. Ph.D. thesis, University of Connecticut, Storrs.

Jakobson, Roman, 1941. *Kindersprache, Aphasie und allgemeine Lautgesetze*. The Hague: Mouton.

Jurka, Johannes, 2013. Subject islands in German revisited. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental Syntax and Island Effects*, pp. 265–285. Cambridge, MA: Cambridge University Press.

Just, Marcel A., Patricia A. Carpenter, & Jacqueline D. Woolley, 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology*, 111(2), pp. 228–238.

Just, Marcel Adam & Patricia A. Carpenter, 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), pp. 122–149.

Kaan, Edith, Anthony Harris, Edward Gibson, & Philip Holcomb, 2000. The p600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, 15, pp. 159–201.

Kayne, Riachard S., 1981. Ecp extensions. *Linguistic Inquiry*, 12, pp. 93–133.

Kayne, Richard, 1976. French relative *que*. In Marta Luján & Fritz Hensey (eds.), *Current studies in Romance linguistics*. Washington, DC: Georgetown University Press.

Kayne, Richard, 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Kayne, Richard S., 2005. Some notes on comparative syntax, with special reference to English and French. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford Handbook of Comparative Syntax*, pp. 3–69. New York: Oxford University Press.

Kazanina, Nina, Ellen Lau, Moti Lieberman, Masaya Yoshida, & Colin Phillips, 2007. The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, 56, pp. 384–409.

Keffala, Bethany, 2013. Resumption and gaps in English relative clauses: Relative acceptability creates an illusion of 'saving'. In Chundra Cathcart, I-Hsuan Chen, Greg Finley, Shinae Kang, Clare s. Sandy, & Elise Stickles (eds.), *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society*, pp. 140–154.

Keller, Frank & Theodora Alexopoulou, 2005. A crosslinguistic, experimental study of resumptive pronouns and *that*-trace effects. In Brun G. Bara, Lawrence Barsalou, & Monica Bucciarelli (eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, pp. 1120–1125.

Keshev, Maayan & Aya Meltzer-Asscher, ms. Grammatical resumption, islands and active dependency formation: Evidence from Hebrew parsing. Tel Aviv Unversity.

Kluender, Robert, 1998. On the distinction between strong and weak islands: A processing perspective. In Peter Culicover & Louise McNally (eds.), *Syntax and Semantics 29: The limits of syntax*, pp. 241–279. San Diego, CA: Academic Press.

Kluender, Robert, 2005. Are subject islands subject to a processing account? In Vineeta Chand, Ann Kelleher, Angelo J. Rodríguez, & Benjamin Schmeiser (eds.), *Proceedings of the 23rd West Coast Conference on Formal Linguistics*, pp. 475–499.

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

Koopman, Hilda & Dominique Sportiche, 1991. The position of subjects. *Lingua*, 85(2/3), pp. 211–258.

Kroch, Anthony, 1981. On the role of resumptive pronouns in amnestying island constraint violations. In *Proceedings of the annual meetings of the Chicago Linguistics Society*, pp. 125–135.

Kunert, Richard, Raquel Fernández, & Willem Zuidema, 2011. Adaptation in Child Directed Speech: Evidence from corpora. In *SemDial 2011: Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue*, pp. 112–119.

Kuznetsova, Alexander, Per Bruun Brockhoff, & Rune Haubo Bojesen Christensen, 2014. lmertest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). `http://CRAN.R-project.org/package=lmerTest`.

Lasnik, Howard & Mamoru Saito, 1992. *Move α: Conditions on its application and put*. Cambridge, MA: MIT Press.

Lewis, M. Paul, Gary F. Simons, & Charles D. Fennig, 2015. *Ethnologue: Languages of the World, Eighteenth Edition*. Dallas, Texas: SIL International. Online verison: http://www.ethnologue.com.

Lukyanenko, Cynthia, Anastasia Conroy, & Jeffrey Lidz, 2014. Is she patting Katie?: Constraints on pronominal reference in 30-month olds. *Language Learning and Development*, 10(4), pp. 328–344.

MacWhinney, Brian, 2000. *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Lawrence Erlbaum Associates, third edition edition.

Manetta, Emily, 2012. *Peripheries in Kashmiri and Hindi/Urdu*. Amsterdam: John Benjamins.

Maxwell, Scott E. & Harold D. Delaney, 2003. *Designing Experiments and Analyzing Data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

McCloskey, James, 2006. Resumption. In Martin Everaert, Henk van Riemsdijk, Rob Goedemans, & Bart Hollebrandse (eds.), *The Blackwell comparion to syntax*, pp. 94–117. Oxford: Blackwell.

McDaniel, Dana & Wayne Cowart, 1999. Experimental evidence for a minimalist account of English resumptive pronouns. *Cognition*, 70, pp. B15–B24.

Menuzzi, Sergio, 2000. *That*-trace effects in Portuguese. *Fórum Lingüistico*, 2.

Mitchell, Don C. & Fernando Cuetos, 1991. The origin of parsing strategies. In Carlota Smith (ed.), *Current issues in Natural Language Processing*, pp. 1–12. Austin, TX: Center for Cognitive Science.

Miyamoto, Edson T., 2002. Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research*, 31(4), pp. 307–347.

Nakao, Chizuru, 2009. *Island Repair and Non-Repair by PF Strategies*. Ph.D. thesis, University of Maryland.

Newmeyer, Frederick, 2004. Against a parameter-setting approach to typological variation. *Linguistic Variation Yearbook*, 4, pp. 181–234.

Newmeyer, Frederick, 2005. *Possible and probable languages: A Generativist perspective on linguistic typology*. Oxford: Oxford University Press.

Nicol, Janet L., Janet Dean Fodor, & David Swinney, 1994. Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, pp. 1229–1238.

Nicol, Janet L. & David Swinney, 1989. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18, pp. 5–20.

Nicolis, Marco, 2008. The null subject parameter and correlating properties: the case of creole languages. In Theresa Bieberauer (ed.), *The limits of syntactic variation*, pp. 271–294. Amsterdam: John Benjamins.

Omaki, Akira & Jeffrey Lidz, 2014. Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, 22(2), pp. 158–192.

Omaki, Akira & Chizuru Nakao, 2010. Does English resumption really help to repair island violations? *Snippets*, 21, pp. 11–12.

Omaki, Akira & Chizuru Nakao, ms. Conditions on resumption strategies and relativized locality violations. `http://ling.umd.edu/~omaki/Research/OmakiNakao-Squib.pdf`.

Omaki, Akira, Imogen Davidson White, Takuya Goro, Jeffrey Lidz, & Colin Phillips, 2014. No fear of commitment: children's incremental interpretation of English and Japanese *Wh*-questions. *Language Learning and Development*, 10(3), pp. 206–233.

Ono, Hajime, Masaya Yoshida, Sachiko Aoshima, & Colin Phillips, 2006. Real-time computation of japanese exclamatives and the strength of locality biases in sentence comprehension. *Cognitive Studies*, 13, pp. 261–287.

Pearl, Lisa & Jon Sprouse, 2013. Syntactic islands and learning biases: Combining experimental syntax and computational modeling to investigate the language acquisition problem. *Language Acquisition*, 20(1), pp. 23–68.

Perfors, Amy, Joshua B. Tenenbaum, & Terry Regier, 2011. The learnability of abstract syntactic principles. *Cognition*, 118, pp. 306–338.

Perlmutter, David, 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Reinhart, and Winston.

Pesetsky, David, 1982. The comp-trace effect and the nominative island condition. *The Linguistic Review*, 1, pp. 297–345.

Pesetsky, David, ms. Complementizer-trace effects. `http://ling.auf.net/lingbuzz/002385`.

Pesetsky, David & Esther Torrego, 2001. T-to-c movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: a Life in Language*, pp. 355–426. Cambridge, MA: The MIT Press.

Phillips, Colin, 2006. The real-time status of island phenomena. *Language*, 82(4), pp. 795–823.

Phillips, Colin, 2013. On the nature of island constraints ii: Language learning and innateness. In Jon Sprouse & Norbert Hornstein (eds.), *Experimental Syntax and Island Effects*, pp. 132–157. Cambridge, MA: Cambridge University Press.

Phillips, Colin, Nina Kazanina, & Shani H. Abada, 2005. ERP effects on the processing of syntactic long-distance dependencies. *Cognitive Brain Research*, 22, pp. 407–428.

Phillips, Colin & Dan Parker, 2014. The psycholinguistics of ellipsis. *Lingua*, 151, pp. 78–95.

Phillips, Colin & Matthew Wagers, 2007. Relating structure and time in linguistics and psycholinguistic. In Gareth Gaskell (ed.), *Oxford Handbook of Psycholinguistics*, pp. 739–756. Oxford: Oxford University Press.

Prince, Ellen F., 1990. Syntax and discourse: A look at resumptive pronouns. In *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society*, pp. 482–497.

Pritchett, Bradley L., 1991. Subjaceny in a principle-based parser. In Robert C. Berwick, Stephen P. Abney, & Carol Tenny (eds.), *Principle-based Parsing: Computation and psycholinguistics*, pp. 301–345. Dordrecht: Kluwer.

Rayner, Keith, 1998. Eye movements in reading and information processing: 20 years of reesarch. *Psychological Bulletin*, 124, pp. 372–422.

Rayner, Keith & Susan A. Duffy, 1986. Lexical complexity and fixation times in reading: effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition*, 14, pp. 191–201.

Rayner, Keith & Alexander Pollatsek, 1989. *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Rayner, Keith, Timothy J. Slattery, Denis Drieghe, & Simon P. Liversedge, 2011. Eye movements and word skipping during reading: Effects of word length and predictability. *Journal of Experimental Psyochology: Human Perception and Performance*, 37(2), pp. 514–528.

Reinhart, Tanya, 1983. *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, Tanya, 1986. Center and periphery in the grammar of anaphora. In Barbara Lust (ed.), *Studies in the acquisition of anaphora*. Dordrecht: Reidel.

Rizzi, Luigi, 1982. *Issues in Italian syntax*. Dordrecht: Foris.

Rizzi, Luigi, 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry*, 17, pp. 501–557.

Rizzi, Luigi, 1990. *Relativizied Minimality*. Cambridge, MA: The MIT Press.

Rizzi, Luigi, 2006. On the form of chains: Criterial positions and ECP effects. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *Wh-Movement: Moving On*, pp. 97–133. Cambridge, MA: The MIT Press.

Rizzi, Luigi, 2013. Locality. *Lingua*, 130, pp. 169–186.

Rizzi, Luigi & Ur Shlonsky, 2007. Strategies of subject extraction. In *Interface + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, pp. 115–160. Berlin: Mouton de Gruyter.

Roberts, Ian & Anders Holmberg, 2005. On the role of parameters in Universal Grammar: a reply to Newmeyer. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, & Jan Koster (eds.), *Organizing Grammar: Linguistic Studies in Honor of Henk van Riemsdijk*, pp. 538–553. Berlin: Mouton de Gruyter.

Ross, John Robert, 1967. *Constraints on variables in syntax*. Ph.D. thesis, MIT.

Safir, Kenneth, 1985. *Syntactic chains*. Cambridge, MA: Cambridge University Press.

Safir, Kenneth & Osvaldo Jaeggli, 1989. *The null subject parameter*. Dordrecht: Kluwer.

Saito, Mamoru, 1985. *Some asymmetries in Japanese and their theoretical implications*. Ph.D. thesis, MIT.

Sakas, William & Janet Dean Fodor, 2012. Disambiguating syntactic triggers. *Language Acquisition*, 19, pp. 83–143.

Sells, Peter, 1984. *Syntax and semantics of resumptive pronouns*. Ph.D. thesis, University of Massachusetts.

Shlonsky, Ur, 1992. Resumptive pronouns as a last resort. *Linguistic Inquiry*, 23, pp. 443–468.

Simpson, Andrew & Tanmoy Bhattacharya, 2000. *Wh*-cp pied piping in Bangla. In Masako Hirotani, Andries Coetzee, Nancy Hall, & Ji yung Kim (eds.), *Proceedings of the 30th Annual Meeting of the North-East Linguistic Society*, pp. 583–596. Amherst, MA: GLSA.

Simpson, Andrew & Tanmoy Bhattacharya, 2003. Obligatory overt *wh*-movement in a *wh-in-situ* language. *Linguistic Inquiry*, 34(1), pp. 127–142.

Simpson, Andrew, Arunima Choudhury, & Mythili Menon, 2013. The derivation and licensing of covert nominals in Bangla, Hindi, and Malayalam. *Lingua*, 134, pp. 103–128.

Simpson, Andrew & Hao Tam Ho, 2008. The comparative syntax of passive structures in Chinese and Vietnamese. In Marjorie K.M. Chan & Hana Kang (eds.), *Proceedings of the 20th North American Conference on Chinese Linguistics*, pp. 825–841. Columbus, OH: The Ohio State University.

Snyder, William, 2001. On the nature of syntactic variation: Evidence from complex predicates and complex word-foramtion. *Language*, 77, pp. 324–342.

Sobin, Nicholas, 1987. The variable status of comp-trace phenomena. *Natural Language and Linguistic Theory*, 5(1), pp. 33–60.

Sobin, Nicholas, 2002. The comp-trace effect, the adverb effect, and the minimal CP. *Journal of Linguistics*, 38, pp. 527–560.

Sobin, Nicholas, 2009. Prestige case forms and the comp-trace effect. *Syntax*, 12(1), pp. 32–59.

Sprouse, Jon, 2008. The differential sensitivity of acceptability judgments to processing effects. *Linguistic Inquiry*, 39(4), pp. 686–694.

Sprouse, Jon, Ivano Caponigro, Ciro Greco, & Carlo Cecchetto, 2013. Experimental syntax and the cross-linguistic variation of island effects in English and Italian. To appear in Natural Language and Linguistic Theory.

Sprouse, Jon, Matthew W. Wagers, & Colin Phillips, 2012a. A test of the relation between working memory capacity and syntactic island effects. *Language*, 88, pp. 82–123.

Sprouse, Jon, Matthew W. Wagers, & Colin Phillips, 2012b. Working-memory capacity and island effects: A reminder of the issues and facts. *Language*, 88, pp. 401–407.

Srikumar, K., 2007. Clausal pied-piping and subjacency. In Josef Bayer, Tanmoy Bhattacharya, & M.T. Hany Babu (eds.), *Linguistic Theory and South Asian Languages. Essays in Honor of KA Jayaseelan*, pp. 53–69. Amsterdam: John Benjamins.

Stowe, Laurie, 1986. Parsing wh-constructions: evidence from on-line gap location. *Language and Cognitive Processes*, 1, pp. 227–245.

Sturt, Patrick, 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, pp. 542–562.

Sussman, Rachel S. & Julie C. Sedivy, 2003. The time-course of processing syntactic dependencies: Evidence from eye-movements during spoken *wh*-questions. *Language and Cognitive Processes*, 4, pp. 211–234.

Sutton, Megan, 2015. *Competence and Performance in the Development of Principle C.* Ph.D. thesis, University of Maryland.

Svolacchia, Marco & Annarita Puglielli, 1999. Somali as a polysynthetic language. In Lunella Mereu (ed.), *Boundaries of Morphology and Syntax*, pp. 97–120. Amsterdam: John Benjamins.

Tanaka, Hidekazu, 2001. Right-dislocation as scrambling. *Journal of Linguistics*, 37, pp. 551–579.

Taraldsen, Tarald K., 1980. On the nominative island condition, vacuous application and the that-trace filter. Bloomington, IN: Indiana University Linguistics Club.

Thompson, Hanne-Ruth, 2012. *Bengali*. Amsterdam: John Benjamins Company.

Tonelli, Livia, Wolfgang U. Dressler, & Renata Romano, 1995. Frühstufen des erwerbs der italienischen konjugation. *Suvremena Lingvistika*, 21(1), pp. 3–15.

Torrego, Esther, 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry*, 15(1), pp. 103–129.

Travis, Lisa, 1984. *Parameters and effects of word order variation*. Ph.D. thesis, MIT.

Traxler, Matthew J. & Martin J. Pickering, 1996. Plausibility and unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, 35(5), pp. 454–475.

Uriagereka, Juan, 1992. The syntax of movement in Basque. In Jon Ortiz de Urbina & Joseba Andoni Lakarra (eds.), *Syntactic theory and Basque Syntax*, pp. 417–445. Bilbao: Euskal Herriko Unibersitatea.

Uriagereka, Juan, 1999. Multiple spellout. In Samuel David Epstein & Norbert Hornstein (eds.), *Working Minimalism*, pp. 251–282. MIT Press.

Uriagereka, Juan, 2007. Clarifying the notion 'parameter'. *Biolinguistics*, 1, pp. 99–113.

Wagers, Matthew W., 2008. *The structure of memory meets memory for structure in linguistic cognition*. Ph.D. thesis, University of Maryland.

Wagers, Matthew W. & Colin Phillips, 2009. Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, 45, pp. 395–433.

Wagers, Matthew W. & Colin Phillips, 2014. Going the distance: memory and control processes in active dependency formation. *The Quarterly Journal of Experimental Psychology*, 167(7), pp. 1274–1304.

Wanner, Eric & Michael P. Maratsos, 1978. An atn approach to comprehension. In Morris Halle, Joan Bresnan, & George Miller (eds.), *Linguistic theory and psychological reality*, pp. 119–161. Cambridge, MA: MIT Press.

Wexler, Kenneth & Maria Rita Manzini, 1987. Parameters and learnability in binding theory. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*. Dordrecht: Reidel.

White, Aaron Steven, 2014. Contextually modulated syntactic variability in child-directed speech. Poster presented at the 13th International Congress for the Study of Child Language.

Yang, Charles, 2002. *Knowlege and Learning in Natural Language*. Oxford: Oxford University Press.

Yang, Charles, 2004. Universal grammar, statistics, or both? *TRENDS in Cognitive Science*, 8(10), pp. 452–456.

Yoshida, Masaya, 2006. *Constraints and mechanisms in long-distance dependency formation*. Ph.D. thesis, University of Maryland.

Yoshida, Masaya, Nina Kazanina, Leticia Pablos, & Patrick Sturt, 2014. On the origin of islands. *Language, Cognition, and Neuroscience*, 29(7), pp. 761–770.

Zaenen, Annie, Elisabet Engdahl, & Joan M. Maling, 1981. Resumptive pronouns can be syntactically bound. *Linguistic Inquiry*, 4(12), pp. 679–682.

Zukowski, Andrea & Jaiva Larsen, 2004. The production of sentences that we fill their gaps. Poster presented at the Annual CUNY Sentence Processing Conference.