ABSTRACT

Title of dissertation: RESPECTING RELATIONS: MEMORY ACCESS AND ANTECEDENT RETRIEVAL IN INCREMENTAL SENTENCE PROCESSING

Dave W. Kush, Doctor of Philosophy, 2013

Dissertation directed by: Professor Colin Phillips
Department of Linguistics

This dissertation uses the processing of anaphoric relations to probe how linguistic information is encoded in and retrieved from memory during real-time sentence comprehension. More specifically, the dissertation attempts to resolve a tension between the demands of a linguistic processor implemented in a general-purpose cognitive architecture and the demands of abstract grammatical constraints that govern language use. The source of the tension is the role that abstract configurational relations (such as *c-command*, Reinhart 1983) play in constraining computations.

Anaphoric dependencies are governed by formal grammatical constraints stated in terms of relations. For example, Binding Principle A (Chomsky 1981) requires that antecedents for local anaphors (like the English reciprocal *each other*) bear the c-command relation to those anaphors. In incremental sentence processing, antecedents of anaphors must be retrieved from memory. Recent research has motivated a model of processing that exploits a cue-based, associative retrieval process in content-addressable memory (e.g. Lewis, Vasishth & Van Dyke 2006) in which relations such as c-command are difficult

to use as cues for retrieval. As such, the c-command constraints of formal grammars are predicted to be poorly implemented by the retrieval mechanism.

I examine retrieval's sensitivity to three constraints on anaphoric dependencies: Principle A (via Hindi local reciprocal licensing), the Scope Constraint on bound-variable pronoun licensing (often stated as a c-command constraint, though see Barker 2012), and Crossover constraints on pronominal binding (Postal 1971, Wasow 1972). The data suggest that retrieval exhibits fidelity to the constraints: structurally inaccessible NPs that match an anaphoric element in morphological features do not interfere with the retrieval of an antecedent in most cases considered. In spite of this alignment, I argue that retrieval's apparent sensitivity to c-command constraints need not motivate a memory access procedure that makes direct reference to c-command relations. Instead, proxy features and general parsing operations conspire to mimic the extension of a system that respects c-command constraints. These strategies provide a robust approximation of grammatical performance while remaining within the confines of a independently-motivated general-purpose cognitive architecture.

RESPECTING RELATIONS:
MEMORY ACCESS AND ANTECEDENT RETRIEVAL IN INCREMENTAL SENTENCE PROCESSING

by

Dave W. Kush

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
2013

Advisory Committee:
Professor Colin Phillips, Chair/Advisor
Professor Norbert Hornstein
Professor Michael Israel
Professor Ellen Lau
Professor Jeffrey Lidz, Co-Advisor

Acknowledgments

Thank you to everyone who has known me well and supported me anyway.
In particular:
Colin Phillips, Jeff Lidz, and Norbert Hornstein, the three of whom have all shouldered the responsibility of advising me. Colin has played an instrumental role in getting me in shape intellectually and physically. The standards he sets are exacting, but trying to meet them has forced me to grow. Jeff has provided encouragement whenever I needed it (no mean feat). He has had the patience to help me coax latent insights from jumbled messes of thought, as well as the decency to tell me when a half-baked idea would go nowhere. His impeccable binding judgments have also come in handy more than once. Norbert rounds out the troika. Norbert has taught me the value of 'productive BS' and of pursuing interesting ideas even when they're sure to be wrong. In addition to being a de facto advisor, Norbert's been a father figure of sorts, dispensing life/book advice, providing food, and even letting us have the car when he was out of town.

Thanks also goes to the other members of my committee:
To Ellen Lau, for listening to me ramble and providing a great example of an approachable, positive, and serious scientist. To Michael Israel, for agreeing to read this dissertation and for employing me sporadically in my first few years of grad school.

Other members of the department deserve acknowledgment, too:
Tonia Bleam has always let me drop in on her unannounced to ~~gripe~~ talk. Valentine Hacquard has not held the torturous process of reading my many drafts against me. Alexander Williams has always been gracious enough to treat me like an equal. Howard Lasnik, Paul Pietroski, and Bill Idsardi have all played their part in making me a clearer thinker. Thanks to Andrea Zukowski for support inside the department and along Sideling Hill.

Thanks to Dave Baggett for funding my first year at Maryland. The department would be a very different place without the Baggett fellowship.[1]

My peers at Maryland were a frighteningly intelligent, capable, and assiduous bunch that got me to step my game up. In addition to being superstars, they're all pretty decent people, too. The members of my unofficial first class, Pedro Alcocer, Brian Dillon, Alex Drummond, Annie Gagliardi, and Shannon Hoerner, all have had an immense impact on my life. Pedro has been a valued friend. I envy his ability to confidently pursue his interests, as well as his PRs on fitocracy. Brian Dillon paved the way for a lot of the research contained here, and, more importantly, was the glue that held us together for so long. Alex Drummond has made me laugh harder than almost anyone I've ever known (often intentionally!) and has been generous enough to make me a coauthor on papers. Shannon, the last of the (not that) old guard, has my gratitude for sticking it out with me

[1]The research reported in this dissertation was supported by an NSF IGERT Grant #DGE-0801465 to the University of Maryland, and an NSF Grant #BCS-0848554 to Colin Phillips.

every step of the full six years (and not running screaming after that completely ill-advised first experiment). Rest assured, the padrino will make it out to Utah.
Wing Yee Chow, Ewan Dunbar, Michaël Gagnon, Brad Larson, Chris LaTerza, Shevaun Lewis, Terje Lohndal — my official classmates — are no slouches either. Brad, who I am not ashamed to say I begged to come to Maryland, has played the easy-going West-coast counterpoint to my East-coast neuroticism for 5 years. Ewan deserves credit for getting me to think bigger, while still sweating the details (also for reawakening my unabashed love for DS9). Shevaun ensured I didn't act indecorously in Europe (in contrast to Ewan and Brad). She has also never failed in providing everything from gourmet eats, vigorous counter-argument, and patience. Terje provided good discussion, insider information, and innumerable Norwegian judgments. Wing Yee deserves thanks for the good cheer, playing the good cop to Shevaun's bad cop in arguments, and for all the Assam tea she (un)knowingly provided me.

More thanks:
To my fellow flag hag, Nora Oppenheim, for being a great co-blogger and friend. To Elika Bergelson, for helping me develop a work ethic and consistently producing entertaining speech errors. To Dan Parker, for rapping about ACT-R and being a frequent conference bunkmate. To Sol Lago, for finally coming around. To Dustin Alfonso Chacon, Kate Harrigan, Alexis Wellwood, and Aaron Steven White for ensuring I never wanted for an interesting conversation. To other, elder, linguists who have contributed an idea, a drink, or their friendship over the last six years: Rajesh Bhatt, Tim Hunter, Tim Hawes, John Hale, Akira Omaki, Bridget Samuels, Matthias Scharinger, Jon Sprouse, Masaya Yoshida, and Matt Wagers. To Sam Epstein and Elaine McNulty for getting me into syntax. To Karthick, Pritha Chandra, Ayesha Kidwai, Usha Udaar, and members of the student bodies at IIT, Delhi and JNU for help with the Hindi experiment.

The following people deserve special thanks for not being linguists:
AJ Annunziata, Edison Barrios (who narrowly qualifies as a non-linguist), Rob Bedoya, Sebastian 'Stump Beefknob' Benthall, Courtney Blanchard, Kate Boulton, Romina Casadei, Katy Gorman, Jorge Guzman, Leila Mays, Alex Mazarredo, Chris Mesaros, Maggie Mesaros, Ric Meyer, Eric Naing, Julian Vu (my House-mate), and Donna Ward. Lily Mendelson deserves special thanks for her kindness and compassion, as well as her efforts to make me a better person.

Finally, thanks to my family, who only grow more important to me with time.
To Aunt Roz, Uncle Rick and Dan (Dan), Heather, Manu, Tanu, Ambrish Chacha. To Aunt Irene, for keeping me up to date on the *New Yorker* and sending the final care-package of tea that kept me and my office-mates blissfully buzzed while dissertating. To Emily Rosenberg, for being my sister. To Catherine Roosevelt, for growing up with me.
To Molly, who brings out my best, endures my worst, and whose ability to live with purpose still amazes me. To my father, who came to terms with my being the wrong kind of doctor and finally got the hang of being an informant. And to my mother, whose unconditional support I have never once needed to be reminded of (though that hasn't stopped her from doing so).

Table of Contents

List of Tables ix

List of Figures x

1 Introduction 1
- 1.1 Starting Out . . . 1
  - 1.1.1 The Challenge . . . 7
- 1.2 Outline of the Thesis . . . 12
  - 1.2.1 Chapter 2: Cue-Based Retrieval and Relations . . . 12
  - 1.2.2 Chapter 3: Local Anaphor Licensing . . . 13
  - 1.2.3 Chapter 4: The LOCAL Feature . . . 14
  - 1.2.4 Chapter 5: Crossover and Principle C . . . 14
  - 1.2.5 Chapter 6: Scope Constraints . . . 16
  - 1.2.6 Chapter 7: Bound-Variable Anaphora . . . 16
  - 1.2.7 Conclusion . . . 18

2 Motivating a Model of Memory Encoding and Retrieval 19
- 2.1 Representation and Encoding . . . 19
- 2.2 Procedures for Recalling Information . . . 20
  - 2.2.1 Tree Traversal/Serial Search v. Direct Access . . . 21
  - 2.2.2 Evidence . . . 24
  - 2.2.3 The Consequence of Trading Structure for Speed: Interference . . . 32
- 2.3 (Retroactive) Interference Effects in Memory Retrieval . . . 32
  - 2.3.1 Computing Feature Match for Retrieval . . . 33
  - 2.3.2 Inhibitory Interference . . . 35
  - 2.3.3 Facilitatory Interference: Illusions of Grammaticality . . . 39
- 2.4 Remembering Relations . . . 45
  - 2.4.1 Exhaustive Encoding is Too Exhausting . . . 47
- 2.5 Summary . . . 50

3 Local Anaphoric Licensing in Hindi 52
3.1 Introduction 52
3.2 Previous Studies on Processing Reflexives 57
3.2.1 Interference-Immunity 57
3.2.1.1 Nicol and Swinney 1989 57
3.2.1.2 Clifton, Frazier & Deevy (1999) 59
3.2.1.3 Badecker & Straub (2002) 60
3.2.1.4 Sturt (2003) 63
3.2.1.5 Xiang et al. (2009) 65
3.2.1.6 Dillon et al. (2013) Local Reflexives/Agreement 66
3.2.2 Interference Susceptibility: King et al. (2012) 68
3.2.3 Summary of Results 70
3.3 A Proxy Feature and Strategies for Its Use 71
3.3.1 Clause Index as Proxy Feature for Local Anaphor Accessibility . 71
3.3.2 Implementing Immunity to Interference: Cue Selection v. Weighting 73
3.3.3 Linear Adjacency: Context-Dependent Cue Selection/Weighting . 77
3.4 SOV Local Anaphors: A Primer on Hindi Reciprocals 80
3.5 Experiments 83
3.5.1 Materials 83
3.5.2 Experiment 1: Modeling Hindi Reciprocal Licensing 90
3.5.2.1 Background: The Model 91
3.5.2.2 Results 96
3.5.2.3 Discussion 98
3.5.3 Experiment 2: Hindi Reciprocal Licensing 99
3.5.3.1 Participants 99
3.5.3.2 Procedure 99
3.5.3.3 Analysis 100
3.5.3.4 Results 101
3.5.4 Discussion 106
3.6 General Discussion 106
3.6.1 Which Cues to Use? Assessing Cue Diagnosticity 108
3.6.1.1 'Counting' Clausemates in Hindi - Implicitly 113
3.7 Conclusion 117

4 Clause-mate Conditions, Binding Principles, and the LOCAL Feature 119
4.1 Introduction 119
4.2 A LOCAL Account of Principle A 120
4.2.1 Local Update 122
4.3 Principle B 124
4.3.1 Previous Work on Principle B In Retrieval 125
4.3.2 Turning a Negative (Constraint) Into a Positive (Feature) 126
4.3.3 Updating LOCAL in Constant Time 132
4.3.4 A Gap in LOCAL Coverage 133
4.4 Summary 137

5 Crossover (and Principle C) 139
5.1 Introduction 139
5.1.1 Crossover Constructions 141
5.1.2 Principle C in Parsing 145
5.2 Experiment 3: Strong Crossover I 150
5.2.1 Participants 150
5.2.2 Procedure 150
5.2.3 Materials 151
5.2.4 Analysis 152
5.2.5 Results 153
5.2.6 Discussion 155
5.3 Experiment 4: Strong Crossover II 158
5.3.1 Materials 160
5.3.2 Experiment 4a: Acceptability Judgment Study 163
5.3.2.1 Participants 163
5.3.2.2 Materials 164
5.3.2.3 Procedure 164
5.3.2.4 Results 164
5.3.3 Experiment 4b: Self-Paced Reading Task 166
5.3.3.1 Participants 166
5.3.3.2 Procedure 166
5.3.3.3 Analysis 166
5.3.3.4 Results 166
5.3.3.5 Comprehension Question Accuracy 166
5.3.3.6 Self-paced Reading Results 167
5.3.4 Discussion 170
5.3.5 Does Antecedent Retrieval Simply Ignore Unintegrated Fillers? . 173
5.4 Experiment 5: Weak Crossover 175
5.4.1 Materials 176
5.4.2 Experiment 5a: Acceptability Judgment Study 179
5.4.2.1 Participants 179
5.4.2.2 Results 179
5.4.3 Experiment 5b: Self-paced Reading 180
5.4.3.1 Participants 180
5.4.3.2 Procedure 181
5.4.3.3 Analysis 181
5.4.3.4 Results 181
5.4.4 Discussion 184
5.4.4.1 Weak Crossover Violation v. Resumptive Pronouns . . 185
5.5 General Discussion 187
5.6 Accounting For Crossover Sensitivity Using the Anti-Clausemate Constraint 189
5.7 Conclusion 193

6 Variable Binding, C-command, and Scope 195
6.1 Introduction . . . 195
6.2 Barker (2012) . . . 198
6.2.1 Possessives . . . 200
6.2.2 Inverse Linking . . . 201
6.2.3 Binding out of DP . . . 202
6.2.4 Binding out of PP . . . 204
6.2.5 Binding out of VP . . . 205
6.2.6 Binding out of Finite Clauses? . . . 206
6.2.7 A Previous Attempt: Accessibility . . . 208
6.3 Ways of Scope-Taking . . . 210
6.3.1 Scope As Order of Composition . . . 212
6.3.2 Scope as C-command at LF . . . 214
6.4 Structural Limits on Scope and Scope Displacement . . . 217
6.5 The Take-home . . . 220

7 Bound Variable Pronouns 221
7.1 Introduction . . . 221
7.2 Previous Work on Processing Bound Pronouns . . . 223
7.2.1 Carminati et al. 2002 . . . 227
7.3 Experiment 6: Bound Variable Pronoun Licensing I: Constraint Sensitivity 233
7.3.1 Materials . . . 234
7.3.2 Experiment 6a . . . 237
7.3.2.1 Participants . . . 237
7.3.2.2 Procedure . . . 238
7.3.2.3 Materials . . . 238
7.3.2.4 Analysis . . . 239
7.3.2.5 Results . . . 240
7.3.2.6 Discussion . . . 240
7.3.3 Experiment 6b: Sentence Judgments . . . 241
7.3.3.1 Participants . . . 242
7.3.3.2 Materials . . . 242
7.3.3.3 Procedure . . . 243
7.3.3.4 Analysis . . . 244
7.3.3.5 Results . . . 245
7.3.3.6 Discussion . . . 246
7.3.4 Experiment 6c. Eye-tracking while reading . . . 247
7.3.4.1 Materials and Procedure . . . 248
7.3.4.2 Data Analysis . . . 249
7.3.4.3 Results . . . 251
7.3.4.4 Discussion . . . 253
7.4 Experiment 7: Bound Variable Pronoun Licensing II: Interference . . . 258
7.4.1 Materials . . . 258
7.4.1.1 Test Conditions . . . 258
7.4.1.2 Control Conditions . . . 261

7.4.2 Experiment 7a: Acceptability Judgment Study 263
7.4.2.1 Participants 263
7.4.2.2 Procedure 263
7.4.2.3 Materials 265
7.4.2.4 Results 265
7.4.3 Discussion 266
7.4.4 Experiment 7b: Eye-tracking while Reading 267
7.4.4.1 Participants 267
7.4.4.2 Procedure 268
7.4.4.3 Analysis 268
7.4.4.4 Results 270
7.4.5 Discussion 276
7.5 General Discussion 278
7.6 Accommodating Sensitivity to the Scope Constraint With Cue-Based Retrieval 279
7.6.1 A Spine-based Retrieval Strategy Is Not Viable 280
7.6.2 ACTIVE-ly Restricting Retrieval 285
7.6.2.1 Squaring ACTIVE with LOCAL 289
7.6.3 Possible Interference as a Result of Failure To Recognize Domain Edges 296
7.7 Conclusion 297

8 Conclusion 298
8.1 Summary 298
8.2 Where's C-Command? 299
8.2.1 Is C-command Still Needed? 302
8.3 Cue-combination (and its Discontents) 303
8.3.1 Reconciling the Power of the Account With Previous Work On Agreement 304
8.4 Remaining Issues 306
8.4.1 Methods of Retrieval 306
8.4.2 On the Origin of Features By Means of Grammatical Specification 308

List of Tables

2.1 Number of Retrievals Required to Exhaustively Encode C-command Relations in a Right-Branching Tree 49

3.1 Summary of Work on Local Anaphoric Licensing 70
3.2 Experiment 1: Predicted Interference for Equal Weight Retrieval 97
3.3 Experiment 1: Model Predicted Error Rates for Diagnostic Cue Retrieval. 97
3.4 Experiment 1: Model Predicted Interference for Diagnostic Cue Retrieval 98
3.5 Experiment 2: Comprehension Question Accuracy 101

5.1 Experiment 3: Comprehension Question Accuracy 153
5.2 Experiment 4: Mean Acceptability Ratings 165
5.3 Experiment 4: Comprehension Question Accuracy 167
5.4 Experiment 5: Mean Acceptability Ratings 179
5.5 Experiment 5: Comprehension Question Accuracy 181

7.1 Experiment 6a: Acceptability Ratings (Test Items) 239
7.2 Experiment 6a: Acceptability Ratings (Binding Control Conditions) . . . 239
7.3 Experiment 6b: Proportion of Quantificational-Reading Responses on Test Conditions. 245
7.4 Experiment 6c: Regions for Items 251
7.5 Experiment 6c: Reading Time Measures 252
7.6 Experiment 7a: Acceptability Ratings for Items 265
7.7 Experiment 7b: Reading Time Measures (Test Conditions) 271
7.8 Experiment 7b: Reading Time Measures (Control Conditions) 275

8.1 Summary: Empirical Results 298
8.2 Summary: Gating Cues and Update Points for Various Dependencies . . . 299

List of Figures

2.1 A *chunk* representation of simple syntactic structure. Figure from Lewis and Vasishth (2005: 380). 20

3.1 Experiment 1: Model Predicted Error Rates for Equal Weight Retrieval . 95
3.2 Experiment 1: Activation Profile for *Ungrammatical-Intervener* Condition 95
3.3 Experiment 1: Activation Profile for *Grammatical-NoIntervener* Condition 95
3.4 Experiment 2: Self-Paced Reading Results 103
3.5 Experiment 2: Un-pooled Post-Reciprocal Region 104
3.6 Experiment 2: Pooled Post-Reciprocal Region 105
3.7 Hindi Reciprocals: Comparison of Modeled and Observed Interference Effects 107

5.1 Experiment 3: Self-Paced Reading Results 154
5.2 Experiment 4: Self-Paced Reading Results (Control Conditions) 168
5.3 Experiment 4: Self-Paced Reading Results (Test Conditions) 169
5.4 Experiment 5: Self-paced Reading Results 182
5.5 Parsing Strong Crossover Sentences With Anti-clausemate Condition . . . 191
5.6 Parsing Weak Crossover Sentences With Anti-clausemate Condition . . . 192

7.1 Experiment 6c: Right-bound Reading Times 254
7.2 Experiment 6c: Total Reading Times 255
7.3 Experiment 7b: First-pass Reading Times 273
7.4 Experiment 7b: Right-bound Reading Times 274
7.5 Experiment 7b: Reread Reading Times 274

Chapter 1: Introduction

1.1 Starting Out

This thesis uses the processing of anaphoric relations as a tool for understanding the way in which linguistic information is encoded and accessed during real-time sentence comprehension. Anaphors are grammatical elements whose interpretation is dependent on a previous item.[1] For example, the reciprocal *each other* in (1-a) is dependent on the noun phrase (NP) *the cyclists*. The pronoun *them* is similarly dependent on *the cyclists* in (1-b).

(1) a. The cyclists rode alongside *each other*.

b. The cyclists hoped no one would push *them* over.

Resolving anaphoric interpretation appears, on the surface, to be an easy task. Natural language is littered with anaphors and users rarely display difficulty in establishing their interpretation. The facility that language users display with this task belies its complexity and the mental computations that support it. I adopt a view of anaphoric interpretation that breaks the process into two stages (e.g. Garrod and Sanford 1988). The first is an *initial retrieval* stage, where appropriate potential antecedents are selected from the set

[1]Natural languages boast a variety of anaphoric elements. In this thesis I restrict my focus to two kinds of anaphora in particular: local anaphors (reciprocals and reflexives), and pronouns.

of previously-seen NPs in memory and recalled. Suppose, for the moment that retrieval access all and only those NPs that are grammatically appropriate potential antecedents. Sometimes, this retrieval yields only one appropriate antecedent, as in (1-a). At other times, retrieval can yield multiple possible antecedents, reflecting genuine ambiguity in context. The examples in (2) illustrate such occasions. The reciprocal in (2-a) could take the *the cyclists* or *the spectators* as an antecedent, depending on its context of utterance.[2] In (2-b), the pronoun *him* can refer to either *the cyclist*, or *the other one*.[3]

(2) a. The cyclists pointed out the spectators to *each other* ...

$\Rightarrow_{PossibleGrammaticalAntecedents}$ {*the cyclists, the spectators*}

b. The cyclist said to the other one that the shorts flattered *him* ...

$\Rightarrow_{PossibleGrammaticalAntecedents}$ {*the cyclist, the other one*}

Determining the actual antecedent of an anaphor from the initial set occurs in a secondary *resolution* stage. At this stage, a unique referent is chosen, its 'fit' evaluated with context,

[2]Suppose there is a group of cyclists in a peloton who intend to remain vigilant against possibly running into groups of spectators distributed haphazardly along a race course. If over the course of the race different bikers spotted groups of on-lookers and alerted the team to their presence, (2-a) would be a felicitous description of the state of affairs, where *each other* picked out *the cyclists*. Alternatively, if a group of cyclists were to show two distinct groups of spectators the location of the other group, *each other* in (2-a) would be interpreted as *the spectators*.

[3]In the former case, an act of vanity is described. In the latter, a compliment is paid.

and covaluation set.[4] The process is complex, taking into account the interactions of a range of syntactic, semantic, and pragmatic information (see, e.g., Greene, McKoon & Ratcliff 1992; Gordon, Grosz & Gilliom, 1993; Gordon & Searce, 1995; Grosz, Joshi & Weinstein 1995; Kehler 2002; Rohde 2008 for discussion of various factors governing anaphoric resolution strategies) The resolution stage will not be discussed in this thesis. Instead we will restrict our attention to the initial retrieval stage.

Anaphors must provide instructions for the retrieval of their potential antecedents, which exist elsewhere in the linguistic or discourse representation. In principle, instructions for retrieval could come in two forms:

i) An anaphor could specify inherent *features* of potential antecedents, or

ii) an anaphor could specify positions potential antecedents could occupy in a representation.

In short, anaphors might specify what retrieval should look for, where in the representation retrieval should look, or both. To illustrate the first case: an anaphor could require that its potential antecedents match it on lexical/morphological features, such as gender and number. For example, the retrieval triggered by the pronoun *him* might use the features *masculine* and *singular* to recall an antecedent. This strategy would recall both

[4] 'Co-valuation' is a cover term for 'interpretive dependence' that subsumes coreference and binding. Examples of these two kinds of interpretive dependence can be exhibited using pronouns. Pronouns may be *co-referent* with a sentence-internal, or discourse-internal referential antecedent (i). Pronouns may also be *bound*. In (ii) the quantified NPs (QPs) bind the pronouns. These instances of QP-pronoun covaluation are not instances of coreference, because QPs are inherently non-referential.

(i) a. The cyclist$_i$ thought Mary would like his$_i$ new shorts.
    b. The cyclist$_i$ bought new shorts$_k$. He$_i$ thought Mary would them$_k$.

(ii) a. Every cyclist$_i$ believed the length of his$_i$ shorts was appropriate.
     b. No cyclist believed that the spectators were laughing at *him*.

NPs in (2-b), but retrieve only *the cyclist*, and ignore *the saleswomen*, in (3).

(3) The saleswomen$_{Fem.,Pl.}$ told the cyclist$_{Masc.,Sg.}$ that the shorts flattered *him* ...

$\Rightarrow_{Masc.,Sg.}$ {*the cyclist*}

Features used in this manner need not be limited to morphological or lexical features. They might also specify other grammatically important information. In some languages, anaphors are subject-oriented. For example, the German local reflexive *sich* must be bound by a subject. Here, the anaphor is unspecified for morphological features, so use of a feature like *[subject]* would be useful in retrieving the antecedent.

(4)

Der Radfahrer$_{[subject]}$	hat	*sich*	vor	den	Zuschauern	verletzt.
The cyclist	has	REFL	in.front.of	the	spectators	hurt

'The cyclist hurt himself in front of the spectators.'

Using features like these imposes a good filter on retrieval. However, it is not perfect. Retrieving candidates exclusively in terms of their inherent features predicts, under some circumstances, that retrieval should consider NPs as potential antecedents for an anaphor that the grammar otherwise disallows.

Consider the sentences below. Retrieving potential antecedents for the reciprocal in (5-a) solely based on the feature *plural* would return both *cyclists* and *spectators*. However, native speakers would agree that *cyclists* cannot be the antecedent for *each other*. Similarly, retrieving objects that match the morphological features of the pronoun in (5-b) would return *the man* as a potential antecedent for the pronoun, despite native speaker intuitions that *the man* is not an acceptable antecedent. The strategy also over-generates candidates when structural features are used. In the German example, retrieving *subject*s returns both the relative-clause and main subjects, though binding by the relative-clause subject is considered ungrammatical by native speakers.

(5) a. The cyclists that the spectators watched cheered for *each other*.

$\Rightarrow_{Plural}$ { the cyclists, *the spectators }

b. The cyclist said that the man spoke with *him*.

$\Rightarrow_{Masc.,Sing.}$ { the boy, *the man }

c. Der Redfahrer$_{[subject]}$, den die Frau$_{[subject]}$ sah, hat *sich* verletzt.
The cyclist who the woman saw has REFL hurt
'The man who the woman saw hurt himself.'

$\Rightarrow_{subject}$ { *the woman, the man }

What rules these feature-matching NPs out as antecedents is their structural position in relation to the anaphor. Structural relations have long been acknowledged to be important in determining possible covaluations between NPs and anaphors (Langacker 1969, Reinhart 1976, Lasnik 1976, Chomsky 1981, a.o.). Chomsky's (1981) *Binding Principles* codified the role of structure in governing the distribution of anaphoric elements and their potential antecedents.[5]

(6) THE BINDING PRINCIPLES (SIMPLIFIED):

Principle A:

[5] Chomsky also included *Principle C* in his Binding Principles, a constraint which ensures that R-expressions (proper names, descriptive NPs, etc.) are not covalued with any NP that c-commands them.

An anaphor must be bound within its local domain.

Principle B:

A pronoun must be free (i.e. not bound) within its local domain.

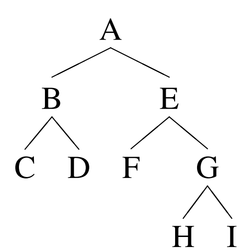
(7) α BINDS β iff:

- α is co-indexed with β, and
- α c-commands β

Chomsky's Binding Principles make direct reference to a generalized relational notion *c-command*.[6]

(8) C-COMMAND:

α c-commands β if the first branching node that dominates α also dominates β



In the tree above, *B* c-commands every node in the tree except *A*. *B* c-commands *F*, for example, because the first branching node immediately dominating *B* (*A*) also

[6]Klima's 'in construction with', and Langacker's (1969) 'command' began the push towards developing abstract geometric relations stated over constituent structure for describing constraints on transformations and dependency creation in natural language. Throughout the 1970s, various command relations were entertained for describing constraints on anaphoric dependencies, including Lasnik's (1976) *kommand*, and Reinhart's (1974) *superiority* - the precursor to modern formulations of c-command (see Barker and Pullum 1990 for a formal review).

dominates F. F does not dominate B because the first branching node that dominates F (E) does not dominate B.

Returning to the problematic cases of (5), we can now explain how the unwanted antecedents are blocked by the Binding Principles. *Spectators* is ruled out in (5-a) by Principle A, because (i) the NP does not c-command the reciprocal (and therefore cannot bind it), and (ii) the NP is not in the local domain of the reflexive (the minimal clause containing the reciprocal). The same explanation rules out *Die Frau* in the German example (5-c). Principle B rules out *the man* in (5-b) because it c-commands the pronoun, but is 'too local' (it is contained in the minimal clause). If the c-command relation is disrupted, however, by embedding *the man* in a possessor, it becomes an acceptable antecedent.

(9) The cyclist$_1$ said that the man$_2$'s mother spoke with him$_{1/2}$

What the above examples demonstrate is that the distribution of grammatical antecedents for an anaphor is subject not only to a condition of feature-concord, but also to the structural *relation* that the anaphor and the antecedent bear to one another in the syntactic representation. In order to align with the grammar, then, a retrieval process would necessarily take both feature and relational information into account.

1.1.1 The Challenge

Many popular memory models implement the idea that $p(R_m)$, the probability of retrieving an item m, is related to the degree to which features of the target memory match a set of retrieval cues that make up a *retrieval probe*, $\mathbb{Q}$.[7]

(10) $p(R_m) \propto \texttt{feature\_match}(m|\mathbb{Q})$

Accounts of retrieval of this kind were initially formulated to model performance in list memory tasks (Gillund & Shiffrin 1984, Hintzmann 1988, Clark & Gronlund 1996). This view of memory retrieval has been adopted by recent models of memory operations that underlie incremental linguistic processing (e.g. Lewis & Vasishth 2005). In such models any information explicitly encoded as a feature on an item in memory is an eligible cue for retrieval. However, some linguistic information is easier to encode as a feature than others. Lexical features like gender and number, and morphosyntactic features that encode grammatical role (e.g. *subject*) or case (e.g. *nominative*) are features that are easily encoded on individual items in memory. Representational relations, such as c-command between nodes in a tree, are not so easily encoded as features (as explained in detail in the next chapter). As a result, relational information cannot be used as an explicit cue to guide retrieval.[8]

Herein lies the challenge. As seen before, the distribution of potential antecedents for an anaphoric item is subject to relational constraints. Unfortunately, relational information

[7]The function `feature_match()` above is a stand-in for a host of different functions used in different models to calculate the degree of *association* between the encoding features on an image in memory and the retrieval probe. The innards of this function will be discussed in the next chapter, but for now this stand-in suffices.

[8]There is another possible way, apart from using c-command relations as direct cues for retrieval, to enforce that relations be respected. A serial search through a linguistic representation guided by local relations that are easily encoded as features can identify items along a c-command path. As I will discuss in the next chapter, such algorithms provide an adequate degree of relational-sensitivity, the time they require to execute is, in the limit, incompatible with the temporal demands of incremental processing, and inconsistent with empirical data on the time-course of retrieval.

for enforcing these constraints is, by hypothesis, not available to the retrieval operation. Without access to this information, it would initially seem that the retrieval mechanism would be incapable of sorting structurally-appropriate NPs from structurally-inappropriate NPs. Assessing if this prediction is borne out forms the empirical backbone of this thesis. I ask:

I. **Are structurally-inappropriate NPs distinguished from structurally-appropriate NPs in initial retrieval?**

I survey the processing of three anaphoric dependencies that have been argued to conform to a c-command constraint: *reciprocal licensing (in Hindi)*, *Bound-Variable pronoun licensing*, and *pronominal licensing in Crossover configurations*. The relational sensitivity of the retrievals associated with these dependencies has not been tested in the past. Contra the predictions of our base model, I show that retrieval apparently does distinguish between those NPs that the grammar rules out with a relational constraint from those the grammars allows.

This alignment between grammatical accessibility and retrieval accessibility must be explained. Given the assumption (motivated in the next chapter) that relations like c-command are not encoded explicitly on items in memory, an alternative explanation for retrieval's ability to pick out appropriate NPs must be offered. One hypothesis, not to be pursued in this dissertation, is that the findings of c-command sensitivity constitute evidence against using a feature-based retrieval mechanism. The alternative hypothesis is that there are methods of achieving the observed sensitivity while adhering to the assumption of a cue-based retrieval operation. Because feature-match is the method by which NPs are activated by retrieval, the first candidate solution is to find a feature with which the distinction can be made.

II. **Are there features that allow us to distinguish appropriate from inappropriate NPs?**

I argue that it is possible to give a feature-based account of the cases where we find a retrieval advantage for structurally-appropriate NPs. In all of the cases where this is an option, the feature exploited is a *proxy feature* that roughly encodes an NP's accessibility. The features themselves do not encode structural information *per se*, but rather some derivative thereof. I propose that the parser can dynamically track and assign NPs features that encode their accessibility relative to their structural position within the current syntactic domain. Determination of accessibility follows from simple heuristic procedures that take only local syntactic relations into account. On this view relational constraints are 'compiled out' into easily encoded features and functions that dynamically update features relative to a given domain. In many cases I cast the functions as *production rules* (Newell 1973, 1990; Lewis & Vasishth 2005) to be maximally explicit.

Feature-based strategies allow us the ability to align the output of retrieval processes more closely with the grammar, without requiring retrieval or encoding to make direct reference to c-command, or other relations. There is no 'c-command feature' that the retrieval operation uses to implement constraints.[9]

[9] In this regard my account differs from the attempts in Alcocer and Phillips (2012) to find an encoding of a generalized c-command feature.

With our proxy features in place, we can ask how effectively the system exploits them. In (most) grammatical theories, relational constraints are categorical. However, sensitivity to the extension of these constraints does not automatically imply that a parsing mechanism will display categorical sensitivity. In many models of memory retrieval, the default assumption is that various sources of information are combined linearly with equal priority or weight (see, e.g. Lewis & Vasishth 2005). This admits the possibility that information like morphological feature-match can effect retrieval of an NP under some circumstances independent of whether that NP is deemed structurally appropriate. This would mean that the information used to distinguish structurally-appropriate and inappropriate NPs would only impose a soft, or gradient constraint on retrieval. With this in mind, our third question becomes:

III. **Does the system restrict access to structurally-inappropriate NPs *categorically*, or in a gradient fashion?**

In a number of cases considered in this dissertation, it appears that the information used to distinguish licit from illicit antecedents imposes a hard or categorical constraint. Thus, we must find a way, within the confines of extant architectures, to impose such categorical constraints.

I explore the option of allowing retrieval to *preferentially weight* a distinctive structural feature *F*, such that mismatch on *F* renders moot feature match between an item and a probe on non-F features. If a probe and an item do not match on F, the item is functionally inert for the purposes of retrieval. I attempt to provide an explanation for categorical effects within the confines of well-motivated linear cue-combination schema (see Van Dyke and McElree 2011).[10]

1.2 Outline of the Thesis

Each of the content chapters of the dissertation is listed below accompanied by its synopsis. Chapters 3, 5, and 7 contain experiments and discussion of novel empirical findings. Chapters 2, 4, and 6 provide background context and theoretical discussion.

1.2.1 Chapter 2: Cue-Based Retrieval and Relations

Chapter 2 briefly lays out foundational assumptions about how items are encoded in memory. It discusses the practical challenges to encoding emergent, non-local relations explicitly on individual items. I argue that exhaustively encoding c-command relations on individual nodes as features is not a viable strategy for enforcing relational constraints because the algorithm for encoding such information cannot be performed in a time-efficient manner. I also argue that temporal considerations lead us to disprefer one conceivable method for searching memory, serial search through a representation, as a plausible candidate for memory access for online anaphoric dependency creation. Finally, the chapter introduces a direct-access, cue-based memory access mechanism adequate to the temporal demands of incremental parsing, reviews past psycholinguistic evidence supporting such a mechanism, and discusses the price of adopting such a retrieval mechanism: similarity-based interference.

[10] For most of the thesis, I adopt Van Dyke and McElree's (2011) proposal that differential weighting of retrieval cues is an adequate method for achieving certain apparently categorical restrictions on retrieval. In the conclusion, however, I discuss the challenges that face any strategy that relies on weighting to yield the correct results.

1.2.2 Chapter 3: Local Anaphor Licensing

The chapter begins the investigation of online sensitivity to relational constraints by taking up the processing of local anaphors. Previous work has furnished evidence that grammatically-accessible antecedents are privileged in antecedent retrieval. This evidence is reviewed. I propose that a simple feature-based explanation can be given for the observed sensitivity. A feature that tracks individual NPs' clause membership is utilized by retrieval mechanisms to implement a rough *clause-mate* restriction on the distribution of an anaphor's antecedents.

More importantly, it is shown that retrieval demonstrates selective categorical sensitivity to the structural position of antecedents for local anaphors. Sometimes retrieval considers only grammatical antecedents, while at other times antecedent retrieval appears to be influenced by the presence of NPs whose morphological features match the anaphor, but whose position in relation to the anaphor should bar them from consideration. These data suggest that though a feature-based account of the grammatical advantage in retrieval is possible, it cannot always be exploited by the parser. The experimental and theoretical contribution of the chapter is to (i) sharpen our understanding of the factors that govern the selective sensitivity, and (ii) propose an explicit account of how it arises.

Based on a review of English data, the selectivity appears to track the linear distance of the anaphor from its verb. By testing local anaphor licensing in Hindi, an SOV language, I tested whether immunity to interference is entirely predicated on linear adjacency to the verb. I show that pre-verbal reciprocals in Hindi display categorical sensitivity, which casts doubt on stating a cross-linguistic generalization of the selective sensitivity in terms of verbal adjacency.

I propose that the selective effects arise due to a context-dependent deployment of the *clause* cue. When the parser can determine that the *clause* feature is all that is required to retrieve an appropriate antecedent for the anaphor, it prefers its use. When it is unable to do so, it resorts to an error-prone strategy that results in gradient constraint application. I show how language-specific properties in combination with general architectural constraints confound, or support, the parser's ability to make this determination in English and Hindi.

1.2.3 Chapter 4: The LOCAL Feature

In Chapter 4 I propose a new feature, LOCAL that can be used to encode an item's accessibility as an antecedent with respect to Principles A and B. LOCAL is a dynamically-updated feature that tracks accessible chunks within the current clause. This function 'compiles out' the locality condition on both anaphor and pronominal licensing in procedural terms.

1.2.4 Chapter 5: Crossover and Principle C

The LOCAL feature is domain-dependent. It encodes structural accessibility relative to a limited structural domain. Not all relational constraints are limited to the local clause. This feature would initially seem ill-suited to be used to encode unbounded c-command constraints. Chapter 5 considers a case of an unbounded c-command constraint on pronominal interpretation: Strong Crossover, which rules out co-valuation between a pronoun and a filler whose gap the pronoun c-commands. Strong Crossover constructions provide a test of c-command sensitivity outside the domain of the local clause. Experiments 3 and 4 show that retrieval exhibits categorical sensitivity to Strong Crossover.

(11) STRONG CROSSOVER

Jane asked *which janitor*$_1$ it seemed that *he*$_{*1/2}$ had spoken with ...

Having established the sensitivity to the constraint, the mechanism by which the sensitivity is achieved is at issue. Experiment 5 tests susceptibility to interference from Weak Crossover-violating fillers, as a means of determining whether the sensitivity to Strong Crossover constructions in Experiments 3 and 4 can be explained by a general avoidance of unresolved fillers as antecedents.

(12) WEAK CROSSOVER

Jane asked *which janitor*$_1$ it seemed that *his*$_{*1/2}$ supervisor had spoken with ...

Experiment 5 shows that Weak Crossover-violating fillers are considered as potential antecedents for a pronoun, suggesting that the feature-matching filler's unresolved status is not the determining factor in the lack of interference in Experiments 3 and 4. Despite the fact that grammars often state the Strong Crossover constraint in unbounded terms, I show how an incremental parser can use the LOCAL feature to achieve sensitivity to Strong Crossover, and how such an account predicts the insensitivity to Weak Crossover.

1.2.5 Chapter 6: Scope Constraints

Resistance to interference in local anaphor licensing and Crossover is achieved via appeal to the LOCAL feature, which conjointly encodes locality and c-command. Chapter 6 introduces the relational constraint on the licensing of bound variable pronouns, sensitivity to which is investigated in Chapter 7. Before moving directly to experiments, Chapter 6 discusses various approaches to formalizing the constraint on bound variable licensing. In traditional Generative Syntax, the constraint has often been characterized as a positive c-command requirement between a quantifier and a pronoun it binds (e.g. Reinhart 1983). However, there are known examples of deviation from this requirement. Barker (2012) argues against a c-command constraint on bound variable pronouns, plumping instead for a *Scope Constraint*: all bound pronouns must fall within the semantic scope of their binders. I argue that a *covert*/LF c-command constraint is tenable, but a surface c-command constraint is not possible. I propose that if the the Scope Constraint is to be characterized in terms of a surface relation, it must be characterized in terms of MCCOMMAND (minimal-clause command).

1.2.6 Chapter 7: Bound-Variable Anaphora

Chapter 7 tests retrieval's sensitivity to the scope constraint on bound variable pronouns. It assesses whether gender-matching quantified noun phrases (QP) that scope over a pronoun (and are grammatically licit binders of the pronoun) are treated differently by retrieval than QPs that do not (and are therefore ungrammatical antecedents for the pronoun).

Two sets of studies are discussed. To answer question I, the first set determines whether an advantage for grammatically accessible QPs is observed over ungrammatical QPs in retrieval. I show that a distinction is indeed made. Experiments 6a & 6b, both judgment experiments, test offline sensitivity to the bound variable constraint. The results show that participants rarely chose a bound reading when the grammar would prohibit it, instead choosing to interpret unbound pronouns as coreferential with a discourse new entity. Experiment 6c used the previous experiments' materials in an eye-tracking study to show that feature-matching quantified phrases are readily retrieved as antecedents for pronouns in their scope, but not for pronouns outside their scope domain (i.e. pronouns across a clausal conjunct). These experiments establish a baseline level of sensitivity to quantifier-scope.

The second set of studies addresses question III. It tests whether the sensitivity to the constraint on bound-variable pronouns is categorical. Experiment 7 demonstrates that feature-matching, but structurally-illicit QPs are not considered during antecedent retrieval. Experiment 7 contrasts the accessibility of referential and quantificational noun phrases embedded inside a relative clause, for the purposes of licensing a pronoun outside the clause. We show that referential NPs are easily retrieved from a non-scoping RC-internal position based on their features, but QPs in the same position are essentially ignored (viz. they do not interfere with retrieval).

Given the fact that (categorical) sensitivity is observed, an implementation procedure is required. A LOCAL-based account is not viable, as it was with Crossover, because Scope Constraint is truly unbounded. I sketch a feature-based recoding of the scope constraint in terms of a feature, ACTIVE, which is used to mark an item's general accessibility to retrieval. An item's ACTIVE feature can be dynamically updated in much the same way as the LOCAL feature, though the relevant domain is not the clause.

Though the grammatical constraint on bound-variable licensing is stated in terms of MCCOMMAND, we see that the online explanation for sensitivity to the constraint does not make reference to the relation.

1.2.7 Conclusion

I discuss various open issues, including the origin of encoding features, the problems associated with assuming a preferential weighting approach to categorical effects and the role that c-command or syntactic relations play in the account of constraint sensitivity.

Chapter 2: Motivating a Model of Memory Encoding and Retrieval

2.1 Representation and Encoding

The work in this thesis adopts the view that items are stored as unitized *chunks*, or *images*, in a content-addressable memory (Kohonen 1980, McElree 2000, a.o.). Chunks are encoded as a set of feature-value pairs (e.g. Nairne 1990). These are usually called 'content features', where the term 'content' is read as picking out any information that can be stated as a one-place predicate. Values for features may either be symbols or pointers to another node (Lewis & Vasishth 2005). In the latter case, the feature-value pair represents a relation between two items (e.g. immediate precedence/subsequence of items in a list might be so encoded).[1]

The basic units of storage relevant for parsing are taken to correspond to a head and the phrase it projects to the exclusion of its complement (Kimball 1973; Lewis & Vasishth 2005). Each chunk has a unique identifying label (its chunk ID). Lexical content

[1]One way of defining content, in contrast to relations, is to specify that content features are attributes drawn from a finitely specifiable set of values. This conception is familiar to linguists. Consider content features such as Case, or grammatical role. Both of these features have a finite set of possible values, limited by the individual grammar of a language (and ultimately by UG). These can include nominative, accusative, ablative, etc. Grammatical roles are similarly circumscribed (subject, object, usw.). Relations, by contrast, cannot be so specified, as they are generated, and therefore potentially unbounded in number. Because relations can obtain between any (new) objects, they do not have this property. Admission of pointers, or node labels, as possible values in the encoding scheme above entails that the system does not limit itself to finitely specifiable feature values. So, the ontology of 'content' features is not limited in this way.

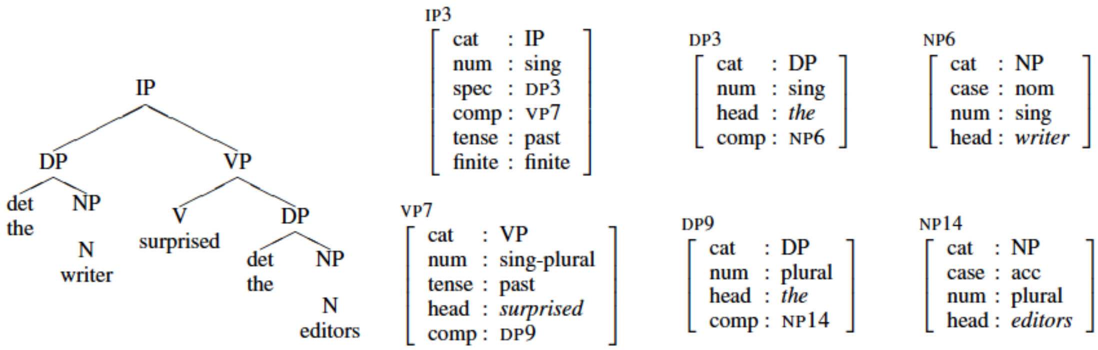


Figure 2.1: A *chunk* representation of simple syntactic structure. Figure from Lewis and Vasishth (2005: 380).

(category, morphological features, etc.), as well as structural features (Case, grammatical role, etc) make up a portion of its features. Local hierarchical relations can also be represented as features on a chunk. For example, local sisterhood/complementation is encoded (unidirectionally). the chunk ID of an element's complement is specified as the value for its *comp* feature.

2.2 Procedures for Recalling Information

We saw above how images are encoded, and the information we assume can and cannot be used in that encoding. With this, we can turn to how images are retrieved from memory. In this section I lay out two possible retrieval strategies: *serial search* and *direct-access retrieval* based on associative match.[2] 'Serial search' is a cover term for a family of search strategies that all have one thing in common: recall of information requires some iterative procedure through a representation. Serial search algorithms are very good at respecting relations, but they do not execute in constant time. Direct access retrieval

[2]For early discussion of different of the range of possible retrieval strategies, see Sternberg (1966, 1975).

does not involve an iterative search through a representation to make contact with potential retrievees. As a result, it executes in constant time. I discuss psycholinguistic evidence that supports the timing predictions of a direct-access retrieval mechanism. Following McElree, Foraker & Dyer (2003); Lewis, Vasishth & Van Dyke (2006), and others, I conclude that the time-course of retrieval online is inconsistent with serial search algorithms. A direct-access mechanism is therefore likely to underlie antecedent retrieval. I then go on to discuss a property of mechanisms like direct access that use associative match to retrieve items: their susceptibility to interference.

2.2.1 Tree Traversal/Serial Search v. Direct Access

Tree structures are abstract data structures that allow representation of hierarchical relations between elements. Tree-traversal algorithms provide guides for methods of iteratively searching a tree (Knuth 1967/1995). Each step in a tree traversal algorithm moves some prescribed distance.[3]

Tree traversal strategies are very good at restricting retrieval to elements within a domain defined in terms of relational properties. For example, it is easy to write an algorithm to search through the path of nodes that c-command an element in a tree. Such an algorithm in the representational confines of our encoding system would iteratively retrieve images based on their complement features, until the target of retrieval was found.

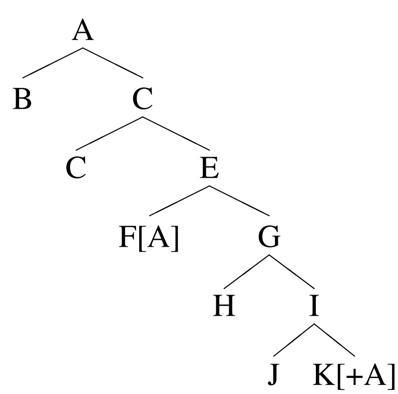
(1) ALGORITHM FOR SEARCHING THROUGH C-COMMAND DOMAIN

```
While current.Node not target:
```

[3] In many implementations of the algorithm, direction of search is specified by some alternative function. An example of this kind of search strategy is the A* search algorithm (Hart, Nilsson,& Raphael 1968).

```
ret = retrieve([comp:  current.Node]);

If ret == target:

    stop

Else:

    current.Node = max.Projection(ret);
```

To illustrate how this would work, consider a search for *F*, a node with the desired feature [A], from origin *K*. First, *J* would be retrieved via the [comp: K] feature. The algorithm would then check if the retrieved node was the target. Since it is not (it does not bear [A]) the search continues. The maximal projection of J,I, would then be retrieved through call of the function `max.Projection(J)`. As *I* does not have the appropriate feature, the search would continue, initiating subsequent retrieval of [comp: I],H. And so forth.



- retrieve(comp: K) $\rightsquigarrow \begin{bmatrix} label : J \\ comp : K \end{bmatrix} \rightarrow$
  check: J == target? **No.** $\rightarrow$
  max.Project(J) $\rightsquigarrow \begin{bmatrix} label : I \end{bmatrix}$

- retrieve(comp: J) $\rightsquigarrow \begin{bmatrix} label : H \\ \\ comp : I \end{bmatrix} \rightarrow$

  check: H == target? **No.** $\rightarrow$

  max.Project(H) $\rightsquigarrow \begin{bmatrix} label : G \end{bmatrix}$

- retrieve(comp: G) $\rightsquigarrow \begin{bmatrix} label : F \\ \\ comp : G \end{bmatrix}$

  check: F == target? **Yes.** $\rightarrow$

  F retrieved!

We see above that the number of steps (retrievals) required to retrieve the target increases linearly with distance between retrieval trigger and target. If individual retrievals take a fixed amount of time, the time required to retrieve of a target item using serial search is expected be proportionate to the distance between target and probe as measured in number of retrievals. If it could be shown that the time-course of item retrieval did not increase as distance between trigger and target increased, this would constitute direct evidence against use of a tree-traversing search procedure.

A separate way to retrieve items from memory forgoes iterative search entirely. On direct-access models retrieval specifies a set of features (or retrieval *cues*, combined into a *probe*) that the desired retrieval target should possess (McElree, Foraker & Dyer 2003, Lewis & Vasishth 2005). The probe contacts all items in memory in parallel. Once memory is contacted, items are activated in proportion to their strength of association or degree of match with the set of retrieval cues.[4] An item's activation level subsequently determines the probability that it will be retrieved.

2.2.2 Evidence

A parallel-access mechanism has been shown to be operative in a number of (though not all) retrieval tasks in laboratory settings. For example, McElree and Dosher (1989 - henceforth M&D) presented experimental evidence for its use in probing list memory. In a series of experiments, they showed that list size and an item's list position do not affect the speed with which the item is retrieved. Experiment 1 was a Speed-Accuracy Trade-off (SAT) study (Wickelgren 1977; Wickelgren, Corbett & Dosher 1980), in which participants were asked to identify items from a study list. Words in study lists were presented one-by-one in RSVP mode for 500ms apiece. Study lists were either 3 or 5 words in length. A brief break followed the presentation of the final word in the list, followed by the presentation of a probe word. Probe words were from the previous list or not. Participants were prompted to make a yes/no decision by button press indicating whether the probe word was or was not in the study list. The time lag of the response prompt varied from 100-1800ms after the presentation of the probe.

M&D found that set size and an item's serial position affected asymptotic accuracy of retrieval. The most recent item in the test lists was retrieved reliably faster than other items in the list, however, the speed of retrieval for other items in the list was not affected by their serial position. That is, items were retrieved at a constant speed, regardless of

[4] In the introduction, the promissory function `feature_match` $(i|Q)$ was used as a stand in for the function that determines how association is calculated. I defer the fleshing out of this function until the next section.

their position and regardless of list length. These results were confirmed in follow-up studies. M&D therefore argued for a direct-access parallel search mechanism that could access items based on their features in constant time, independent of their position in a representation.

There is evidence that a direct access mechanism is utilized for retrieval in incremental sentence processing. Studies have shown that the speed with which a target image is retrieved does not scale with the distance between the retrieval trigger and target. I first discuss demonstrations that linear distance between trigger and target does not influence retrieval speed, then go on to discuss the evidence that hierarchical distance does not affect retrieval speed either.

For the processing of wh-dependencies, McElree (2000) argued that the speed of retrieval of a filler at its resolution site was not influenced by the linear distance between resolution site and overt position of the filler. Using the SAT procedure, McElree contrasted speeds of filler retrieval at the final verb in sentences similar to the set given below. Linear distance was manipulated via center-embedding relative clauses onto the subject of the RC associated with the filler.

(2) a. This was the book that the editor admired.

b. This was the book that the editor who the receptionist married admired.

c. This was the book that the editor who the receptionist who quit married admired.

Retrieval dynamics were not affected by the length manipulation, but asymptotic accuracy was, with lower accuracies correlating with longer filler-gap distances. The lack of dynamics differences indicates uniform speed of retrieval across differing linear distances.

Subsequent studies have also demonstrated that linear distance does not influence retrieval dynamics for the resolution of other dependencies. In a series of studies Martin and McElree (2008, 2009, 2011) show that the retrieval of antecedents for elided constituents proceeds at uniform speed. Unlike retrieval of wh-filler, retrieval of an antecedent for an elided verb-phrase does not target active, unintegrated constituents. Verb phrases retrieved for ellipsis resolution have already been integrated, nor are they prospectively marked as being potential antecedents for later dependency creation.

Consider the task of recovering elided material. In constructions like English verb-phrase ellipsis (VPE), it is often the case that comprehenders encounter an underspecified, or dummy verbal auxiliary, whose local environment does not provide sufficient bottom-up information to convey the speaker's meaning. The interpretation of this dummy verbal auxiliary is dependent on a previously seen VP; *did*, or *was* in (3). In both instances the meaning of the 'missing' VP is filled in by copying information from the previous VP (4).

(3) a. John ate an apple, and Mary did, too.

b. John was a total glutton, and Mary was, too.

(4) a. John ate an apple, and Mary <ate an apple>, too.

b. John was a total glutton, and Mary was <a total glutton>, too.

Martin and McElree (2009) found that retrieval of an antecedent VP for resolution of VPE was subject to retroactive interference, but that dynamics differences did not vary as a function of the linear distance between the acceptable antecedent and the ellipsis site.

Martin and McElree (2011) further investigated ellipsis retrieval processes by looking at the resolution of sluiced structures. The term 'sluicing' (Ross 1969; Merchant 2001; Chung, Ladusaw & McCloskey 1995) refers to constructions in which the clause of an embedded question has been elided. Like VPE, semantic resolution of the elided phrases requires retrieval of a previous constituent's syntactic and/or semantic content. Unlike VPE, a dummy verbal auxiliary is not left behind. Instead the wh-word associated with the embedded question remains. The constituent retrieved (by hypothesis) in cases of sluicing is an entire preceding clause.

(5) *Michael typed something*, but he didn't tell me what. $\rightsquigarrow$

Michael typed something, but he didn't tell me what <he typed>.

When a sluiced structure is encountered online, its antecedent clause must be retrieved from memory. Martin and McElree compared the retrieval of an antecedent for a sluice, manipulating the distance of the appropriate antecedent from the sluice site (*Recent/Distant*), and whether another, semantically inappropriate VP was present in the sentence. Semantically inappropriate VPs were intransitive and therefore incapable of providing a verb that could take the sluiced remnant as an object. Below are the relevant example items from their overall experiment.[5]

[5]Unacceptable versions of each sentence (in which *which* served as the cue for sluicing retrieval) provided matched baseline conditions against which participants' sensitivity to the grammatical sentences could be compared.

(6) a. RECENT DOUBLE VP, SLUICING, ACCEPTABLE/UNACCEPTABLE

Michael slept and studied, but he didn't tell me what/*which.

b. DISTANT DOUBLE VP, SLUICING, ACCEPTABLE/UNACCEPTABLE

Michael studied and slept, but he didn't tell me what/*which.

c. RECENT SINGLE VP, SLUICING, ACCEPTABLE/UNACCEPTABLE

In the morning, Michael studied but he didn't tell me what/*which.

d. DISTANT SINGLE VP, SLUICING, ACCEPTABLE/UNACCEPTABLE

Michael studied in the morning, but he didn't tell me what/*which.

Martin and McElree found that speed of retrieval did not differ across conditions, regardless of (i) the distance between appropriate antecedent for the sluice and the wh-phrase, and (ii) the presence of the second VP. Accuracy, on the other hand, was affected significantly by both distance and presence of an additional verb. The results demonstrate that speed of retrieval of an unmarked constituent from memory is not affected by the constituent's *linear distance* from the probe.[6] However, just like McElree's (2000) findings, the results do not speak to the issue of whether *hierarchical distance* between probe and antecedent impacts retrieval dynamics. In all conditions above the structural distance between probe and VPs remains constant.[7]

McElree, Foraker, Dyer (2003) attempted to generalize the linear distance findings

[6]The study can arguably be seen as providing evidence for (partial-match) similarity-based interference (to be discussed in the next section). The additional verb *sleep* in the conjoined VP conditions was interfered with retrieval (presumably due to the fact that it bears a *verb* feature), despite the fact that it is not a suitable antecedent for the sluice interfering Vs were all intransitive, and therefore incompatible with the object wh.

[7]To accurately test the effects of hierarchical distance in VPE, raising clauses could be interpolated between probe and antecedent.

- Michael read something, but {∅ | *it seemed that*} he couldn't remember what.

to hierarchical distance. In their first experiment, the authors used filler-gap constructions very similar to those of McElree (2000), with one change. In order to test the effect of hierarchical distance the researchers interpolated complement clauses between filler and gap. Three conditions were tested: a condition in which the gap and filler were separated by one clause (7-a), a condition in which they were separated by two clauses (7-b), and one by three clauses (7-c).

(7) a. It was the scandal that the celebrity relished.

b. It was the scandal that the model believed the celebrity relished.

c. It was the scandal that the model believed that the journalist reported that the celebrity relished.

Once again, no results of distance were found on the filler's retrieval speed. Structural distance did influence asymptotic accuracy in the same direction as in McElree (2000) longer distances between filler and verb resulted in lower overall retrieval accuracy.

These results are encouraging, but they alone do not suffice to conclusively establish the insensitivity of retrieval speed to hierarchical distance. Interpreting the results as evidence of constant-time, direct access to a filler regardless of its structural distance from the gap requires the assumption that the filler was being retrieved its surface position in the representation. This assumption requires some motivation.

Some models of dependency creation assume that unintegrated constituents are kept in some privileged state, or register (see Wanner & Maratsos' 1978 idea of a *hold cell*). The invariance in retrieval dynamics observed in the foregoing studies might simply reflect direct access to the contents of such a store.[8]

In order to rule out this alternative interpretation and have support for direct access to previously-seen constituents in parsing, it is necessary to show that the filler is not in a privileged state. This could be done by demonstrating that retrieval at a gap site displays properties of general retrieval from an undifferentiated memory store.

There is some (modest) empirical evidence for this. Van Dyke and McElree (2006) showed that items that are inactive for the purposes of gap-filling interfered with the retrieval of the filler. Recall that Van Dyke and McElree (2006) found interference from semantically-related NPs in a memory load list. These NPs would arguably not occupy a hold cell, so direct access to the contents of such a cell would not encounter interference from these items. Yet, interference was observed, suggesting that retrieval of the filler accessed a general memory store.[9]

In sum, the experiments considered above motivate the idea that the retrieval of items from memory in real-time language processing is not affected by distance between a retrieval probe and target. This implicates the use of a direct-access retrieval mechanism and argues against the use of a serial search procedure for retrieval, at least *for some tasks*. None of them speak, directly, to the issue of whether retrieval of antecedents for anaphoric

[8]Successive cyclic reactivation of the wh-filler at every clause boundary could also lie behind the observed lack of distance effects.

[9]Another way to test the hold-cell hypothesis would be to look from interference from already discharged fillers in filler-gap resolution.

- It was the scandal that the model next to John believed (that the journalist reported) that the celebrity relished.
- It was the scandal that the model that John admired believed (that the journalist reported) that the celebrity relished.

McElree (2000) has some evidence that inactive fillers interfere with filler retrieval for active gap-filling, but the effects are rather small.

elements proceeds in constant time. Some studies have been conducted investigating pronominal resolution in SAT (e.g. Foraker & McElree 2007), but these studies have not tested the time-course of antecedent retrieval as a function of antecedent-pronoun distance. To my knowledge, there is one study that has investigated the dynamics of reference resolution using SAT: Dillon et al.'s (submitted) investigation of reflexive licensing in Mandarin Chinese. Dillon and colleagues show that resolution of the Mandarin anaphor *ziji*, which allows local and non-local antecedents, is faster when the anaphor's antecedent is local (in the same clause as the anaphor), than when it is distant (separated by more than one clause). These results could be taken as evidence that a structure-guided search mechanism is employed for retrieving referential antecedents. They are also consistent with a direct-access mechanism that employs to distinct retrievals. Suppose that the processor prefers local binding of the anaphor to long-distance binding. If there is such a preference, one would expect that the processor would attempt to resolve *ziji*'s reference locally. Encountering *ziji* would, on this account, trigger a retrieval of the local subject. If this subject was ineligible to antecede the anaphor, a second retrieval would be required to retrieve the next subject up. We see here that differences in time-course of retrieval needn't always entail tree-traversal or structure-guided search. As Dillon et al.'s findings are open to interpretation, I will move forward assuming that the method for retrieving antecedents is one and the same as the method for retrieving items for other long-distance dependency creation in incremental parsing.[10]

[10] See Alcocer (2011) for a different perspective.

2.2.3 The Consequence of Trading Structure for Speed: Interference

Though it provides rapid retrieval, the cue-based, direct-access model has its drawbacks. One is that it is not sensitive to implicitly encoded relations the way that tree-traversal algorithms were. An unavoidable outcome of cue-based retrieval process that always retrieves the item in memory that matches the retrieval probe best is that grammatically-inaccessible items may be activated based on their content features. As we saw in the introduction, ungrammatical candidates are retrieved as potential antecedents if only inherent cues are used. This type of unwanted *interference* is a hallmark property of retrieval from content-addressable memory. In the next section I discuss how interference arises in more detail and review evidence of interference in sentence-processing tasks.

2.3 (Retroactive) Interference Effects in Memory Retrieval

Within the memory literature, there are two main types of interference discussed: proactive, and retroactive interference (see, e.g., Anderson & Neely 1996). Proactive interference refers to instances when a previously-seen item influences the encoding of a new memory. Retroactive interference occurs when previously-seen, non-targeted memories affect the retrieval of a target memory. For the purposes of this thesis I will have nothing to say about proactive interference because it plays little role in any of the questions that concern me here. I focus exclusively on retroactive interference, and its origins. Given this, I will refer to retroactive interference simply as 'interference' from this point forward. The reader should keep this in mind. Within a cue-based model of retrieval, retroactive interference effects arise when the features of an unintended item resonate (fully or partially) with the feature contents of a retrieval probe.

2.3.1 Computing Feature Match for Retrieval

Most models of direct-access retrieval assume that the probability of retrieving an image from memory is proportional to its activation (A_i in the equations below). Functions to compute activation combine the strengths of association between individual cues ($Q_{j...n}$) in a retrieval probe and a particular image (I_i) to calculate that image's activation value ($S(Q_j, I_i)$ in the examples below).[11] Individual cues can be associated with a weight (w_j) that scales a cue's contribution to an item's overall activation level. The point of variation across models is the manner in which individual cue strengths are combined : the cue-combinatorics scheme. One way of combining cue strengths is with a linear, or additive, function. On such a model each cue contributes directly to an image's degree of activation, independent of the strengths of other cues in the retrieval probe. Given this assumption, the model allows activation of items that partially match the retrieval probe. (8) is an example of an additive combinatorics scheme adapted from Lewis and Vasishth's (2005) model. Cues can also be combined in a non-linear, or multiplicative fashion. In such models the contribution of individual features is not independent. As a result, they can exhibit sensitivity to conjunctions of features in the retrieval probe, which makes activation of partial matches less likely. Below is an example of such a scheme, adapted from Gillund & Shiffrin (1998).

[11] Gillund & Shiffrin (1984), Hintzman (1984, 1988), Nairne (1990), Raajimakers & Shiffrin (1981) and many others have used retrieval mechanisms with this general property in their models of recognition and recall - see Clark & Gronlund (1996) for a review. Functions in fully implemented models also incorporate a baseline activation term, which will be discussed in the next chapter, in the computation of overall activation at retrieval time. Many also include a stochastic noise term.

(8) LINEAR/ADDITIVE CUE COMBINATION

$$A_i = \sum_{j=1}^{n} w_j S(Q_j, I_i)$$

(9) MULTIPLICATIVE CUE COMBINATION

$$A_i = \prod_{j=1}^{n} S(Q_j, I_i)^{w_j}$$

To date, the most explicit parsing model that incorporates a memory and retrieval component (Lewis & Vasishth 2005) assumes an additive cue-combinatorics schema. For the purposes of this thesis, I will also assume that a linear cue-combinatorics schema is used to calculate associative match (see Van Dyke and McElree 2011). There are two reasons for this. First, the assumption makes it easy to account for evidence of activation of items that partially, but not fully, match the retrieval probe (a phenomenon which will be discussed next). Second, a linear cue-combinatorics schema is routinely used in modeling associative match effects across a number of cognitive domains (see Wolfe 1998; Treisman & Gormican 1988; Treisman & Souther 1985 for analogs of feature-match in the visual domain, and Trommershäuser, Landy, & Körding 2011 for recent perspectives across a number of modalities), and therefore has more independent theoretical support.

This system allows multiple items to be accessed and activated simultaneously on the basis of their feature content. They also permit interference. Interference occurs when feature-match exerts influence on the retrieval process. This influence comes in two kinds: *inhibitory* and *facilitatory* interference. I discuss evidence that both kinds of interference occur in incremental sentence processing below.

2.3.2 Inhibitory Interference

Inhibitory interference is the principal kind of interference studied in the psychology of memory. Inhibitory interference arises when smooth access to the intended target of retrieval is disrupted by one or more items in memory that share (some) features with the intended target. Within psycholinguistics, a number of studies have found evidence for similarity-based interference during the computation of dependencies that require retrieval. In all the studies discussed in this subsection, the trigger for inhibitory interference is the same: A grammatically illicit item that matches in (some subset of) features with the retrieval cues exerts a disruptive influence on the retrieval of a grammatically available item.

In a number of studies, Gordon and colleagues (Gordon, Hendrick, & Johnson 2001, 2004; Gordon, Hendrick, Johnson & Lee 2004; Gordon, Hendrick, & Levine 2002) have argued that similarity-based interference plays a role in explaining a number of effects in sentence processing. One example is the differing processing difficulties of computing subject and object relative clauses. The authors showed that manipulating the referential properties of nouns intervening between a filler and its gap-site affected the degree of integration difficulty at the gap site. If the filler (*the banker*) was more similar to an intervening referent (*the barber*), processing was more difficult at the gap site than when the the intervener was less similar (*you, everyone*). The complexity of later referents did not affect processing of a subject RC gap, as the gap site was encountered prior to the other referent.

(10) a. The banker that {the barber/a barber/Joe/you/everyone} praised climbed the mountain.

b. The banker that praised {the barber/a barber/Joe/you/everyone} climbed the mountain.

In a series of studies Van Dyke and collaborators have found evidence that structurally-illicit NPs engender inhibitory interference based on their syntactic and semantic features. Van Dyke and Lewis (2003) showed an effect of interference from *syntactic* feature-match. In the sentences below, the authors compared RTs at the verb region *was complaining* in the two conditions. By assumption, retrieval of the grammatical subject (*the resident*) is required at the VP, because the subject will have been displaced from the focus of attention by the linearly-intervening RC.

(11) a. The worker was surprised that the resident who was living near the dangerous warehouse *was complaining* about the investigation.

b. The worker was surprised that the resident who said that the warehouse was dangerous *was complaining* about the investigation.

In (11-a), the NP *the warehouse* is not a syntactic subject. In (11-b) it is the subject of the intervening RC. On the assumption that subject retrieval makes use of a subject feature, interference is predicted in (11-b), but not in (11-a). This is what the authors found. RTs at the verb phrase were significantly higher in (11-b), than in (11-a).

Van Dyke and McElree (2006) found evidence that extra-sentential semantic associates held in working memory could also interfere with the processing of filler-gap resolution.

(12) a. {TABLE/SINK/TRUCK}

It was the boat that the guy who lived by the sea *sailed* in two sunny days.

b. {TABLE/SINK/TRUCK}

It was the boat that the guy who lived by the sea *fixed* in two sunny days.

In (12-a) there is no semantic association between the verb *sailed* and any of the nouns in the memory load list. However, there is arguably an association between all of the items in the load list and the verb (12-b), since all three nouns are objects that can be *fixed*. The authors found an effect of this semantic association in increased RTs at the critical verb in (12-b) conditions compared to their counterparts in (12-a). This effect was not observed when participants read the sentences without the memory load lists, suggesting that the slowdown observed in load conditions is, in fact, related to retrieval interference and not to spurious effects of sentence complexity.

Van Dyke (2007) compared the relative strength of similarity-based interference from syntactic and semantic cues. In a 2x2 design, Van Dyke manipulated the semantic similarity of an intervening NP to the grammatical subject of a verb (*resident*) by manipulating its animacy (*warehouse / neighbor*). As in Van Dyke and Lewis (2003), the syntactic position of the intervening NP was also manipulated - it was either a subject of the RC-internal verb or the object of a PP - to manipulate syntactic similarity.[12]

[12]Colin Phillips points out that the 'syntactic' factor here, which manipulates grammatical roles, could

(13) The worker was surprised that the resident ...

a. who was living near the dangerous warehouse — LoSyn/LoSem

b. who was living near the dangerous neighbor — LoSyn/HiSem

c. who said that the warehouse was dangerous — HiSyn/LoSem

d. who said that the neighbor was dangerous — HiSyn/HiSem

... was complaining about the investigation.

Across a number of tasks, Van Dyke found evidence for inhibitory syntactic interference. HiSyn conditions displayed longer RTs in critical regions and lower accuracy on comprehension questions. Evidence from semantic interference was observed in some measures of an eye-tracking study, but was not observed in a study where participants read a sentence and were then asked to give an index of their understanding of the sentence (the "Got It?" task of Frazier, Clifton, & Randall, 1983).

Van Dyke and McElree (2011) showed that semantic features such as *animacy* can also contribute to inhibitory interference in both an SAT and an eye-tracking study. The verb *compromised* requires an animate subject, thus retrieval for its subject would presumably use *animacy* as a cue for retrieval. McElree and Van Dyke manipulated the animacy of a structurally-inappropriate intervening NP (*witness/motion*).

(14) The attorney who the judge realized had declared that {*the witness / the motion*} was inappropriate *compromised* (during the negotiations).

also be seen as a 'semantic' factor because the difference in grammatical roles corresponds to a difference in semantic roles.

The researchers found evidence for greater interference from the animate NP than the inanimate NP, as measured by asymptotic accuracy in the SAT study, as well as total reading times in the eye-tracking study. Importantly, in a follow-up experiment the authors found that the effects of semantic interference disappear when the intervening animate NP is placed in a non-subject position.

(15) The attorney who the judge realized had rejected {*the witness / the motion*} in the case *compromised* (during the negotiations).

This effect could be taken as evidence for multiplicative cue-combinatorics. Here, it would appear that the conjunction of a mismatching syntactic feature and a matching semantic feature does not result in interference, whereas matching on both syntactic and semantic cues does.

2.3.3 Facilitatory Interference: Illusions of Grammaticality

The hallmark of inhibitory interference is slower, or more difficult, processing relative to cases where no interferer is present. Facilitatory interference is the opposite: there are cases where it appears that the presence of a non-target image in memory that matches the probe (on some portion of the retrieval cues) eases processing. Within the context of dependency processing, the conditions under which facilitatory interference arises differ from those of inhibitory interference. Whereas inhibitory interference occurs when the conditions for the computation of a grammatical dependency are met, facilitatory interference occurs when the cost of processing an grammatically unlicensed item is reduced by the presence of a grammatically unavailable but feature-matching item in memory. The effect of facilitatory interference can be seen as an *illusion of grammaticality*, a case where interference (temporarily) tricks the parser into considering an ungrammatical dependency grammatical (Phillips, Wagers & Lau 2011).

Facilitatory interference has been argued to be a case of mis-retrieval (e.g. Vasishth et al 2008; Wagers et al. 2009). Retrieval of the feature-matching intervener eases processing, by either preventing retrieval failure or preventing retrieval of an item with even lower feature-match. Facilitatory interference can be seen as a kind of unavoidable consequence of a memory access mechanism designed to always choose the chunk with the highest degree of match with the probe - regardless of other factors.

To date, two phenomena have been argued to be instances of facilitatory inteference: subject-verb agreement attraction (e.g. Wagers et al. 2009; Dillon et al. 2013; inter alia, and illusory negative polarity item (NPI) licensing (Drenhaus, Frisch & Saddy, 2005; Vasishth et al, 2008).[13]

Agreement attraction occurs when a verb agrees in features with an NP other than the grammatical subject (in nominative-accusative languages - see, e.g. Kimball & Aissen 1971; Bock & Miller 1991; den Dikken 2001). This most often occurs when the NP with which it erroneously agrees intervenes between it and its grammatical subject. For example, in the sentence below, the main verb *were* agrees in number with the NP *the cabinets*, rather than the grammatical subject *the key*.

(16) The key to *the cabinets were* on the table.

[13]Recent evidence from Sloggett (2013) also suggests that facilitatory interference may be observed in Case licensing as well.

This phenomenon has been well-documented in language production experiments in English (Bock and Miller 1991; Eberhard, Cutting & Bock 2005, Vigliocco & Nicol 1998, Franck, Vigliocco & Nicol 2002; Staub 2010; Gillespie & Pearlmutter 2011), as well as in other languages (Vigliocco & Franck 2001; Anton-Mendez, Nicol & Garrett 2002; Franck et al 2002; Hartsuiker, Schriefers, Bock & Kikstra 2003; Badecker & Kuminiak 2007). More important for our purposes, the effect is also robust in comprehension studies (Clifton et al 1999; Pearlmutter, Garnsey and Bock 1999; Häussler & Bader 2009, Wagers et al. 2009, Dillon et al. 2013).

Wagers et al. (2009) showed that reading times at a verb that mismatched its grammatical subject in number features were lower when another preceding NP matched the features on the verb. That is, (17-c) and (18), were read more slowly than (17-d) and (18-d), respectively.

(17) a. The key to the *cell* unsurprisingly *was* rusty ...
b. The key to the *cells* unsurprisingly *was* rusty ...
c. *The keys to the *cell* unsurprisingly *were* rusty ...
d. *#The keys to the *cells* unsurprisingly *were* rusty ...

(18) a. The musician who the reviewer praises so highly...
b. The musicians who the reviewer praises so highly...
c. *The musician who the reviewer praise so highly...
d. *#The musicians who the reviewer praise so highly...

Wagers et al. (2009) argued that this effect should be modeled as an instance of facilitatory interference. Encountering a verb that mismatches the predicted number features for the grammatical subject triggers a retrieval for a subject with the features seen on the verb. For example, encountering *were* in (17-d), triggers a probe for a plural-marked subject. In such a case, the intervening NP (*cells*) will be erroneously retrieved some portion of the time, due to its partial-match with the probe.

Interference arises from NPs in a number of positions. Wagers et al. (2009) showed this effect of facilitatory interference arises on main verbs from structurally-inappropriate NPs embedded within a PP attached to a verb's grammatical subject (17), or on RC-internal verbs from a their unintegrated head. Dillon et al. (2013) also showed that RC-internal NPs could interfere with the processing of main verbs.

The second dependency licensing process that has been used to argue for facilitatory interference is NPI-licensing (e.g. Vasishth et al 2008). Drenhaus et al. (2005), Vasishth et al. (2008), Xiang, Dillon & Phillips (2009) and Parker & Phillips (2011, et seq.) have all shown that negative polarity item (NPI) licensing is subject to 'illusions of grammaticality'. NPIs are lexical items or phrases such as 'ever', 'any', 'anymore', or 'a red cent' that must fall within a semantically-licensing context. Often, this licensing requirement is translated into the restriction that NPIs must fall within the scope of a downward-entailing operator (Ladusaw 1979), or in the c-command domain of a negative element (a more restrictive paraphrase of Ladusaw's scope restriction).

An illustrative example is given below. *Ever* is licensed in (19-a) below because it is in the scope of the negative QP *no cyclist* (as it is in the c-command domain of that item). It is not in the scope of a negative item in (19-b), so the sentence is judged unacceptable.

(19) a. No cyclist will ever stop wearing tight shorts.

b. *A/The/Some cyclist will *ever* stop wearing tight shorts.

The presence of a potentially-licensing quantifier does not improve acceptability if it does not scope over the NPI.

(20) *The cyclist that no spectator could help staring at will ever stop wearing tight shorts.

These offline judgments are quite robust. However, under time pressure they seem to be susceptible to noise. Drenhaus et al., 2005 showed that the presence of a negative item in the wrong position raises the likelihood that participants judge unlicensed NPIs as acceptable in German. In a speeded acceptability task, participants were asked to rate sentences containing the German NPI *jemals* (roughly equivalent to English *ever*). The negative quantifier *kein* (*no*) was either present and c-commanded the NPI (21-a), was present, but did not c-command the NPI (21-b), or was completely absent from the sentence (21-c). Participants were fast and accurate in judging the licensed and unlicensed (negation absent) sentences (accuracy rates are presented next to the conditions below), but were slower, and less accurate in judging the sentences where negation was present, but did not c-command the NPI. Accuracy rates in the acceptability task are given to the right of each condition below.

(21) a. Kein Mann, der einen Bart hatte, war jemals glücklich. 85%

No man who had a beard was every happy.

b. *Ein Mann, der keinen Bart hatte, war jemals glücklich. 70%

*A man who had no beard was ever happy.

c. *Ein Mann, der einen Bart hatte, war jemals glücklich. 83%

*A man who had a beard was every happy.

A general susceptibility to illusory NPI-licensing has since been shown to be a rather robust phenomenon. Vasishth et al. (2008) found the same pattern in self-paced reading times in German, while Xiang, Dillon & Phillips (2006, 2009) found evidence for illusory licensing in English using judgment measures, self-paced reading, and ERP. Parker & Phillips (2011 et seq.) have also found the phenomenon in English, in speeded-acceptability and self-paced reading.

Vasishth and colleagues (Vasishth et al. 2008; Lewis, Vasishth, & van Dyke, 2006) have argued that the illusion of grammaticality is due to mis-retrieval of the structurally-inaccessible negation as a licensor for the NPI due to partial cue-overlap. They have argued that encountering an NPI triggers a retrieval of its licensor, which must be [+negation]. Some proportion of the time, cuing for an element with this feature erroneously retrieves the non-commanding *kein*, leading to facilitated processing of an otherwise unlicensed NPI.[14]

To conclude this section, we have seen empirical evidence for both inhibitory and facilitatory interference effects in online sentence processing, supporting the hypothesis that cue-based, associative match drives retrieval.

[14]This interpretation has not gone unchallenged. For example, Xiang, Dillon and Phillips (2009) argued that illusory NPI licensing was better viewed as a consequence of over-active pragmatic accommodation, rather than mis-retrieval. For the moment, I will not discuss these objections further, but I do return to considering the proper interpretation of illusory NPI-licensing in the conclusion chapter.

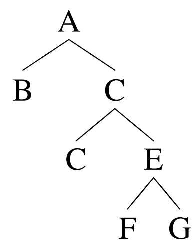
2.4 Remembering Relations

The Binding Principles make explicit reference to c-command, but as we saw in Figure 2.1, it is often assumed that an individual node's c-command relations are not encoded as features. Since relations like c-command are not encoded as features, c-command information is not immediately accessible for use as a cue in retrieval.[15] If we desire a system that made c-command information directly accessible, it would have to be explicitly encoded. At times, such encoding schemes have been dismissed under the assumption that they would undermine a principled distinction between relations and content. One reason for dismissing the claim is that specification of relational information requires use of two-place predicates, but the system has the representational capacity for one-place predicates. But, two-place predicates can be curried/schönfinkeled to act as one-place predicates. This process is already used to represent other relations as feature-value pairs in encoding schemes like Lewis & Vasishth's. For example, the *comp* feature encodes a selection/sisterhood relation between a node and its complement by providing the label of the complement node (or a pointer thereto) as the *value* of that feature. Encoding one relation in this way creates the capacity to encode *all* (binary) relations in the same manner. Using the same tactic, the two-place *X c-commands Y* relation could be curried to create the one-place feature [C-COMMANDS: Y], which would be predicated of X. One could easily imagine an encoding scheme that tagged every individual node with an exhaustive

[15] That non-local relations are only implicitly encoded in the representation scheme above is sometimes discussed as a unique failing of a content-addressable/multi-image encoding. In truth, the problem cuts across most encoding/representational schemes. C-command is similarly implicit in a tree representation of a syntactic parse as it is in the chunk-based representation. The relation can be reconstituted, or read off the entire representation, but not off any individual node itself.

list of its c-commanders via features like this.[16]

Take the small, right-branching tree below.



An exhaustive encoding of c-command relations as feature-value pairs on the node *B* might look like this:

$$\text{(22)} \qquad \mathrm{B} = \begin{bmatrix} commands : C \\ commands : D \\ commands : E \\ commands : F \\ commands : G \end{bmatrix}$$

Under such an encoding scheme, it would be easy to impose a c-command constraint on retrieved items. Retrieving a list of c-commanders for node G, for example, would only require specifying that anything retrieved bear the feature [C-COMMANDS: G]. Given that such an exhaustive encoding of c-command is possible in principle, one must ask why it is not exploited. The answer, I argue, lies in functional constraints on encoding operations.

[16]A system that exhaustively encoded c-command relations on each node would be similar in spirit to Frank & Vijay-Shanker's (1999) *primitive c-command* approach to syntactic representation, in which c-command played the role of a primitive, constitutive relation.

2.4.1 Exhaustive Encoding is Too Exhausting

People understand sentences in real time. The task imposes temporal constraints on all parsing. One way to formalize the notion of an algorithm's run time is *time complexity analysis* (see Sipser 2006). An algorithm's time complexity, commonly stated in big O notation, provides a way of quantifying the number of steps required to run the algorithm as a function of the length of input length (n).

Two examples:

1. Determining whether a number represented in binary is odd or even proceeds in *constant time* - $O(1)$. Regardless of the length of binary input string be it *101*, or *1010101010101010*, only one operation is necessary: checking the final digit.

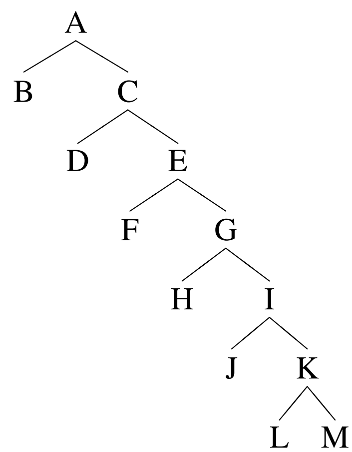
2. Checking, iteratively, whether the first digit in a list matches any other digit in the same list proceeds in *linear time* - $O(n)$. In the worst case, the number of operations it will take to check the list is equal to the number of items in the list.

Algorithms with time complexity $O(n)$ or greater take longer to run as a the size of the input list grows. For parsing on the whole, this makes intuitive sense. The time to process a sentence scales with its length. When restricting focus to an individual step in incremental parsing, we have the opposite intuition. Primitive operations should be insulated from the length of the full representation. We can state this intuition as a principle.

(23) CONSTANT-TIME PRINCIPLE

Operations, whenever possible, should execute in constant time (they should be $O(1)$)

Should we accept the principle, we have a foothold in explaining why an exhaustive encoding scheme of the kind proposed above is not a viable option. Exhaustive encoding requires an algorithm that will fail to adhere to (23). Let's see why by considering the procedure required to provide an exhaustive specification of c-command features on every node in the right-branching tree structure below.



Assume that the nodes in the tree above are encountered incrementally, from left to right. Upon encountering an item, e.g. node *B*, an image must be created. Suppose that the feature-value pairs encoded on *B*'s image can draw from all information available at the time B is processed. Information about the specific items B c-commands are not available at this time, because they have not been encountered yet. Therefore, assignment and encoding of the feature [C-COMMANDS: X], for any later item X, can only be assigned *retrospectively*. Upon encountering X, *B* must be retrieved if it is to be marked [C-COMMANDS: X].

For a right-branching structure, where precedence corresponds to c-command, we

list size	1	2	3	4	5	6	...	n
# retrievals at new item	0	1	2	3	4	5	...	*n-1*
absolute retrievals	0	1	3	6	10	15	...	$\frac{n^2-n}{2}$

Table 2.1: Number of Retrievals Required to Exhaustively Encode C-command Relations in a Right-Branching Tree

can state the algorithm for assigning [C-COMMANDS] features to commanding items.[17]

1. Input `current.Item`.

2. Construct feature [C-COMMANDS: `current.Item`]

3. For every preceding image *i*, retrieve *i*, assign [C-COMMANDS: `current.Item`] to *i*

The problem with this strategy is immediately apparent. Exhaustively encoding command relations between the next item in a list and its predecessors requires retrieval of all predecessors to update their images with the new feature. Therefore, every new item will initiate *n-1* retrievals, where *n* is its ordinal position in the list of terminal nodes.

Because the number of retrievals scales with the number of items in the list, encoding will not be a constant-time operation. Each step of encoding will be $O(n)$. Complexity will cascade throughout the system. The algorithm to exhaustively encode an entire structure will be $O(n^2)$. Note such a complex encoding algorithm makes a very clear prediction: if retrievals take some amount of time (fixed or otherwise), the minimum amount of time required to complete each iterative step of the parse should increase by at least the amount of time taken to complete a retrieval. This does not comport with the evidence: no studies have shown that comprehenders reliably slow down as they move through a sentence. Thus,

[17]In more complicated structures containing left-branches, precedence does not map perfectly to c-command, so the retrieval procedure would have to be amended. A right-branching structure represents the *easiest case* for the algorithm.

the encoding scheme should be rejected based on its disagreeable level of complexity and its unsupported predictions.[18]

Note that encoding local (hierarchical) relations is not plagued by the same problem. The feature [comp: X] can be filled in simply by copying the label of a node's sister. Regardless of the size of the representation, the number of operations remains the same. The same holds for all information displayed in the encodings of Figure 2.1. Lexical information is encoded as a set of inherent features on an image as a reflex of lexical access, while syntactic features, such as encoding of grammatical role, or Case, can be determined in the same local manner, regardless of the size of the preceding representation.

2.5 Summary

In this brief section we have discussed the representational and architectural assumptions that are the basis for later chapters. To recap:

i) Linguistic representations are stored in memory as sets of discrete images, or chunks encoded as feature-value pairs.

ii) Retrieval proceeds via a direct, parallel-access mechanism.

iii) Retrieval based on associative match is prone to interference from images that overlap

[18] Alcocer & Phillips (2012) discuss another way of encoding c-command exhaustively, by using the feature [C-COMMANDED-BY], instead of [C-COMMANDS]. The authors show that the encoding algorithm proceeds in constant time. However, the authors argue that encoding c-command relations in terms of 'commanded-by' does not provide a useful set of retrieval cues. Under a 'c-commanded-by' encoding scheme, node F in the tree above would have the features [*c-commanded-by: E*,*c-commanded-by: D*,*c-commanded-by: C*,*c-commanded-by: B*,*c-commanded-by: A*]. With such an encoding scheme, retrieval of all of the c-commanders of F could not be called using a singular cue. This encoding scheme does not treat the c-commanders of F as an equivalence class for the purposes of retrieval. Using so many cues would lead to increased partial-match interference. It would also entail that a target could never fully match the retrieval probe in all its features.

in features with the target and probe.

iv) Non-local relations like c-command are not explicitly encoded as features, due to functional constraints on encoding.

v) Relations like c-command cannot be used as retrieval cues, because they are not encoded as features on memory images.

Keeping these assumptions in mind, return to the issue of antecedent retrieval for anaphora and enforcing relational constraints thereon. With respect to enforcing categorical restrictions of the type enforced by the grammar, our retrieval procedure appears to offer little help. For example, (d) would lead us to expect some degree of interference from feature-matching, structurally-inappropriate NPs on the basis of partial feature-match even in the presence of 'c-commanding' features. Assumption (e) complicates matters further. Without c-command features a simple account of pronoun retrieval, e.g. one based solely on the morphological features of the trigger, would have no way of distinguishing between eligible and ineligible NPs at all. Interference from grammatically illicit NPs would be rampant. I show see that this prediction is not borne out.

Chapter 3: Local Anaphoric Licensing in Hindi

3.1 Introduction

Some anaphoric expressions in natural language require their antecedents to fall within structurally-delimited domains. These are called ‘local anaphors’. The set of local anaphors and their precise locality restrictions are subject to some variation cross-linguistically, but many languages (such as English) have two types of local anaphors: reflexives (1-a), and reciprocals (1-b).

(1) a. John likes *himself.*

b. The men like *each other*.

Reflexives and reciprocals require that their antecedent sit ‘above’ them in the syntactic tree and that their antecedents be sufficiently *local*. Items whose feature composition matches that of the reflexive or reciprocal cannot antecede the element if these conditions are not met. The italicized NPs in (2) match the local anaphors in features, but they do not c-command them. They are therefore ineligible antecedents. The matching phrases in (3) are too far from the anaphors to serve as antecedents because they are separated by a clause boundary.

(2) a. *The woman [who *the man* met] was full of *himself*.

b. *The woman [who *the men* met] was full of *each other*.

(3) a. **John* said Mary likes *himself*.

b. **The men* said Mary likes *each other*.

In this chapter I investigate local anaphor licensing as the first step in assessing retrieval's performance in respecting relational grammatical constraints. The three questions stated at the outset of the thesis can be reformulated into questions specifically concerned with local anaphor licensing:

I. Does antecedent retrieval for local anaphors distinguish structurally-accessible antecedents from structurally inappropriate feature-matching NPs?

II. If so, how is this distinction achieved without making reference to structural relations like c-command? What is the feature, or strategy, the system exploits?

III. If a distinction is made, is it categorical, or gradient? Do structurally inappropriate feature-matching NPs interfere with retrieval at all?

Unlike the dependencies that I consider in later chapters, a good deal of work has been done on the processing of local reflexives, which has established, or begun to establish answers to these questions. As concerns the first question, there is ample evidence that grammatically-accessible antecedents are privileged in retrieval of antecedents. I review this evidence in the next section.

With regards to the second question, there is not a uniform consensus. Some authors (e.g. Dillon et al. 2013) have proposed that antecedents for reflexives are distinguished using ‘structural features’, as opposed to morphological features, though they have remained vague as to the exact identity of those features. A number of possibilities exist. In this chapter I advocate one option: a feature that tracks the clause membership of an NP can be exploited to make the rough distinction.[1]

To come to the third general question of interest in the thesis: Do grammatically-inaccessible feature-matching NPs interfere with retrieval, or are they categorically ruled out? In reviewing the literature we find selective categorical sensitivity to the structural position of antecedents for local anaphors. In English, when the reflexive trigger is adjacent to its verb, retrieval of antecedents does not appear prone to interference. (e.g. Dillon et al. 2013) Interpolating distance between an English reflexive and its verb, though, increases susceptibility to interference (King et al. 2012).

The experimental portion of this chapter seeks to determine the source of the selective effect. By testing local anaphor licensing in Hindi, an SOV language, we are able to test whether immunity to interference is entirely predicated on linear adjacency. Pre-verbal reciprocals in Hindi resist interference, a fact which casts doubt on the strongest version of a linear adjacency account. I propose that the selective effects arise due to a context-dependent cue generation[2] procedure: when the parser can determine that use of an entirely diagnostic cue is appropriate, interference effects do not arise. I argue that a rational cue generation procedure would use the following rule to determine a local anaphor’s retrieval

[1] I revise this proposal in Chapter 4 so that the same feature is used in the account of both local anaphors and sensitivity to Binding Principle B.

[2] Cue generation procedures have received very little attention in recent accounts of retrieval in incremental processing. Lewis, Vasishth & Van Dyke (2006: 448) briefly mention the procedure and justify it so: ‘the cues that drive the retrieval are not simply features of the lexical item but are grammatically derived from the current word and context…’. Yet, the authors devote no other discussion to the process by which it occurs. In this chapter I will be more explicit about what I take the process to entail.

probe:[3]

(4) When generating a cue set for a local anaphor A,

- If A's only clause-mate is the local subject:
  - Prefer the diagnostic feature [clause] over [ϕ_A]
- Else:
  - Exhibit no preference for [clause], [ϕ_A], where ϕ_A are A's morphological features

Rational performance is *bounded* by the architecture of memory and the parsing mechanism (see Simon 1957, Gigerenzer & Selten 2002). I show how language-specific properties in combination with general architectural constraints make the rational ideal more or less difficult to achieve in English and Hindi. It is argued that the Hindi parser can implement the optimal cue generation procedure by making use of information implicitly encoded in its incremental parsing state. In English, on the other hand, this information is not available. In English, the parser uses a more error-prone heuristic to determine the cue generation procedure - which reduces to a linear distance metric. While linear distance from the verb is referenced in the English explanation, it is given no direct role in the larger characterization of the cross-linguistic problem of local anaphor licensing.

The structure of the chapter is as follows. First, I review previous experimental work on the consideration of structurally inappropriate feature-matching NPs licensing during the real-time licensing of object reflexives. I propose a feature-based method by

[3] The term 'prefer' in the statement below is an intentional hedge. There are a number of ways to cash out a 'preference' for a particular cue in retrieval, as I later show.

which retrieval could make the required distinction between structurally inappropriate and appropriate NP antecedents. This feature could be exploited to avoid interference categorically in at least two distinct ways, each of which are discussed. In light of previous findings that suggests that anaphor antecedent retrieval is not absolutely immune to interference, I conclude that the application of the procedure for avoiding interference must be conditionalized on some kind of information, or subject to a decision process. I show how conditionalizing the procedure on linear adjacency between anaphor and verb can achieve the right facts in English. Conditionalizing the procedure entirely on linear adjacency makes the strong prediction that retrieval for local anaphors in SOV languages should always display interference from structurally inappropriate feature-matchers. This prediction is tested using Hindi reciprocals, anaphors which precede their verbs, but are otherwise similar to English reflexives in terms of the structural constraints they are subject to (see e.g. Bhatt & Dayal 2007). Section 3.5 presents the results of simulations intended to model the differing predictions of interference effects in Hindi reciprocal licensing made by an account that predicates lack of interference on linear order. These results are then compared with behavioral results from a self-paced reading experiment. The Hindi results disconfirm the hypothesis that interference immunity is directly predicated on the anaphor directly following its verb in linear order. In section 3.6 I provide my alternative explanation for selective susceptibility to interference that covers the cross-linguistic data.

3.2 Previous Studies on Processing Reflexives

3.2.1 Interference-Immunity

This section reviews a number of previous studies concerned with the processing of reflexives and the retrieval of their antecedents. Prior work has focused on finding evidence for or against interference from, or temporary consideration of, structurally illicit NPs during antecedent retrieval.

3.2.1.1 Nicol and Swinney 1989

In a cross-modal priming task, Nicol and Swinney (1989) tested the activation of grammatical and ungrammatical potential antecedents for anaphoric DPs (both pronouns and reflexives). On the assumption that (i) consideration of an NP as a potential antecedent involves its retrieval and (ii) retrieval and activation of this potential antecedent would result in facilitative priming of its lexical associates, participants were shown semantic associates of all three nouns that preceded the anaphor at the probe point indicated by the asterisk in the sentences below.

(5) The boxer told the skier that the *doctor* for the team would blame *himself* * for the recent injury.

(6) *The boxer* told *the skier* that the doctor for the team would blame *him* * for the recent injury.

Priming scores, which represent the difference between the lexical decision reaction time for a semantically unrelated control word and the decision time for a semantic associate of one of the preceding words, are given below. Larger priming effects indicate faster RTs on semantic associates than on unrelated controls. Priming effects are indicative of antecedent consideration.

Referent	Priming Effect at Reflexive	Priming Effect at Pronoun
DISTANT$_1$ (Boxer)	-1	43
DISTANT$_2$ (Skier)	11	58
LOCAL (Doctor)	104	-21

In the reflexive condition there were no priming effects associated with the DPs that were structurally ineligible to antecede the reflexive. RTs to semantic associates of these nouns did not differ significantly from the unrelated control conditions. Semantic associates of the reflexive's grammatical antecedent (*doctor*), on the other hand, were processed significantly faster than unrelated controls in this position. The reverse was found in the pronoun conditions. Semantic associates of the local subject did not show priming, but associates of both non-local DPs did. The pronoun results stand against explaining the priming effects in the reflexive in terms of recency or linear proximity.

Nicol and Swinney concluded that structurally inappropriate NPs were not retrieved as potential antecedents for reflexives based on the absence of priming effects. The rest of the studies considered in this section have drawn the same conclusion from reading-time studies. Here it is useful to distinguish between two types of interference: inhibitory interference, and facilitatory interference (intrusion). Inhibitory interference occurs when the presence of multiple items in memory with overlapping cue sets cause difficulty with the processing of an otherwise grammatically-licensed element. A large majority of the previous work on interference effects in sentence processing has been concerned with finding such effects (see chapter 2). Facilitatory interference effects are of a different kind. Instances of facilitation involve eased processing of an otherwise *ungrammatical* element in the presence of a feature-matching, but inaccessible item.

3.2.1.2 Clifton, Frazier & Deevy (1999)

In one experiment[4] which was part of a larger set investigating inhibitory interference effects from feature-matching inaccessible NPs on dependency creation subroutines. Clifton, Frazier & Deevy (1999) argued that reflexive antecedent retrieval was not subject to similarity-based interference from feature-matching inaccessible items. In an eye-tracking-while-reading experiment the authors monitored the effect of an inaccessible feature-matching NP on reading times at or after a reflexive. They compared these reading times to conditions in which the feature-matching NP was absent. The authors also used identical sentences with the reflexive replaced by a proper name to give a baseline measure of complexity effects independent of antecedent retrieval. The authors divided the sentences into regions, indicated by the slashes below, and analyzed effects on a region-by-region basis.

(7) a. The son of the fireman/ hurt himself/ in a bad accident

[4] An additional experiment in Clifton, Frazier and Deevy (1999) looked at the influence of additional strucutrally-inaccessible NPs on the processing of reflexives. This experiment is not of direct relevance here as the inaccessible NPs did not match the reflexive in morphological features.

b. The son/hurt himself/ in a bad accident

c. The son of the fireman/ hurt Fred/ in a bad accident

d. The son/ hurt Fred/ in a bad accident.

In the reflexive region the authors found main effects of subject complexity and the reflexive/proper name manipulation, but an interaction indicative of inhibitory interference was not observed. Although these results are consistent with lack of interference, the study does not provide baseline conditions of non-interference against which to compare the processing of the multiple-match conditions. Therefore, the study provides only weak evidence, if any, for privileged access to grammatical antecedents.

3.2.1.3 Badecker & Straub (2002)

In a series of three experiments contained in a larger set of experiments on interference in referential processing, Badecker & Straub (2002), henceforth (B&S), show that antecedent retrieval for local reflexives does not reliably exhibit signs of interference from ungrammatical, feature-matching NPs. In Experiment 4, B&S compared RTs immediately at and following a reciprocal anaphor (*each other*) in two conditions. In the *Single-Match*, condition the local grammatical antecedent for the reciprocal bore the appropriate plural feature to license the anaphor, but the matrix subject did not. In the *Multiple-Match* condition, both matrix and local subject were plural.

(8) a. The attorney thought that the judges were telling each other which defendants had appeared as witnesses before.

b. The attorneys thought that the judges were telling each other which defendants had appeared as witnesses before.

The authors found that region-by-region comparisons between the two conditions in the regions following the reciprocal showed no significant effect of inhibitory interference from a structurally inappropriate plural NP. B&S note that collapsing across the four regions following the reciprocal yielded an overall pairwise difference between the conditions: RTs in the *Multiple-Match* condition are higher on average than in *Single-Match* conditions, which B&S suggest may be due to inhibitory interference. In their Experiment 5 B&S investigated the interference susceptibility of reflexive antecedent retrieval to non-c-commanding possessors. As in the previous study, comparisons were made between a *Single*- and *Multiple-Match* condition. These are shown below in ((9)). *Pronoun* conditions served as a baseline to establish that possessors can, in fact, serve as antecedents for pronouns, so long as no grammatical constraint blocks the coreference.

(9) REFLEXIVE SENTENCES

a. Jane thought that Bills brother owed *himself* another opportunity to solve the problem.

b. Jane thought that Beths brother owed *himself* another opportunity to solve the problem.

(10) PRONOUN SENTENCES

a. Jane thought that Bills brother owed *him* another opportunity to solve the problem.

b. Jane thought that Beths brother owed *him* another opportunity to solve the problem.

B&S found no pairwise differences between Reflexive *Single-* or *Multiple*-match conditions, suggesting no interference from the structurally inappropriate NP. Pronoun sentences displayed the expected pairwise split. A feature-matching NP possessor was retrieved as an antecedent for a pronoun, as evidenced by decreased RTs in comparison to the condition in which no feature-matching NP was available to serve as antecedent.

B&S conducted one further study to test susceptibility of reflexive retrieval to interference. In experiment 6, they tested whether feature-matching extra-clausal experiencers would induce interference. The relevant comparison is below.

(11) a. It appeared to John that Bill owed *himself* another opportunity to solve the problem.

b. It appeared to Jane that Bill owed *himself* another opportunity to solve the problem.

Once again, B&S failed to find any evidence of inhibitory interference. No significant differences in RTs emerged between the two conditions in any post-reflexive region. As with the previous study these results support privileged access to grammatically-accessible antecedents. On the whole, the results point towards categorical sensitivity to the constraint, as well. However, B&S take the slight inhibitory effects found by summing over all post-reciprocal regions to be grounds for equivocation. I am inclined to interpret those results, if at all, as late effects that do not reflect interference in early retrieval.

3.2.1.4 Sturt (2003)

Sturt (2003) investigated the sensitivity of reflexive antecedent retrieval to non-local, feature-matching antecedents using the eye-tracking-while-reading methodology. Sturt presented participants with a series of three-sentence vignettes containing 2 referents: an individual named with a proper name in the first sentence (*John, Jennifer*), and a descriptive, gender-stereotyped NP which was the subject of an embedded clause in the second sentence (*the surgeon*). A critical reflexive was introduced as the object of the embedded verb in the second sentence. Grammatically, the only licit antecedent for the reflexive was the embedded clause subject (*the surgeon*). Gender-match between the reflexive and the non-local NP, as well as between the gender-stereotyped local NP was varied. Of the studies considered so far, Sturt's is the only study whose design permitted a test of facilitatory interference.

(12) a. Accessible-match/inaccessible-match

Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked himself with a used syringe needle. There should be an investigation soon.

b. Accessible-match/inaccessible-mismatch

Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked himself with a used syringe needle. There should be an investigation soon.

c. Accessible-mismatch/inaccessible-match

Jonathan was pretty worried at the City Hospital. He remembered that the surgeon had pricked herself with a used syringe needle. There should be an investigation soon.

d. Accessible-mismatch/inaccessible-mismatch

Jennifer was pretty worried at the City Hospital. She remembered that the surgeon had pricked herself with a used syringe needle. There should be an investigation soon.

Sturt found that early measures did not show a pattern consistent with interference from nonlocal, feature-matching NPs. First-fixation and first-pass reading times were faster in the reflexive region when the reflexives gender matched the stereotyped gender of the local, grammatically licit antecedent (e.g., *surgeon . . . himself*) than when they did not (e.g., *surgeon . . . herself*). However, no significant differences were observed as a function of gender-match between the nonlocal NP and the reflexive.

Patterns that showed some effect of the inaccessible NP emerged in later reading measures, but they remain open to interpretation. Second-pass or Re-read times in the spillover regions showed elevated reading times in the condition where the nonlocal NP matched the reflexive, but the local NP did not match (in stereo-typed gender). Also, conditions where both local and non-local NP matched the reflexives were read the fastest overall.

These findings were confirmed in a second experiment, wherein Sturt manipulated the linear position of the local and nonlocal NP in relation to a critical reflexive. In the second experiment, the nonlocal, structurally inappropriate NP intervened between the grammatical antecedent and the reflexive.

(13) {John |Jennifer} was pretty worried at the City Hospital. The surgeon who treated Jennifer had pricked {himself | herself} with a used syringe needle. There should be an investigation soon.

First-fixation and first-pass measures showed the same pattern as the first experiment. Moreover, whereas late effects of gender-match were observed in the first study, there was no indication of a late-effect of gender-match of the illicit NP in second-pass reading times in experiment 2. Sturt's findings also provide support for a grammatical preference and categorical sensitivity in early retrieval.

3.2.1.5 Xiang et al. (2009)

As part of a larger comparison of intrusive licensing effects, Xiang, Dillon & Phillips (2009) compared average EEG responses to unlicensed reflexives in the presence and absence of an intrusive (inappropriate) licensor. Examples of their three experimental conditions are given below.

(14) Congruent

The tough soldier [that Fred treated in the military hospital] introduced *himself* to all the nurses.

(15) Intrusive

The tough soldier [that *Katie* treated in the military hospital] introduced *herself* to all the nurses.

(16) Incongruent

The tough soldier [that *Fred* treated in the military hospital] introduced *herself* to all the nurses.

Xiang and colleagues found no statistically significant differences between Intrusive and Incongruent conditions, suggesting that the presence of the feature-matching inaccessible name (*Katie*) did not facilitate processing of the ungrammatical reflexive.

3.2.1.6 Dillon et al. (2013) Local Reflexives/Agreement

Dillon et al. (2013) conducted a controlled comparison of susceptibility to interference between subject-verb agreement and reflexive dependencies. In their materials, the authors varied the feature-match between a structurally inappropriate NP (*manager(s)*) inside a relative clause and the tail of either a reflexive or agreement dependency. Inaccessible NPs were either singular or plural, while accessible subjects were held singular. In reflexive conditions, a downstream reflexive either matched in features with the accessible subject, or the inaccessible intervener. In agreement conditions the main verb of the sentence either agreed with the accessible subject, or with the inaccessible NP.

(17) REFLEXIVE CONDITIONS

a. The new executive who oversaw the middle manager apparently doubted *himself* on most major decisions.

b. The new executive who oversaw the middle managers apparently doubted *himself* on most major decisions.

c. *The new executive who oversaw the middle manager apparently doubted *themselves* on most major decisions.

d. *The new executive who oversaw the middle managers apparently doubted *themselves* on most major decisions.

(18) AGREEMENT CONDITIONS

a. The new executive who oversaw the middle manager apparently *was* dishonest about the companys profits.

b. The new executive who oversaw the middle managers apparently *was* dishonest about the companys profits.

c. *The new executive who oversaw the middle manager apparently *were* dishonest about the companys profits.

d. #The new executive who oversaw the middle managers apparently *were* dishonest about the companys profits.

Consistent with previous findings from Wagers et al. (2009), the authors found a reliable effect of facilitatory interference in agreement conditions. Processing of a plural verb that did not agree with the grammatical subject was eased when the intervener was plural. No such effect was found in reflexive conditions. Processing of a reflexive that did not agree in features with the local subject was not facilitated by the number-marking on the intervener. The authors interpret their effects (as do I) as indicative of not only a grammatical preference, but also as evidence that feature-matching inaccessible NPs do not interfere with retrieval.

3.2.2 Interference Susceptibility: King et al. (2012)

King et al. (2012) argued that the absence of interference in previous studies should be taken as evidence of a confound of linear order, not evidence that local anaphor licensing is always immune to interference. King et al. tested this by monitoring indices of interference when licensing anaphors in positions directly adjacent to, and linearly distant from, the main verb. King et al. employed a stereotypical gender-mismatch paradigm similar to Sturt (2003). The stereotypical gender of the grammatical antecedent for the reflexive (*bricklayer*) could either match or mismatch the reflexive. Additionally, the gender of a structurally inappropriate NP was also so manipulated. The distance between the verb and the critical reflexive was manipulated while holding the semantics of the constructions constant by using dative/benefactive alternations. In predicate-adjacent conditions, the reflexive appeared immediately following the verb as part of a double-object cluster. In predicate-separated conditions, the predicate was marked as an indirect object or benefactor by a preposition, and distanced from the verb itself by the direct object and preposition.

(19) PREDICATE-ADJACENT

The bricklayer who employed Gregory/Helen shipped **himself/herself** sacks of mortar…

(20) PREDICATE-SEPARATED

The bricklayer who employed Gregory/Helen shipped sacks of mortar to **himself/herself** ...

In a self-paced reading experiment the authors found longer RTs in post-reflexive regions RTs in conditions where the accessible antecedent did not match the reflexive in gender, but no effect of the inaccessible NPs gender. These results were found irrespective of position of the reflexive. In a subsequent eye-tracking study the authors observed differing patterns of interference dependent on the reflexives distance from the predicate with which it was associated. In predicate-adjacent reflexive conditions, the SPR pattern of results were replicated in first-fixation, gaze-duration and go-past (right-bound) reading times. There was no effect of the inaccessible NP's gender. In predicate-separated conditions, a different pattern was observed. Conditions where the inaccessible NP matched the reflexive in gender were read at speeds comparable to conditions where the grammatical antecedent matched. This suggests that the susceptibility of reflexives to similarity-based retrieval interference increases as the distance between the reflexive and the verb increases.

The data therefore indicate that categorical sensitivity observed in previous studies is a selective, rather than an absolute effect. However, they are also consistent with previous findings in that they show grammatically-accessible antecedents are distinguished from inaccessible NPs.

Study	Accessibles Distinguished?	Interference?	Anaphor Position
Nicol& Swinney (1989)	Y	N	Adjacent
Clifton et al. (1999)	Y?	N?	Adjacent
Badecker & Straub (2002)	Y	N?	Adjacent
Sturt (2003)	Y	N	Adjacent
Xiang et al. (2009)	Y	N	Adjacent
Dillon et al. (2013)	Y	N	Adjacent
King et al. (2012)	Y	N/Y	Object / Oblique

Table 3.1: Summary of Work on Local Anaphoric Licensing

3.2.3 Summary of Results

The results of the brief survey of previous studies is in Table 3.1.

The foregoing studies are in unanimous agreement that access to grammatical antecedents is privileged in comparison to inaccessible antecedents. There is also widespread agreement that object reflexive licensing is not susceptible to interference. The next section discusses a way of implementing the ability to distinguish appropriate NPs from inappropriate NPs using a feature.

Moving outside the domain of simple object reflexives, King et al.'s predicate-separated reflexive findings show that the immunity to interference is not absolute. It thus appears that categorical immunity to interference is somehow conditionalized. In light of this fact, the next section also discusses a decision process that would allow for selective application of the categorical constraint.

3.3 A Proxy Feature and Strategies for Its Use

3.3.1 Clause Index as Proxy Feature for Local Anaphor Accessibility

The most prominent account of reflexive licensing in Generative Syntax is Chomsky's (1981) *Principle A*.

(21) *Principle A*

An anaphor must be bound in its governing category.

(22) α *binds* β iff:

i) α c-commands β

ii) α is *co-indexed* with β

(23) *Governing Category*

The minimal maximal projection containing the anaphor and an accessible SUBJECT.

A feature that encodes 'Principle A' accessibility on individual items is not possible or plausible because in order to define a Principle A feature, one would need to make reference to c-command. As we saw in the previous chapter, the computational costs for encoding c-command relations online are prohibitive.

While the distribution of reflexives is often discussed in the context of Chomsky's (1981) Principle A, an alternative characterization of the constraint is widely acknowledged

(Lees & Klima 1963, Bruening 1999, a.o.): the *clause-mate* restriction.[5]

(24) *Clause-Mate Constraint*

A reflexive and its antecedent must occur in the same clause.

Unlike Principle A, the clause-mate constraint does not make reference to an item-to-item relation like c-command. An NP can be a member of a clause, irrespective of whether there are any other NPs in the clause. To implement a clause-mate constraint in a feature-based approximation strategy, we need a method for identifying clauses. If we suppose that the control state of the parser maintains a counter that increments with every successive clause, we can use this counter to provide a clause-index feature.[6] Its assignment function is:

(25) $$\textsc{Clause}(x) = \begin{cases} \mathit{current.ClauseIndex} & \text{x is DP} \\ \text{NULL} & \text{otherwise} \end{cases}$$

Once we have this CLAUSE feature in place, imposing a clause-mate constraint reduces to the trivial task of enforcing clause-match between target and probe.

It is important to note that the implementation discussed above achieves apparent sensitivity to Principle A by eschewing appeal to relational information like c-command. As noted in the past, a simple clause-mate restriction absent some notion of structural prominence (e.g. c-command) is inadequate to fully cover the observed data. For example, a simple clause-mate restriction would not be able to distinguish between the grammatical

[5] It should be noted that *governing category* most often evaluates to *the minimal clause containing the anaphor*.

[6] Alcocer and Phillips (2012) also make use of an index counter in their functions to assign proxy c-command features.

and ungrammatical sentence in (26) because in both sentences the reflexive is co-indexed with a clause-mate.

(26) a. $John_i$ likes $himself_i$.

b. $Himself_i$ likes $John_i$

Relational information is required to rule out the (b) example, where the the reflexive is not c-commanded by its antecedent. The current system does not have the representational power to enforce such a constraint.

3.3.2 Implementing Immunity to Interference: Cue Selection v. Weighting

Before discussing how selective interference effects arise, we should first discuss the mechanisms by which immunity to interference could arise under any circumstances. Immunity to interference from feature-matching inaccessible NPs is challenging for the following reason. As discussed in chapter 2, many theories assume (in some form or another) that probability of retrieval is proportional to an item's level of activation, which in turn is (partially) governed by associative feature-match. Many models (including the ACT-R model of Lewis & Vasishth 2005, which we use in later sections) assume that cues are combined *linearly*, with *equal weights*, to determine activation. In Lewis & Vasishth's model, for example, activation of chunk i at retrieval time (A_i) is computed according to the general equation below, where B_i is i's baseline activation, w_j is the weight assigned to cue j and S_{ji} is the strength of association between feature j and chunk i.

(27)

$$A_i = B_i + \sum_j w_j S_{ji}$$

If we suppose that the retrieval probe for reflexives includes morphological features of the reflexive (gender, number) in addition to the structural feature CLAUSE, we still expect partial activation of structurally inappropriate, feature-matching NPs. Take the processing of the reflexive *himself* in (28)

(28) *The cyclists [$_{Clause:2}$ that collided with the man] embarrassed *himself*.

Suppose that strength of association $S_{j,man}$ is assumed to be 1 under match, weights are kept constant across all cues, and baseline activation value of any given chunk (B_i above) is abstracted away from. Under these assumptions, the activation value of *the man* in (28) would be calculated as below.

(29)
$$\begin{aligned} A_{man} &= w_{Gender}{*}S_{Gender,man} + w_{Number}{*}S_{Number,man} + w_{Clause}{*}S_{Clause,man} \\ A_{man} &= (.33{*}1) + (.33{*}1) + (.33{*}0) \\ A_{man} &= .66 \end{aligned}$$

The only other NP in the sentence *the cyclists* would receive less activation, as it does not match on number (or gender) features.

(30)
$$\begin{aligned} A_{cyclist} &= w_{Gender}{*}S_{Gender,cyclist} + w_{Number}{*}S_{Number,cyclist} + w_{Clause}{*}S_{Clause,cyclist} \\ A_{cyclist} &= (.33{*}0) + (.33{*}0) + (.33{*}1) \\ A_{cyclist} &= .33 \end{aligned}$$

If we assume, that the NP with the highest activation value is retrieved, we expect retrieval of *the man*. We also expect that retrieval of this NP should facilitate processing of the reflexive *himself* relative to its processing in sentences with no feature-matching antecedent (see Wagers et al. 2009).

There are two methods for tackling this problem, each of which targets a different assumption.[7] The first possible response denies that morphological features are used in the cue set for retrieving reflexives. Dillon et al. (2013) advocate this approach. Jettisoning morphological cues from the retrieval structure guts their ability to interfere. Activation would therefore be driven entirely by the CLAUSE feature, permitting retrieval to ignore any non-clause-mates of the reflexive.

The second approach, which I dub the *Preferential Weighting* strategy, permits morphological cues into the retrieval structure, but vitiates their strength by assigning them less weight than CLAUSE. If cue weights are forced to sum to one, then increases in the weight of one cue proportionally decrease the weight assigned to other cues in the set.[8]

(31) SUM-TO-ONE CONSTRAINT ON WEIGHTS

$$\sum_j w_j = 1$$

The weight assignments under a preferential weighting scheme would simply need to guarantee that the activation resulting from matching on morphological features always

[7]For our purposes, I will discuss these strategies as though they were functionally equivalent. Depending on modeling decisions, however, the two can have considerably different consequences. For example, if one imposes a partial-match penalty on chunks in memory, the second option discussed provides higher activation for feature-matching NPs than non-matching NPs. The first strategy would not distinguish between matching and non-matching NPs.

[8]Ratcliff & McKoon (1995) impose such a constraint on weights in their cue-based retrieval model of lexical access.

fell below some minimum threshold for retrieval. For example, assume that retrieval preferentially weighted the *clause* cue, assigning it 90% of the weight mass and leaving 10% of the weight mass to be split among the other cues.

(32) PREFERENTIAL WEIGHT DISTRIBUTION

$w_{Clause} = 0.9$

$w_{Gender} = 0.05$

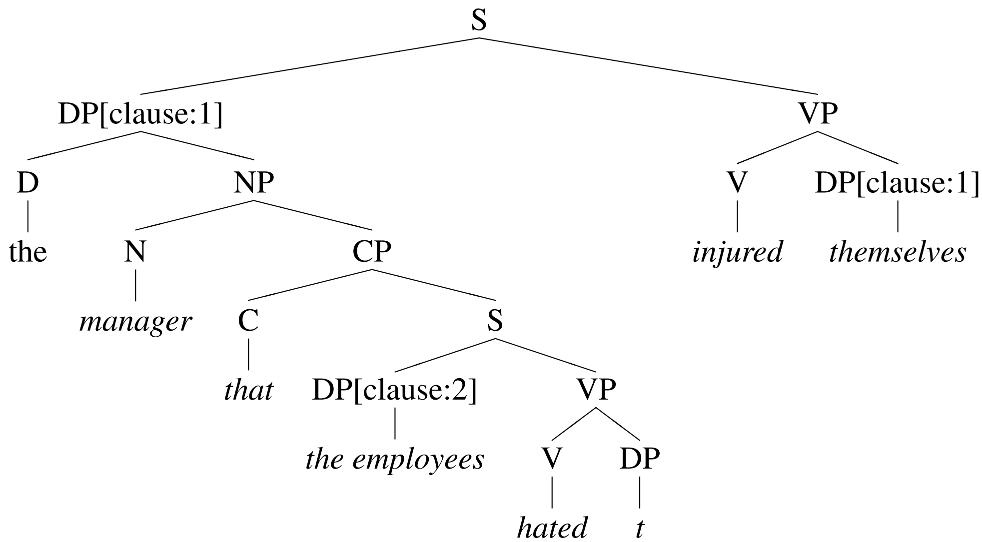
$w_{Number} = 0.05$

With this weighting in place, the activation profiles of the two NPs reverse. *The cyclist* is now the preferred target of retrieval, and *the man* is considerably less activated. In principle, the weight distribution could be further skewed to lower activation of *man* even more.

(33)
$$A_{man} = w_{Gender}{*}S_{Gender,man} + w_{Number}{*}S_{Number,man} + w_{Clause}{*}S_{Clause,man}$$
$$A_{man} = (0.05{*}1) + (0.05{*}1) + (0.9{*}0)$$
$$A_{man} = .1$$

(34)
$$A_{cyclist} = w_{Gender}{*}S_{Gender,cyclist} + w_{Number}{*}S_{Number,cyclist} + w_{Clause}{*}S_{Clause,cyclist}$$
$$A_{cyclist} = (0.05{*}0) + (0.05{*}0) + (0.9{*}1)$$
$$A_{cyclist} = .9$$

Now to illustrate how these strategies produce sensitivity when the reflexive is verb-adjacent. Consider the syntactic structure of a sentence predicted to engender interference effects using morphological cues. In the structure below, the reflexive *themselves* requires a plural antecedent in a structurally appropriate position. Suppose that the matrix subject (a structurally appropriate, but feature-mismatching, antecedent for the reflexive) bears the diagnostic feature [clause:1].



If the probe provides the cue [+plural] for retrieval, and the weights of cues are equal, the plural embedded subject is predicted to interfere. Similarly, interference would arise if the retrieval probe included structural cues as [subject], because these cues are shared by both appropriate subject and inappropriate subject.

However, using [clause:1] as the sole cue, or preferentially weighting the cue, would retrieve the grammatically-accessible antecedent - despite it not bearing the right morphological features. Little to no observable interference from the embedded NP *employees* is expected.[9]

3.3.3 Linear Adjacency: Context-Dependent Cue Selection/Weighting

Above we saw two strategies for guaranteeing privileged and exclusive retrieval of the grammatically-accessible subject through preferential use of the diagnostic clause

[9] Note, that this explanation also extends to cover the finding from Sturt (2003) that extra-clausal DPs do not interfere with the retrieval of a reflexive. Once again, the structurally illicit feature-matching NP would not bear the right clause feature, so it would not be retrieved. Moreover, a strict clause-based retrieval strategy as a first pass for finding antecedents for reflexives can easily accommodate the lack of interference from non-local feature-matching NPs in licensing the chinese reflexive *ziji* (Dillon et al. submitted)

feature. I now discuss how these strategies might be influenced by verbal adjacency.[10]

Suppose that preferential use of a cue is conditioned on its diagnosticity: its ability to uniquely identify an appropriate antecedent. We saw above that use of the clause feature is entirely diagnostic for retrieving the grammatically-accessible antecedent of an anaphor when the anaphor is verb-adjacent. When a reflexive is distanced from the verb, however, the diagnosticity of the *clause* feature often decreases. For example, if an object intervenes between subject and reflexive, as it does in (35), the object should also bear the *clause*

---

[10]The explanations for interference resistance above amended traditional assumptions concerning cue generation or cue-combination. There are, in principle, two ways in which verbal proximity could ensure privileged access to the structurally-accessible subject without resorting to changes to the system. First, an explanation could be given in terms of the structurally-accessible subject's *baseline activation*. Second, an explanation could be contrived that avoids retrieval entirely (and therefore interference). I discuss each below.

Baseline Activation

A chunk's activation is determined by two terms: an activation boost for matching items in memory, and *i*'s baseline activation (B_i) at the time of retrieval.[11]

- $B_i = ln[\sum_m t_m^{-d}]$

B_i's contribution to the activation equation is to increase the probability that a previously-retrieved item will be subsequently retrieved. An account of the lack of interference for verb-adjacent reflexives could use baseline activation in the following way to suppress the contribution of interference-prone morphological features.

If baseline activation of the grammatical subject were high enough, it would be reliably retrieved, irrespective of feature-match. The more recent the last retrieval of a particular memory chunk, the higher its baseline activation level. Grammatical fidelity in retrieval varies as a function of the appropriate antecedent's level of baseline activation. Residual activation of the subject from its recent retrieval by the verb for purposes of integration guarantees that the grammatically licit subject is the most highly activated NP immediately following the verb. Since the reflexives in question trigger a retrieval directly after the verb, the licit subject is most likely to be picked. If however, distance between the verb and the reflexive were increased, and the residual activation was thereby allowed to drop, accuracy of retrieval would be predicted to decline.

No Need for Retrieval

A second explanation that uses verbal proximity to privilege access to the local subject does not involve baseline activation. It has been argued by a number of psychologists of memory that participants possess a privileged window of processing called the *focus of attention* (see e.g., Cowart 2004, McElree 2006, Wagers and McElree 2009, a.o.). This subset of memory holds items currently being processed, and allows rapid access to its information. Retrieval involves calling items from long-term storage into focus of attention, where they can be integrated and processed with current words or phrases. Focus of attention is often thought to contain the last element processed (or some small number of recently processed items). If we suppose that the local subject occupies the focus of attention after being recently retrieved for subject-verb integration, it should remain in focus of attention when the reflexive is input.[12] If the local subject is in focus of attention, it can be immediately considered, requiring no retrieval of a potential antecedent from long-term storage - thereby obviating the need to use interference-prone retrieval cues.

feature.

(35) John$_{[clause:1]}$ bought [the sacks of mortar]$_{[clause:1]}$ for himself.

If we suppose that the cue-generation process underlying reflexive retrieval always selects the *maximally diagnostic* cue set, it would exhibit a preference for *clause* features when no other clause-mates exist, but not otherwise. For the sake of simplification, assume that verbal-adjacency correlates perfectly with the presence or absence of other clause-mates.[13] We could conditionalize preferential use of the cue on linear adjacency to reflect this strategy.

(36) DECISION STRATEGY FOR REFLEXIVE ANTECEDENT RETRIEVAL (EXCLUSIVE CUE SELECTION):

```
If reflexive is verb-adjacent:
    Probe ⇝ [clause]
Else:
    Probe ⇝ [clause, morphological features]
```

(37) DECISION STRATEGY FOR REFLEXIVE ANTECEDENT RETRIEVAL (WEIGHTING):

```
If reflexive is verb-adjacent:
    Weight [clause] disproportionately higher than [morphological
```

[13] This assumption is patently wrong, witness:

- The cyclist won't stop talking to himself.

```
        features].
    Else:
        Weight [clause], [morphological features] equally.
```

3.4 SOV Local Anaphors: A Primer on Hindi Reciprocals

Strategic use of the diagnostic feature strategy conditioned on linear adjacency alone predicts that when a local anaphor does not immediately follow its verb morphological features in the retrieval probe should engender interference. This therefore predicts that, all things equal, all local anaphoric licensing in SOV languages should be susceptible to interference because they uniformly precede their verbs. I show that this prediction is not borne out in the coming section by looking at the processing of Hindi reciprocals. Before doing so so, it is necessary to briefly discuss the syntax of Hindi and its local reciprocals.

Hindi is an SOV language internal to the clause. In unmarked word order, a verb is preceded by all of its associates (arguments, or adjuncts).

(38) John Mary-ko pyaar kartaa hai.
John Mary-ACC love does is
'John loves Mary.' S-O-V

(39) John Mary-se naach rahaa hai.
John Mary-WITH dance PROG is
'John is dancing with Mary' S-ADJUNCT-V

Hindi reciprocals are similar to English reflexives because, apart from word order variation, they follow licensing conditions similar to local anaphors in English. They must be bound within their minimal clause (40). Their antecedent must also be a plural NP.

(40) a. *The kids* hit *each other*.

b. BaccoN$_i$-ne ek-dusre$_i$-ko maar-aa.
kids-ERG one-another-ACC hit-PFV
'Kids hit each other.'

(41) a. **The kids* said that Mary hit *each other*.

b. *BaccoN$_i$-ne kah-aa ki Mary-ne ek-dusre$_i$-ko maa-raa.
kids-ERG say-PFV C Mary-ERG one-another-ACC hit-PFV
'The kids said Mary hit one another.'

Reciprocals in Hindi, just as local anaphors in English, cannot be bound by non-c-commanding items.

(42) a. *At [*the kid*s' party], John saw each other.

b. *John-ne *baccoN*-ki parTi me *ek-dusre*-ko dekhaa.
John-ERG kids-GEN party in one-another-ACC saw
'John saw each other at the kids' party.'

The restriction above cannot be attributed to subject-orientation on the part of the reciprocals, because reciprocals can be bound by non-subjects that c-command them. In (43), the ACC-marked object *thaliyoN* can bind the reciprocal *ek-dusre*, because it c-commands the reciprocal (as read off linear order). Bhatt and Dayal (2007) also note that scrambling of objects can also license binding of a reciprocal - so long as the object is scrambled to a c-commanding position (44).

(43) Bacce-ne thaaliyoN$_i$-ko ek-dusre$_i$ -ke-upar rakhaa.
kids-ERG plates-ACC one-another atop set
'The kids set the plates on top of each other.'

(44) [Anu aur Ramaa]-ko$_i$ [ek-dusre-ke$_i$ baccoN]-ne dekh-aa.
Anu and Ramaa-ACC one-other-GEN kids-ERG see-PFV
'[Each others]$_i$ kids saw [Anu and Ramaa]$_i$. .

Unlike English reflexives, but very similar to English reciprocals, Hindi reciprocals enjoy a wider distribution of environments where they may appear: they are not as strictly limited to argument positions of verbs. Like English reciprocals, they may appear as possessors (45).[14]

(45) Bacce-ne aadmiiyoN$_i$-ko ek-dusre$_i$ kii tasviir dikhaa-yaa.
kids-ERG men-DAT one-another GEN picture show-PFV
'The kids showed the men each other's photos.'

Thus, Hindi reciprocals provide a minimal test dependency for comparison with English reflexives to investigate the role of verbal adjacency in guiding antecedent retrieval for local anaphora. If the appearance of immunity to similarity-based interference observed

[14] In English reciprocal possessors have been argued to be logophors (Pesetsky 1995, Pollard & Sag 1992, Janke & Neeleman 2009, Drummond & Kush in press).

- a. *John and Mary hoped that the psychologist would explain their weaknesses to each other.
  b. John and Mary hoped that the psychologist would explain each others weaknesses to them. (Janke and Neeleman 2009, 37, (90))

Hindi reciprocals do not pattern like English reciprocals in this regard. They do not permit logophoric readings in possessor positions.

- a. *John aur Mary-ne kahaa thaa ki psychologist ek-dusre-ko photo dikhaa degaa.
  John and Mary-ERG said was that psychologist each-other-DAT photo show give-FUT
  'John and Mary said that the psychologist would show the photo to one another.'
- *John aur Mary-ne kahaa thaa ki psychologist ek-dusre-ki photo khiinchegaa.
  John and Mary-ERG said was that psychologist each-other-GEN photo take-FUT
  'John and Mary said that the psychologist would take each other's photo.'

in English is parasitic on verbal adjacency, we should expect similarity-based interference effects to emerge in the processing of Hindi reciprocals.

3.5 Experiments

This section assesses whether selective interference effects are conditioned on linear adjacency. Experiment 1 presents the results from two computational simulations that model the interference effects predicted in Hindi reciprocal licensing if selective interference were dependent on linear order or if linear order is not the determinant of the effect. Experiment 2 provides behavioral results from a self-paced reading experiment on the processing of Hindi reciprocals which the model predictions are compared to.

3.5.1 Materials

Materials manipulated the feature-match between two NPs and a subsequent critical reciprocal *ek-dusre*. The first NP, the *main subject* of the sentence, was a c-commanding grammatically appropriate potential antecedent for the reciprocal. The second NP, the *intervener*, was a non-c-commanding NP positioned between the main subject and the reciprocal. 24 test items were arranged in a 2x2 factorial design crossing factors GRAMMATICALITY and INTERVENER. In *Grammatical* conditions the main subject bore [plural] marking, and was therefore a grammatical antecedent of the reciprocal. In *Ungrammatical* conditions the main subject was singular.[15] The factor INTERVENER had two levels: *Intervener* and *NoIntervener*. In *Intervener* conditions, the intervener matched the reciprocal in

[15] Plurality was unambiguously signaled by the plural objective case ending *oN*

number features. In *NoIntervener* conditions the intervener was marked singular.

To create the conditions under which retroactive interference could obtain and to maintain consistency with the design of previous studies the potential interferer was placed between the main subject and the reciprocal. To rule out one of the possible confounds related to linear proximity, the design also had to ensure that reactivation of the grammatical subject was not triggered before the reciprocal.

The structures used in the experimental materials differed from those used in most studies investigating interference from structurally inaccessible NPs in one important regard. In studies such as Dillon et al (2013), structurally inaccessible NPs are positioned in RCs attached to the matrix subject. Such a design is not possible in Hindi, where RC-modification of subjects is highly marked.

(46) *?Aadmii [jis-ko mai-ne dekhaa] dukaan jaa rahaa thaa.
man-ACC CORR.OBL 1sg-ERG see-PFV store go PROG was
'The man I saw was going to the store.'

(47) ?*Aadmi [jo paniir khaataa hai] daal bhii khaataa hai.
man CORR paneer eat-IMPF is lentils also eat-IMPF is
'The man who eats cheese also eats lentils.'

Instead, use of a correlative construction is required for subject modification. (See Bhatt 2003 and references therein for discussion of the correlative in Hindi).

(48) Jo aadmi paniir khaa-taa hai, vo daal bhii khaa-taa hai.
Which man paneer eat-IMPF is 3sg lentils also eat-IMPF is
'The man who eats cheese, he also eats lentils.'

(49) Jis aadmi-ko mai-ne dekhaa, vo aadmii dukaan jaa rahaa thaa.
Which.OBL man-ACC 1sg-ERG see-PFV 3sg man store go PROG was
'The man I saw, he was going to the store.'

Correlative constructions are unsuitable for the purposes of the present study because of the linear order of NPs. In correlatives the modifying correlative clause precedes the matrix clause, which contains a demonstrative pronoun coreferential with the individual introduced in the correlative phrase. The potential interferer, introduced inside the correlative phrase, would precede the pronoun coreferent with the grammatical subject. On the assumption that the demonstrative pronoun would reactive its associate (the head of the correlative), the grammatical subject would be the most recently activated NP before the reciprocal was encountered, thus increasing the likelihood of retrieval of the grammatical antecedent and reducing the potential for interference. The interferer must fall between the subject and the reciprocal. Given that Hindi is an SOV language, the number of potential host sites for the Interferer-introducing RC is increased.

The intervener was embedded inside a possessor contained within a locative phrase that preceded the critical reciprocal. Our design made use of the Hindi *wala* construction as the embedding environment. In Hindi, a nominalizing suffix *wala* can be right-adjoined to an infinitival verb phrase, to create a NP meaning *one who Xes*.

(50) samosa khaa-na
somosa eat-INF
'to eat (a) samosa.' →

(51) samosa khaa-ne wala
somosa eat-INF wala
'the samosa-eating one.'

These constructions can adjoin to common nouns, functioning as complex adjectives. In this regard, they are similar to pre-nominal infinitival RCs. For the purposes of this paper, I henceforth refer to them as *wala*-RCs.

(52) [t_i samosa khaa-ne wali] nurse_i
samosa eat-INF *wala* nurse
'The samosa-eating nurse/ the nurse who is eating samosa'

The example above shows that the infinitival verb in the *wala*-RC licenses objects (*samosa*). In the current study potential interferers are introduced as the object of the infinitival verb. The *wala*-RC was attached to a possessor NP internal to a pre-reciprocal locative phrase.

(53) [t_i samosa khaa-ne wali] nurse_i-ke steshan me
samosa eat-INF *wala* nurse-GEN station in
'In the samosa-eating nurse's station.'

Embedding the potential interferer inside a *wala*-RC, which was itself embedded inside a genitive possessor of a locative phrase resulted in the interferer being multiply-embedded, ensuring that it could not c-command the reciprocal.

(54) [Subj] [[RC Int] AdvP] [reciprocal] [args] V

Example materials for the experiments are given below. I also include a syntactic tree that illustrates the assumed structural relations that obtain between relevant elements in the sentence, glossed in English for convenience.

(55) a. *Grammatical-NoIntervener*

DoctoroN-ne mariz-ki dekhbaal-karne wali nurse-ke steSan me
Doctors-ERG patient-GEN care-doing wala nurse-GEN station in
ek-dusre ke-bare-me gupt-rup-se baat kii.
one-another about secretly chat did

'The doctors talked about each other in the station of the nurse caring for (the) *patient*.'

b. *Grammatical-Intervener*

DoctoroN-ne marizoN-ki dekhbaal-karne wali nurse-ke steSan me
Doctors-ERG patients-GEN care-doing wala nurse-GEN station in
ek-dusre ke-bare-me gupt-rup-se baat kii.
one-another about secretly chat did

'The doctors talked about each other in the station of the nurse caring for (the) *patients*.'

c. *Ungrammatical-NoIntervener*

Doctor-ne mariz-ki dekhbaal-karne wali nurse-ke steSan me
Doctor-ERG patient-GEN care-doing wala nurse-GEN station in
ek-dusre ke-bare-me gupt-rup-se baat kii.
one-another about secretly chat did

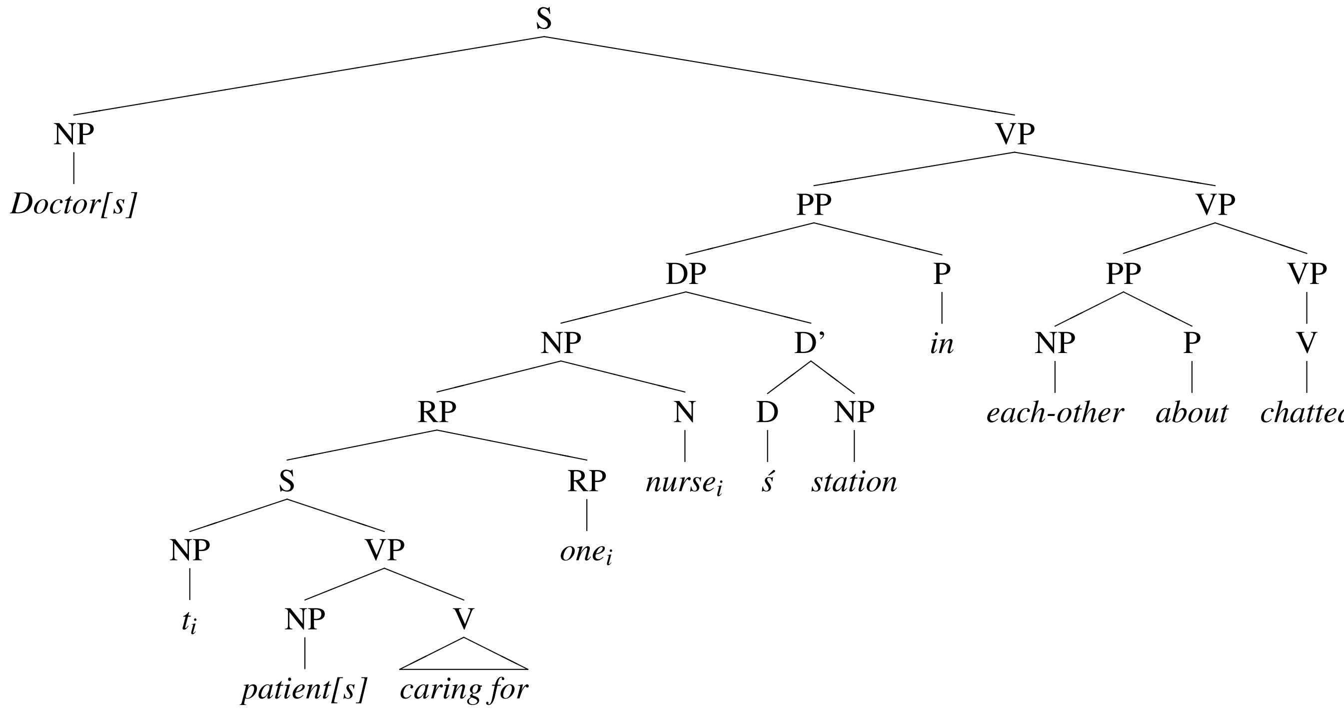
'*The doctor talked about each other in the station of the nurse caring for (the) *patient*.'

d. *Ungrammatical-Intervener*

Doctor-ne marizoN-ki dekhbaal-karne wali nurse-ke steSan me
Doctor-ERG patients-GEN care-doing wala nurse-GEN station in
ek-dusre ke-bare-me gupt-rup-se baat kii.
one-another about secretly chat did

'*The doctor talked about each other in the station of the nurse caring for (the) *patients*.'

(56)



Avoiding Potential Confounds

One concern that arises given the use of plural anaphors relates to the possibility of split antecedence. Plural anaphoric pronouns are capable of taking aggregates of previously-seen referents as their antecedents. These referents can be in distinct syntactic positions. This is seen in the example below. The plural pronoun *them* can pick out the set comprised of *John* and *Mary*, even though the two are introduced separately.

(57) John$_i$, who Mary$_j$ secretly met behind the dumpster, was worried that someone had seen them$_{i+j}$

If reciprocals were able to participate in split reference, it might be possible to coerce a plural referent for the reciprocal even when the grammatical subject were singular. This would mean that the *Ungrammatical-NoInterference* condition would not serve as a a good baseline of retrieval failure. Fortunately, split antecedence is not available for *ek-dusre*, as shown in ((58)). Two singular clause-mates that precede the reciprocal cannot serve as a joint antecedent.

(58) *?John-ne Mary-ko ek-dusre-ko dikhaa-yaa.
John-ERG Mary-ACC on-another-DAT show-PFV
'John showed Mary to each other.'

The above concerns relate to the possibility of finding a grammatical antecedent for the reciprocal other than the matrix subject. We must also show that the intervening NP whose number-marking is manipulated has the appropriate feature composition to interfere. That is, we must show that it has the right features to be an antecedent for the reciprocal, were it in a structurally appropriate position. One way in which the potential intervener in the materials differs from the matrix subject is that interferers were either accusative-marked, or genitive-marked. If NPs marked with these features were prohibited from serving as antecedents uniformly, the design would be confounded. It can be shown, however, that both accusative- and genitive-marked NPs can bind anaphors, so long as they are in the appropriate structural configuration.

Hindi reciprocals are not subject-oriented. They can be bound by objects, so long as the objects precede, c-command, and are sufficiently local.

(59) John-ne thaaliyoN-ko ek-dusre ke baraabar rakh-aa.
John-ERG plates-ACC one-another GEN equal place-PFV
'John set the plates next to one another.'

Thus, accusative case-marking on the potential interferer does not hinder its ability to interfere.

Genitive-marked NPs can also antecede a local anaphor when they c-command it (see Mohanan 1994 for further discussion of Genitive subjects).

(60) Un-ka ek-dusre-ko dekhnaa ka irada thaa.
They-GEN one-another-ACC see GEN desire was.
'They wanted to see each other. (lit. Theirs was a desire to see each other.)'

It would thus appear that Case-marking alone cannot rule out consideration of the potential intervener. It seems plausible that the only cue for retrieval is the plural number feature. We may therefore set aside concerns about potential confounding factors in the design.

3.5.2 Experiment 1: Modeling Hindi Reciprocal Licensing

Experiment 1 furnishes predictions of retrieval interference based on the implementation of two distinct retrieval strategies. The first model, the *equal weight* model, provides the predicted retrieval profiles if the features CLAUSE and *plural* are combined equally to retrieve the antecedent of a reciprocal in Hindi. The second model, the *diagnostic cue* model, provides the profiles predicted if only the diagnostic *clause* feature is used to retrieve the antecedent. This second model can be used as the guide for what to expect if the strategies to resist interference discussed in the previous section are not conditionalized on verbal adjacency. These results provide a general baseline against which to compare the behavioral results from the self-paced reading experiment reported later.

3.5.2.1 Background: The Model

The model used for the experiment is a modified version of Lewis & Vasishth's (2005) ACT-R model of sentence processing. ACT-R is a general cognitive architecture that has been used to model a wide range of phenomena and behavior in cognitive psychology (Anderson 1990). The model proposed by Lewis & Vasishth is a fully-implemented model that specifies more than just a retrieval component. The model used here focuses solely on retrieval latencies, abstracting away from the other aspects of Lewis and Vasishth's (2005) parsing model.[16]

The view of memory access here is a rational one, in the technical sense (Anderson 1989; Anderson & Milson 1989; Anderson & Schooler 1991). It assumes that retrieval adopts a strategy that will minimize retrieval error in the limit, given the nature of the task. It also assumes a proportional, monotonic relation between retrieval latencies and behavioral measures of processing time (in this case RTs - Anderson & Milson 1989). As mentioned above, for every retrieval, the model chooses the chunk with the highest activation as the winner.

Activation of a chunk i, A_i, is given by the equation in (61). B_i is chunk i's baseline activation. S_{ji} is the strength of association between cue j and chunk i. PM_{ki} in the equation below is a term that penalizes partial matches. The term ε introduces stochastic noise.

[16]Other components include a left-corner parser and a module devoted to lexical access.

(61)

$$A_i = B_i + \sum_j w_j S_{ji} + \sum_k PM_{ki} + \varepsilon$$

S_{ji} is calculated according to the equation in (62), where S is a parameter specifying the maximum strength of association allowed.[17]

(62)

$$S_{ji} = S - ln(fan_j)$$

The baseline activation of a chunk i, B_i, is determined according to the function (63).

(63)

$$B_i = ln[\sum_m t_m^{-d}]$$

Above, d is the decay rate of a chunk's activation in memory. Activation values for all chunks in memory are calculated according to the probe's features and their baseline activation. The chunk with the highest activation will have the shortest retrieval latency (T_i), as calculated according to the equation below. The chunk with the shortest retrieval latency is the chunk that is retrieved in simulations.

(64)

[17]The fan effect $ln(fan_j)$ is a term that reduces the strength of activation between item i and of a given feature j proportional to the number of items in memory bearing feature j (Anderson 1974; Anderson & Reder 1991).

$$T_i = Fe^{-A_i}$$

The equations above contain a number of free parameters whose settings could impact the results of the simulation. For ease of implementation, and to remain consistent with prior modeling work, we adopted the default parameter settings assume in previous models (Anderson & Lebiere 1998; Lewis & Vasishth 2005). The decay rate d was set to 0.5, the associative strength S to 1.5, and the latency factor to 0.14.

I report two measures of interest, generated by the model: predicted rate of error and predicted retrieval latency. Rate of error refers to the proportion of simulations in which the retrieval initiated by the probe (the reciprocal) did not return the grammatical subject. In the model, the chunk with the highest activation level is the one that is retrieved, so if a some chunk has a higher activation level than the grammatically appropriate chunk, then that chunk is retrieved. The second measure I report is the amount of predicted interference, which represents the difference between estimated retrieval latencies between *Intervener* and *NoIntervener* conditions within GRAMMATICALITY. I report the difference in latencies between the two *Ungrammatical* conditions and between the two *Grammatical* conditions. With these measures we get an estimate of the effect of manipulating the plural-marking on the intervener.

A retrieval schedule for relevant NPs was hard-coded in the simulations. In the simulations, the reciprocal functions as the retrieval probe. The retrieval schedule used assumes that no retrievals precede the critical retrieval at the reciprocal. The lack of a prior retrieval history in this particular implementation may at first seem at odds with the general assumption of the ACT-R parser that retrieval is a frequent process that underlies many cases of structural attachment. I opted not to specify a history of prior attachment retrievals, however, for the following reason. Retrievals are generally assumed to coincide with incidence of structural attachment (Lewis & Vasishth 2005, Lewis, Vasishth & Van Dyke 2006). In constructing a phrase, prior chunks associated with that phrase, or the local syntactic context into which it is being integrated, are retrieved. In the example sentences, no structure building operations occurring in the region intervening between the main subject and the reciprocal requires retrieval of the subject. Presumably, once the subject chunk has been constructed, the only chunks constructed before the reciprocal are the pre-nominal RC and the locative adjunct. In the intervening region, the only NPs that are retrieved are the intervener (*patient(s)*) and the head of the RC (*nurse*). Since retrievals increase activation, intermediate retrievals in test sentences should only serve to strengthen, or raise the baseline activation of the non-target NPs: the intervener, or the RC-head, which would simply increase their probability of retrieval. By excluding intermediate retrievals, we actually reduce the probability of interference and misretrieval. This means that the results reported below represent, if anything, a conservative estimate of the magnitude interference.

One run of the model takes into account the outcome of 1000 Monte Carlo simulations. Each simulation output the most probable chunk at retrieval time and the latency of retrieving that chunk. Error was calculated by observing how many times the correct chunk was retrieved. Retrieval latencies across simulations were averaged by condition.

	NoIntervener	*Intervener*
Grammatical	.01	.07
Ungrammatical	.07	.32

Figure 3.1: Experiment 1: Model Predicted Error Rates for Equal Weight Retrieval

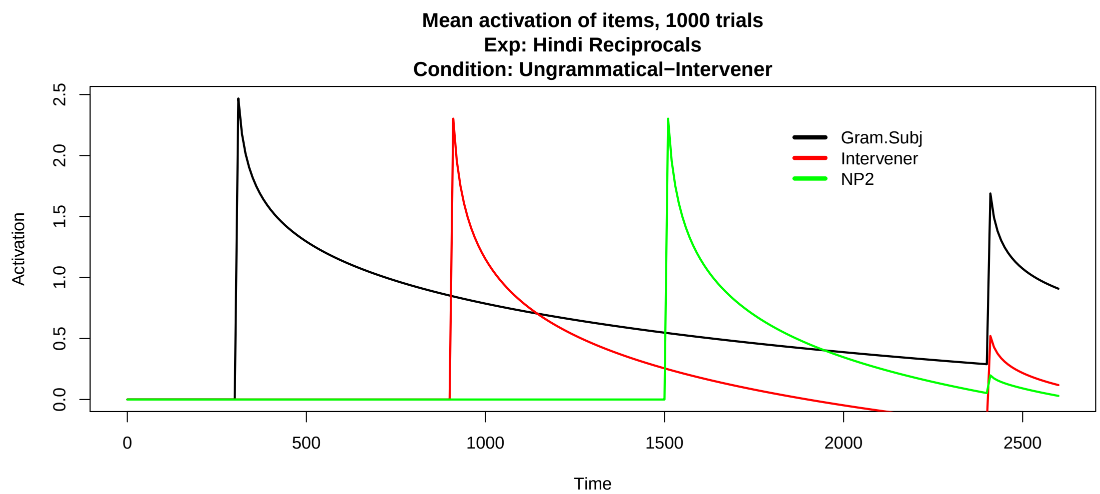


Figure 3.2: Activation Profile for *Ungrammatical-Intervener* Condition with Equal Weights.

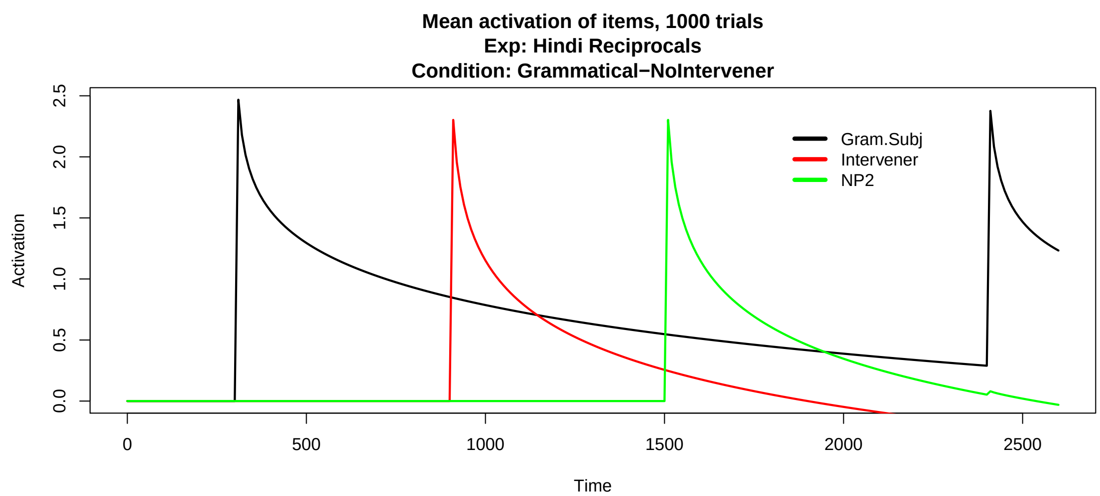


Figure 3.3: Experiment 1: Activation Profile for *Grammatical-NoIntervener* Condition with Equal Weights

3.5.2.2 Results

Equal Weight Model

A slight pairwise difference in error-rate is observed between *Grammatical* conditions. Error was lowest overall in the *Grammatical-NoIntervener* condition, reflecting relatively consistent access to the feature-matching, grammatical subject. In the *Grammatical-Intervener* condition, error rates are numerically higher, reflecting a slightly elevated chance of retrieving the inappropriate intervener. A larger pairwise difference is observed between *Ungrammatical* conditions. In the *Ungrammatical-NoIntervener* condition, error rates are low, comparable to rates in *Grammatical-Intervener* conditions. The *Ungrammatical-Intervener* condition displays significantly higher rates of erroneous retrieval.

Figure 3.2 presents the activation values of each individual NP in the sentence plotted over time averaged over 1000 simulations for the *Ungrammatical-Intervener* condition. Spikes in activation values correspond to points when an NP was accessed (either because it was created at that point, or because it was retrieved and therefore reactivated). The retrieval coinciding with the reciprocal was scheduled around 2400ms. The height of each individual spike in the plot at this point provides a rough estimate of the likelihood that each NP would be retrieved at that time. Figure 3.2 shows that although the structurally appropriate antecedent bears an overall advantage in that it is most likely to be retrieved, other NPs are still activated at the time of retrieval or are retrieved on some portion of model runs. Compare the activation profiles in figure 3.2 to those in figure 3.3, which plots activation values for the *Grammatical-NoIntervener* condition. No interference is observed, as evidenced by the lack of activation spikes for either non-target NP at retrieval

	Interference Effect
Grammatical	33
Ungrammatical	-124

Table 3.2: Model Predictions: Interference Effects (in ms.) for Equal Weight Retrieval.

	NoIntervener	*Intervener*
Grammatical	.03	.03
Ungrammatical	.04	.05

Table 3.3: Experiment 1: Model Predicted Error Rates for Diagnostic Cue Retrieval.

time.

A large facilitatory interference effect is predicted between *Ungrammatical* conditions. Misretrieval of the partially matching plural intervener in the *Ungrammatical-Intervener* condition is expected to drive faster retrieval latencies in that condition, relative to the condition in which no feature-matching antecedent is present for the reciprocal. Facilitation is predicted because the greater degree of feature-match between intervener and probe results in higher absolute activation of the interferer, which in turn speeds retrieval latency. In *Ungrammatical-NoIntervener* conditions, on the other hand, the intervener is just as likely to be retrieved - its activation relative to the activation of the matrix subject is still higher, but its absolute activation level is lower, because it matches the probe on fewer features. A slight inhibitory effect is predicted between *Grammatical* conditions.

Diagnostic Cue Model

Error rates are consistently low and comparable across all conditions reflecting the fact that the clause feature reliably retrieves the structurally appropriate subject and no morphological features are present to interfere with the process.

	Interference Effect
Grammatical	5
Ungrammatical	0

Table 3.4: Model Predictions: Interference Effects (in ms.) for Diagnostic Cue Retrieval.

The model predicts that there should be no observable effect of interference.

The exact differences in latency should not be over-interpreted or taken to reflect exact values that the model is committed to in the behavioral data. What is more important is the relative, proportional differences in predicted interference effect. Here the model's predictions are very clear. An effect of facilitatory interference is expected within *Ungrammatical* conditions. *Ungrammatical-Intervener* conditions are expected to exhibit lower retrieval latencies, and hence, lower RTs in the region indexing retrieval relative to *Ungrammatical-NoInterference* conditions. This is driven by higher rates of interferer-retrieval. Within *Grammatical* conditions, the facilitatory interference effect is expected to be non-existent, or a slight amount of inhibition is expected. It is important to note that lower retrieval latencies are expected in the *Grammatical-Interference* case for the same reason that facilitation is expected in the *Ungrammatical-Intervener* condition: erroneous retrieval of the interferer.

3.5.2.3 Discussion

The ACT-R model predicts the use of morphological cues in an equally-weighted retrieval probe should engender facilitatory interference in *Ungrammatical-Intervener* conditions, relative to *Ungrammatical-NoIntervener* conditions. When diagnostic cues are used exclusively, on the other hand, no interference from inaccessible feature-matching

NPs is predicted.

3.5.3 Experiment 2: Hindi Reciprocal Licensing

Experiment 2 tested whether structurally inappropriate feature-matching NPs interfere with reciprocal antecedent retrieval. The results were then compared with the predicted interference profiles for both the diagnostic cue and equal weight models furnished by the simulations above.

3.5.3.1 Participants

32 self-reported native speakers of Hindi were recruited from the student bodies of IIT, Delhi and Jawaharlal Nehru University in New Delhi. Participants were compensated Rs. 300 for their participation.

3.5.3.2 Procedure

Participants were run on one of two laptop PCs using the Linger software package (Doug Rohde, MIT) in a self-paced word-by-word moving window paradigm (Just, Carpenter, & Woolley, 1982). Each trial began with a sentence masked by dashes appearing on the screen. Letters and punctuation marks were masked, but spaces were left unmasked so that word-boundaries were visible. As the participant pressed the spacebar, a new word appeared on the screen as the previous one was re-masked by dashes. All text appeared in Devanagari font.

A comprehension question followed each sentence. Participants were instructed to read sentences at a natural pace and to respond to the comprehension questions as accurately as possible. Participants responded to questions using the f-key for 'yes' and the j-key for 'no'. If the question was answered incorrectly the word *galat* ('incorrect/wrong') appeared briefly in the center of the screen. Each participant was randomly assigned to one of the lists and the order of the stimuli within the presentation list was randomized for each participant.

3.5.3.3 Analysis

Data from 5 participants whose mean accuracy on comprehension questions fell below 60% were excluded from further analysis and data from one participant were excluded due to failure to comply with experimental guidelines. This resulted in the data of 26 subjects being used for later analysis. In the remaining data, RTs that fell above an upper threshold of 3000ms, or below a minimum threshold of 100ms were excluded from analysis. This resulted in an average of 1.8% data loss across conditions.

Statistical analyses were carried out using linear mixed effects regression (Baayen, Davidson & Bates 2008). Experimental fixed effects were the factors GRAMMATICALITY (whether the matrix subject was plural or singular) and INTERVENER (whether the NP embedded inside the *wala*-RC matched the reciprocal in features, or did not), and their interaction. A simple difference sum coding scheme was used for coding fixed effects (+0.5/-0.5 for the levels *Ungrammatical/Grammatical* within GRAMMATICALITY, and +0.5/-0.5 for the levels *NoIntervener/Intervener*). Reported coefficients reflect the magnitude of the difference between levels of a given factor in milliseconds. P-values were

	NoIntervener	*Intervener*
Grammatical	.70	.72
Ungrammatical	.73	.71

Table 3.5: Experiment 2: Comprehension Question Accuracy

estimated by MCMC sampling using the LanguageR() package (Baayen 2008).

Sentences were split into the following regions:

(65) /{That | Those} Doctor(s) ERG / Patient(s) ACC / care-giving nurse / GEN station inside / each-other / of / with / secret / manner in / chat/ did /.

3.5.3.4 Results

Comprehension question accuracy averaged 71.4%.

No significant differences were found in average accuracy rate across conditions (logistic mixed effects model, all zs < 1).

Reading Time Results

Pre-reciprocal region:

No significant effects were found in the pre-reciprocal region (all ts < 1).

Reciprocal region:

No significant effects were found immediately in the reciprocal region.

Post-reciprocal genitive region:

A significant main effect of GRAMMATICALITY was observed in the region immediately following the reciprocal (β = -46.01, s.e. = 20.79, t = 2.2, p $<$ .05), with no effect of Interference ($t < 1$). This was driven by shorter RTs in the *Grammatical* conditions, relative to *Ungrammatical* conditions. Despite the fact that there was a numerical difference between the two *Grammatical* conditions, the lower RTs in the *Grammatical-Intervener* condition did not drive a significant pairwise difference ($t < 1$). No significant differences were observed between *Ungrammatical* conditions.

Reciprocal+2 region:

No main effects achieved significance, but a GRAMMATICALITY x INTERFERENCE interaction was marginally significant ($\beta = 52.55$, s.e. = 31.29, t = 1.68, $p < .10$), driven by elevated RTs in the *Ungrammatical-Intervener* condition. The pairwise comparison between *Ungrammatical* conditions was also marginal (t = 1.86, $p < .10$). These effects may reflect increased difficulty in the *Ungrammatical-Intervener* condition, a numerical effect towards which there was a trend in the previous region.

Reciprocal+3-5 region:

No significant effects were observed (all ts < 1).

Final verb region:

A main effect of GRAMMATICALITY was observed in the sentence-final verb region (β=-77.51 ,s.e.= 43.48, t= -1.78, $p < .10$), with *Grammatical* sentences being read faster than *Ungrammatical* sentences. This suggests that regardless of previous interference, the grammatical subject could be reliably reaccessed at the verb.

Pooling Post-Reciprocal Post-Position Regions

The analysis above decomposes the post-position associated with the reciprocal into two distinct regions: the genitive morpheme, found in all complex post-positions, and

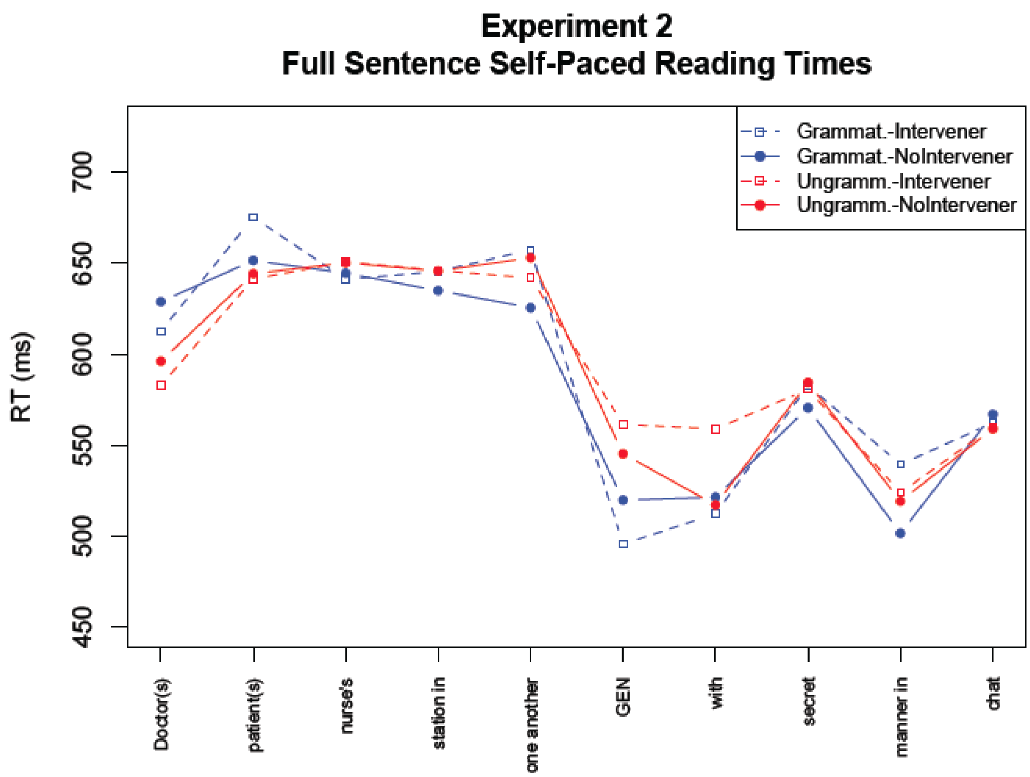


Figure 3.4: Experiment 2: Self-Paced Reading Results

the content word associated with the post-position. This division reflects the order of presentation in the experiment - the words were presented as distinct units. Analysis was also performed pooling the two words as a single region to reflect the intuition that these words are often thought of as a single unit (a complex post-position). When the two regions were pooled, the effect of GRAMMATICALITY remained significant (β = -33.35, s.e. = 12.72, t = -2.62, $p < .05$). *Grammatical* conditions were read significantly more quickly than *Ungrammatical* conditions on average. The GRAMMATICALITY x INTERVENER interaction which was significant in the reciprocal+2 region became marginally significant in the pooled analysis (β = 46.37, s.e. = 25.45, t =1.82, $p < .10$). RTs in the *Ungrammatical-Intervener* condition are marginally slower than in the *Ungrammatical-NoIntervener* condition.

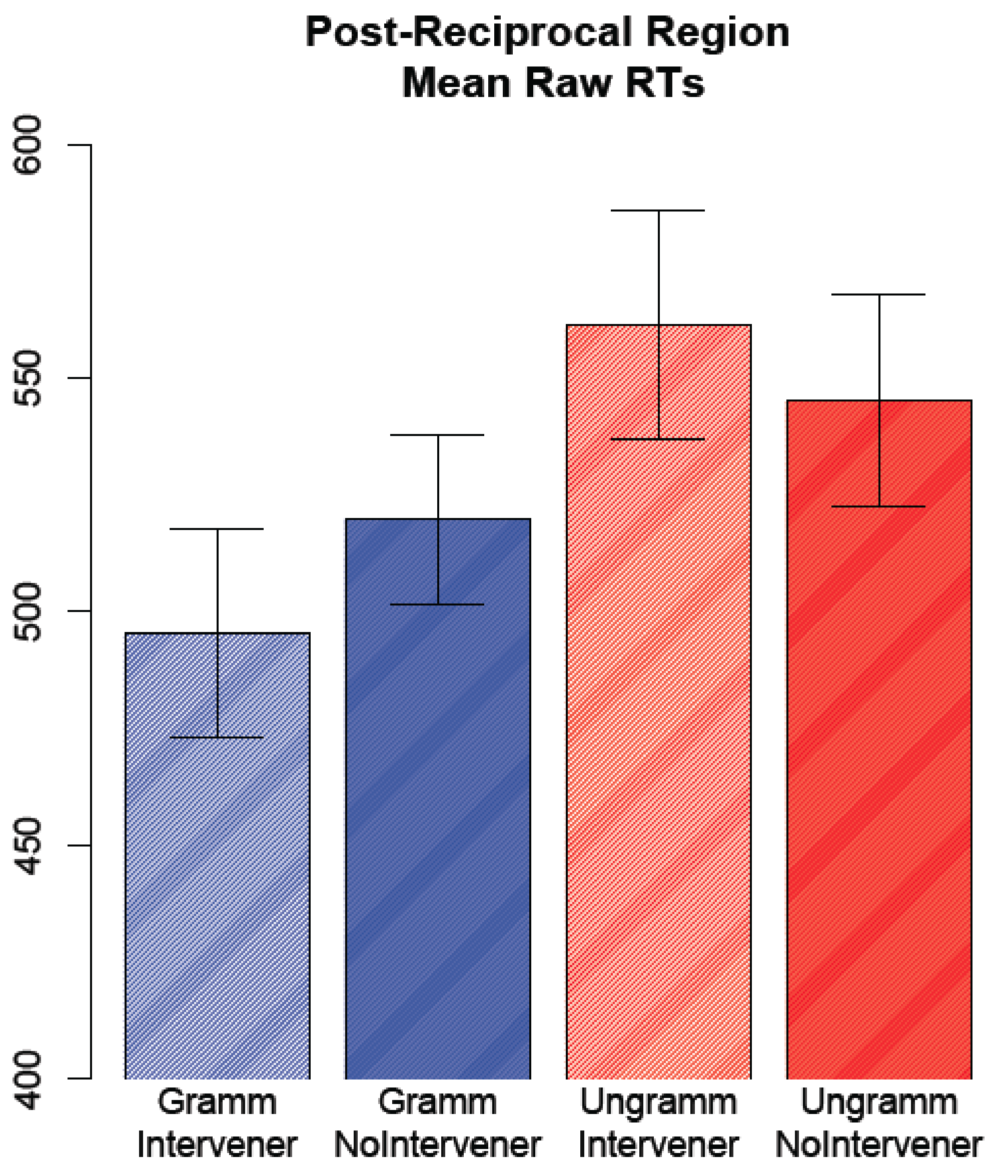


Figure 3.5: Experiment 2: Un-pooled Post-Reciprocal Region

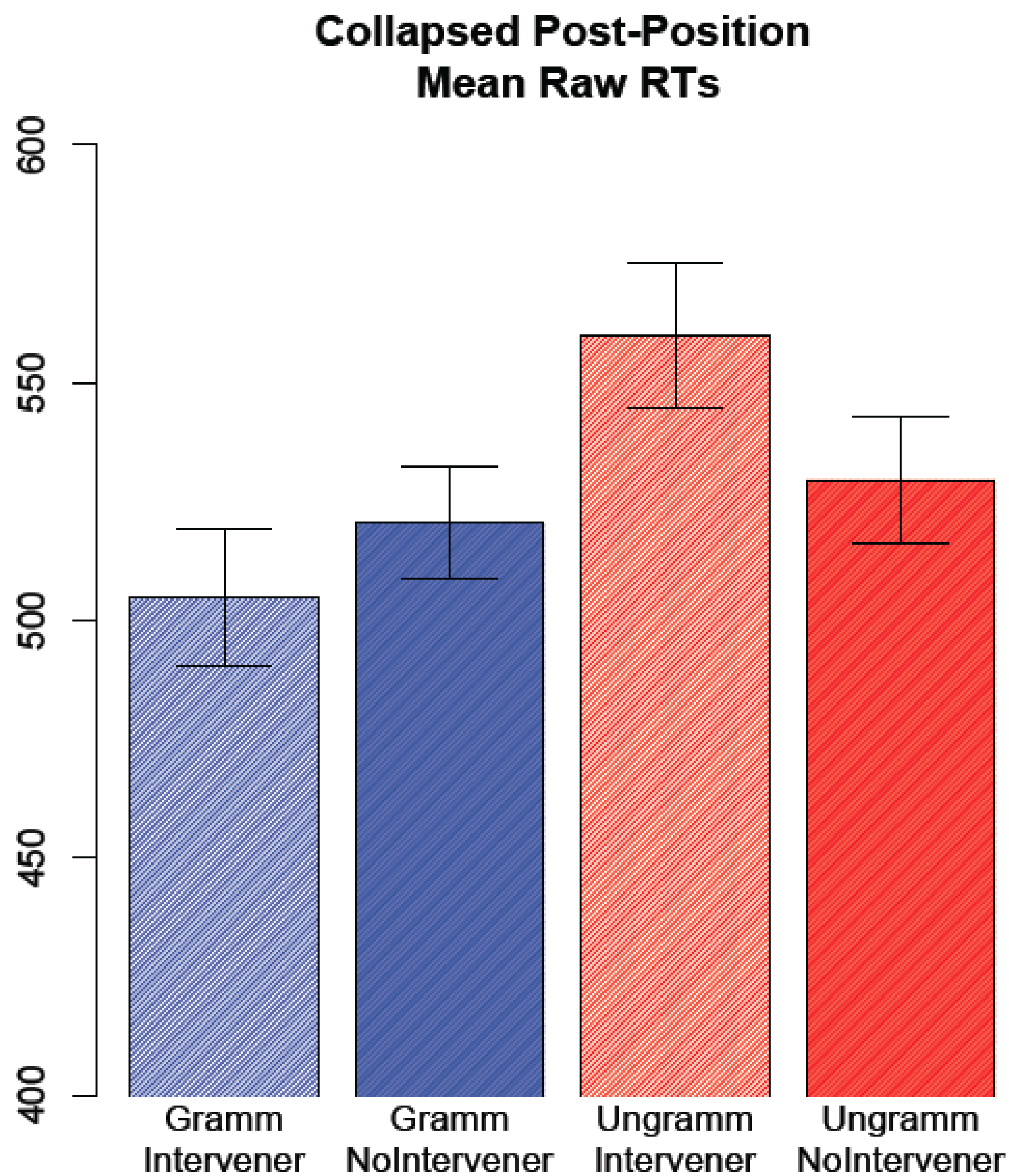


Figure 3.6: Experiment 2: Pooled Post-Reciprocal Region

3.5.4 Discussion

The study sought to determine whether structurally inappropriate feature-matching NPs interfere with the retrieval of an antecedent for a local reciprocal that precedes its verb.

Faster RTs in the post-reciprocal region were observed in conditions in which a structurally appropriate feature-matching antecedent was available. This finding is consistent with earlier work on local anaphor licensing, which demonstrated a retrieval advantage for grammatical antecedents.

The presence of a feature-matching intervener neither facilitated nor inhibited processing immediately following the reciprocal. In subsequent regions, elevated RTs were observed in the condition where only the intervener matched the reciprocal in features. The early behavioral results do not display the characteristic patterns of either kind of interference. Inhibitory interference is not observed - processing of the grammatically licit reciprocal in *Grammatical-Intervener* condition is not hampered by the presence of an inaccessible, feature-matching NP. Nor do the reaction time profiles indicate facilitatory interference, which would predict lower RTs in the *Ungrammatical-Intervener* condition relative to its *NoIntervener* counterpart.

3.6 General Discussion

Modeling results from Experiment 1 suggested that unweighted use of morphological cues in addition to the diagnostic clause cue would result in facilitatory interference effects. Such interference effects were not predicted by a model that preferentially used the

	Equal Weight	Diagnostic Cue	Behavioral Results
Ungrammatical	-124	0	15
Grammatical	33	5	-20

Figure 3.7: Hindi Reciprocals: Comparison of Modeled and Observed Interference Effects

diagnostic clause cue. Results from Experiment 2 suggest that feature-matching distractors do not interfere with local reciprocal licensing in Hindi.

The data appear to align better with the predictions of the diagnostic cue model, which predicted a main effect of GRAMMATICALITY, but no facilitation within *Ungrammatical* conditions.

It appears that retrieval is able to use a diagnostic cue strategy when licensing Hindi reciprocals.[18] Antecedent retrieval for a reciprocal either (i) uses a full complement of retrieval cues (including morphological cues) to access the local subject, with the diagnostic cue [clause] weighted such that it overwhelms interference from morphological cues, or (ii) retrieval of the reciprocal's antecedent exploits [clause] exclusively to retrieve the subject. The Hindi data suggest that linear adjacency is not the determinant of interference immunity. Therefore, a different strategy for determining the appearance of selective interference effects is motivated.

[18]Two of the three proposed methods to use linear adjacency to decrease interference fair especially poorly with these results. The baseline activation account, which assumes that recent activation of the grammatical subject drives the result receives little support. There is no reason to assume that the local subject was recently reactivated immediately before the reciprocal was encountered, as there was no bottom-up input that required its processing. For the same reason, the no-retrieval account is similarly dispatched: there is no reason to suspect that the appropriate subject occupies the focus of attention directly ahead of the reciprocal.

3.6.1 Which Cues to Use? Assessing Cue Diagnosticity

As noted above, the diagnosticity of [clause] as a retrieval cue varies as a function of the anaphor's grammatical context. When the subject is *ek-dusre* or *himself*'s sole clause-mate, the cue accurately singles the subject out.

(66) Subject$_{[Clause:1]}$ ... [NP$_{[Clause:2]}$...] *himself*$_{[+Clause:1]}$

(67) Subject$_{[Clause:1]}$ ... [NP$_{[Clause:2]}$...] *ek-dusre*$_{[+Clause:1]}$ ...

However, the cue loses diagnosticity in the presence of other clause-mates. Objects, or other arguments may precede a reciprocal.

(68) a. TeacheroN-ne$_{[clause:1]}$ baccoN-ko$_{[clause:1]}$ *ek-dusre* ki tasviir dikhaayii.
Teacher-ERG kids-ACC GEN picture showed

'The teachers showed the kids each other's photos.'

b. TeacheroN-ne$_{[clause:1]}$ baccoN-ko$_{[clause:1]}$ *ek-dusre* ke saamnee DaaNTaa.
Teachers-ERG kids-ACC GEN before scolded

'The teachers scolded the kids in front of each other.'

The same goes for English reflexives.

(69) a. John$_{[clause:1]}$saw himself in the vanity.

b. John$_{[clause:1]}$showed Bob$_{[clause:1]}$to himself in the vanity.

Above we used linear distance from the verb as a heuristic way of assessing whether [clause] would be a diagnostic cue. To encompass the Hindi data this strategy must be revised. Stepping back, we can see that linear distance from the verb is not what determines a [clause]'s diagnosticity, but rather whether another clause-mate intervenes between the anaphor and the subject. In the sentence below, the anaphor's only NP clausemate is the local subject, even though it is linearly distanced from the verb.

(70) John kept talking on and on to himself.

Therefore, rather than using linear-distance as the metric, an *optimal* cue selection procedure would instead follow the rule below, if it wished to maximize the diagnosticity of its cue set.

(71) Prefer [clause] for retrieving a local anaphor's antecedent, *unless other clause-mates of the anaphor intervene.*

This can be stated as an idealized decision rule. Two versions are given below, to reflect that one may either adopt to use the exclusive cue selection, or preferential weighting strategy.

(72) a. RULE FOR MAXIMIZING DIAGNOSTICITY OF CUE SET (SELECTION VERSION)

```
For an anaphor A, in clause C:
If number of verbal arguments in C preceding A == 0:
```

```
    Probe ⇝ {CLAUSE}
Else:
    Probe ⇝ {CLAUSE, ϕ}
```

b. RULE FOR MAXIMIZING DIAGNOSTICITY OF CUE SET (PREFERENTIAL WEIGHTING)

```
For an anaphor A, in clause C:
If number of verbal arguments in C preceding A == 0:
    Preferentially weight {CLAUSE}, over ϕ
Else:
    Weight cues in probe evenly.
```

19

This rule is simple enough to write, and simple enough to implement in the predicate-adjacent reflexive case. In the sentence below, when the parser encounters the reflexive next to a verb, I assume that the verb is still in the focus of attention because it was the last item processed (Cowan 2004; McElree 2006).[20]

[19] One might ask why including morphological features in the probe set in the cue set is preferable to omitting them and allowing some degree of indeterminacy. Adding morphological features to the retrieval probe does not guarantee that the cue set would uniquely identify one previously-seen NP. In the sentence below, adding morphological features would not help in retrieving one antecedent over the other.

- John showed Bill to himself in the vanity.

While, it is true that in cases of true syntactic ambiguity adding morphological features may be of no avail, there are obvious cases, where the strategy would yield benefits. For example, when previous NPs do not match in features.

- John$_i$ showed Mary$_k$ to himself$_{i/*k}$ in the vanity.
- Mary$_i$ showed Bob$_k$ to himself$_{*i/k}$ in the vanity.

[20] In ACT-R, the equivalent of the focus-of-attention is a limited capacity buffer.

(73) The managers that the employees hated embarrassed *themselves* . . .

FOCUS: *embarrassed*

I assume that the processor can *infer* from the fact that a verb is in the focus of attention/recent buffer that no other clause-mates have yet been encountered. An inference rule can be stated to codify this.

(74) INFERENCE RULE FOR PREDICATE-ADJACENT LOCAL ANAPHORS:

```
For anaphor A in input buffer:
If V in focus of attention and → subject is the only clause-mate
of A.
```

The cue generation procedure can use this inference rule in conjunction with one of the rules in (72) to determine the appropriate course of action to retrieve the antecedent.[21]

In the simple case, the English cue generation can make the optimal choice. Outside of the predicate-adjacent case, though, the inference/production rule schema above would be of no use. When the reflexive is not-predicate adjacent, the verb would no longer occupy the focus of attention/buffer. Instead, the immediately preceding word would occupy the position (in the case of King et al.'s stimuli, it would be the benefactive preposition *for*). The best that the cue generation can do is to adopt a heuristic strategy that tracks whether a verbal argument *could* have come before a reflexive based on the immediate environment

[21]Production rules stated in terms of the contents of a privileged buffer, or in terms of the current state of the system are commonplace in explicit models of cognitive functioning (see e.g. Newell 1973, 1990). It is often the case that in such systems '[a]ll procedural knowledge is represented as production rules. . . ' (Lewis & Vasishth 2005: 380).

the reflexive is encountered in.[22]

(75) INFERENCE RULE FOR PREDICATE-SEPARATED LOCAL ANAPHORS:

For anaphor A in input buffer:

If V not in focus of attention and $\rightarrow$ *assume there are non-subject clause-mates*

This inference rule produces the right results in English cases. If a verb is not in the focus of attention, morphological cues would be used in the probe in such a manner as to allow interference in accordance with (72). It is clear that the rule schema above would not work in capturing the Hindi facts. The verb would not sit in the focus of attention at the time the reciprocal was input because the verb would not have been encountered. Nevertheless, we observe no interference in Hindi, leading us to believe comprehenders can make the judgment.

Our challenge lies in finding some source of information for the Hindi parser to exploit to assess that no clause-mate has been encountered prior to the reciprocal. One could imagine reading such information directly off a tree structure. However, most models assume no unitary syntactic tree is maintained as a coherent object in declarative memory (e.g. Lewis and Vasishth 2005).

[22]Note that this general rule predicts interference when reflexives are contained in PPs, but no clause-mate besides the subject exists. This prediction should be tested:

- The manager(s) [that the employee(s) hated] spoke to themselves all day.

The account also predicts that the processing of emphatic reflexives should be subject to interference, even though they are typically argued to be local-subject oriented:

- The manager(s) [that the employee(s) hated] filled out the TPS reports *themselves*.

I would like to tentatively propose that the information required for the Hindi production rule is implicitly encoded in the control state of the parser, to which the cue-generation procedure has access.

3.6.1.1 ‘Counting’ Clausemates in Hindi - Implicitly

Lewis & Vasishth’s mocdel incorporates a left-corner (LC) parser (Aho & Ullman 1972; Johnson-Laird 1983), which makes predictions of immediate structural constituents based on a stored context free grammar. An LC parser uses the following rule to parse:

(76) LEFT-CORNER RULE

Given an input (terminal) and a goal category, if there exists a grammar rewrite rule of which the input is a left corner, replace the input with the left-hand side (LHS) of that rewrite rule; repeat this process with the LHS nonterminal symbol until no further replacements are possible. (Lewis & Vasishth 2005)

An LC parser makes use of a *goal stack*, which stores nodes that are predicted based on a (set of) phrase-structure rewrite rules triggered from bottom-up input. Whenever the LC rule is used to predict structure, non-terminal nodes are added to the goal stack. I propose that retrieval can reference the contents of the goal stack in Hindi to determine whether a non-subject clause-mate has been encountered prior to the reciprocal-associated retrieval.

Consider how the grammatical sentence (77) in Hindi is parsed incrementally by an LC parser up to the time when the reciprocal is encountered. In the diagram below I abstract away from the internal structure of the locative phrase which contains the intervener, collapsing it simply into the atomic *Locative PP*. First *Doctors*-ERG is encountered, which leads the parser to project an NP, and subsequently expand the S rule to predict a VP. This adds *VP* to the goal stack. Accommodating the locative phrase next, requires processing an adjunct structure. The PS rule associated with adjunct PPs does not permit postulation of a later V head.

(77) Doctors-ERG { In the station of the nurse } each-other about chatted.

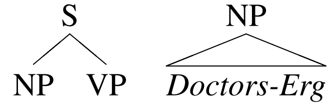
- RULES:

S $\rightarrow$ NP VP	PP $\rightarrow$ NP P
VP $\rightarrow$ NP-ACC V	VP $\rightarrow$ PP VP;
NP $\rightarrow$ N-ERG	NP $\rightarrow$ N-ACC;
N $\rightarrow$ Doctors, Kids . . .	

- INPUT: *Doctors*-ERG

  GOAL STACK: []

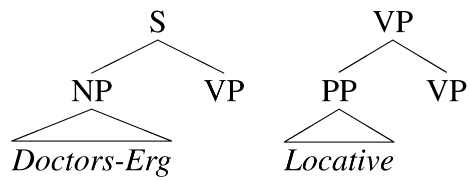
  ACTIONS: *Project NP, S, VP*, *Add VP to Goal Stack*



- INPUT: *Locative-PP*

  GOAL STACK: [VP S]

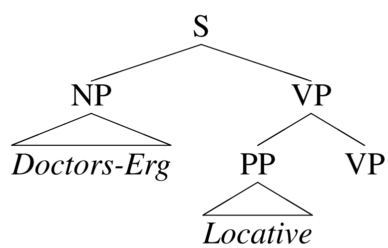
ACTIONS: *Project VP → PP VP*



- INPUT: *Reciprocal*

  GOAL STACK: [VP S]

  ACTIONS: . . .

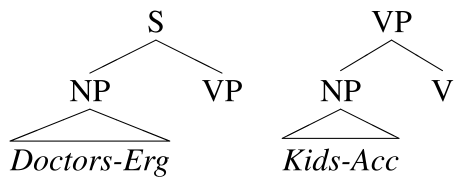


Because the rule *VP → PP VP* does not allow prediction of a V, when the reciprocal is encountered after the locative phrase, the goal stack is [VP, S]. If the reciprocal were encountered after an internal argument of the verb were encountered, however, the goal stack would be different. Encountering a direct object leads the parser to predict a verb head. Therefore, if the reciprocal were encountered *after* the object, the goal stack of the parser would be [V S]. Assuming the first step from the last parse, we can now describe the changes encountering an object would have on the parse state.

- INPUT: *Kids-Acc*

  GOAL STACK: [VP S]

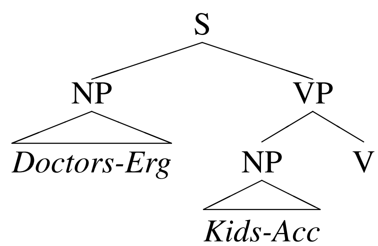
ACTIONS: *Project NP-*ACC*, Predict V from VP, Add V to Goal stack*



- INPUT: *Reciprocal*

  GOAL STACK: [V S]

  ACTIONS: *Project NP-*ACC*, Predict V from VP, Add V to Goal stack*



Because the goal stack differs depending on the syntactic environment, we can say its contents implicitly represent the knowledge of whether an object has been observed. I propose that the cue generation procedure uses the inference rule below to determine which cues to use based on the goal stack:

(78) INFERENCE RULE FOR HINDI CLAUSE-MATE DETERMINATION

`If` *V* `in` GOAL STACK $\rightarrow$ `an internal argument is present.`

In English, this inference strategy is not possible, given the ordering of parsing. Because reflexives (and reciprocals) follow their verbs, a V will never sit in the goal stack when a local anaphor is encountered. This is because the goal stack does retain any nodes that it has already found confirmatory evidence for (Johnson-Laird 1983).

The account outlined above makes one prediction: If a non-subject clause-mate NP *were* introduced before the reciprocal in Hindi, interference would emerge. We would therefore expect interference in (79-b) in the presence of the indirect object *ma-ko*, but not in its absence.

(79) a. *DoctoroN-ne (ma-ko) baccoN-ki party me ek-dusre ke-bare-me bataayaa.
Doctors-ERG mother-ACC kids-GEN party in one-another about told
'The doctor told the mother about each other at the kids' party'

b. *Doctor-ne (ma-ko) baccoN-ki party me ek-dusre ke-bare-me bataayaa.
Doctor-ERG mother-ACC kids-GEN party in one-another about told
'The doctor told the mother about each other at the kids' party'

I leave testing this prediction to future work.

3.7 Conclusion

In this chapter I reviewed evidence demonstrating antecedent retrieval for local reflexives in English is able to distinguish structurally appropriate from feature-matching inappropriate NPs. I argued that this sensitivity could be brought about using a *clause* index feature. I also argued that this feature could be used in one of two ways to guarantee access to the local subject and minimize interference:

i) Use of the *clause* cue to the exclusion of morphological cues (also proposed by Dillon et al. 2013)

ii) Preferential weighting of *clause* relative to morphological features.

For the purposes of this thesis, I will henceforth adopt the preferential weighting approach as the method for minimizing the effects of interference-prone features in the retrieval probe.

I also argued that the explanation of selective immunity to interference observed across studies should not be explained as a direct effect of linear adjacency. Immunity to interference arises across languages and word orders when the cue generation procedure can choose a maximally diagnostic cue set. Interference arises when cue generation cannot make this determination, either because a diagnostic cue set is not possible, or when architectural restrictions conspire to prevent access to the appropriate information. The strategies used to distinguish environments of diagnosticity from non-diagnosticity were discussed, and were argued to differ across languages. English uses a local heuristic strategy to approximate the ideal cue generation behavior, whereas Hindi uses information about the control state of the parser to make a more informed decision.

Chapter 4: Clause-mate Conditions, Binding Principles, and the LOCAL Feature

4.1 Introduction

In the previous chapter a *clause* feature was used to provide an account of how cue-based retrieval might distinguish feature-matching NPs that were grammatically-appropriate potential antecedents from grammatically inappropriate ones for the purposes of local anaphor licensing. In this chapter I refine the account. I propose a new feature, LOCAL, that subsumes the role played by the clause feature in the previous account. More importantly, I show that the utility of positing such a feature extends beyond cases of local anaphor licensing. I argue that the same feature can be employed to cover retrieval sensitivity to Principle B. As will be argued in later chapters, this LOCAL feature also serves as the lynchpin in other instances of retrieval's grammatical compliance.

The structure of the chapter is as follows. The next section introduces the LOCAL feature and shows how it can be applied to achieve Principle A-sensitive retrieval. A dynamic updating function is introduced as a means for allowing the distribution of LOCAL to track accessible chunks within a given structural domain (the current clause). This function 'compiles out' the domain restriction on local anaphor licensing in procedural terms. The following section shows how LOCAL can provide an account of Principle B sensitivity in retrieval. The account enshrines the complementarity of Principles A and B discussed by formal grammarians in the binary valuation of the LOCAL feature. Various minor modifications to the account are discussed before the last section, which concisely summarizes the account.

4.2 A LOCAL Account of Principle A

Attempts to flesh out how exactly a clause feature that could be used as needed for local anaphor licensing could be assigned raises some empirical issues. If we assume that all NPs bear the clause feature of their *minimal container clause*, the non-c-commanding NP *John* in (1) would be considered a clause-mate of the anaphor. This would predict that it should be accessible to retrieval, contrary to what the grammar would allow.

(1) *John$_i$'s mother likes himself$_i$

This problem arises because the set of NPs that c-command a reflexive within a minimal clause form a subset of the NPs that can possibly appear within that minimal clause. This suggests that assignment of the feature relevant for local anaphor licensing is subject to more stringent restrictions. For example, suppose that only NPs directly attached to the main spine of the current clause were marked with the clause feature.[1] For clarity, assume that this new feature is called LOCAL. The assignment function for LOCAL would be as

[1] A spine can be (informally) defined as a path through the right-most, non-adjunct daughters of a series of connected phrases. Tracking spines provides a way of encoding a general proxy for surface c-command feature (see Alcocer and Phillips 2012 for further discussion of spinal encoding).The notion of a *spine* has been used in both formal syntax (e.g. Uriagereka 1999), as well as in computational linguistics (e.g. Nederhof 1993).

below.

(2) LOCAL ENCODES SPINE-MATE

$$\text{LOCAL}(\text{x,current.Clause}) = \begin{cases} 1 & \text{if x is a DP on the spine of current.Clause} \\ 0 & \text{if x is any other DP} \\ \textit{NULL} & \text{if x is not a DP} \end{cases}$$

A different characterization/assignment function would forgo appeal to the structural notion ‘spine’ entirely, opting instead to assign LOCAL:1 to arguments of the current clause’s verb.

(3) LOCAL ENCODES CO-ARGUMENTHOOD

$$\text{LOCAL}(\text{x,curr.Clause}) = \begin{cases} 1 & \text{x is an argument of curr.Clause's verb} \\ 0 & \text{if x is any other DP} \\ \textit{NULL} & \text{if x is not a DP} \end{cases}$$

On this analysis, retrieval of LOCAL:1-marked items would implement a *co-argument* restriction on local anaphor licensing (co-argument restrictions on anaphoric licensing in some form or another are well-established in the syntax literature, see Hellan 1988; Dalrymple 1993; Reinhart & Reuland 1993; Pollard & Sag 1994; or Safir 2012a).[2]

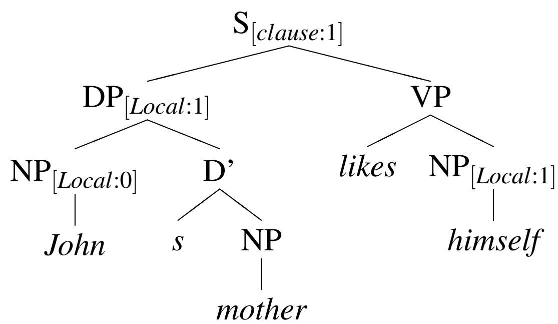
The distribution of LOCAL in a mono-clausal sentence is illustrated below:

[2]Co-argument-based assignment of LOCAL extends the reach of the LOCAL strategy even further. There are known cases where it appears that a local anaphor can be bound by a preceding NP contained in a PP. These cases of *binding out of PP* appear to be incidences of binding in the absence of c-command (Van Riemsdijk and Williams 1986; Pollard & Sag 1992; Drummond & Kush to appear, a.o.).

- Mary spoke to John$_i$ about himself$_i$.

Under the argument approach to LOCAL assignment, *John* would bear LOCAL:1, and would therefore be an accessible antecedent for the reflexive in 2. The spine-based strategy could also be modified slightly to allow assignment of LOCAL: 1 to NPs inside (certain) selected PPs attached to the main spine.

(4)



With this feature a simple retrieval strategy for a local anaphor's antecedent can be stated.

(5) LOCAL ANAPHOR RETRIEVAL

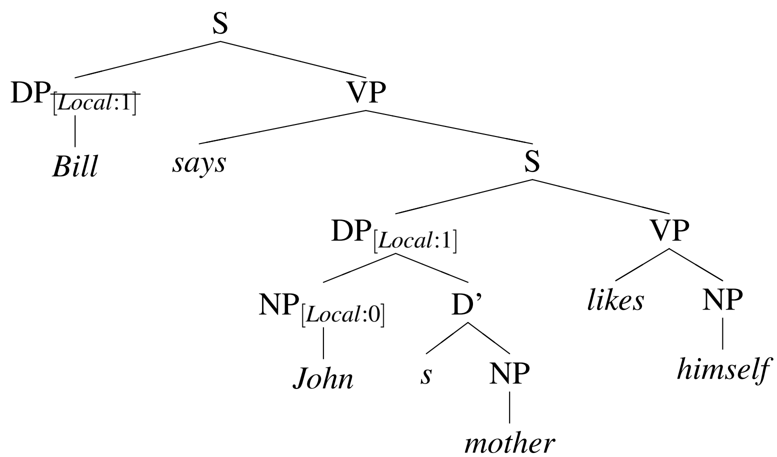
A local anaphor's antecedent must be [LOCAL:1].

The spine-mate system collapses the locality and c-command constraint into a single feature. NPs that are 'spine-mates' would be marked LOCAL: 1, whereas other NPs would not be. For example, an NP contained within a PP would not be marked LOCAL: 1, as it would not be connected directly to the main spine.

4.2.1 Local Update

As the account currently stands, LOCAL would distinguish eligible from ineligible NPs for the purposes of anaphor licensing in simple mono-clausal sentences, but it would not always be successful in multi-clausal sentences. This is because the account does not yet have the ability to distinguish NPs that are local relative to the current clause from those that were local in a previous clause. At present, all previously-seen argument NPs will always bear LOCAL:1, as illustrated below.

(6) Bill says John's mother likes him.



Bill would be predicted accessible to retrieval in the sentence above, because it would match the LOCAL:1 retrieval probe. This outcome is undesirable. Bill bears the LOCAL feature because it was an argument of the current clause's verb at the time of its encoding. It is only after the second clause has become the current clause that the LOCAL value on *Bill* is mis-assigned. What we require to avoid this outcome is an iterative procedure that updates LOCAL clause features whenever the parser posits a new clause. A procedure like this would require the power to recall all the constituents of the previous clause, and rewrite the value for their LOCAL feature.

(7) LOCAL UPDATE RULE:

```
If a clause-boundary is identified:
    While any item in memory is marked LOCAL:1:
        q = retrieve([LOCAL: 1]);
```

```
set q to [LOCAL: 0]
```

In effect, this update and encoding procedure allows retrieval to keep a moving window of the local domain (the current clause).

This reformulation of the feature-based account of local anaphor licensing may at first seem unnecessarily baroque. The new account posits a simplistic LOCAL feature and a computationally costly update function. The previous account used a clause-index feature – a more economical solution in that it does not require any update function. In the next section I will argue that the LOCAL-based account is to be preferred to the clause-index account because it can be generalized to binding phenomena beyond Principle A licensing in a way that the simple clause-index strategy cannot. More specifically, a LOCAL-based account provides a foothold in explaining Principle B sensitivity. Thus, it would appear that the ends (a uniform account of the Binding Principles) justify the (computationally costly) means.

4.3 Principle B

Requiring that a local anaphor and its antecedent match on LOCAL:1 enables implementation of Principle A as a c-commanding clause-mate or co-argument condition. I argue that the feature also provides the basis for implementing Principle B. It has long acknowledged that local anaphors and pronouns are in (roughly) complementary distribution.

(8) a. Mary said John$_i$ likes himself$_i$/*him$_i$.

b. John$_i$ said Mary likes hims$_{i}/k$/*himself$_i$.

Some authors have assumed that this complementarity arises because the same underlying mechanism determines the distribution of anaphors and pronouns (e.g. Lees & Klima 1963; Reinhart & Grodzinsky 1993; Hornstein 2001; Drummond, Kush & Hornstein 2011). If the two phenomena are inextricably bound together in the grammar, or if the explanation of their distributions make reference to the same notions, pursuing a fully transparent implementation of grammatical principles would lead us to the conclusion that the two should be tightly connected in our implementation.

4.3.1 Previous Work on Principle B In Retrieval

The history of work on interference effects and antecedent retrieval is varied. Badecker and Straub (2002) and Kennison (2003) have argued that feature-matching NPs within a pronoun's local domain interfere with the retrieval of antecedents for that pronoun. In sentences like the one below, Badecker and Straub (2002) reported signs of inhibitory interference contingent on feature-match between the local subject *Jane/Bill* in the post-pronoun region.

(9) John thought *Jane/Bill* owed him another chance.

Contra Badecker and Straub (2002), other authors have argued that retrieval respects Principle B (Nicol 1988; Clifton, Kennison & Albrecht 1997; Lee & Williams 1999; Chow, Lewis & Phillips in prep.). Across a number of studies, Chow, Lewis and Phillips (in prep.) consistently found no evidence of interference from feature-matching NPs in the same local domain as the pronoun. For example, in their second experiment of five, Chow, Lewis and Phillips did not find an interaction effect suggestive of interference when the local noun *producer* matched the pronoun *him* in gender.

(10) {**Ethan/Paige**} revealed that the {**producer / dancer**} had doubted {*him*} *even after several successful performances.*

Regardless of whether the hallmarks of interference are observed online, the previous studies all agree that there is a retrieval advantage for those NPs that Principle B considers a grammatically antecedent for a pronoun in comparison to inaccessible NPs. This distinction must be made somehow. Whether the attribute that distinguishes accessible from inaccessible NPs lends itself to imposing a categorical constraint on interference is a separate issue.

4.3.2 Turning a Negative (Constraint) Into a Positive (Feature)

If Principle A can be thought of, more or less, as a clause-mate/co-argument constraint, Principle B becomes (roughly) an 'anti-clausemate' condition. If Principle A retrieval were thought of as strategy that enforced match between an antecedent and reciprocal's clause-index, Principle B might be thought of as a constraint that enforced a mismatch between those features.

(11) ANTI-CLAUSEMATE CONDITION:

For current clause x, permit all NPs whose clause feature is not x.

As with Principle A, reliance on a clause feature to implement Principle B faces serious problems. Though simple to state informally, a negative constraint like an anti-clausemate condition cannot be easily implemented with retrieval cues because of a fundamental assumption that underlies most, if not all, functiosn that govern associative strength: they are matching functions. Cue-based retrieval can activate an item in proportion to its feature-overlap, but it cannot enforce a mismatch. Because mismatch cannot be enforced, the only way to implement an anti-clausemate condition using a clause-index feature would be to explicitly specify all of the possible clauses that an antecedent could be in as individual retrieval cues. An algorithm for generating a cue set like this is trivial. By taking the current clause's index, one could populate a list of all previous clauses by iteratively decrementing the clause counter. This would provide a list of clause indices that could each be used as an individual cue. As a production rule:

(12) AN UNPROMISING ALGORITHM TO IMPLEMENT AN ANTI-CLAUSEMATE CONSTRAINT IN RETRIEVAL

```
To retrieve an antecedent for pronoun P:
    get current.Clause index;
    create list possible.Clauses = range(0,current.Clause);
    For x in possible.Clauses:
        create cue q = [clause:  x]
        add q to Probe
    retrieve using Probe
```

The algorithm would generate the the retrieval probe in (14) for *him* in the sentence below.

(13) $[_{Clause:1}$ John said $[_{Clause:2}$ Mary thought $[_{Clause:3}$ that Jake liked him]]].

(14) $$him \rightsquigarrow \begin{bmatrix} Clause:1 \\ Clause:2 \\ Number:Sing. \\ Gender:Masc. \end{bmatrix}$$

Such a probe would encounter a number of challenges, many of which relate to the fact that the number of clause cues in the sentence would grow as the number of clauses grows. As clause cues proliferate, the individual strength of each cue decreases. Inversely, the possibility of interference increases. More problematically, were we to require a categorical constraint, there is no single cue that could be weighted to cancel out the effects of other potentially interfering features. Finally, as the number of clause cues grows, the retrieval strategy would lose the ability to ever find a complete match because images cannot be in superposition. No single image can possibly bear all of the clause features on a probe. Even though *John* has the appropriate features to antecede the pronoun, the cue set above would not consider the image a complete match because it does not bear the [Clause:2] cue. Therefore, as noted by Dillon (2011), it might seem initially impossible to enforce anything like the relevant constraint on the distribution of possible antecedents for pronoun as an initial filter on retrieval.

An account that uses the LOCAL feature to implement Principle B, on the other hand, does not run into the same challenges as a clause-index account. Above, LOCAL was defined as a binary feature assigned exclusively to DPs. Arguments in the current clause are LOCAL:1, whereas all other DPs are LOCAL:0. The account of Principle A enforced match on LOCAL:1. In order to implement an anti-clausemate constraint, LOCAL:1 pronouns could require that their antecedents *match* a retrieval probe that specified a value of '0' for LOCAL. We can make a first-pass at stating this retrieval strategy.

(15) ANTI-CLAUSEMATE RETRIEVAL (non-final):

Pronouns must have antecedents marked LOCAL:0

With this strategy, the LOCAL feature, and its update function, we have the ability to cover a broad swath of the traditional Principle B effects. Consider the examples below.

(16) a. John likes him.

b. John's mother likes him.

c. John said Mary likes him.

In (16-a), both *John* and *him* will be marked LOCAL:1 because they are spine-mates, or co-arguments. The pronoun *him* positively specifies that its antecedent must bear LOCAL:0, which *John* does not match.

$$him \rightsquigarrow \begin{bmatrix} Local : 0 \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$$

The case of (16-b) is different. As a possessor, *John* is neither a spine-mate of the pronoun, nor a co-argument. It will therefore bear LOCAL:0; it has the proper features to antecede the pronoun. So too does *John* in (16-c). Though *John* would initially be tagged LOCAL:1 during the incremental parsing of the sentence, the update function would ensure that it was recoded as LOCAL:0 at the onset of the embedded clause.

Though a majority of the Principle B cases are covered by a rigid requirement that pronouns take LOCAL: 0 antecedents, some important cases are left out. Pronouns in possessor position may take antecedents within their minimal clause. That is, LOCAL: 0 pronouns may take LOCAL: 1 antecedents, unlike LOCAL: 1 pronouns.

(17) a. John$_i$ loves his$_i$ mother.

b. John$_{[Local:1]}$ loves his$_{[Local:0]}$ mother.

These pronouns may also take LOCAL: 0 antecedents.

(18) John$_{[Local:0]}$ said Mary loves his$_{[Local:0]}$ mother.

Thus, we must amend the account to permit LOCAL: 0 pronouns to take antecedents with either value for LOCAL. Allowing the probe for LOCAL: 0 pronouns to remain underspecified for a LOCAL value achieves this. With this, we are able to state the cue-generation rule for anti-clausemate retrieval, which is conditioned on the LOCAL value of the retrieval trigger.

(19) ANTI-CLAUSEMATE RETRIEVAL (final):

When retrieving an antecedent for a pronoun P, use the following features for LOCAL:

•If P is LOCAL: 1 $\rightsquigarrow$ Probe [LOCAL: 0]

•If P is LOCAL: 0 $\rightsquigarrow$ Probe [LOCAL: NULL]

Before concluding this section, we should note here the choosing the right assignment function for LOCAL is an empirical matter that can be settled by investigating cases of possible co-valuation. Recall that different assignment functions would make LOCAL a feature that either a notion of 'spine-mate' or 'co-argument'. In the first case only NPs attached directly to the main spine were marked LOCAL: 1. In the second case, only those NPs that were selected arguments of the current clause's verb. Each of these accounts makes differing predictions regarding slightly more complicated cases of co-valuation than have been considered. For example, a strict spine-mate approach to LOCAL would predict that *John* would be accessible to antecedent retrieval for the pronoun *him* below, because *him* would be marked LOCAL: 0. This prediction does not comport with native judgments. However, both the clause-mate and the co-argument accounts would block retrieval of *John* as an antecedent for *him* because both NPs would be marked LOCAL: 1.

(20) *Mary spoke to John$_i$ about him$_i$.

The co-argument formulation also blocks *John*from being retrieved in (21). The pronoun would be marked LOCAL: 1 and would therefore require a LOCAL: 0 antecedent.

(21) a. John said to him that ...

b. John spoke with/about him.

4.3.3 Updating LOCAL in Constant Time

Consider one issue related to the assignment and maintenance of LOCAL. LOCAL is a dynamically-updated feature. It must be over-written at the end of a clause, a procedure which requires retrieval of multiple items. One might worry that this dynamic updating would introduce the problem of linear time complexity. The more DPs there were in the most recent clause, the more DPs there are that require retrieval for update at the clause edge. This situation is quite reminiscent of the problem of encoding c-command exhaustively. The two cases differ, importantly, in one respect.

The number of retrievals required to exhaustively encode c-command at a given step is potentially unbounded, since the number of previously seen (and therefore potentially c-commanding) nodes is unbounded. In the case of the retrievals required to update LOCAL, however, there is a strict upper bound imposed by the grammar on the number of possible retrievals there could ever be. The number of arguments in a clause is limited by the grammar - by the number of θ-roles a verb can assign. In most cases a retrieval of no more than two previously-seen items would ever be needed. As long as there is a finite upper-bound on the number of retrievals required by the update procedure, it is still a constant time operation.[3]

[3]Ewan Dunbar (p.c.) points out the interesting counter-intuitiveness of this reasoning. Currently, the constraint rules anything in that is finitely-bounded, without imposing a clear limit on what that upper-bound should be. This means that procedures with an astronomically high, but still finite, number of steps would still be admitted by the constraint (counter to what out intuitions tell us about efficient parsing routines). This suggests, as Chomsky (1955:150) put it, the relevant distinction 'is not the difference between finite and infinite, but the more elusive difference between too large and not too large.'

4.3.4 A Gap in LOCAL Coverage

Use of LOCAL leaves one conspicuous gap in the system's coverage of Principle B phenomena: the ability to handle NP-internal pronouns. At present, the account does not have the ability to rule out retrieval of local c-commanding NPs in picture NPs or NPs with PP adjuncts. *The man* would be retrievable in either of the configurations below, despite being an illicit antecedent.

(22) a. The man$_i$'s picture of *him*$_{*i/j}$ frightened John.

b. The man$_i$ next to *him*$_{*i/j}$ talked to John$_j$.

Because the pronoun is neither on the main spine, nor an argument of the verbal predicate, it would be marked LOCAL: 0. As such, it would not impose a constraint on the LOCAL value of its antecedent. Regardless of how *the man* is marked, it should be accessible.

(23) a. The man$_{[Local:0]}$'s picture of *him*$_{[Local:0]}$ frightened John.

b. The man$_{[Local:1]}$ next to *him*$_{[Local:0]}$ talked to John.

Two options offer themselves in the face of this inadequacy. We can either accept that the account is deficient in this regard without any modification, or we can attempt to modify the feature-assignment and update function to accommodate the data.

Following the first route leads us to expect interference in these configurations. Preliminary evidence against this prediction comes from Runner, Sussman & Tanenhaus (2006) who used evidence from visual world studies to argue that interference is not observed on pronouns in picture NPs.

The second alternative is to attempt to provide an account using the LOCAL feature. A pronoun can only enforce a constraint on its antecedents if it is LOCAL: 1. At the moment, LOCAL references a clause-based notion of 'spine' - the local domain is always the current clause. Suppose instead that while parsing complex NPs the local domain were re-evaluated (temporarily) as the 'spine' of the complex NP itself. If this were sothe system could provisionally assign LOCAL: 1 to anaphors within a complement of N, which would be later overwritten at the point which the end of the constituent was encountered. For example, the incremental parse of the constituent and the update of the function are outlined below.

(24) The man$_{[Local:1]}$ next to ...

The man$_{[Local:1]}$ next to him$_{[Local:1]}$ ... *Provisional Assignment*

The man$_{[Local:1]}$ next to him$_{[Local:1]}$ was ... *End of Phrase*

The man$_{[Local:1]}$ next to him$_{[Local:0]}$ was ... *Rewrite* LOCAL

Above, we assume that retrieval occurs before *was* is encountered. As the pronoun bears LOCAL: 1, it requires a LOCAL: 0 antecedent. Therefore, *the man* is not eligible.

One further modification is required to accommodate the picture NP cases. Pronouns internal to picture NPs are blocked from taking the NP's possessor as an antecedent (25).

(25) a. John$_i$'s picture of him$_{*i/j}$ upset Bill$_j$.

b. Bill$_j$ saw John$_i$'s picture of him$_{*i/j}$.

Above it was argued that possessors of NP are marked LOCAL: 0 so as to allow coreference/binding between subject and an possessor of a local NP in, say, object position.

(26) John$_{[Local:1]}$ likes his$_{[Local:0]}$ mother.

However, if possessors are not LOCAL: 1, they should be accessible to the retrieval triggered by the provisionally LOCAL: 1-marked anaphor within the NP. The solution to this problem is to also allow provisional assignment of LOCAL: 1 to possessors of picture NPs. This assignment reflects the often remarked-upon fact that picture NPs display the traits of predicative environments of which the possessors are the subject. Importantly, in order to maintain the bindability of possessors by their local subject, we cannot assume that possessors are uniformly marked LOCAL: 1. Rather, marking a possessor as LOCAL: 1 must be contingent upon it being part of a predicative NP. This would require encountering the head of the NP to assess whether this marking is necessary. Upon first encountering the possessor, the parser would assume that it was LOCAL: 0. Only after encountering the head noun would the possessor be re-encoded. This provisional LOCAL marking could then persist until the end of the picture NP, where it would then be erased.

(27) The man$_{[Local:0]}$'s ... *Treat Possessor as 'Non-local'*

The man$_{[Local:0]}$'s picture ... *Identify predicative NP*

The man$_{[Local:1]}$'s picture ... *Rewrite* LOCAL

The man$_{[Local:1]}$'s picture of him$_{[Local:1]}$ ... *Assign* LOCAL:1 *to NP-internal pronoun*

The man$_{[Local:1]}$'s picture of him$_{[Local:1]}$ was ... *Reach NP edge*

The man$_{[Local:0]}$'s picture of him$_{[Local:0]}$ was ... *Reassign* LOCAL: 0

This final rewriting step is required to render the possessor ineligible to bind later LOCAL: 1 clause-mates, and to allow it to antecede subsequent pronouns. If this step were not taken, the co-valuation in (28) would be potentially considered, and the co-valuation in (29) would be ruled out - both contrary to native speaker intuitions.

(28) Mary have John$_i$'s picture of Bob to himself$_{*i}$.

(29) Mary have John$_i$'s picture of Bob to him$_i$.

The above discussion shows how one might go about augmenting the procedure for LOCAL assignment so that retrieval accessibility could align with grammatical descriptions of anaphoric relations internal to complex and picture NPs. Whether to adopt the more complicated system depends on two factors: (i) empirical evidence demonstrating that antecedent retrieval is, in fact, immune to interference from ungrammatical NPs in these configurations, and (ii) aesthetic preference. The procedure described is significantly more complicated than other encoding strategies, and requires more retrievals and intermediate updating than some might be willing to accept. As such I leave determining whether this strategy is necessary to later work.

4.4 Summary

In this chapter I introduced the LOCAL feature, which could be used as a rough proxy to identify argument/c-commanding NPs in a clause.

(3) $\text{LOCAL(x,curr.Clause)} = \begin{cases} 1 & \text{x is an argument of curr.Clause's verb/on the curr.Clause's s} \\ 0 & \text{if x is any other DP} \\ NULL & \text{if x is not a DP} \end{cases}$

The account was supplemented by an update procedure that changed NPs that were previously valued 'local' to 'non-local' NPs. The successive-cyclic update of the LOCAL feature produces a system that provided an ability to distinguish the co-arguments/c-commanders of the current clause from all other NPs.[4]

(30) LOCAL UPDATE RULE (paraphrased):

At the start of a new clause, rewrite all instances of LOCAL: 1 as LOCAL: 0.

Using this dynamically updated feature, cue-based retrieval strategies were easily devised that could that recoded both Principles A and B.

(5) LOCAL ANAPHOR RETRIEVAL

A local anaphor's antecedent must be [LOCAL:1].

(19) ANTI-CLAUSEMATE RETRIEVAL:

When retrieving an antecedent for a pronoun P, use the following features for LOCAL:

[4] As seen during the discussion of complex and picture NPs, this update procedure could be further adapted to provide a more flexible definition of the local domain, though such elaboration should be first empirically motivated.

- If P is LOCAL: 1 $\rightsquigarrow$ Probe [LOCAL: 0]

- If P is LOCAL: 0 $\rightsquigarrow$ Probe [LOCAL: NULL]

Using LOCAL as the basis for anaphoric retrieval in a cue-based system was argued to be superior to using a clause-index system because it provided a unified solution to both Principles A and B.

Chapter 5: Crossover (and Principle C)

5.1 Introduction

In the previous chapter it was argued that a unified account of Principle A and B sensitivity in retrieval could be achieved through the use of the LOCAL feature. The LOCAL feature identifies a set of NPs along the command path within a bound domain (the local clause). LOCAL provides a limited window onto local c-command relations. It does not provide a way to track c-command relations for items beyond the current clause. For example, the feature does not provide a way to distinguish between non-local NPs that c-command items in the current clause from those that do not c-command into the clause. In both of the sentences below, *the politician* is marked LOCAL: 0, despite its different position.

(1) a. The politician$_{Local:0}$ said that people liked him.

b. The woman that the politician$_{Local:0}$ consulted with said that the people liked him.

Some c-command constraints are not bounded in their domain of application; they require computation of c-command relations across a potentially unbounded number of clauses.

Strong Crossover (Postal 1971, Wasow 1972) and Principle C are two such constraints. Because they are unbounded, these constraints would appear to be outside the ambit of the LOCAL strategy.

The studies in this chapter show that antecedent retrieval is sensitive to the Strong Crossover constraint (and Principle C, if Strong Crossover is reducible to Principle C as has been argued). Feature-matching wh-fillers that c-command a target pronoun are considered as antecedents only when covaluation of the wh-filler and the pronoun is not ruled out by Strong Crossover. Though antecedent retrieval categorically respects Strong Crossover, I also show that it does not respect Weak Crossover. Retrieval's differential sensitivity to these two constraints suggests that the observed pattern of sensitivity tracks whether the critical pronoun c-commands the wh-filler's resolution site. Despite the prima facie challenge that the constraints pose to a LOCAL-based strategy, I provide an account of the observed sensitivity in terms of the LOCAL feature. The key idea behind this solution lies in an incremental parser's ability to treat the application of the Strong Crossover constraint as an instance of the anti-clausemate procedure used for Principle B.

The structure of the chapter is as follows. Below I discuss Crossover constructions (Postal 1971, Wasow 1972, a.o.). I discuss Strong Crossover's relation to Binding Principle C (Chomsky 1981), and the relevance of these constraints to the questions of this thesis. Three experiments are discussed. Their results show that c-command constraints appear to be relevant for characterizing the NPs that retrieval has access to, and those that it does not. In the final discussion, I provide an explicit account of the use of LOCAL in capturing Strong Crossover/Principle C effects.

5.1.1 Crossover Constructions

There is a family of constructions in which a quantifier scopes over a pronoun, but cannot bind it: Crossover constructions. Crossover effects (first discussed in Postal 1971, Wasow 1972) refer to the inability to covalue a pronoun and an NP in a variety of configurations. Traditionally, Crossover has been invoked to explain the inability for the pronouns in (2) to be coindexed with the quantifier phrase and relative pronoun, respectively. Crossover comes in two 'varieties': Strong, and Weak. In Strong Crossover constructions, the pronoun is usually in argument position, whereas in Weak Crossover examples it is embedded inside another phrase.

(2) a. STRONG CROSSOVER

*The man who$_i$ he$_i$ likes looked in the mirror.

*ix. x is a man and x likes x and x looked in the mirror.

$\approx$ 'The man who liked himself looked in the mirror.'

b. WEAK CROSSOVER

*?The man who$_i$ his$_i$ mother likes looked in the mirror.

*ix. x is a man and x's mother likes x and x looked in the mirror.

$\approx$ 'The man whose mother liked him looked in the mirror.'

Crossover has also been used to explain the unavailability of certain types of bound variable anaphora. As discussed in Chapter 3, May's (1977, 1985) Quantifier Raising operation covertly moves QPs to their minimal tensed clauses to establish scope. As a result, the QP in (3) would covertly move to a position that c-commanded the pronouns, thus bringing about the right configuration for the former to bind the latter. If an element's covert c-command position is the only factor determining its binding capabilities, raised QPs should be able to bind pronouns that they do not linearly precede if they scope over them. Yet, we see that the predicted reading of this co-indexation is unavailable - the sentences in (3) cannot be read as synonymous with the paraphrases listed below them. A general Crossover constraint was therefore motivated to rule these constructions out as well.[1]

(3) a. *He$_i$ likes every man$_i$.

*∀x. x is a man & x likes x

≈ Every man likes himself.

b. *?His$_i$ mother likes every man$_i$

*∀x. x is a man & x's mother likes x.

≈Every man's mother likes him.

Most authors assume that the strength of Strong crossover is an additive effect. First, a general constraint blocks configurations in which a pronoun precedes the base position of its binder.[2] This is the *Crossover Constraint* that is common to both Strong and Weak

[1] I should note that sentences of this kind are not just a problem for approaches with QR. Any theory that links a QP's binding domain to its scope domain (and not its surface command domain) will have this problem. For example, Categorial Grammars must also invoke a separate principle to rule out Crossover sentences (see Steedman 2003, 2012).

[2] In this discussion I gloss over important historical details of the development of the modern conception of Crossover.

Crossover. This constraint is the only one that is operative in ruling out Weak Crossover constructions.[3] In addition to the Crossover constraint, many researchers have assumed that Strong Crossover configurations run afoul of an additional constraint: Binding Principle C (Chomsky 1976; Chomsky 1981; Shan and Barker 2006; a.o.)[4]. Principle C bars co-indexation between an item and any non-pronominal expression in its c-command domain.

(4) Principle C

An R-expression must be free (it must not be bound).

Principle C rules out binding between the co-indexed phrases in the examples below. In (5-a),(5-b), we see that co-indexation between a pronoun and a name it c-commands is out, whether the two elements are local to one another, or the relation spans a clause boundary. (5-b) shows that this ban on co-indexation under c-command also holds of epithets and their c-commanders. If the pronoun does not c-command the R-expression, as in (6), coreference is acceptable.

(5) a. *He$_i$ likes John$_i$.

b. *He$_i$ promised that Jim$_i$ would go to the store.

c. *The man$_i$ said that the jerk$_i$ wanted to fight.

(6) a. His$_i$ mother likes John$_i$.

[3]For the purposes of this discussion, it is not important exactly how this Crossover Constraint is stated.

[4]The first idea that Crossover and Principle C could be treated as one traces back to Chomksy 1976, where it was first observed that a wh-trace behaves like a name with respect to disjoint reference rules. Thanks to Jeff Lidz for reminding me of this fact.

b. His$_i$ mother promised that Jim$_i$ would go to the store.

c. The man's$_i$ enemy said that the jerk$_i$ wanted to fight.

The same configurations occur in Strong Crossover. In (7-a), the QP cannot bind the pronoun because Principle C is violated at surface structure. In (7-b), the pronoun does not c-command the wh-quantifier on the surface, but it does c-command its base position, or trace.

(7) a. *He$_i$ likes every man$_i$.

b. *Who$_i$ did he$_i$ promise that t$_i$ would go to the store?

Similarly, if the pronoun follows the co-indexed phrase but precedes that phrase's gap, binding is still blocked. This is the case in (2), where the relative pronoun in an intermediate position cannot be linked to the subject pronoun. Because c-command is crucial for defining the illicit relation between the quantifier and pronoun in Strong Crossover constructions, the constraint provides a test of retrieval's sensitivity to relational information.

To my knowledge no work has been done on the processing of Crossover constructions in online sentence-processing. Previous papers have investigated online effectiveness in deploying Principle C to prospectively bar coreference between a pronoun and an R-expression downstream. Testing prospective application of Principle C is orthogonal to assessing retrospective retrieval sensitivity to Principle C. There has been no work that has tested the Principle C sensitivity of retrieval. Below I discuss previous work on Principle C in parsing that is relevant to our work here.

5.1.2 Principle C in Parsing

While most work on Principle C in the formal syntax literature has been based on informal acceptability judgments, the constraint has also been confirmed in controlled laboratory experiments. Gordon and Hendrick (1997) presented evidence from a number of judgment studies that showed that comprehenders recognized configurations under which coreference between pronouns and R-expressions was licit, and those that Principle C blocked. However, the authors showed that additional factors also modulated acceptability, above and beyond simple c-command configurations. They found, for example, that the role of the antecedent NP influenced its availability for later coreference. Antecedents in subject position were less likely to enter into coreferential relations that violated Principle C than NPs in object position.

Some incremental processing studies that tested adult comprehenders' ability to deploy Principle C as a constraint on coreference in real time have been concerned with the resolution of cataphora, or backwards anaphora. The work has investigated the conditions under which a previously-encountered pronoun, usually embedded within a fronted subordinate phrase, is interpreted as coreferent with a linearly subsequent NP.[5]

Cowart and Cairns (1987) used structural ambiguity resolution as a test of the parser's forward application of Principle C in resolving an antecedent for a cataphor. In the sentences below the phrases *charming babies* is ambiguous between a referential NP and gerund parse, but only the gerund parse is compatible with the singular auxiliary.

[5] Children's sensitivity to Principle C has also been an active topic of research in language acquisition (e.g. Crain & McKee 1985; Lust & Clifford 1986; Guasti & Chierchia 1999/2000; Kazanina & Phillips 2001; Leddon & Lidz 2005).

(8) a. Whenever they lecture during the procedure, charming babies is ...

b. Whenever you lecture during the procedure, charming babies is ...

The authors tested whether a cataphoric pronoun (*they*) would be treated as coreferential with a subsequent NP (*charming babies*). Cowart and Cairns reasoned that the need for a sentence-internal antecedent for the cataphor would increase the likelihood that the ambiguous phrase would be analyzed as a referential noun phrase in (8-a), because only a referential noun phrase could act as an antecedent for the pronoun *they*. This preference should not be observed in (8-b), where the indexical pronoun *you* would not require an antecedent.

The index of whether the parser chose the referential NP parse was the reading time at the auxiliary verb *is*. If the referential NP parse were chosen, the parser would expect a plural-marked verb and would slow down at the singular verb. Increased RTs in (a) conditions, relative to (b) would therefore indirectly indicate an attempt at cataphora resolution.

The authors included a pair of test sentences that used the same manipulation to test Principle C sensitivity . In (9) the pronoun c-commands the ambiguous phrase (*visiting uncles*). Therefore, *visiting uncles* should not be considered as an antecedent for the cataphor, and there should be no bias to parse the phrase as a referential NP instead of a gerund. This predicts no differences in processing of the singular verb between conditions.

(9) a. If they want to believe that visiting uncles is ...

b. If you want to believe that visiting uncles is ...

In a word naming task, the authors found longer response latencies to repeat the verb *is* in (8-a) relative to (8-b), but no corresponding pairwise difference between (9-a) and (9-b). They concluded that the parser's search for a cataphor's antecedent results in a bias towards the noun phrase analysis only if the antecedent is in a grammatically accessible position.

While Cowart and Cairn's results suggest that the parser does not consider NPs in positions blocked by Principle C, they were neither completely conclusive nor exhaustive. The effects found were only significant by participants, not items, and their task did not provide indices of processing at any stage other than the critical verb. Finally, as pointed out by Kazanina et al. (2007), the test sentences are all subject to constraints that reduce the probability of reference, both semantic and syntactic. In (8-a), *charming babies* can be coreferent with *they*, but it is an implausible antecedent given the context. More appropriate control conditions would be ones that are not subject to any constraint.

Kazanina et al. (2007) followed up on this work exploiting the *gender-mismatch* paradigm of van Gompel and Liversedge (2003). Van Gompel and Liversedge found that comprehenders generate an expectation for a feature-matching referent in the next upcoming NP position in a sentence after encountering a cataphor in a preposed modifier. When the subsequent NP mismatches the cataphor in gender, a slowdown (dubbed the *gender-mismatch effect*) ensues. Whether this search reflects a simple linear strategy, or one that is structurally-guided was not determined by the authors.

In a series of three studies Kazanina et al. (2007) varied the gender match between a cataphor and the first linearly subsequent NP. The researchers also varied whether the cataphor c-commanded the NP in question, thus manipulating the NPs eligibility to antecede the cataphor according to Principle C. The sentences below are examples from their third experiment. In the materials c-command between the sentence-initial cataphoric pronoun and the R-expression was manipulated via embedding; if the pronoun was the matrix subject, as in (10-a) and (10-b) it c-commanded the R-expression, whereas if it was a genitive modifier (as in (10-c) and (10-d)) it did not.

(10) a. He$_i$ chatted amiably with some fans while *the talented, young quarterback* signed autographs for the kids, but Steve$_i$ . . .

b. She$_i$ chatted amiably with some fans while *the talented, young quarterback* signed autographs for the kids, but Carol$_i$ . . .

c. His$_i$ managers chatted amiably with some fans while *the talented, young quarterback$_i$* signed autographs for the kids, but Carol . . .

d. Her$_i$ managers chatted amiably with some fans while *the talented, young quarterback* signed autographs for the kids, but Carol$_i$ . . .

. . . wished the children's charity event would end soon so she could go home.

RTs beginning at the R-expression *the talented, young quarterback* were of interest. A mismatch effect was found between the two conditions in which the cataphor did not c-command the R-expression in the critical region, with faster RTs in the gender-matching condition (10-c) than in (10-d). The authors found no corresponding difference between the first two conditions. reflecting the confirmation of the expectation and discharge of the dependency. This asymmetric gender-mismatch effect suggested that coreference was considered between an NP and a matching pronoun that preceded it, only if the pronoun did not c-command the NP. In light of these findings, the authors concluded that Principle C could be used as a filter on prospective dependency creation.

Kazanina et al (2006) suggested the prospective nature of the Principle C dependency is responsible for the parser's structural fidelity. The authors analogized this sensitivity to the parser's respect for Island environments during active gap-filling procedures. Previous studies have shown that the parser suspends active gap-filling routines within island boundaries (Stowe, 1986; Traxler & Pickering, 1996; Wagers & Phillips, 2009; Yoshida, 2006). In both cases of filler-gap resolution and cataphoric pronominal processing it appears that the parser recognizes that domains in which the dependency it is currently pursuing should not be considered. It would seem that the crucial aspect of both of their processing requires creating an active expectation for a later element (either an NP, in the case of cataphora, or a gap).

Wagers and Phillips (2009) offered a rationale for the grammatical fidelity of prospective dependency creation: Prospective dependencies respect long-distance structural relations, because dependency formation procedures do not call upon an interference-prone retrieval system. By contrast, retrospective dependencies are subject to ungrammatical interference because their processing necessarily recruits memory mechanisms that cannot employ structural information required for grammatical computation. According to this proposed division, we would therefore expect retrieval not to respect Principle C.

5.2 Experiment 3: Strong Crossover I

Experiment 3 probed retrieval's sensitivity to Strong Crossover/Principle C by testing whether a feature-matching filler whose gap site is c-commanded by a pronoun interferes with the antecedent retrieval process for that pronoun.

5.2.1 Participants

32 students from the University of Maryland community (mean age 20.1 years, 24 female) participated for course credit.

5.2.2 Procedure

Participants were run on a desktop PC using the Linger software package (Doug Rohde, MIT) in a self-paced word-by-word moving window paradigm (Just, Carpenter, & Woolley, 1982). Each trial began with a sentence masked by dashes appearing on the screen. Letters and punctuation marks were masked, but spaces were left unmasked so that word-boundaries were visible. As the participant pressed the spacebar, a new word appeared on the screen and the previous one was re-masked.

A comprehension question followed each sentence. Participants were instructed to read sentences at a natural pace, while avoiding skimming, and to respond to the comprehension questions as accurately as possible. Participants responded to questions using the f-key for 'yes' and the j-key for 'no.' Participants were notified if they answered incorrectly. Each participant was randomly assigned to one of the lists, and the order of the stimuli within the presentation list was randomized for each participant.

5.2.3 Materials

Experimental materials consisted of 24 sets of 4 conditions organized in a 2x2 factorial design crossing the factors GENDERMATCH and CROSSOVER. GENDERMATCH had two levels: the critical pronoun could either *Match* the head noun of the wh-phrase in morphological gender, or *Mismatch* the head noun. The main subject always mismatched the critical pronoun in gender, so as to never be an eligible antecedent. The second factor, CROSSOVER, controlled whether the wh-phrase was associated with a gap that preceded (and c-commanded) the critical pronoun or followed (and was c-commanded by) the pronoun. In *NoCrossover* conditions, the wh-phrase was in the appropriate position to antecede the critical pronoun, whereas in the *Crossover* conditions the wh-phrase's gap fell in the scope of the critical pronoun, precluding such an antecedence relationship from being established despite the wh-phrase preceding the pronoun linearly.

An example test set is given in (11).

(11) a. MISMATCH-NOCROSSOVER

Olivia said which **midwife** in the crowded ward had thought that *he* already spoke with Jeff about the diagnosis.

b. MATCH-NOCROSSOVER

Olivia said which **doctor** in the crowded ward had thought that *he* already spoke with Jeff about the diagnosis.

c. MISMATCH-CROSSOVER

Olivia said which **midwife** in the crowded ward Kaitlyn had thought that *he* already spoke with about the diagnosis.

d. MATCH-CROSSOVER

Olivia said which **doctor** in the crowded ward Kaitlyn had thought that *he* already spoke with about the diagnosis.

A gender-mismatch effect was predicted within *NoCrossover* conditions. Lower RTs, indicative of facilitated processing, at or after the critical pronoun were expected in the *Match* condition condition because the matching wh-phrase can antecede the pronoun. In the *Mismatch* condition longer RTs reflecting processing difficulty are expected in the same region as the pronoun has no feature-matching antecedent in the sentence.

Strong Crossover sensitivity was measured by the presence or absence of a similar gender-mismatch effect within *Crossover* sentences. If retrieval respects Principle C/Strong Crossover, no gender-mismatch effect was expected. Despite matching the pronoun's gender features the filler should be inaccessible, leaving the pronoun without an antecedent. If, however, retrieval does not respect the constraint a gender-mismatch effect was predicted within the *Crossover* conditions.

5.2.4 Analysis

RTs that exceeded 2.5 standard deviations from the mean by region and condition were excluded from the analysis. 4 participants, whose overall accuracy was below 80% on comprehension questions were excluded from later analysis. One item was excluded from analysis due to a typo.

Statistical analysis was carried out using a linear mixed effects regression. The

	Crossover	*NoCrossover*
Match	.91 (.03)	.90 (.03)
Mismatch	.92 (.02)	.89 (.04)

Table 5.1: Experiment 3: Comprehension Question Accuracy

experimental fixed effects were the factors MATCH (whether the wh-filler matched the pronoun in gender) and CROSSOVER (whether the filler's gap site preceded the pronoun *NoCrossover*, or followed it *Crossover*), as well as their interaction. Coding of fixed effects followed a simple difference sum coding scheme (*Match* conditions coded as -0.5, *Mismatch* conditions as 0.5; *NoCrossover* conditions -0.5, *Crossover* as 0.5). Reported coefficients reflect the magnitude of the difference between levels of a given factor in milliseconds. P-values were estimated for the linear models via MCMC sampling using the pvals.fnc() function in the Language R package (Baayen 2008).

5.2.5 Results

Comprehension Questions

The mean comprehension question accuracy for experimental items across participants and items was 90.6%, and did not differ significantly across conditions (logistic mixed effects model, all zs < 1).

Self-paced Reading

Region-by-region self-paced reading times are reported in Figure 5.1. For the purpose of presentation, conditions are grouped into comparison pairs by level within the factor *Crossover*.

At the pre-pronominal complementizer region a significant GENDERMATCH x

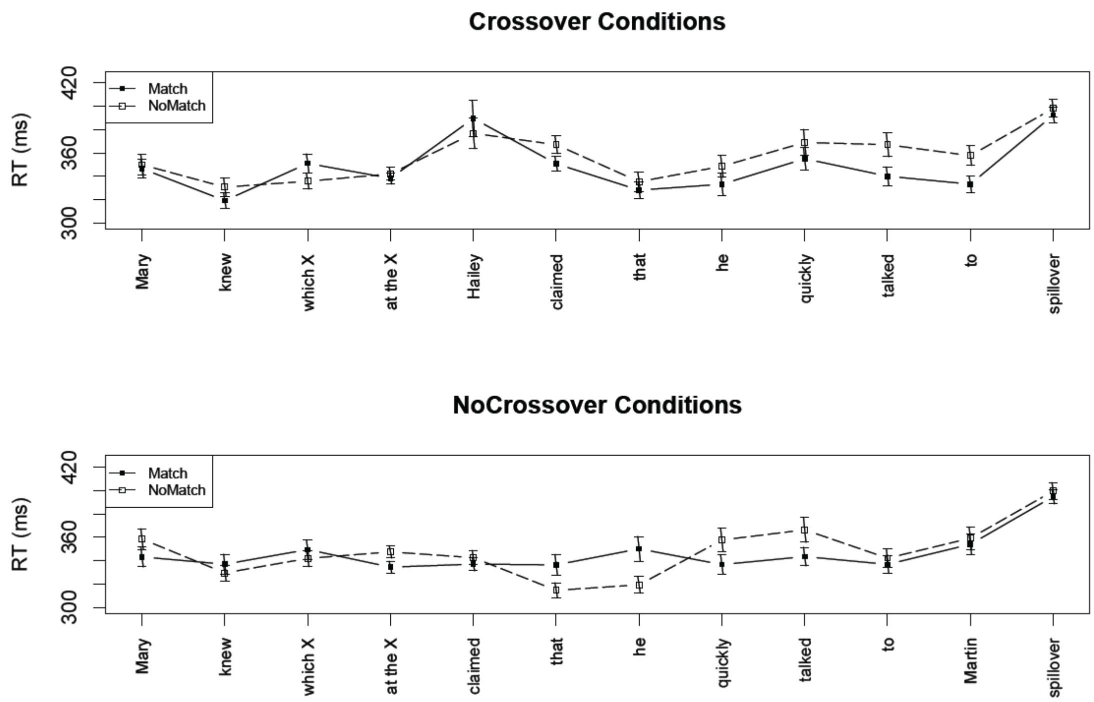


Figure 5.1: Experiment 3: Self-Paced Reading Results. Error bars indicate standard error.

CROSSOVER interaction effect was observed (β = 27.1, s.e. = 12.5, t = 2.17). This effect reflects a pairwise difference within *NoCrossover* sentences, with longer RTs in *Match-NoCrossover* condition (β = -20.6, s.e. = 9.37, t = -2.19). No significant pairwise difference was observed in this region between *Crossover* sentences ($t < 1$).

Prior to the pre-critical region, there were no significant effects of GENDERMATCH, CROSSOVER, or their interaction.

At the critical pronoun no significant main effects were observed ($ts < 1$), but a MATCH x CROSSOVER interaction was significant (β = 45.0, s.e. = 14.7, t = -3.06, $p < .01$). Again, this interaction was driven by significantly higher RTs in the *Match-NoCrossover condition*, relative to the *Mismatch-Crossover* condition (β = -26.7, s.e. = 10.2, t = -2.62, $p < .05$). Elevated RTs in the *Match-Crossover* condition could be attributed to lingering effects of the same pairwise split observed in the complementizer region, and are most likely orthogonal to the question of interest.

The post-pronoun adverbial region displayed a significant effect of GENDERMATCH (β = 20.0, s.e. = 7.97, t = 2.51, $p < .05$) , as well as a marginally significant main effect of CROSSOVER (β = 15.6, s.e. = 7.97, t = 1.89, $p < .10$). The effect of GENDERMATCH was driven by faster RTs in *Match* conditions relative to their *Mismatch* counterparts. Pairwise comparisons revealed this effect was significant in *NoCrossover* conditions (β = 22.3, s.e. = 10.4, t = 2.13, $p < .05$), but not *Crossover* conditions (t = 1.55). This effect appears to be a continuation of the effect observed in the pre-complementizer region.

In the verb region a main effect of GENDERMATCH was significant (β = 25.2, s.e. = 7.52, t = 3.35, $p < .01$). In this region, both pairwise comparisons were significant: *Crossover* (β = 25.5, s.e. = 10.3, t= 2.49, $p < .05$), and *NoCrossover* (β = 24.3, s.e. = 11.0, t= 2.22, $p < .05$).

A numerical pattern similar to the previous region was found in the next, post-verbal, region. *Match* conditions were read faster than *Mismatch* conditions. This effect of GENDERMATCH was significant (β = 16.5, s.e. = 6.15, t = 2.68, $p < .05$). However, this was driven by pairwise differences in the *Crossover* conditions only (β = 25.2, s.e. = 8.11, t = 3.11, $p < .01$).

5.2.6 Discussion

The study sought to test antecedent retrieval's Strong Crossover sensitivity as indicated by its susceptibility to interference from grammatically inappropriate feature-matching wh-fillers.

A pairwise effect was observed between the *Mismatch-NoCrossover* and *Match-NoCrossover* conditions at the pre-critical and critical pronoun region. *Mismatch-Crossover* conditions were read significantly faster than their *Match* counterparts. A later sustained effect of GENDERMATCH was found in the adverbial region immediately following the critical pronoun and in the next region after. In the adverbial region, the main effect of GENDERMATCH was driven by a significant difference between *NoCrossover* conditions. Pairwise effects between *Crossover* conditions emerged later in the sentence, where differences between *NoCrossover* sentences decreased.

The pairwise effect of gender match between *NoCrossover* conditions is in the opposite direction than initially predicted: faster RTs, indicating success of antecedent retrieval for the pronoun, were predicted in the *Match-NoCrossover* condition relative to the *Mismatch-Crossover* condition, where no feature-matching antecedent was present for the pronoun.

The results found are equivocal with respect to the questions of interest to the study. Interpretation of the results is confounded by baseline problems beginning in the region preceding the critical pronoun. As mentioned above, though a significant interaction effect is observed at the critical pronoun region, it is likely that this effect does not reflect processing relevant to our hypotheses because the RT pattern responsible for the interaction effect emerges prior to the pronoun region. Presently, I have no concrete explanation for this effect other than it is spurious noise. Prior to the regions where the two *NoCrossover* conditions diverge, the two are identical save for the gender of the filler head noun. In the *Mismatch-NoCrossover* condition, both filler and matrix subject share gender features.

Filler and matrix subject do not share features in the *Match-Crossover* condition. If this difference is somehow responsible for the RT patterns, I would expect it to have had an effect in the *Crossover* conditions, as well, where the same gender manipulation was also operative.

The interaction effect at the critical pronoun complicates interpretation of later results, because it does not permit a clean baseline against which to make later comparisons. On the one hand, the pairwise difference between the NoCrossover conditions that was reversed in the subsequent region might initially speak in favor of Crossover-sensitivity. Alternatively, lingering effects from the pre-pronominal complementizer region may obscure important pairwise differences between the NoCrossover conditions.

Though the effect was not immediately apparent from the raw RTs, one potentially relevant difference between the *Crossover* and *NoCrossover* conditions occurs in the post-pronoun region. In the *Mismatch-Crossover* condition, there was a sharp increase in RTs at the post-pronoun region, relative to the previous regions. This change in RT profile was not observed in any other condition. This might suggest that readers did encounter difficulty at the pronoun, and that that difficulty resulted in longer RTs in the subsequent region. These longer RTs might have been tempered by the unusually fast RTs in the previous region. If this interpretation is correct, the larger absolute change in RTs across regions would indicate Strong Crossover sensitivity.

The significance of the pairwise difference in the *NoCrossover* conditions, and the fact that they were not significant in *Crossover* conditions might provide suggestive evidence that retrieval does distinguish between fillers in accordance with Strong Crossover. The results could not be used to support a categorical sensitivity to Strong Crossover because of the trend towards a similar pairwise difference in the *Crossover* pairs.

RT profiles in subsequent regions suggest effects of later influence of GenderMatch, irrespective of Crossover. A persistent, though initially non-significant, numerical difference between Crossover conditions emerged at the propositional attitude verb, three regions before the critical pronoun. This divergence continued through the critical region, until the end of the sentence. This pairwise difference contributed to an overall effect of GENDERMATCH throughout the critical region, generally elevating average RTs in the *Mismatch-Crossover* condition. At the post-pronoun verb region, the difference achieved statistical significance, but it is unclear whether these effects were driven by late consideration of the illicit gender-matching filler.

Because I take the effects in the *NoCrossover* condition in the pre-critical and critical pronoun region to be spurious noise I believe it best not to attempt to divine anything from the results of Experiment 3. Experiment 4 was designed as a follow-up to Experiment 3 to minimize the pre-pronominal baseline differences, as well as address other concerns with the design.

5.3 Experiment 4: Strong Crossover II

Experiment 4 was a second test of Strong Crossover sensitivity that followed the same logic as Experiment 3, but used improved materials. In Experiment 3, we saw mixed evidence for the idea that prospective application of Principle C (or a crossover constraint) blocked immediate retrieval of a structurally-inappropriate, but feature-matching wh-phrase as a potential antecedent for a pronoun. In the region immediately following the critical pronoun, the RT profile consistent with the predictions of a Principle C-sensitive retrieval regime was observed. Yet, there were also factors which confounded clean interpretations and prevented simple conclusions. Unexpected baseline differences obscured early comparison between *Match-NoCrossover* and *Mismatch-NoCrossover* conditions. Similarly, interpretation of apparent late effects of GENDERMATCH in *Crossover* conditions could either reflect consideration of the filler as an antecedent for the pronoun in violation of Strong Crossover, or they could reflect separate baseline differences between the conditions that emerged as a numerical trend earlier in the sentence.

Experiment 4 was designed to control for a number of potential confounding factors in Experiment 3. First, the design was intended to hold referential complexity constant across conditions (which the previous experiment failed to do). Gibson (1997) has argued that the difficulty of processing a filler-gap dependency can be related to the number of new referents introduced between the filler and its gap. In Experiment 3 the number of referents that fell within the filler-gap path was not constant across conditions. This confound was eliminated in Experiment 4. Finally, in Experiment 3, 75% of the test sentences contained pronouns that lacked sentence-internal antecedents. If participants were able to subconsciously identify test sentences based on their form, they might have been able to predict the presence of an unheralded pronoun. For example, a participant might have developed a strategy to interpret an upcoming pronoun as referring to an extra-sentential entity, upon encountering a *wh-filler*. The materials in Experiment 4 introduced additional control conditions to balance the number of acceptable and unacceptable test sentences and decrease the likelihood that a pronoun would follow a wh-filler. This would discourage participants from developing implicit expectations for upcoming pronouns based on properties of the test sentences.

5.3.1 Materials

One complication that Experiment 4 set out to address had to do with the number of referents in the test sentences. In Experiment 3, though sentences were lexically matched across conditions, the number of referents differed between *Crossover* and *NoCrossover* conditions. In the former set of conditions, 3 referents preceded the pronoun: the matrix subject, the wh-phrase, and the subject of the second clause. In the latter there were only two: the matrix subject and the wh-phrase. This distinction could impact the ability to make clear cross-condition comparisons of the processing of the pronoun for two reasons: (i) in *NoCrossover* conditions, the wh-dependency spans an additional referent, thereby increasing processing load asymmetrically (see Gibson 1997), (ii) the number of referents that partially match the pronoun in retrieval features (such as [+referential]) might affect cue diagnosticity, or could affect expectations regarding the continuation of the sentence, thereby indirectly shifting expectations for coreference resolution.

To address the potential confound of referent number, Experiment 4 changed the predicate in the intermediate clause within the Crossover conditions. Raising predicates (*it seemed that, it appeared that*) were used, because they permit expletive subjects that do not add to the referential load or complexity of the intervening clause. Predicates in NoCrossover sentences were kept as manner-of-saying, or propositional attitude verbs.

One other change between Experiment 3 and Experiment 4 concerned the proportion of unacceptable, or deviant, sentences in the test set. In Experiment 3 75% of the test sentences were unacceptable - either due to gender mismatch, or crossover violation. If subconscious recognition of the test structures influences their processing, the ungrammatical bias of the test sentences might induce non-standard parsing strategies or routines.

In Experiment 4, two control conditions were added to the set of test sentences. These can be seen in (12-e) and (12-f). Control sentences were identical to *Match* test items except for one change: the critical pronoun in test sentences was replaced with a proper name (*Donna* in the example sentences. This change (i) balanced the proportion of acceptable/unacceptable sentences with similar structures and (ii) disrupted the ability to reliably predict a pronominal subject in the most deeply embedded clause. *Match* sentences were selected as templates for these controls.

Finally, two additional changes were made to the template for materials from Experiment 3. In Experiment 3, wh-phrases were modified by a prepositional phrase (which midwife *in the crowded ward*). No such PPs intervened between filler and gap in Experiment 4 to reduce potential processing load associated with the dependency. A second change lengthened the post-gap region (e.g. the phrase *in the cafeteria* was added to the end of the sentence).

The 24 items from Experiment 3 were taken and changed in the manner described above. In two item sets, the head noun of the wh-filler was changed to ensure gender mismatch. For example, the noun *janitor* in item 12 from Experiment 3 was changed to *maintenance man* in Experiment 4, as seen below in (12). 12 additional item sets were created resulting in a total of 36 experimental items arranged in a 2 x 2 + 2 factorial design.

An example item set is given below.

(12) a. MATCH-CROSSOVER

Jane asked **which maintenance man** it appeared that *he* already spoke with regarding the food-fight in the cafeteria.

b. MISMATCH-CROSSOVER

Jane asked **which lunch-lady** it appeared that *he* already spoke with regarding the food-fight in the cafeteria.

c. MATCH-NOCROSSOVER

Jane asked **which maintenance man** had said that *he* already spoke with Jim regarding the food-fight in the cafeteria.

d. MISMATCH-NOCROSSOVER

Jane asked **which lunch-lady** had said that *he* already spoke with Jim regarding the food-fight in the cafeteria.

e. CONTROL-CROSSOVER

Jane asked **which maintenance man** it appeared that *Donna* already spoke with regarding the food-fight in the cafeteria.

f. CONTROL-NOCROSSOVER

Jane asked **which maintenance man** had said that *Donna* already spoke with Jim regarding the food-fight in the cafeteria.

The logic of the previous experiment's predictions applied to Experiment 4. A pair-wise gender-mismatch effect is expected between *NoCrossover* test conditions, indicating sensitivity to the presence or absence of a intra-sentential antecedent for the pronoun. Crossover sensitivity was to be assessed based on the presence or absence of a similar pairwise difference in *Crossover* conditions. A gender-mismatch effect within *Crossover* sentences would indicate interference from the gender features of the ungrammatical filler, and therefore insensitivity to Strong Crossover. On the other hand, a lack of gender-mismatch effect would indicate that retrieval does not access feature-matching fillers in violation of the Strong Crossover constraint.

5.3.2 Experiment 4a: Acceptability Judgment Study

An acceptability judgment study was run to (i) experimentally verify Crossover judgments and (ii) make sure that baseline difference in Experiment 3 were not driven by some unforeseen unacceptability or difficulty with *NoCrossover* conditions.

5.3.2.1 Participants

12 participants completed an offline acceptability judgment study. 7 participants came from the University of Maryland Linguistcs participant pool. These participants completed the judgment study task after having completed an unrelated self-paced reading experiment as part of a one-hour session. Participants received course credit. The remaining 5 participants were recruited through the Amazon Mechanical Turk (AMT) marketplace and paid $4.00 for their participation. For AMT participants eligibility was determined in two ways. First, AMT automatically restricted access to the experiment to those with IP addresses corresponding to US-internal locations. Second, participants were required to take a Native speaker evaluation test and to score above a threshold of 7/10.

5.3.2.2 Materials

36 test sentences were interspersed among 60 fillers of comparable length and complexity for a total of 96 sentences.

5.3.2.3 Procedure

Presentation used the IBEX farm online experimental presentation platform, developed and managed by Alex Drummond (www.spellout.net/ibexfarm/docs). Sentences were presented one at a time centered on the screen. Below each sentence was a series of 7 boxes corresponding to values on a 7-point acceptability scale. Each box was marked with a number, 1 -7, in ascending order. Next to the left-hand side of the scale was a label 'bad', whereas the right-hand side of the scale bore the label 'good'. Participants could record a response by either clicking a box, or pressing a number on their keyboard. As in previous studies, participants were given practice sentences prior to beginning the exercise. Two practice sentences drew participants' attention to an unheralded pronoun, and explained how the pronoun made the sentence unacceptable.

5.3.2.4 Results

Average acceptability judgments by condition are given in Table 5.2.

A linear-mixed effects model with fixed effects for CROSSOVER and GENDERMATCH and random intercepts for Subject and Item was fit to the data in non-control conditions. The model revealed a significant main effect of CROSSOVER (β = -1.1, s.e. = .13, t = -8.56, $p < .001$), a main effect of GENDERMATCH (β = -.64, s.e. = .18, t = -3.52,

COND	Avg. Rating	S.E.
Match-Crossover	2.14	.18
Mismatch-Crossover	2.15	.18
Match-NoCrossover	4.34	.22
Mismatch-NoCrossover	3.05	.23
Control-Crossover	3.73	.23
Control-NoCrossover	4.85	.22

Table 5.2: Experiment 4: Mean Acceptability Ratings

$p < .01$), and a significant CROSSOVER x GENDERMATCH interaction ($\beta = .66$, s.e. = .18, $t = 3.60$, $p < .001$). Pairwise comparisons between *Match-* and *Mismatch-Crossover* conditions were not significant ($t < 1$). A pairwise comparison between *Crossover* conditions was significant ($t = -4.85$, $p < .01$).

These results confirm that participants do not regard covaluation of a *wh*-phrase and a feature-matching pronoun in a Crossover configuration acceptable. Importantly, although there is a main effect of CROSSOVER independent of GENDERMATCH, the lowered acceptability of the pronoun-containing *Crossover* sentences cannot be reduced to the complexity of the construction in general. Control conditions serve as an independent baseline of syntactic complexity, as well. *Control-Crossover* conditions are significantly less acceptable than *Control-NoCrossover*, but they are still more acceptable than either of the *Crossover* test conditions.[6]

[6] Pairwise comparisons were conducted between the *NoCrossover-Control* condition and the *NoCrossover* test conditions individually. A significant difference was found only between *NoCrossover-Control* and the *Mismatch-NoCrossover* condition ($t = 5.58$, $p < .001$; *Match* condition $t = 1.59$, p n.s.). Pairwise comparisons between *Control-Crossover* and *Match-Crossover*, as well as between *Control-Crossover* and *Mismatch-Crossover* conditions, were significant (Match, $t = 5.50$, $p < .001$; *Mismatch* $t = 5.38$, $p < .001$).

5.3.3 Experiment 4b: Self-Paced Reading Task

Having established that comprehender's judgments of test items align with the judgments of Crossover presented in the formal syntax literature, their online behavior is now at issue.

5.3.3.1 Participants

30 students from the University of Maryland community participated for course credit.

5.3.3.2 Procedure

The procedure for Experiment 4b was identical to that of Experiment 3.

5.3.3.3 Analysis

Data treatment was the same as in Experiment 3.

5.3.3.4 Results

5.3.3.5 Comprehension Question Accuracy

The mean comprehension question accuracy for experimental items across participants and items was 94.8%, and it did not differ significantly across conditions (logistic mixed effects models, all z values < 1).

	Crossover	*NoCrossover*
Match	.94 (.02)	.95 (.01)
Mismatch	.94 (.01)	.92 (.02)
Control	.94 (.02)	.97 (.01)

Table 5.3: Experiment 4: Comprehension Question Accuracy

5.3.3.6 Self-paced Reading Results

Control Conditions

A pairwise difference between control conditions was significant in the predicate region ($t = -2.14$, $p < .05$). *Control-NoCrossover* was read significantly more slowly than the *Control-Crossover* condition. In the last three regions of the sentence *Control-Crossover* sentences were read significantly more quickly than in *NoCrossover-Control* sentences (ts = -2.48, -2.17, -2.73; all ps $< .05$).

Test Conditions

No effects emerged at the complementizer region immediately preceding the critical pronoun in any of the comparison pairs.

At the critical pronoun region, a significant main effect of CROSSOVER was observed (β= 9.54, s.e. = 4.08, t = -2.20, p <.05). *NoCrossover* sentences were read more slowly on average than *Crossover* sentences. A marginally significant GENDERMATCH x CROSSOVER interaction was also found in this region (β= -7.89, s.e. = 4.08, t = -1.94, p <.10). This interaction reflects the fact that the *Match-Crossover* condition was read slightly more slowly than *Mismatch-Crossover*, but slightly faster than *Mismatch-*

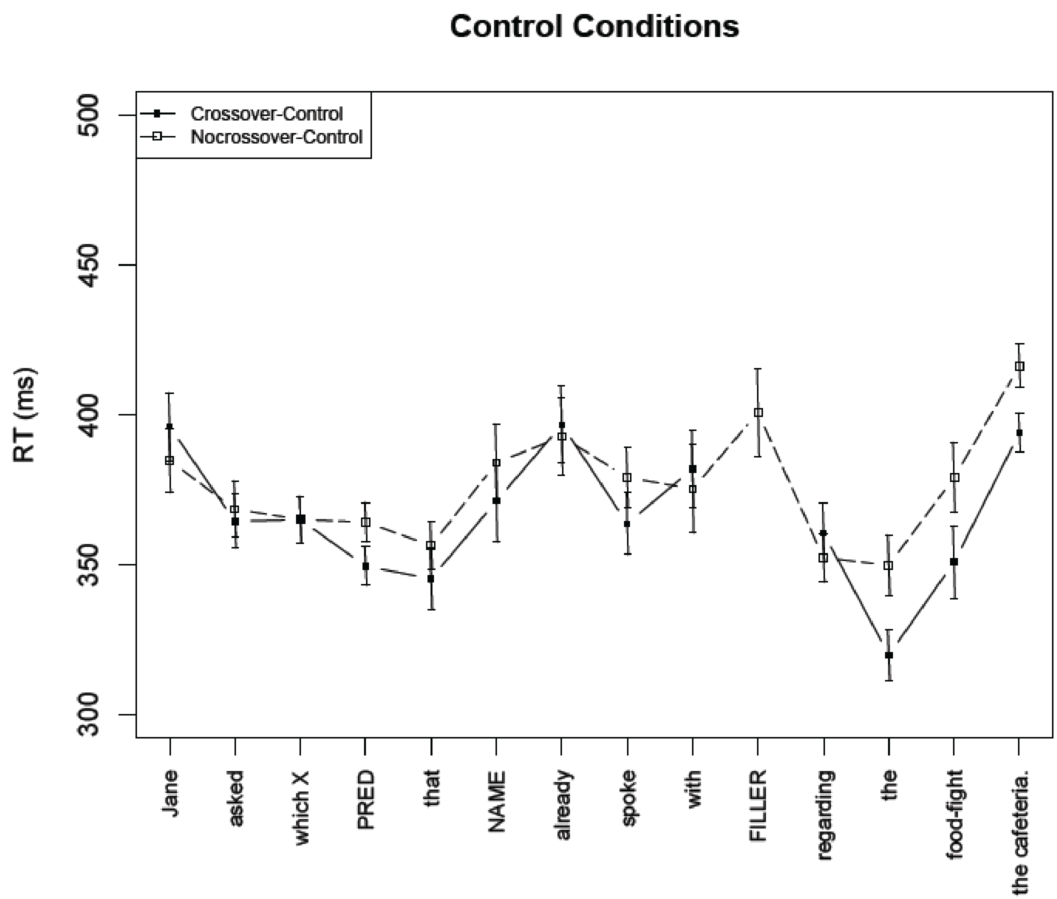


Figure 5.2: Experiment 4 Control Conditions: Self-Paced Reading Results. Error bars indicate standard error.

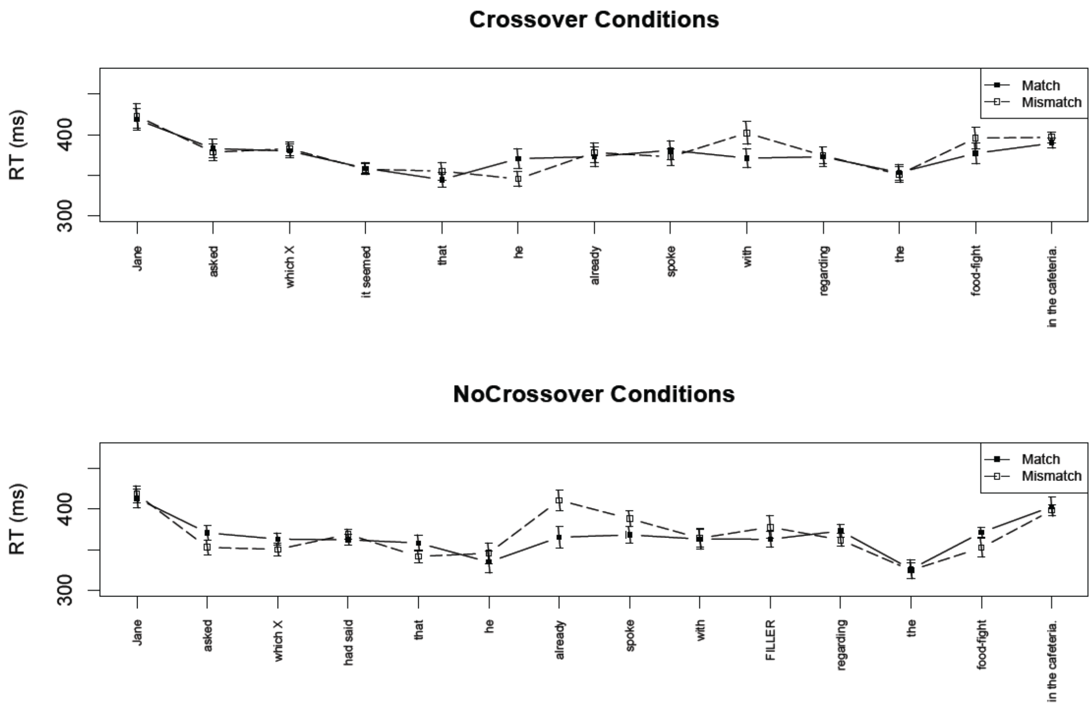


Figure 5.3: Experiment 4 Test Conditions: Self-Paced Reading Results. Error bars indicate standard error.

NoCrossover. *Crossover* conditions differed marginally in pairwise comparisons ($t = -1.80$, $p < .10$), but not in *NoCrossover* sentences ($t < 1$).

In the adverb region directly following the critical pronoun, a main effect of GENDERMATCH was significant ($\beta = 12.92$, s.e.$= 4.91$, $t = 2.63$, $p < .05$). This effect was driven by a pairwise difference between *Match* and *Mismatch* conditions within *NoCrossover* sentences. *Mismatch-NoCrossover* sentences were read significantly more slowly than *Match-NoCrossover* ($|\text{difference}| = 48$ms). This difference resulted in a significant GENDERMATCH x CROSSOVER interaction ($\beta = 12.10$, s.e.$= 4.92$, $t = -2.46$, $p < .05$).

No significant main effects were found at the region following the adverb (all ts $<$ 1), nor interaction effects (ts $<$ 2), in spite of the numerical pattern that mimics the RT profile observed in the previous region.

In the post-verbal region an effect of CROSSOVER achieved significance ($\beta = 11.26$, s.e. = 4.97, t = 2.27, $p < .05$). This effect reflects overall longer mean RTs in *Crossover* conditions, driven asymmetrically by elevated RTs in the *Mismatch-Crossover* condition relative to *Mismatch-NoCrossover*. The numerical trend towards higher RTs in the *Mismatch-Crossover* condition, relative to its *Match* counterpart was only marginal in the pairwise comparison (t = 1.78, $p < .10$).

5.3.4 Discussion

The experiment tested whether wh-fillers that matched a pronoun in morphological features would be accessed by antecedent retrieval in violation of the Strong Crossover constraint.

Within *NoCrossover* conditions, participants retrieved the grammatical feature-matching filler as an antecedent for a pronoun, and exhibited difficulty at the pronoun when no feature-matching filler was present. In *Crossover* conditions, the feature-matching filler was not retrieved as an antecedent for a pronoun. Processing of the pronoun did not differ as a function of the gender-match between filler and pronoun.

Two additional effects were observed in *Crossover* conditions. First, the pronoun was read marginally more slowly when it matched the filler in gender than when it did not. Second, at the gap site reading times were faster when the grammatically illicit filler matched the pronoun in gender than when it did not.

The results immediately following the critical pronoun are consistent with the conclusion that retrieval is Strong Crossover sensitive. The lack of a gender-match effect in *Crossover* conditions suggests that antecedent retrieval did not contact the matching structurally inappropriate filler.

The two additional effects merit further discussion. The marginally significant increase in RTs at the critical pronoun in the *Match-Crossover* condition is not easily characterized as either facilitatory or inhibitory interference. Although the slowdown might suggest inhibition, it does not pattern with other cases of inhibitory interference which emerge when grammatically unavailable constituents reduce the accessibility of an otherwise grammatical item. In *Match-Crossover* sentences, no grammatical intra-sentential antecedent exists for the pronoun. The configurations tested were most conducive to producing facilitatory interference - processing of the otherwise unheralded pronoun should have been *eased* by mis-retrieval of the matching filler, but these effects were not observed.

One possible interpretation of the effect is that it reflects a post-retrieval interpretive operation (though this interpretation would go against the general line of argument to be pursued). On this interpretation, mis-retrieval of the filler occurs, but its ungrammaticality as a potential antecedent for the pronoun is realized relatively quickly, leading to temporary difficulty. One argument against this hypothesis relates to the time-course of the effect relative to the effect of *Match* in *NoCrossover* conditions. In *NoCrossover* conditions we observe the effect of successful retrieval and interpretation after the pronoun. All things equal, we would expect indices of early interpretation to be simultaneous. The time-course of the slowdown observed in *Match-Crossover*, then, is too early to be an interpretive effect.

Kazanina & Phillips (2010) found a superficially similar GENDERMATCH effect in their conditions investigating the backwards application of the *poka*-constraint, a language-specific constraint in Russian on pronominal coreference similar to Principle C insofar as it restricts coreference between a pronoun and a subsequent name. In conditions where no constraint blocked coreference between a cataphor and a subsequent subject, a gender mismatch effect was observed when the cataphor's gender was manipulated. In conditions where Principle C ruled out coreference between cataphor and the proximate subject, no effect of the cataphor's gender was found following the subject. However, in the *poka*-constraint condition, an effect of gender match emerged: longer reading times were found in the *Gender-Match* condition than in the *Gender-Mismatch* condition. The authors argued that the gender-match effect in the constraint condition reflected constraint application that followed initial attempts to link matching cataphor and subject. The *poka*-constraint of Kazanina and Phillips' also differs from the one tested here in another important regard that makes the initial consideration-late constraint approach feasible is that the Russian sentences were still consistent with a coreferential interpretation between pronoun and name at the point when the matching name was encountered. It is only after the case/thematic role of the argument is confirmed that coreference becomes impossible. For this reason, initially considering coreference is grammatical at the point when the name is encountered. In the Crossover cases considered here, there is no stage at which covaluation between filler and pronoun would ever be grammatical. Therefore, initial consideration of the relation is less likely. Due to the different directionality of constraint application in our study and Kazanina and Phillips', as well as the differences in constraint application, I believe that the slight GENDERMATCH effect observed in our studies cannot be attributed to the same source as Kazanina and Phillips' match effect.

The second effect – the facilitated reading times in the *Match-Crossover* condition at the gap site, relative to the *Mismatch-Crossover* condition – may suggest late consideration of the Crossover-violating interpretation. Previous studies have found that initial grammatical-sensitivity can give way to later effects suggesting violation of grammatical principles. Sturt (2003) found that early indices of initial sensitivity to Principle A in antecedent retrieval, which were then followed by patterns that suggested comprehenders considered relations between reflexives and grammatically-illicit feature-matching NPs.[7]

5.3.5 Does Antecedent Retrieval Simply Ignore Unintegrated Fillers?

Despite the small puzzles presented by the pairwise effects that emerge in *Crossover* conditions, the evidence against early interference from Crossover-violating fillers seems relatively sound. Now we must consider the manner by which this apparent sensitivity is obtained. According to formal grammatical description, this sensitivity is achieved via reference to the c-command relation between the pronoun and the filler's resolution site. However, we would ideally like to achieve this sensitivity without direct recourse to relations. Two general possibilities exist for explaining the results observed. The first is to find a method to encode the relevant structural information for distinguishing appropriate from inappropriate fillers using a feature. The second possibility is to use a feature that encodes a 'crossover constraint' that does not make reference to structural appropriateness *per se*. Here I explore one family of explanations that falls in the second category.

[7] The late effect might also be taken to be an indication of eased integration or retrieval of the filler as a result of earlier interference. Suppose that, despite my arguments, that the feature-matching filler was erroneously retrieved at the pronoun. This mis-retrieval would result in an activation boost for the filler in *Match-Crossover* cases, which would thereby speed the constituent's retrieval at the gap site.

Suppose that a filler's unintegrated status rendered it inaccessible for antecedent retrieval. There are a number of ways this could be brought about. Fillers could occupy a distinct position in memory inaccessible to antecedent retrieval operations. Wanner and Maratsos' (1978) 'hold cell' is one such distinct working-memory buffer that is separate from long-term procedural/episodic memory. On the assumption that access to this special hold cell is restricted to filler retrievals only, the retrieval would be insulated from potential interference from the filler. This particular implementation runs up against evidence that wh-fillers are not stored separately (see the discussion of the issue in chapter 2), and the emerging view that the role of distinct buffers should be limited (e.g. Nairne, 2002; McElree, 2006; Jonides et al., 2008).

A separate implementation that does not make a distinction between the storage site for fillers and other items in memory would preferentially weight, or 'gate' retrieval access to items based on, the feature [filler], which could tag unintegrated wh-phrases. This type of approach is a mechanistic implementation of a general Crossover constraint. One such constraint has been proposed by Shan and Barker (2006), who give an account of Crossover strictly in terms of *evaluation order*. A quantifier is not eligible to bind a pronoun unless its reconstructed position linearly precedes the pronoun's. In our implementation linear order would be taken as proxy for evaluation order.

If the filler's unintegrated status is the determining factor for the apparent lack of interference in strong Crossover sentences, we would expect all filler-pronoun antecedence relations to be categorically blocked. Because the unintegrated filler hypothesis makes no reference to the structural position of the pronoun, we should be able to vary the pronoun's position on this account and not expect any modulation of interference. On the other hand, if the immunity to interference is related to Principle C sensitivity, we would expect changing the structural position of the pronoun in relation to the gap site would have an effect on interference. The next experiment tests these hypotheses.

5.4 Experiment 5: Weak Crossover

Experiment 5 tested whether a feature-matching filler would interfere with antecedent retrieval for a pronoun that preceded, but did not c-command, the filler's gap site. The experiment was intended to test the hypothesis that the lack of interference in Experiment 4 (and Experiment 3) could be attributed to the filler's unintegrated status, instead of Strong Crossover sensitivity. The logic of the experiment is as follows: If the feature-matching filler's inaccessibility in Experiment 4 stems from its being unintegrated, manipulating the critical pronoun's c-command relation with the filler's gap should not impact retrieval processes at the pronoun. The feature-content of the filler should be equally invisible to retrieval. If, however, the inaccessibility of the filler was due to Strong Crossover/Principle C sensitivity, disrupting the c-command relation between pronoun and gap should render the filler accessible. Weak Crossover constructions are the appropriate test of these hypotheses because the constructions are identical to Strong Crossover constructions in all respects save the c-command status of the pronoun and gap site.

As noted By Wasow (1972) and subsequent researchers, Weak Crossover sentences in which the pronoun does not c-command the gap are less deviant than sentences in which the pronoun does c-command the gap. Though occurrence in natural speech does not serve as a completely reliable indicator of relative acceptability, the occasional occurrence of weak crossover constructions does provide suggestive evidence that the constructions may be slightly less deviant than Strong Crossover constructions.[8]

(13) He was the type of man with *whom his* work would always come first.

(Agatha Christie, *Remembered Death*, p.58 from Wasow 1972, p.137)

That such sentences are occasionally encountered might be suggestive evidence that comprehenders do entertain a relation between pronoun and quantifier.[9] However, anecdotal evidence of their occurrence does not provide much insight into the processing of these structures. It is perfectly consistent with the facts to assume that such relations are initially blocked in left-to-right parsing, but are considered in later processing.

5.4.1 Materials

Materials from the previous experiment were changed minimally to create Weak Crossover configurations. Subject pronouns from the previous materials were replaced with NPs with pronominal possessors (*he* in the previous experiment becomes *his supervisor* in the current experiment). This is seen in (16). In this configuration the pronouns do not c-command the filler's gap site, as they are embedded inside a DP.

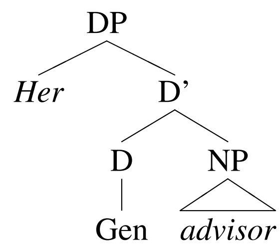
[8] This description is somewhat anachronistic. Wasow's description of the facts did not make reference to the notion of c-command.

[9] Wasow notes that a cline of acceptability exists even between questionable weak Crossover constructions, with bare wh-fillers at the low end of acceptability, and contentful, complex filler phrases being more acceptable.

i) a. **Who* did the woman [*he* loved] betray?
   b. ?*Which man* did the woman [*he* loved] betray?

ii) a. **What* did the artist [who painted *it*] refuse to sell?
    b. ?*Which picture* did the artist [who painted *it*] refuse to sell?

(14)



Examples in (15) show that the pronoun does not c-command out of the possessor position. Possessors of NPs in subject position can be coreferential with NPs in object position, indicating lack of Principle C effects (and therefore lack of c-command).[10] Similarly, NPs in possessor position cannot bind reflexives, which also indicating lack of c-command.

(15) a. His$_i$ mother likes John$_i$.

b. *John$_i$'s book pleases himself$_i$.

Proper names used in the embedded subject position in the materials of previous experiments were replaced with definite NPs whose head noun matched the possessed NP in test conditions, as shown in (16-e) and (16-f). An additional auxiliary was also added between the critical subject NP and the verb to lengthen the post-pronoun critical region.

(16) a. MATCH-CROSSOVER

Jane asked which maintenance man it appeared that *his supervisor* had already spoken with regarding the food-fight in the cafeteria.

[10]As Gordon & Hendrick (1997) note, however, these covaluations are still dispreferred.

b. MISMATCH-CROSSOVER

Jane asked which lunch-lady it appeared that *his supervisor* had already spoken with regarding the food-fight in the cafeteria.

c. MATCH-NOCROSSOVER

Jane asked which maintenance man had said that *his supervisor* had already spoken with Jim regarding the food-fight in the cafeteria.

d. MISMATCH-NOCROSSOVER

Jane asked which lunch-lady had said that *his supervisor* had already spoken with Jim regarding the food-fight in the cafeteria.

e. CONTROL-CROSSOVER

Jane asked which maintenance man it appeared that *the supervisor* had already spoken with regarding the food-fight in the cafeteria.

f. CONTROL-NOCROSSOVER

Jane asked which maintenance man had said that *the supervisor* had already spoken with Jim regarding the food-fight in the cafeteria.

If fillers are uniformly inaccessible, reaction time profiles should parallel those observed in Experiment 4. However, if the inaccessibility of the filler in Experiment 4 was linked to the c-command relation between pronoun and gap site, the feature-matching filler should be accessible as an antecedent for the genitive pronoun and an effect of gender-match within *Crossover* should be observed.

COND	Avg. Rating	S.E.
Match-Crossover	2.86	.18
Mismatch-Crossover	2.54	.18
Match-NoCrossover	4.36	.21
Mismatch-NoCrossover	3.11	.21
Control-Crossover	3.88	.17
Control-NoCrossover	4.50	.17

Table 5.4: Experiment 5: Mean Acceptability Ratings

5.4.2 Experiment 5a: Acceptability Judgment Study

5.4.2.1 Participants

12 participants (mean age = 33.6, 6 male) were recruited through the Amazon Mechanical Turk (AMT) marketplace, and paid $4.00 for their participation. Participants' IP addresses were restricted to US-internal locations and participants were deemed eligible only if they scored above a threshold on a native speaker evaluation test commonly used in laboratory settings.

5.4.2.2 Results

Average acceptability judgments by condition are given in Table 5.4.

A comparison of test conditions revealed main effects of MATCH and CROSSOVER (β = -0.82, s.e. = 0.16, t = -5.04, $p < .001$; β = -1.00, s.e. = 0.17, t = -5.94, $p < .001$, respectively). The MATCH x CROSSOVER interaction indicative of Weak Crossover sensitivity was also observed (β = -1.03, s.e. = 0.34, t = -3.06, $p < .01$). The numerical difference between the two *Crossover* conditions was not significant in pairwise comparison (β = -0.31, s.e. = 0.20, t = 1.64, $p > .10$), but the difference between *NoCrossover* conditions was significant (β = -1.34, s.e. = 0.26, t = -5.20, $p < .001$).

Pairwise comparison revealed that the effect of CROSSOVER was significant in *Control* conditions (β = -0.63, s.e. = 0.19, t = -3.24, $p < .01$). *Control-Crossover* sentences were reliably judged to be less acceptable than *Control-NoCrossover* sentences. This preference for *NoCrossover* sentences is also observed in the test conditions.

The results confirm the generalization from the formal syntax literature that Weak Crossover constrains pronoun interpretations. The general pattern of acceptability scores roughly parallels the pattern found for Strong Crossover configurations in Experiment 4a. Interestingly, apart from the numerical difference between *Match-Crossover* and *Mismatch-Crossover*, the results do not provide much support for the general view that Weak Crossover sentences are somewhat more acceptable than Strong Crossover sentences.

5.4.3 Experiment 5b: Self-paced Reading

5.4.3.1 Participants

30 participants from the University of Maryland community participated in the experiment in exchange for course credit. 2 participants' data were excluded for not meeting the threshold score on a native speaker assessment test (7/10 correct responses). The data of one participant was excluded due to history of cognitive impairment. One additional participant's data were excluded due to low accuracy on comprehension questions. Data from the remaining 26 participants was analyzed as discussed below.

	Crossover	*NoCrossover*
Match	.81 (.02)	.86 (.02)
Mismatch	.81 (.02)	.88 (.02)
Control	.82 (.03)	.86 (.03)

Table 5.5: Experiment 5: Comprehension Question Accuracy

5.4.3.2 Procedure

Presentation used the IBEX farm online experimental presentation platform, and followed the same procedure as Experiment 4a.

5.4.3.3 Analysis

Participants whose average accuracy on comprehension questions fell below an 80% cut-off threshold on all items (fillers and test items) were excluded from analysis. This resulted in the exclusion of one participant from later analysis. One item was excluded due to a typo. All other analysis steps were identical to the previous experiments.

5.4.3.4 Results

Comprehension Questions Overall comprehension question accuracy was 84%, with accuracies in individual conditions reported below.

A logistic mixed-effects model reveals a main effect of CROSSOVER in test conditions (β = z = -2.24, p $<$ .05), but no other main effects (all zs $<$ 1). Comprehension question accuracy in *Crossover* sentences, irrespective of GENDERMATCH, was significantly lower than in *NoCrossover* sentences. A numerical trend in this direction is also observed in Control conditions, but the pairwise difference was not significant. This indicates constructional difficulty is behind the effect of *Crossover*, and not the presence of a crossed-over pronoun.

Self-paced Reading Times

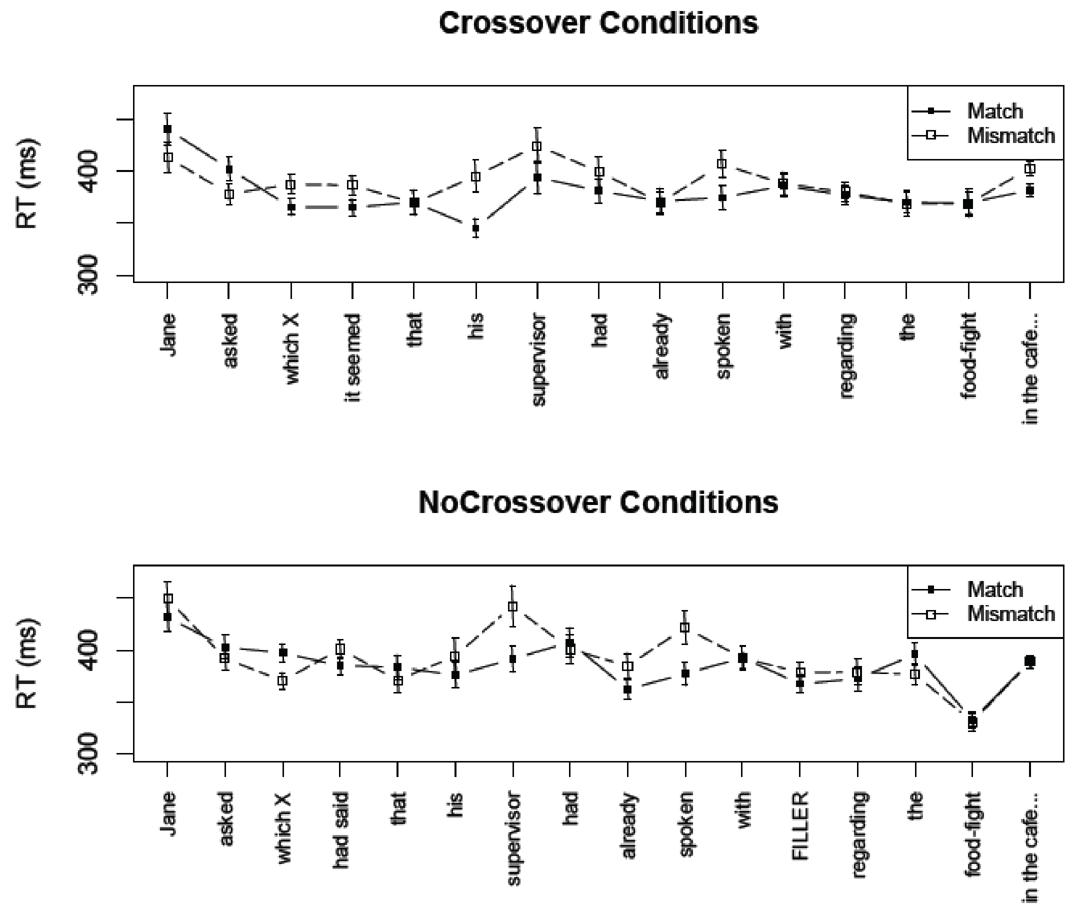


Figure 5.4: Experiment 5: Self-paced Reading Results

In the complementizer region immediately preceding the critical pronoun, there were no significant effects (all ts $<$ 1).

In the critical genitive pronoun region, a main effect of GENDERMATCH was significant (β = 35.3, s.e. = 12.5, t = 2.83, p $<$ .01). This effect was due to lower mean RTs in *Match* conditions in comparison to their *Mismatch* counterparts. This effect was disproportionately driven by the significant pairwise difference between *Crossover* conditions (β = 49.5, s.e. = 15.8, t = 3.13, p <.01). Pairwise comparison between *NoCrossover* conditions did not reveal a significant effect (β = 19.9, s.e. = 18.4, t = 1.08). Effects of CROSSOVER and the MATCH x CROSSOVER interaction failed to achieve significance (β = -15.1, s.e. = 12.5, t = -1.20; β = 31.2, s.e. = 25.0, t = 1.25).

At the post-pronoun nominal region, a main effect of GENDERMATCH was again observed (β = 42.2, s.e. = 15.4, t = 2.81, p < .05). As in the preceding region, the effect was driven by lower average RTs in *Match* conditions than *Mismatch* conditions. However, the pairwise difference between conditions follows a different pattern than in the previous region. In this region, a significant pairwise difference emerges between *NoCrossover* conditions (β = 51.6, s.e. = 21.9, t = 2.36, p < .05), but not between *Crossover* conditions (t = 1.51). No other effects were significant (*t*s < 1).

A GENDERMATCH effect persists when both pronoun and post-pronoun region are pooled (β=38.6, s.e. = 9.77, t=3.95, p < .01), and both pairwise comparisons reach significance: within *Crossover* (β= 40.84, s.e. = 13.23, t = 3.09, p < .01) and within *NoCrossover* (β= 35.6, s.e. = 14.2, t = 2.51, p < .05).

No significant effects were found in either adverbial or auxiliary region.

At the intransitive verb, there was a main effect of GENDERMATCH (β= 35.9, s.e. = 11.2, t= 3.20, p < .01). Pairwise effects of GENDERMATCH were found in both *Crossover* condition pairs and *NoCrossover* condition pairs (β= 29.6, s.e. = 14.2, t= 2.08, p < .05; β= 43.8, s.e. = 17.1, t= 2.56, p < .05, respectively). In both pairs, *Match* conditions were read significantly faster than their *Mismatch* counterparts. No other significant effects were found (all *t*s < 1).

5.4.4 Discussion

A gender-mismatch effect was observed in both the *NoCrossover* and *Crossover* pairs. *Match* conditions were read significantly ore quickly than their *NoMatch* counterparts immediately at the critical pronoun or in the subsequent region. These pairwise conditions persisted as non-significant trends in the next two regions, before re-emerging at the post-pronoun verb.

The results suggest that comprehenders access a feature-matching wh-filler as a potential antecedent if that filler does not c-command the filler's gap. Access to the feature-matching filler is allowed in spite of Weak Crossover. These results are inconsistent with explanation of the results from Experiment 4 that makes all unintegrated fillers inaccessible. More generally, it is inconsistent with an account that posits the same mechanism to explain the deviance of Crossover structures (both strong and weak) as ruled out by a 'processing constraint' that privileges left-to-right semantic evaluation (e.g. Shan and Barker 2006).

Comprehension question results suggest that *Crossover* sentences presented comprehenders with more difficulty than *NoCrossover* configurations. This effect was independent of the gender of the filler, and was also observed in the *Control-Crossover* condition, where no pronoun was present. This suggests that the locus of difficulty was not in processing the unheralded pronoun. Rather, it would be easiest to attribute the decrement in acceptability to the length of the dependency computed in *Crossover* sentences. Within *NoCrossover* conditions, the wh-filler was resolved immediately after being encountered, whereas *Crossover* conditions required its storage and maintenance across two clauses.

The effect of GENDERMATCH at the intransitive verb in both conditions can be seen as a continuation of the pattern in the pronoun and post-pronoun region. The fact that a pairwise difference in both *Crossover*, as well as *NoCrossover*, conditions suggests that the effect at the verb cannot be related to the process of gap-filling, as the wh-filler is already integrated by the time the embedded verb is encountered in *NoCrossover* conditions. I suggest that the best way to make sense of the effect is to consider it a continuation of the previous gender-mismatch effect from the pronoun and post-pronoun region that is attenuated at the auxiliary and the adverb. Having registered difficulty at the the subject, it could be that comprehenders accelerate through processing the auxiliary and adverbial in order to find the next content word in the sentence: the verb. At the verb, however, their difficulty is not resolved, so effects re-emerge.

5.4.4.1 Weak Crossover Violation v. Resumptive Pronouns

One concern that might arise in the interpretation of the mismatch effects at the pronoun in Experiment 5 relates to the possibility that comprehenders initially mis-analyzed the pronoun as a resumptive pronoun, instead of a bound pronoun. Resumptive pronouns have been argued to stand in for gaps in positions from which syntactic movement is prohibited, or degraded (Kroch 1981; Chao and Sells 1983; Prince 1990; Demirdache 1991, among many others). For example, it has been noted that fillers can be linked to resumptive pronouns inside Islands, where filler-gap dependencies are blocked.

(17) a. *Which man$_i$ did Mary say that she had met the lady who met t$_i$?

b. ??Which man$_i$ did Mary say that she had met the lady who met him$_i$?

As extraction of possessors from their containing DPs is impossible in English (see Ross 1967), resumptive use is licensed in these environments.

(18) a. John said that he met Samantha's mother.

b. *Whose did John say that he met t_i mother?

c. ??Which girl did John say that he met her mother?

Zukowski and Larsen (2004) have noted that resumptive pronouns in possessor position are produced with some regularity in production, in particular when distance between extracted filler and extraction site is increased. If we suppose that comprehenders are willing to posit resumptive pronouns in positions where they might be encountered in a natural language setting, it might seem possible that the participants in Experiment 3 analyzed the pronoun in Experiment 5 as a resumptive. Rather than parsing the Weak Crossover sentence (19-a) as intended, the comprehenders would have instead committed to a parse related to the one in (19-b):

(19) a. Jane asked which maintenance man it seemed that *his supervisor* had spoken with ...

b. Jane asked which maintenance man's supervisor it seemed ___ had spoken with...

If this is the case, the gender mismatch effect observed indicates not consideration of a bound reading in violation of the crossover constraint, but instead the attempt to resolve a resumptive pronoun. A resumptive analysis is not available incrementally in the

Strong Crossover items from Experiment 4, as the pronoun occupies a position from which extraction is licit. That is, simple subject extraction never licenses resumptives.[11]

(20) Which maintenance man did it seem (*he) had spoken ...

The principal evidence that speaks against this hypothesis comes from the absence of diffculty at the gap-position in *Match-Crossover* sentences. On analogy to standard gap-filling procedures, we can assume that consideration of a resumptive pronoun analysis would *discharge* the filler. This would entail that active-gap filling would be turned off, and encountering the filler's *true gap* downstream would cause difficulty. If the filler had been discharged at the pronoun in *Match-Crossover* sentences, reading times following the preposition indicating the true gap site should have been elevated, indexing surprise. It would appear to the parser that the subcategorization frame of the preposition was violated, because there is no overt object for it.

5.5 General Discussion

The results show that antecedent retrieval does not consider feature-matching fillers that violate Strong Crossover. However, antecedent retrieval does allow access to feature-

[11]The presence of the complementizer 'that' in the Strong Crossover items might lead one to suspect a resumptive strategy would be possible for the purposes of repairing a *that-trace* effect. It has long been known that in (many dialects of) English, extraction of a subject immediately following a complementizer results in unacceptability:

- a. Who does John think __ likes Mary?
  b. Who does John think that __ likes Mary?

If comprehenders are sensitive to that-trace effects, one might expect they would possibly entertain a resumptive analysis following a complementizer. However, this possibility is challenged by acceptability judgment findings by Keller & Alexopoulou (2005) who argued that use of a resumptive could not 'save' a that-trace violation.

matching fillers in violation of Weak Crossover. The distinction that retrieval draws appears to track whether the pronoun that triggers the retrieval c-commands the filler's gap site.

Experiment 3 presented equivocal evidence in support of Strong Crossover sensitivity. It appeared that there was slight evidence that retrieval distinguished grammatically accessible from inaccessible fillers, however interpretation of these effects was complicated by spurious fluctuations in reading times in the pre-critical region.

In Experiment 4, there were clear differences in the effects of gender features between *Crossover* and *NoCrossover* conditions. Gender mismatch effects were found, as expected in *NoCrossover* conditions immediately following the critical pronoun, but no such effects were observed in *Crossover* conditions. This suggests that antecedent retrieval was largely immune to interference from feature-matching, but Principle C-violating fillers.

Experiment 5 showed that comprehenders violate Weak Crossover by retrieving gender-matching fillers as antecedents for pronouns that precede the fillers' gap sites. This shows that initial retrieval for pronominal antecedents does not observe a general crossover constraint, and that unintegrated wh-fillers are accessible to general antecedent retrieval. These results argue in favor treating the lack of interference in Experiment 4's Strong Crossover conditions as a function of the c-command sensitivity of retrieval.

As in previous experiments, it is important to consider whether there is a plausible method for obtaining the observed results without requiring retrieval to make reference to relational information. The simplest candidate mechanism, that the filler's unresolved status somehow prevented it from being retrieved, was ruled out by Experiment 5. As discussed at the start of the chapter, an explanation based on the LOCAL feature might initially appear inappropriate given the unbounded nature of Strong Crossover/Principle C.

The problem posed by unbounded constraints to LOCAL-based strategies is a real one, but I argue that in the particular case of Crossover/Principle C, a retrieval strategy that exploits the LOCAL feature is capable of making the appropriate distinctions. I argue that although Crossover/Principle C must be stated as an unbounded constraint over static representations in the formal grammar, an incremental parser can employ a clause-bounded strategy to restrict retrieval in the necessary way. Essentially, Strong Crossover/Principle C sensitivity can be achieved using the same mechanisms as were used to produce Principle B sensitivity.

5.6 Accounting For Crossover Sensitivity Using the Anti-Clausemate Constraint

The idea behind the LOCAL account of differential sensitivity to Strong and Weak Crossover is that fillers whose gap site a pronoun c-commands are treated as clause-mates of the pronoun by an incremental parser that pursues an active gap-filling strategy. Because pronouns require antecedents that are not clause-mates, covaluation between the filler and pronoun is blocked. In cases of Strong Crossover, the filler and pronoun are judged to be clause-mates, in Weak Crossover cases they are treated as non-clause-mates. In order for this solution to work, it is necessary to specify why and how the filler would be considered a clause-mate of the pronoun.

Much work has established that incremental parsers pursue an active gap-filling strategy (Frazier 1978, Frazier, Clifton & Randall 1983, Stowe 1986, Frazier & Flores-d'Arcais 1989, Traxler & Pickering 1996; Aoshima, Phillips & Weinberg 2004). Parsers attempt to resolve a gap in the closest argument position possible, before receiving bottom-up confirmation of a definite gap site. I posit that by predictively positing a gap site in the closest clause, a parser also commits to encoding the current filler as a constituent of the current clause. This entails that active fillers should always bear the feature LOCAL: 1. If active fillers are marked LOCAL:1, they will be inaccessible to antecedent retrieval for a LOCAL:1 pronoun, in accordance with the retrieval strategy for pronouns discussed in Chapter 4, reprinted below.

(21) ANTI-CLAUSEMATE RETRIEVAL:

When retrieving an antecedent for a pronoun P, use the following features for LOCAL:

- If P is LOCAL: 1 $\rightsquigarrow$ Probe [LOCAL: 0]
- If P is LOCAL: 0 $\rightsquigarrow$ Probe [LOCAL: NULL]

As an illustration, consider the step-wise description of the incremental parse through a Strong Crossover example from Experiment 4 given in Figure 5.5.

In accordance with the active gap-filling strategy the fronted *wh*-filler *which janitor* should be marked with the feature [LOCAL: 1]. After encoding the filler, the parser moves through the sentence word-by-word until it reaches the complementizer *that*, which informs the parser that a new clause has been reached. Usually, according to the LOCAL UPDATE RULE, all items in memory marked LOCAL:1 would be revalued as LOCAL: 0. We can assume that fillers are not revalued at clause edges and are considered LOCAL until they are integrated, on the assumption that the parser maintains the expectation that they

...which janitor ① it seemed that ② he ③ ...

① - input: *which janitor* - encode: $\begin{bmatrix} Local : 1 \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$
② - input: *that* - increment clause counter
③ - input: *he* - cue retrieval: $\begin{bmatrix} Local : 0 \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$ - RETRIEVAL FAILS!

Figure 5.5: Parsing Strong Crossover Sentences With Anti-clausemate Condition

will be discharged in the (new) current clause. Moving forward, the parser encounters the pronoun, triggering antecedent retrieval. Retrieval enforces the anti-clausemate restriction using [LOCAL: 0]. The filler bears [LOCAL: 1], so it is blocked from retrieval.

A similar parse occurs with Weak Crossover sentences, but the filler accessed by retrieval. In Weak Crossover sentences everything proceeds as in the example sentence above up to the pronoun retrieval. Because the pronoun in Weak Crossover sentences is a genitive pronoun and not a member of the clause's main spine, it should not be marked LOCAL:1. Instead it should bear LOCAL:0. The retrieval probe, according to the retrieval rule above, should therefore accept NPs with either LOCAL:1, or LOCAL:0 values. Therefore, the filler is predicted to be accessible for the same reason that *his* and *John* can be co-valued in the sentence *His mother loved John*.

. . . *which janitor* ① *it seemed that* ② *his* ③ . . .

①
- input: *which janitor*
- encode: $\begin{bmatrix} Local : 1 \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$

②
- input: *that*
- increment clause counter

③
- input: *his*
- cue retrieval: $\begin{bmatrix} Local : \text{NULL} \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$
- retrieve: *which janitor* $\begin{bmatrix} Local : 1 \\ Gender : Masc. \\ Number : Sing. \end{bmatrix}$

Figure 5.6: Parsing Weak Crossover Sentences With Anti-clausemate Condition

5.7 Conclusion

This chapter antecedent retrieval's sensitivity to both Strong and Weak Crossover was tested, as a final test of retrieval's ability to respect constraints on anaphoric dependencies stated in terms of c-command. Returning to the three questions of interest we began with, we can ask whether retrieval distinguishes between feature-matching fillers that can grammatically serve as antecedents, and those that are blocked by Strong Crossover. Here, our results provide relatively clear evidence that the accessible and inaccessible fillers are treated differently by retrieval. We found no evidence of facilitatory interference from ungrammatical items in Strong Crossover-violating constructions, suggesting that the effect is categorical, rather than gradient.

With Weak Crossover constructions, we found entirely different results. At the earliest possible point of observation it appeared quite obvious that retrieval directly accessed feature-matching fillers irrespective of whether they were grammatically-accessible, or not. These results suggest that online approximation of the Strong Crossover constraint is not due to the inaccessibility of fillers as antecedents.

Though the distinction between Weak Crossover and Strong Crossover is one of c-command in grammatical terms, I argued that the differential effects do not require a retrieval mechanism that is itself c-command-sensitive. Instead, I sketched a possible account of the sensitivity based on the LOCAL feature, which was previously motivated as a feature to be used in implementing clause-bounded relational constraints. Perhaps counterintuitively, though Principle C and Strong Crossover are taken to be constraints stated over potentially unbounded distances by formal grammarians, the method used to rule out interference from Principle-C/Crossover-violating fillers is an entirely local computation. It makes reference to a local domain (the current clause) and rules out the co-valuation of a pronoun and filler via an anti-clausemate constraint.

Chapter 6: Variable Binding, C-command, and Scope

"Scope! Scope! Scope! Scope! ... Polemoscope! Scope! ... Telescope! Scope! ..." - Bauhaus, 'Scopes'[1] (as sung by Alexander Williams)

6.1 Introduction

The constraint on reflexive licensing can be cashed out in terms of a clause-mate condition easily coded as a content feature. A stronger test of general relational sensitivity requires testing application of a constraint that applies over an *unbounded* domain. If there is no constraint on the domain of application, translating the local domain of the trigger and the intended item to be retrieved will be of no use.

Bound variable pronouns (BVPs) are a suitable binding dependency for this test, because their licensing is not subject to a locality constraint. A quantified noun phrase (henceforth, QP) can bind a feature-matching pronoun across a potentially unbounded number of clauses.

(1) No linguist$_1$ said (Mary thought) Jill liked him$_1$.

The distribution of bound variable pronouns is also relationally constrained. Though

[1] From: Bauhaus. (1980). *In the Flat Fields*. 4AD Records.

feature-match is a necessary condition for binding, it is not sufficient. The pronoun must be in the right position vis-a-vis the QP. Reinhart (1983) held that the operative constraint was one of c-command.[2]

(2) BOUND ANAPHORA CONDITION

Quantified NPs and *wh*-traces can have anaphoric relations only with pronouns in their c-command syntactic domain. (Reinhart 1983, p.122)

This c-command constraint, which is taken to hold of surface structure configurations (Büring 2005), captures the fact that QPs embedded inside RCs cannot bind pronouns outside that RC ((3)). It also explains why QPs can't scope across sentential conjuncts ((4)).

(3) a. No researcher$_1$ that the linguists disagreed with thought they would try to convince him$_1$.

b. *The linguists that no researcher$_1$ disagreed with thought they would try to convince him$_1$.

(4) a. Every senator$_1$ at the party thought that he$_1$ would have no trouble getting elected.

b. *Every senator$_1$ was at the party and he$_1$ was worrying about getting elected. [Culicover and Jackendoff, 1997, p. 204]

[2] This contrasts with coreference between a pronoun and a referential noun phrase, which does not require c-command.

Though the constraint on bound variable licensing is often stated in terms of c-command, there is not a complete consensus that this characterization is correct. Recently, Barker (2012) has criticized the notion of a c-command constraint on bound-variable anaphora. Instead of a c-command constraint, Barker maintains that bound pronouns must obey a weaker condition: the Scope Condition.[3]

(5) Scope Condition

A pronoun must be in the scope of a quantifier to be bound by it.

While c-command is apparently a sufficient condition for scope, it is not a necessary condition. To prime intuitions Barker notes that in traditional cases of quantifier scope ambiguity, X neednt c-command Y (overtly) in order to scope over Y. The existential in the example below can take wide scope, despite being c-commanded by the universal.

(6) Every woman loves some man.

$\forall >> \exists$

$\exists >> \forall$

If a c-command constraint is required to test retrieval's relational sensitivity, then Barker's arguments would appear to threaten BVP's candidacy as a test dependency. In this brief chapter I discuss Barker's objections to a c-command constraint, and various ways to make the scope constraint more precise. I argue that Barker's objections do not undermine BVP as an eligible test dependency. The take-home point is simple. Whether

[3] In this regard, Barkers proposal is very similar to that of Safir (2004a,b).

c-command or ‘semantic scope’ is the relevant relation that must hold between a QP and pronoun in order to license binding, the constraint is still inherently *relational*, i.e. its statement requires reference to emergent relations between two items in a representation.

The remainder of the chapter is broken into 3 parts. The first reviews Barker’s counterexamples to the (surface) c-command constraint. I agree that the majority of Barker’s examples are legitimate cases of binding in the absence of surface c-command.[4] I then move on to discuss the semantics of quantifier scope, and two methods for computing it: one that does not make reference to c-command, and one that does make reference to c-command (at LF). I conclude by discussing the structural constraints on scope that bear responsibility for making any constraint on BVP inherently relational.

6.2 Barker (2012)

Barker (2012) presents a number of examples where a quantifier can take scope over a pronoun, despite not c-commanding the pronoun at surface structure. These configurations also license BVP in the absence of surface c-command. The majority of these examples had already been discussed in the literature prior to Barker’s paper, but Barker was the first to assemble them all in a coherent salvo against a surface c-command constraint.

Barker’s alternative *Scope Condition* simply requires that a pronoun be in the semantic scope of a quantifier in order to be bound by it. As Barkers paper is offered as a general critique, and not a specific endorsement of a particular approach to scope, he does not outline an explicit method for computing scope. Rather, he offers an operational test of

[4] I do argue that Barker’s argument that QPs can bind pronoun out of finites clause are not true instances of binding.

scope-taking that presumably holds across various implementations.

(7) Operational test for scope

A quantifier can take scope over a pronoun only if it can take scope over an existential inserted in the place of the pronoun. (Barker 2012: 619)

To illustrate this test in action, take the two sentences below, both provided by Barker. In the first sentence *each woman* can scope over the embedded subject pronoun, in the second, it cannot.

(8) Each$_1$ woman denied that she$_1$ met the shah.

(9) The man who traveled with each$_1$ woman denied that she$_{*1}$ met the shah.

The binding possibilities parallel the ability of *each woman* to scope over an indefinite in the pronouns place.

(10) Each woman denied that someone met the shah.

$\forall >> \exists$

(11) The man who traveled with each woman denied that someone met the shah.

$*\forall >> \exists$

I now review Barker's counterexamples, one-by-one. For each counter-example

I demonstrate that (i) surface command does not obtain between QP and pronoun, (ii) the covariation between QP and pronoun is an actual instance of binding, and (iii) the operational test for scope does confirm the QP scopes over the pronoun.

6.2.1 Possessives

Barker points out, following Higginbotham (1980) and Reinhart (1983:177178), that quantificational possessors can bind pronouns that fall in the command path of their DP containers.

(12) a. [Everyone$_1$s mother] thinks he$_1$s a genius.

b. [No one$_1$s mother-in-law] fully approves of her$_1$

c. [[[Everyone$_1$s mother]s lawyer]s dog] likes him$_1$.

The QP does not c-command the pronoun, as shown by the lack of Principle C effects in the constructions.

(13) a. [His$_1$ mother] thinks John$_1$ is a genius.

b. *He$_1$ thinks John$_1$ is a genius.

We can show that these sentences display the behavior of true instances of BVP using the *Identity Substitution* test. It is well known that replacing a BVP in a sentence by an occurrence of the QP that binds it doesn't preserve the truth conditions of the original sentence ((14)a).[5] This differs from cases of coreference, where such replacement does

[5]The two sentences in ((14)a), while both true, are obviously not equivalent. The first statement concerns syntacticians' self-confidence, while the second concerns syntacticians' estimations of their colleagues.

preserve truth conditions ((14)b).

(14) a. No syntactician$_1$ thinks he$_1$ is wrong. $\neq$ No syntactician thinks no syntactician is wrong.

b. John$_1$ thinks he$_1$ is wrong. $\approx$ John thinks John is wrong.

Binding out of possessive examples behave like true cases of binding. Replacing the pronoun in ((12)) changes the truth conditions of the sentence.

(15) [Everyone$_1$s mother] thinks he$_1$ is a genius. $\neq$ [Everyones mother] thinks everyone is a genius.

Barkers test for scope displays the expected parallels between binding and quantifier scope.

(16) a. [Everyones mother] thinks someones a genius.

b. (?)[No ones mother-in-law] fully approves of someone.

c. [[[Everyones mother]s lawyer]s dog] likes someone.

6.2.2 Inverse Linking

Inverse-linking, a phenomenon in which a QP inside a PP modifier of another QP can bind a pronoun outside its c-command domain, was first discussed by May (1977, 1985).

(17) [Someone from every$_1$ city] hates it$_1$.

Above, *every city* can bind the pronoun *it*. Principle C effects do not arise in these configurations.

(18) [Someone next to him_1] coughed on John_1.

Inverse-linking constructions pattern with true cases of variable-binding with respect to Identity Substitution.

(19) [Someone from every_1 city] hates it. $\neq$ [Someone from every_1 city] hates every city.

Finally, the test for scope shows that *every city* scopes over the rest of the sentence. Below, hot dog vendors covary with cities, implying that *every city* scopes over the object of *hates*.

(20) [Someone from every_1 city] hates *a different hot dog vendor*.

6.2.3 Binding out of DP

QPs contained within PP modifiers of N can apparently bind pronouns outside their c-command domain. The test for scope shows that these QPs outscope their containers.

(21) a. The policemen turned a citizen of each_1 state over to it_1s governor. (Gawron and Peters 1990:163)

b. [The cost of each_1 item] was clearly marked on it_1s label.

(22) a. The policemen turned a citizen of each_1 state over to *a different sheriff*.

b. [The cost of each$_1$ item] was clearly marked on *a different price-tag*.

One concern with the above examples is that they feature the quantifier *each*, which has been shown to have extraordinary scope-taking properties not shared by other QPs. These judgments persist even when the QP is changed from *each* to *every*.

(23) a. The policemen turned a citizen of every$_1$ state over to it$_1$s governor.

b. [The cost of every$_1$ item] was clearly marked on it$_1$s label.

(24) a. The policemen turned a citizen of every$_1$ state over to *a different sheriff*.

b. [The cost of every$_1$ item] was clearly marked on *a different price-tag*.

Finally, these examples also pattern like true cases of binding according to identity substitution.

(25) a. The policemen turned a citizen of every$_1$ state over to it$_1$s governor. $\neq$ The policemen turned a citizen of every$_1$ state over to every state$_1$'s governor.

b. [The cost of every$_1$ item] was clearly marked on it$_1$s label. $\neq$ [The cost of every$_1$ item] was clearly marked on it$_1$'s label.

6.2.4 Binding out of PP

Binding without c-command also appears possible out of a number of PPs, be they sentence-initial, arguments of a verb, or adjuncts.[6,7]

(26) a. [After the name of every$_1$ student] will be added his$_1$ place of residence.

b. Our staff keeps a watchful eye [on every$_1$ situation] and on it$_1$s developments.

Identity Substitution shows that these cases are true cases of binding.

(27) After the name of every$_1$ student will be added his$_1$ place of residence. $\neq$ After the name of every student will be added every student's place of residence.

The test for scope confirms that the QP does scope out of the PP.

(28) After the name of every$_1$ student will be added a single star.

$\forall >> \exists$

[6]Note, as discussed in Kazanina et al. (2006) and elsewhere, this kind of binding only occurs when the PP does not obligatorily reconstruct into a position commanded by the pronoun its QP binds.

[7]Barker notes that PPs functioning as adverbial adjuncts also license such binding, particularly in the frame *before* ... *QP* ..., *after* ... *pronoun*. The following are naturally occurring examples from Barker (2012:624):

- [After unthreading each$_1$ screw], but before removing it$_1$, make sure to hold the screw in place while seperating [sic] the screw from the driver.
- ... [after seeing each$_1$ animal] but before categorizing it$_1$ on the computer or recording it$_1$ on their response sheet.
- ... [after fetching each$_1$ pointer], but before dereferencing it$_1$.
- These processors use branch prediction techniques to forecast the code path that will be followed [after each$_1$ branch instruction], but before it$_1$s execution

6.2.5 Binding out of VP

VPs are also among the containers out of which QPs can bind pronouns they do not overtly c-command (All examples come from Barker 2012, unless otherwise noted).

(29) a. We [will sell no_1 wine] before it_1s time. [syntax lore (ad for Paul Masson)]

b. John [left every_1 party] angry at the person who had organized it_1. (Kayne 1994:71)

c. A book [was given to every_1 boy] by his_1 mother. (Harley 2003:64)

That the QP does not c-command the pronouns in the examples above is evidenced by the fact that parasitic gaps are grammatical in the same configurations. Parasitic gaps are subject to an anti-command constraint at surface structure (Engdahl 1983).

(30) Which wine will we sell before tasting?

Identity Substitution shows these are true cases of binding.

(31) We will sell no_1 wine before it_1s time. $\neq$ We will sell no_1 wine before no wine's time.

QPs may also scope out of a VP into an adjunct, shown again using Barker's test (with a subset of the examples from above).

(32) A book [was given to every_1 boy] by *a different person*.

$\forall >> \exists$

6.2.6 Binding out of Finite Clauses?

Barker discusses one final case which he takes to offer further proof that QPs can bind in the absence of c-command. These examples all involve the Q *each* , which, Barker notes, is generally more permissive with respect to the type of binding it allows than other quantifiers. In these examples, *each* appears to scope out of a finite clause, a feat usually assumed to be impossible for QPs (May 1977, Heim 1982).

(33) a. But the actual thinking seems to be [that each_1 person owns his own body], and that he_1 may not alienate his own body, by selling it, and that no one may buy…

b. [That each_1 person is a unique individual] and that he_1 alone can work out his own individuality?

c. It is only nowadays, when strife prevails [that each_1 person needs his neighbour], and that he_1 [n]eeds to pray for peace.

Despite the apparent binding here, it appears that binding out of finite clauses is not a general phenomenon. The judgments cannot be reproduced with other quantifiers, such as *every*, despite its similarity to *each*.

(34) a. ??But the actual thinking seems to be [that every_1 person owns his own body],

and that he_1 may not alienate his own body, by selling it, and that no one may buy…

b. *It is only nowadays, when strife prevails [that $every_1$ person needs his neighbour], and that he_1 [n]eeds to pray for peace.

Moreover, Identity Substitution shows that these cases do not behave like true instances of binding. Substituting the pronoun that covaries with QP with another instance of the QP preserves the original meaning of the sentence.

(35) a. But the actual thinking seems to be that $each_1$ person owns his own body, and that he_1 may not alienate his own body … ≈

But the actual thinking seems to be that $each_1$ person owns his own body, and that each person may not alienate his own body …

b. It is only nowadays, when strife prevails that each $person_1$ needs his neighbour, and that he_1 [n]eeds to pray for peace. ≈

It is only nowadays, when strife prevails that each $person_1$ needs his neighbour, and that each person [n]eeds to pray for peace.

So, it would appear that cases of apparent binding out of finite clauses are not a general phenomenon. As such, I assume that these examples should not be assimilated into the set of phenomena our theory of bound variable licensing should account for. They should instead be taken as further examples motivating a distinct account of the scope-taking properties of *each*.

Summing up, it appears that any account of BVP must be able to handle not only cases where c-command obtains between QP and pronoun, but also the following 5 phenomena.

1. Possessor Binding

2. Inverse Linking

3. Binding out of {VP, DP, and PP}.

6.2.7 A Previous Attempt: Accessibility

One approach to accommodating (some) cases of variable binding without surface c-command was proposed by Higginbotham (1980 et seq.). Higginbotham argues that binding QPs must be accessible to the pronouns they bind, I repeat Büring's (2005) formulation of *accessibility* below:

(36) ACCESSIBILITY (Büring 2005, building on Higginbotham 1980, 1983, 1985, 1987)

NP_1 is *accessible* to NP_2 iff:

- NP_1 c-commands NP_2 from a non-derived/A position (*direct accessibility*)
- NP_1 binds a trace accessible to, or within an NP accessible to, NP_2

(37) CONDITION ON VARIABLE BINDING (Büring's *Crossover Filter*)

NP_1 can bind NP_2 only if it is accessible to NP_2

This notion of accessibility rules in many of the cases of binding in the absence of surface c-command. In ((38)a), the QP is indirectly accessible to the pronoun *him* because it is contained in the DP *nobody's mother*, which c-commands (is directly accessible to) the pronoun. The QPs in ((38)b,c) are similarly indirectly accessible, as their container NPs c-command the pronouns.

(38) a. *Nobody*'s mother-in-law likes *him*.

b. The cost of *every item* is stamped on *its* underside.

c. Someone in *every city* hates *it*.

Accessibility, unfortunately, does not cover all of the problematic cases. Binding out of VP and PP are predicted unacceptable. The NP *no wine* is not *directly accessible* to *its*, because it does not overtly c-command into the *before* adjunct (as shown by the fact that parasitic gaps are possible in the configuration - Engdahl 198x). Similarly, the DP *name of every student* does not c-command the pronoun *his* from within the sentence-initial PP.

(39) a. We sell *no wine* before *its* time.

b. Which wine did we sell *t* before having tried *e*?

(40) [After the name of every$_1$ student] will be added his$_1$ place of residence.

Accessibility cannot cover all cases encompassed by the Scope Constraint because it is not coextensive with scope. Only a portion of scope-taking QPs are *accessible*.[8] In

[8] To be fair, *Accessibility* wasn't designed to cover all violations of a surface c-command constraint on BVP, only those that involved QPs that scope out of a container DP.

order to get a handle on what the scope constraint actually entails, we now discuss how quantifier scope can be computed.

6.3 Ways of Scope-Taking

A QP can be seen as a function that takes a property of individuals and returns a proposition with a bound variable. A quantifier's scope domain corresponds to the argument that the quantifier is passed the level of interpretation. In the sentence below *eat cheese* is the scope domain of the quantifier, because $\lambda x.x\_eats\_cheese$ is the quantifier's argument.

(41) *Every man* eats cheese.

$every\_man(\lambda x.x\_eats\_cheese)$

The scope constraint just requires that a pronoun be inside the quantifier's argument so that it can be bound.

(42) *Every man* likes his mother.

$every\_man(\lambda x.x\_likes\_x's\_mother)$

Barker's main objection to surface c-command constraints on BVP stems from the observation that a quantifier's surface c-command domain and its scope domain are not (always) identical. Barker (2002) calls this 'scope displacement'. For example, a quantifier in object position can take scope over items in its containing clause, despite not c-commanding those items. Consider the sentence below, in which *every apple* takes the open predicate $\lambda x.John\_ate\_x$ as its argument. *John* would appear to be in the scope domain of *every apple*, despite not being c-commanded by it at surface structure.

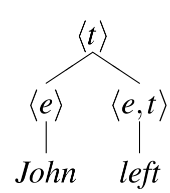
(43) John ate every apple.

$every\_apple(\lambda x.John\_ate\_x)$

Scope displacement arises as a result of competing syntactic and semantic needs. QPs pattern distributionally just like ordinary NPs with respect to the syntactic positions they occupy. Yet, QPs are interpreted semantically in a very different way from ordinary NPs.

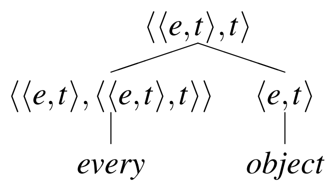
In a simple type-driven semantics (Montague 1974, Heim and Kratzer 1998, a.o.) there are two simple types: $\langle e\rangle$ - the type corresponding to individuals and objects in the world, and $\langle t\rangle$ - the type corresponding to truth values. Higher types, denoting functions from type to type, can be created recursively. For example, the types $\langle e,t\rangle$ can be created for one-place predicates, $\langle e,\langle e,t\rangle\rangle$ for two-place predicates and so on. Constituents of type $\langle t\rangle$ can be created by composing items of the right types via function application, as shown in the sentence below.

(44)



Extending a basic system like this to accommodate quantified noun phrases complicates matters. Mostowski (1957), Lindström (1966), Barwise & Cooper (1981) and many others argued that quantificational determiners should have a fixed type $\langle\langle e,t\rangle,\langle\langle e,t\rangle,t\rangle\rangle$. Quantifiers always compose with predicates of type $\langle e,t\rangle$, and return something of type $\langle\langle e,t\rangle,t\rangle$. Such quantifiers are called 'generalized quantifiers'.

(45)



The problem that generalized quantifiers introduce is this. Because of their fixed type, generalized quantifiers want to compose with properties of individuals and return a proposition. QPs are found, however, in a number of positions where they cannot directly compose with properties of the right type. That is, they are often found in positions where their sister constituent is not $\langle e,t\rangle$. The solution to this type-mismatch problem is relatively simple. To avoid type-mismatch, the grammar must contrive a method for getting the quantifier to compose with a constituent corresponding to a property of individuals. This can be done in a few ways.

6.3.1 Scope As Order of Composition

One solution to the problem of type-mismatch engendered by generalized quantifiers is to allow *type-raising* operations. Such operations adjust the type individual elements in a sentence, allowing them to compose in 'non-standard' ways. This allows for direct, *in-situ* composition. Analyses of quantifier interpretation in frameworks such as Combinatorial Categorial Grammar (Steedman 2003, 2011) go this route.[9] This is also the tack advocated by Barker.

Under a CCG analysis *order of composition*, rather than c-command, determines scope. For example, for *every apple* to take scope over the rest of the sentence, it must compose (at surface structure) with the open predicate $\lambda x.John\_ate\_x$, which corresponds with the CCG constituent of type *S/DP*. Lexically, *John* is type DP. The verb *eat* has the type *(S\DP)/DP*, which states that it composes with a DP argument to its right, to return a constituent of type *S\DP* (something which composes with a DP to its left to return a sentence). Given their lexical types, *John* and *ate* cannot compose together to form a constituent of the right type for later composition with the QP. This is fixed by raising the type of *John* to *((S/DP)/(S\DP)/DP)*. *John* is lifted - it goes from denoting an entity, to denoting a function $\lambda P.\lambda x.P(jx)$.

(46)

$$\dfrac{\dfrac{\dfrac{John}{(S/DP)/((S\backslash DP)/DP)} \qquad \dfrac{ate}{(S\backslash DP)/DP}}{S/DP} \qquad \dfrac{every\_apple}{S\backslash(S/DP)}}{S}$$

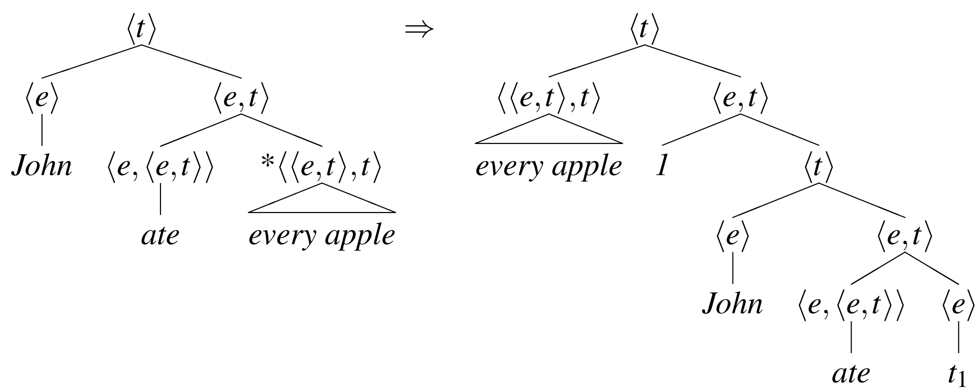
Here, *every apple* need not c-command the rest of the sentence, because the evaluation order specifies that the QP will take the entire sentence as its argument. Note that this solution allows the quantifier to compose with a constituent denoting a property of individuals (λx. John ate x), by fashioning a ‘non-traditional’ constituent (*John likes*).

[9] Hendricks (1987) and Partee (1986) advocate another way of using raising operations to permit direct, or in-situ composition. I will not discuss the particulars of these accounts here, but note that while the proposals do differ in some theoretical details, they are similar to the CCG approach in general form.

6.3.2 Scope as C-command at LF

Another way to handle type-mismatch without resorting to type-raising is to assume that quantifiers are not interpreted in their base position. May (1977 a.o.) argued that quantifiers must move to a position adjoined to their minimal container S at a post-Surface Structure level *Logical Form* through a process of quantifier raising (QR). Movement of the quantifier leaves a trace (of type *e*) later abstracted via *predicate abstraction*, which renders the erstwhile sentence a predicate with an unsaturated argument position (type $\langle e,t \rangle$ - see Heim and Kratzer 1998), with which the raised quantifier can happily compose.

(47)



On this theory, a quantifier's scope at the level of interpretation is read off of its syntactic position at LF. Scope at the interpretive level is *equivalent to c-command at LF*. If scope is just equal to c-command at LF, we can restate Barker's *Scope Condition* on bound variable anaphora in terms of c-command (Bach 1980, Lasnik and Saito 1984). [10]

[10] An objection to computing binding possibilities based on LF c-command relations comes from Lasnik (1999), who argued that new c-command relations established via covert movement do not feed new binding relations. One argument Lasnik used involves binding possibilities in raising-to-object constructions. Lasnik argues that the position of the anaphor wrt. the particle *out* in the sentences below indicate whether the anaphor has been overtly raised into the matrix clause, or whether it has remained in the lower clause. Lasnik

(48) SCOPE CONSTRAINT (LF-COMMAND VERSION)

A pronoun must be in the c-command domain of the quantifier at LF

All of Barker's counter-examples yield to this analysis; binding out of possessives, inverse-linking, and binding out of DP, VP, and PP. In each case, the quantifier that binds the pronoun raises to a position at LF that c-commands that pronoun. Example sentences, and their corresponding LFs are below.[11]

---

assumes that even if the anaphor is left low in overt syntax, it still must raise in covert syntax. Thus, at LF, the position of the anaphors in both sentences will be the same. This predicts that the anaphor should be in the appropriate configuration to be licensed at LF. Nevertheless, the sentence is ungrammatical. Lasnik therefore concludes that covert movement cannot feed binding possibilities.

- The boys made themselves/each other out to be idiots.
- *The boys made out themselves/each other to be idiots.

This analysis has been challenged by van Craenenbroeck and den Dikken (2006), who argue, based on an analysis in Lasnik (2001), that covert movement is not required in the latter case. If the anaphor does not move covertly, it will not be in the right position to be licensed on the surface, or at LF.

Lasnik (1999, 2001) also argues against covert movement feeding binding based on the inability of post-verbal associates of the expletive *there* in existential constructions to bind anaphors c-commanded by spec,TP.

- The DA proved [two men$_1$ to have been at the scene] during each other$_1$s trials.
- *The DA proved [there to have been two men$_1$ at the scene] during each other$_1$s trials

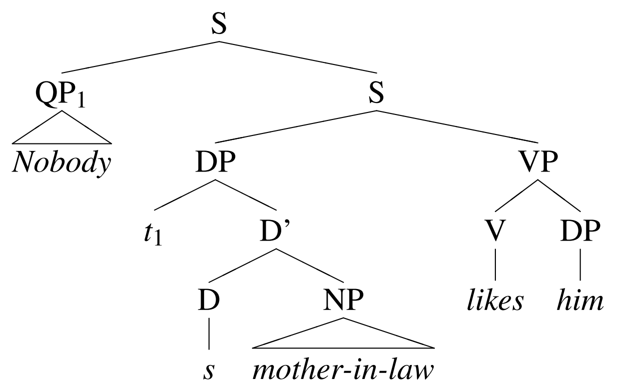
Lasnik's argument rests on the assumption that the post-verbal associate raises to spec,TP at LF. If such raising occurred, the associate would c-command the reciprocal at LF. Recent work has suggested, however, that raising of the associate at LF need not be required. The motivation for moving the associate into subject position was thought to be necessary to check the agreement features on T. However, other ways of achieving this agreement without moving the associate exist. Hornstein and Witkos (2003), Hornstein (2009, 139) propose a doubling-analysis of existential *there*, which allows *there* to bear (and therefore check) the ϕ-features on T. Alternatively, in a syntax that permits Agree (Chomsky 2000, 2001, 2008), the associate need never raise to check the features on T. T could simply probe its sister and find the right features. What these results show is that the case against covert movement feeding binding relations is not particularly strong. Allowing covert movement to feed binding relations also significantly cleans up the interpretation of some facts discussed by Ruys (1993) and Fox (2000). Ruys and Fox note that *him* in the sentence below may be bound by the QP *every professor*, if the QP scopes over *a different student*.

- A different student [liked every professor$_1$] and [wanted him$_1$ to be on his committee].

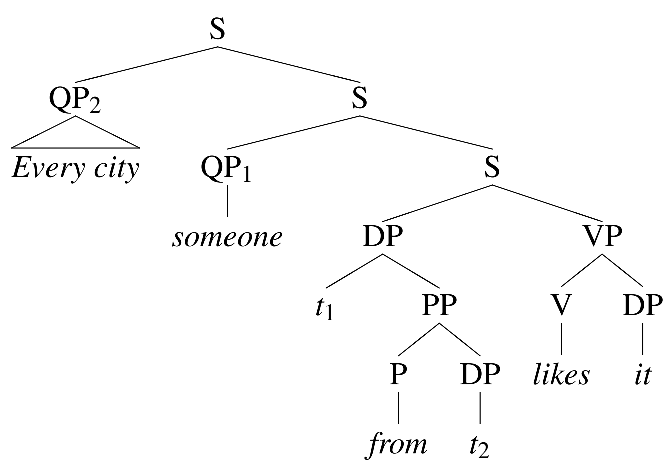
Ruys (1993) assumes that the pronoun *him* above is a kind of resumptive pronoun. This analysis is complicated by the fact that resumptive pronouns are usually thought to be traces of *movement*, and, as such, must be c-commanded by their binders at surface structure. If we assume, however, that the pronoun is base-generated, and then bound at LF, this problem disappears.

[11] A note: Specifying the right LF of inverse-linking examples has been a point of continued debate. Here, I have chosen the simplest analysis possible, which allows the lower QP to move out of the larger containing

(49) *Nobody*'s mother likes *him*.

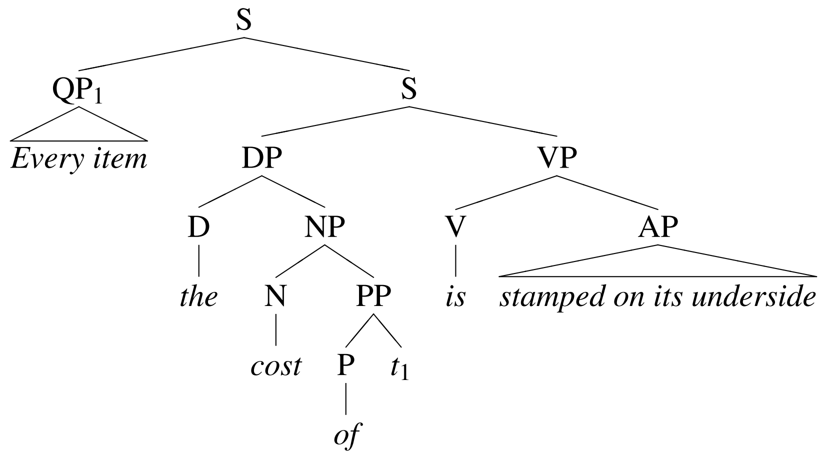


(50) Someone from *every city* likes it.

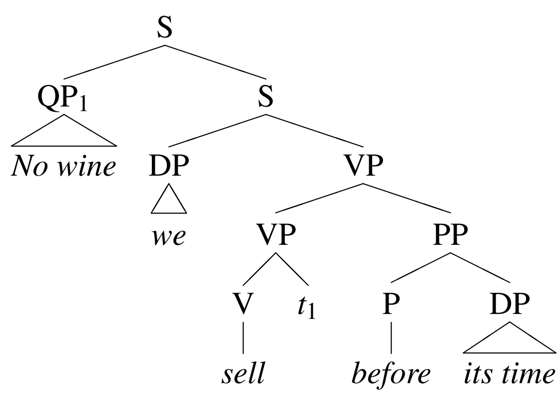


(51) The cost of *every item* is stamped on *its* underside.

subject. Larson (1985), May (1985) and others have argued that the lower QP may not move out of the container NP. May (1985), for example, argues that the lower QP simply adjoins to the container. However, May redefines the definition of c-command to allow an adjoined QP of this kind to c-command out of the phrase to which it is adjoined.



(52) We sell *no wine* before *its* time.



6.4 Structural Limits on Scope and Scope Displacement

As yet, the methods for scope computation are relatively free. The methods do not provide us with a simple way to explain the following contrast. In ((53)a) an inverse scope reading is possible: *every professor* can scope over *a student* when the two are in the same clause. However, inverse scope is not possible in ((53)b), when *every professor* is in a lower clause.

(53) a. A student likes every professor.

$\forall >> \exists$

$\exists >> \forall$

b. A student said John liked every professor.

$*\forall >> \exists$

$\exists >> \forall$

Take the challenge this data presents to a simple theory that licenses scope via movement. Other types of movement are not clause-bounded (as evidenced by the ability to extract wh-words in (54)), so why should scope assignment be?

(54) Who did a student say John liked?

The same problem faces direct-compositional approaches: it is not immediately obvious why *A student said John liked* cannot be composed to yield the property $\lambda x.\exists y.student(y)\&y\_said\_John\_$

These data suggest a QP's scope has an upward bound of its minimal finite clause. The response of many theorists is to simply state the clause-bounded restriction by fiat (May 1977, Heim 1982, Barker 2002).

(55) SCOPE BOUNDEDNESS:

A QP cannot scope above its minimal finite clause.

This restriction introduces a degree of relational sensitivity to the scope constraint that would otherwise not exist. Without such a restriction, *any* QP could potentially scope over *any* pronoun. We can define a relation that encompasses all possible cases of scope, called MINIMAL CLAUSE COMMAND, MCCOMMAND for short.[12]

(56) MCCOMMAND

α MCCOMMANDS β iff an element in the minimal clause containing α dominates β

With MCCOMMAND, we can restate the *Scope Condition* so that it wears its relation-sensitivity on its sleeve.

(57) SCOPE CONDITION

A pronoun must be MCCOMMANDED by a QP to be bound by it.

MCCOMMAND captures the QP's ability can scope *down* arbitrarily far into a series of clauses dominated by its current clause but not into clauses it does not dominate. This is illustrated in (4), repeated below.

(4) a. Every senator$_1$ at the party thought that he$_1$ would have no trouble getting elected.

b. *Every senator$_1$ was at the party and he$_1$ was worrying about getting elected. [Culicover and Jackendoff, 1997, p. 204]

In order to compute the eligibility, the pronoun must know whether its current clause is dominated by the QP's minimal clause. In short it requires reference to the relative position of the pronoun vis-a-vis the QP.

[12] This relation is very similar to Langacker's 1969 *command*, and Lasnik's (1976) *kommand*.

To belabor the point: It is *relative* position that determines scope, rather than, say, the QP's absolute position. A QP inside an RC cannot bind a pronoun external to that RC, because the structural position of the pronoun in relation to the QP is not one that allows the QP to take scope. It is not, however, due to some constraint that generally forbids embedded QPs from taking scope over *any* pronoun. A QP inside an RC can certainly bind an RC-internal pronoun.

(58) a. *The ladies [that *no man*$_1$ liked] admired *him*$_1$.

b. The ladies [that *no man*$_1$ believed liked *him*$_1$] admired the cactus.

6.5 The Take-home

Above we reviewed Barker's arguments against a surface c-command constraint and in favor of a scope constraint. After we discussed two ways of determining a quantifier's scope domain without making reference to surface c-command. The first relied on a series of type-raising operations to allow a quantifier to be fed a pre-composed chunk of structure while remaining in-situ. The second method computed semantic scope as the outcome of syntactic derivational scope (LF c-command), as opposed to surface c-command relations.

Regardless of how scope was computed, it was shown that Scope Boundedness makes computation of scope between a QP and a pronoun an inherently *relational* computation. This was drawn out by stating the Scope Constraint in terms of a relation MCCOMMAND. Thus, the *Scope Constraint* on BVP is a relational constraint. As a long-distance dependency subject to a relational constraint, BVP is a prime candidate to test retrieval's relational sensitivity.

Chapter 7: Bound Variable Pronouns

7.1 Introduction

In this chapter I test whether retrieval of antecedents for pronouns accesses feature-matching QPs in violation of the *Scope Condition* on Bound Variable Pronouns (BVPs). As argued in the previous chapter, the scope condition is a positive relational constraint stated over a potentially unbounded distance. The version of the scope condition put forward in the previous chapter was stated in terms of the relation MCCOMMAND, a descriptive surface relation compatible with either approach to the condition discussed. The condition and the definition of MCCOMMAND are stated below.

(1) SCOPE CONDITION

A pronoun must be MCCOMMANDED by a QP to be bound by it.

(2) MCCOMMAND (minimal clause command)

α MCCOMMANDS β iff an element in the minimal clause containing α dominates β

MCCOMMAND is the type of emergent relational property not well-handled by the multiple image encoding and direct cue-based retrieval scheme assumed in this dissertation.

The chapter begins with a brief summary of previous experimental work on bound-variable readings, and the extent to which comprehenders consider dependencies that violate relational constraints. I argue that what little previous work has considered online application of relational constraints in accessing QPs as antecedents cannot, in fact, speak to the issue. Experiments 3 and 4 provide a preliminary test of general sensitivity to the Scope Condition. I show that retrieval distinguishes between feature-matching QPs that c-command a critical pronoun (and therefore MCCOMMAND it), and those feature-matching QPs that do not. While the first experiments demonstrate a sensitivity to the constraint, they do not conclusively rule out the possibility of interference from structurally inappropriate QPs. Experiment 6 provides this evidence. The experiment uses a standard interference design to show access to grammatically illicit QPs is categorically restricted.

The results suggest that 'retrieval' is 'sensitive to the scope condition'. The scare-quotes in the previous statement reflect my belief (as stated in the introduction) that it is potentially misleading to discuss sensitivity of retrieval, per se, to the relational constraint. A less tendentious description of the facts is that a closer alignment exists between the set of NPs deemed accessible by the grammar and those accessible to retrieval than initially predicted on a simplistic view of the parser/memory interface. The results motivate a change to some component of the system at large, but they do not necessarily entail that the *retrieval process* makes any constraint-specific accommodations. The latter third of the chapter explores a method for shifting responsibility for lack of interference away from the retrieval process insofar as possible. It is shown that the scope condition can be 'compiled out' into a feature that is dynamically updated in a manner sensitive to a local structural environment.

7.2 Previous Work on Processing Bound Pronouns

The majority of prior work on bound variable pronoun interpretations has dealt with how comprehenders resolve potential ambiguity of pronouns in context. In general, two issues have pre-occupied researchers. The first issue concerns whether comprehenders prefer *referential*, or *quantificational* antecedents, if given a choice. The second question is whether bound readings are preferred over coreferential readings when antecedent is held constant and the grammar potentiates both. These issues, though interrelated and often run together, are logically distinct. I discuss each briefly.

Pronouns can, grammatical and discourse context permitting, have quantificational or referential antecedents. Thus, when encountering a pronoun in a sentence, a parser faces a choice point: should it prefer a quantificational antecedent, a referential antecedent, or should it remain agnostic? Cunnings, Patterson and Felser (2011) present evidence that the parser is equally content to take a referential or quantificational antecedent, so long as distance between pronoun and potential antecedent is controlled. In one experiment, the researchers manipulated a pronoun's gender match with both a preceding QP, and a proper name. The first experiment showed longer reading times at the pronoun when the proper name did not match in gender.

(3) a. The squadron paraded through town. *Every soldier* who knew that *James/Helen* was watching was convinced that *he* should wave as the parade passed. The entire town was extremely proud that day.

b. The squadron paraded through town. *Every soldier* who knew that *James/Helen* was watching was convinced that *she* should wave as the parade passed. The entire town was extremely proud that day.

In a second experiment, the authors reversed the order of QP and proper name, to test whether the effects above reflected a general preference for the coreferential dependency, or rather a linear bias to link the pronoun with the closest matching potential antecedent. The researchers found that when the order of antecedents was reversed, no preference for the referential potential antecedent was observed. Mismatch between the pronoun and the close QP lead to increased RTs.

(4) a. The squadron paraded through town. It looked to *James/Helen* that *every soldier* was completely convinced that *he* should wave as the parade passed. The entire town was extremely proud that day.

b. The squadron paraded through town. It looked to *James/Helen* that *every soldier* was completely convinced that *she* should wave as the parade passed. The entire town was extremely proud that day.

These results suggest that, all things equal, comprehenders do not exhibit a bias towards choosing an antecedent with whom a coreference relation must be established (a non-commanding referential NP), and an antecedent that requires binding (a QP). They also suggest that the same factors govern accessibility of QPs and referential NPs in antecedent retrieval.

The second question addressed by previous work concerns whether comprehenders display a general behavioral preference for binding or coreference in interpretation when the pronoun's antecedent is held constant. A referential antecedent that c-commands its pronoun is not restricted to establishing a coreference relation with that pronoun. It may also bind the pronoun. Thus, the relation between *the cyclist* and *his* in the sentence below can be represented two ways grammatically. (5-a) corresponds to the coreferential reading, whereas (5-b) corresponds to the bound.

(5) The cyclist thought his shorts were too short.

   a. The cyclist$_i$ thought his$_i$ shorts were too short.

   b. The cyclist(λx. x thought x's shorts were too short)

In single sentences, the distinction between the two readings is not apparent, but in cases of ellipsis resolution the contrast comes out quite strongly. The sentence below can have two readings, a 'strict' reading, or a 'sloppy' reading. These are paraphrased.

(6) The cyclist thought his shorts were too short, and the official did, too.

   a. STRICT: The cyclist$_i$ thought his$_i$ own shorts are too short, and the referee thinks the cyclist$_i$'s shorts were too short.

   b. SLOPPY: The cyclist$_i$ thought his$_i$ own shorts are too short, and the referee$_k$ thinks his$_k$ own shorts were too short.

If we assume that VP ellipsis resolution in these coordinated constructions involves copying the preceding VP's denotation into the ellipsis site, the strict/sloppy readings of the pronoun in the second VP can be linked to the presence of coreferential/bound dependency in the first VP.

(7) The cyclist thought his shorts were too short, and the official did, too.

a. The cyclist$_i$ thought his$_i$ own shorts were too short, and the referee thinks the cyclist$_i$'s shorts were too short.

b. The cyclist$_i$(λx. x thought x's shorts were too short), and the referee$_k$(λx. x thought x's shorts were too short).

Starting with Avrutin (1994), a number of papers argued that children preferred to resolve pronouns as bound, rather than coreferential (Foley et al. 1997; Guo, et al. 1996). Shapiro and Hestvik (1995) argued, however, that adults preferred coreferential interpretations to bound interpretations in real-time processing. More recently Frazier and Clifton (2000) argued that a general preference for bound pronouns is observed within ellipsis contexts, but comprehenders do not exhibit a general preference for bound or coreferential interpretation outside of ellipsis contexts.

Recalling the distinction made by Garrod and Sanford (1988), the studies discussed above are concerned with preferences that comprehenders exhibit in the resolution process of pronominal interpretation, not the retrieval stage (Garrod and Sanford's *bonding* stage). Deciding between a binding dependency or coreference relation for an antecedent-pronoun pair presupposes retrieval of the appropriate antecedent. As such, the studies are largely orthogonal to our purposes.

7.2.1 Carminati et al. 2002

Carminati et al. (2002) has been the only study in the psycholinguistics literature to date that directly tested sensitivity to relational constraints on QPs as potential antecedents for pronouns. The study did not focus on interference effects, but rather sought to show that the long-supposed c-command constraint was inconsistent with bound interpretations considered by comprehenders online. The authors compared the difficulty of processing a pronoun when the only feature-matching phrase in the discourse was a QP that did not c-command the pronoun, a QP that c-commanded it, or a referential NP. Carminati and colleagues Experiment 1 varied the referentiality of a feature-matching phrase (*Every/The British soldier*), and the c-command relation between the phrase and a later pronoun *he*. The pronoun was either embedded in a subordinate clause c-commanded by the phrase (8), or placed in a separate conjunct, as in (9) In (8) the feature-matching QP c-commands the pronoun; in (9) it does not.

(8) a. Every british soldier thought that he killed an enemy soldier.
b. The british soldier thought that he killed an enemy soldier.

(9) a. Every british soldier aimed and then he killed an enemy soldier.
b. The british soldier aimed and then he killed an enemy soldier.

The authors found that reread times at the pronoun in QP sentences were longer than in Referential sentences. The authors found no differences that suggested that participants exhibited more difficulty processing the pronoun in sentences with non-c-commanding

QPs than in sentences with c-commanding QPs. The authors took these results to indicate that participants retrieved QPs equally easily for pronouns, regardless of c-command.[1]

In a second experiment, the researchers tested whether accessing matching QPs as antecedents for a pronoun across a sentence boundary incurred a cost. C-command is a sentence-internal relation, therefore the antecedents in this experiment were similarly not c-commanded by their antecedents. Test items were short discourses in which the subject of the first sentence was the feature-matching antecedent. Antecedents were either QPs with the quantifier *every* or *each*[2], or proper names.[3] The materials with QP antecedents constitute instances of a grammatical phenomenon called 'telescoping' (Roberts 1989, Poesio & Zucchi 1992, a.o.), in which a pronoun apparently co-varies in interpretation with a QP across a sentence.

(10) {John Frederick | Each executive | Every executive} went home. He broiled a steak. He ate dinner.

[1]Though no significant interaction effect was found with sentence-type, though there were two numerical trends that suggested a pairwise difference between the quantificational conditions not observed between referential conditions. In the target region elevated First-pass reading times were observed in the Quantifier-Complement condition, where the quantifier c-commands the pronoun. Similarly, in the pre-critical region, elevated reread times were also observed in the same region. This suggests that on two measures, the one condition in which c-command mediated binding was required, reading times were disproportionately higher. The authors did not provide pairwise comparisons between the two quantificational conditions in these regions and measures, so it is difficult to draw any firm conclusions either way, though it is at least suggestive that the conditions were, at some level, distinguished.

[2]Quantifiers were manipulated in this way due to the intuition that 'each' more readily allows telescoping than does *every* (Poesio & Zucchi 1992).

[3]The researchers also manipulated an additional factor in the experiment: order of sentences in the discourse. Poesio & Zucchi (1992) and others have noted that the felicity or availability telescoping depends in part on whether the following discourse is 'predictable', or somehow a part of a contextually salient 'script'. Discourses could proceed in an order that was somehow 'typical' or in the reversed order. As no differences between atypical and typical conditions were found in the relevant regions, only effects from the typical conditions are reported.

The authors found that conditions with proper name antecedents were read reliably faster across all measures and regions. No indication of difference between the *every* and *each* conditions at the pronoun was found, despite the commonplace belief that *each* more readily admits telescoping than *every*. It is difficult to interpret the differences between referential and quantificational conditions with respect to pronoun interpretation. The results are consistent with a number of interpretations. The difference could reflect difficulty stemming from treating the pronoun as unheralded. Alternatively, the pronoun could display longer RTs due to the complexity of positing a covariation interpretation. The results are unable to distinguish between these two possibilities because it does not provide an independent baseline condition containing an unheralded, mismatching pronoun.[4]

Carminati et al. interpreted the evidence as support for the idea that binding does not require c-command between antecedent and pronoun. This argument is flawed. The flaw lies in the failure to distinguish distinct types of co-varying pronominal interpretations. The authors conflate grammatical cases of Telescoping (which does not require c-command), with cases of true variable binding. Viz., they assume that lack of interpretive difficulty at a pronoun in non-command conditions implies a binding relationship was established between QP and pronoun. However, the formal semantics literature is in (near) agreement: Telescoping is not binding, despite superficial appearances.[5] Since telescoping is not

[4]The sentence-order manipulation of Experiment 7 showed an effect only within proper name conditions (atypical orders resulted in longer reading times in the second sentence across all measures), not in either of the quantificational conditions.

[5]Various alternative analyses have been offered in the theoretical literature, the details of which are not relevant for this paper. The most recent account of telescoping, Anderssen's (2011) analysis, holds that the apparently bound pronouns are not bound by the quantifier, but rather, they are interpreted as e-type pronouns (Evans 1977, 1980; Heim 1990; Elbourne 2005). The basic idea is that the pronouns are actually more akin to definite descriptions. Apparent bound readings of the pronoun are not achieved through binding of the pronoun by the QP. Instead, a covert quantificational adverb quantifies over minimal situations containing exactly one individual which the pronoun can be coreferential with. The sentence in (i) is assigned the reading below it under Anderssen's e-type analysis:

binding, we cannot conclude anything about the processing of binding dependencies based on observation of Telescoping dependencies.

If Telescoping conditions are to be rejected as evidence for or against real-time implementation of the scope constraint, Carminati and colleagues' second experiment clearly does not bear on the current study. I also argue that their first study falls under the same criticism because it also employed constructions that are demonstrably telescoping constructions. Three tests show that the constructions used by Carminati and colleagues' first experiment are instances of Telescoping.

1. *Identity Substitution*. A QP cannot be substituted in for a pronoun it binds and preserve truth conditions (11).[6]

(11) No cyclist said he bought short shorts. $\neq$
No cyclist said no cyclist bought short shorts.

Carminati's experimental items preserve their meaning under substitution.

(12) Every British soldier picked up his rifle and then he aimed. $\approx$
Every British soldier picked up his rifle and then every British Soldier aimed.

(i) Every degree candidate walked up. He received his diploma and went backc to his seat.
$\forall x$, x a degree student. x walked up. For all minimal situations within the larger situation of candidates walking up, there exists a unique candidate that received his diploma and went back to his seat.

[6] See Büring (2005) for discussion of this test.

2. *Negative Quantifier Replacement*. True instances of binding permit negative quantifiers. [7]

(13) Everybody's mother loves him.

Nobody's mother loves him.

Carminati's items do not permit substitution of negative quantifiers, even when minimal semantic changes are made to ensure sensibility.

(14) Every British soldier picked up his rifle and then he aimed. $\approx$

*No British soldier picked up his rifle and then he ran from the battle field.

3. *Quantificational Independence*. Pronouns in Telescoping constructions can differ in quantificational force from their antecedents, something not possible with pronouns in traditional binding configurations (see Anderssen 2011). *Many a* is not a universal quantifier. Pronouns that covary with *many a* in telescoping constructions are nevertheless interpreted with universal force, allowing (15)a to be paraphrased as (15)b.

[7] Telescoping is occasionally observed with negative quantifiers like *no*, but the semantics of such constructions are clearly different from those of ones in which *no* takes scope over a pronoun. The pronoun in the second clause is interpreted as if bound by an upward-entailing universal (such as *every*), rather than bound by *no*. This is an instance of *quantificational independence* discussed in the next test.

- No man can be friends with a woman. He always wants to have sex with her.
  $\approx$ No man can be friends with a woman. Every man always want to have sex with [the woman he's friends with].
  $\neq$ No man can be friends with a woman. No man always wants to have sex with [the woman he's friends with].

(15) a. Many a student walked to the stage. He received his diploma from the Dean and returned to his seat. $\rightsquigarrow$

b. Many a student walked to the stage. Every student who walked up to the stage received his diploma from the Dean and returned to his seat.

This is not observed with true instances of binding.

(16) Many a student said that he was eager to graduate.

*Many students said all the students were eager to graduate.

The constructions used by Carminati and colleagues' permit quantificationally independent pronouns, once again patterning with telescoping constructions, and not true cases of binding.

(17) Many a British soldier he picked up his rifle and then he fired. $\approx$

Many a British soldier picked up his rifle and then every soldier that picked up his rifle fired.

Thus, neither of the experiments discussed from Carminati et al. (2002) bear directly on the issue of retrieval of ungrammatical feature-matching NPs. What Carminati and colleagues' results show is not a misalignment between interpretive constraints imposed by the grammar, and constraints that are implementable by a parser. Rather, the results point to a tight correspondence between conditions under which the grammar licenses an anaphoric dependency, and when the parser considers one. Insofar as a co-varying interpretation is readily available, processing (as indexed by the measures) will not be disrupted. This, in itself, is a worthwhile result from it we learn about the resilience of the parser in the face of pronominal ambiguity and its sensitivity to the conditions that license particular types of discourse anaphora. Interesting and informative as the results are, they are not evidence against the idea that retrieval of pronominal antecedents accesses quantificational antecedents in violation of a relational constraint. In order to show that a relational constraint is flouted online, one must select a test case that unambiguously obeys the relational constraint.

7.3 Experiment 6: Bound Variable Pronoun Licensing I: Constraint Sensitivity

Experiment 6 compared the effect of the position and the type of potential antecedents on the processing of pronouns. Specifically, it tested whether c-command (and therefore MCCOMMAND) facilitates linking a pronoun to a quantificational antecedent in a manner that it does not affect establishing a link to a referential antecedent. Experiment 6 was designed to test whether antecedent access is sensitive to relations, but could not test whether antecedent access mechanisms are sufficiently sensitive to gate interference from structurally inappropriate QPs. The experiment used materials that were designed to facilitate bound variable interpretations for the pronoun, while excluding the possibility of Telescoping or other related interpretations.

7.3.1 Materials

The experiment used a 2x2 factorial design, manipulating the factors REFERENTIALITY and CONTINUATION, as illustrated in (18). All test items consisted of a pair of clauses linked by a conjunction or connective. A singular pronoun appeared at the beginning of the second clause, and hence needed to find an antecedent in the first conjunct. The first clause began with a noun phrase that mismatched the gender of the pronoun and was followed by a second noun phrase that matched the gender of the pronoun, the *antecedent*. A subsequent noun phrase in the first clause included a possessive pronoun that matched the gender of the antecedent. This possessive ensured that the antecedent was a highly accessible referent when readers reached the end of the first clausal constituent. Potential antecedents were embedded to prevent any advantage in retrieval probability associated with first-mention or being the main subject

The factor REFERENTIALITY manipulated the determiner of the feature-matching potential antecedent. In the *Referential* conditions the determiner was definite, and in the *Quantificational* conditions the determiner was the quantifier *any*. We chose to use any, rather than quantifiers such as *every* or *each* that have been used in previous studies (e.g., Carminati et al. 2002), because downward-entailing quantifiers such as *any* or *no* have a more constrained distribution as pronoun antecedents and in quantifier scope and do not license the additional pronoun interpretations that are sometimes found with upward-entailing quantifiers such as *every* (see, e.g., Behgelli & Stowell 1997). The quantifier *any*, with a preceding negation to license it, was chosen over the quantifier *no* in order to maximize the naturalness of the example sentences.

(18) a. Kathi didn't think *any janitor* liked performing his custodial duties, *but* he had to clean up messes left after prom anyway.

b. Kathi didn't think *any janitor* liked performing his custodial duties *when* he had to clean up messes left after prom.

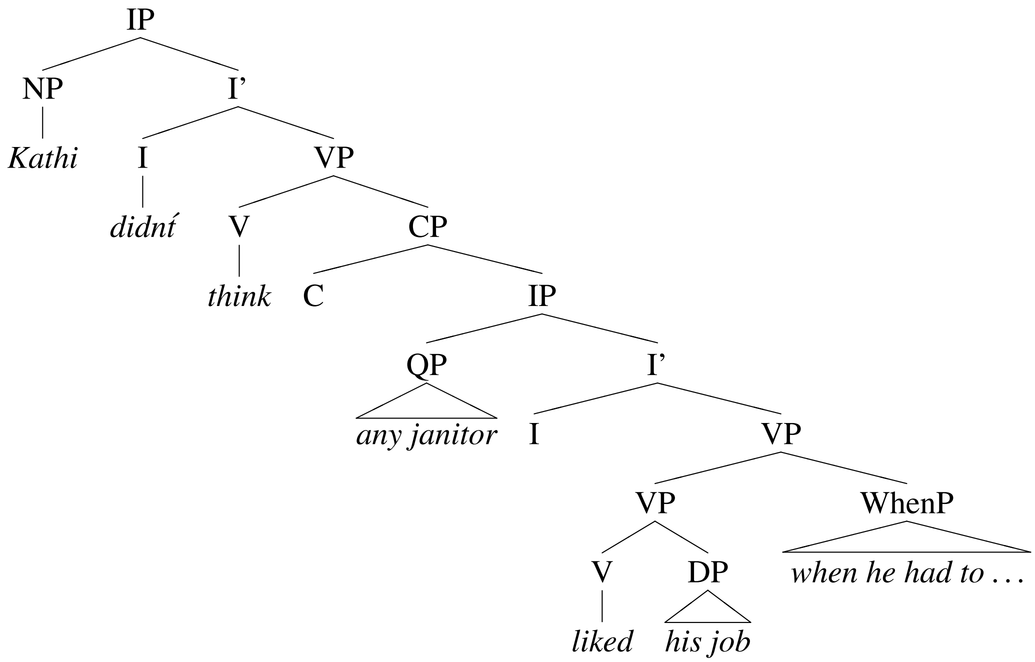
c. Kathi didn't think *the janitor* liked performing his custodial duties, *but* he had to clean up messes left after prom anyway.

d. Kathi didn't think *the janitor* liked performing his custodial duties *when* he had to clean up messes left after prom.

The factor *Continuation* manipulated the attachment height of the clause containing the critical pronoun. The pronoun was either contained in a temporal adverbial clause introduced by *when*, which attaches to the embedded VP of the first clause, or inside a clause introduced by *but*, which attaches to the root of the sentence. The coordinator *but* was chosen over *and*, because *but* is less ambiguous in its attachment options. By manipulating attachment height of the clause containing the pronoun, we manipulated whether a c-command relation holds between the pronoun and feature-matching NP, without changing the position of the feature-matching NP.

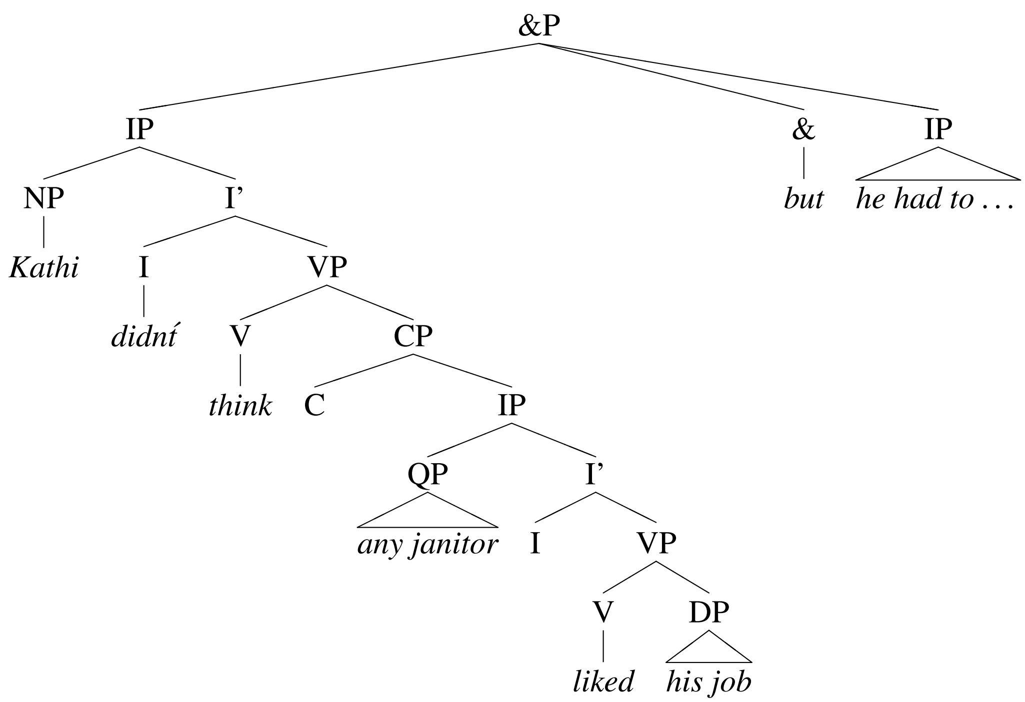
(19)

WHEN CONDITIONS



(20)

BUT CONDITIONS



Irrespective of the connective, referential potential antecedents should always be accessible to retrieval, because they are in a position from which the grammar allows them to corefer with the pronoun. In the *Quantificational-When* condition, the feature-matching QP scopes over the pronoun (under c-command), therefore it is a grammatical antecedent for the pronoun. In *Quantificational-But* conditions, however, the pronoun is not in the scope of QP, and cannot be bound by it. If retrieval of potential antecedents is sensitive to relational information, a REFERENTIALITY x CONTINUATION is predicted. Difficulty is predicted at the pronoun in *Quantificational-But*, relative to the other conditions, due to lack of antecedent. If, however, retrieval is not sensitive to the constraint/structural distinction, we predict that morphological feature-match should drive retrieval alone, thus predicting no pairwise difference between *Quantificational* conditions above and beyond the contribution of the effect of *Continuation*.

7.3.2 Experiment 6a

As a first step in testing the effect of structure on resolving bound variable pronouns I sought to verify the structural generalization in offline ratings. Experiment 6a does so using offline acceptability judgments.

7.3.2.1 Participants

16 participants (mean age = 34.9, 7 male) were recruited through the Amazon Mechanical Turk (AMT) marketplace and paid $4.00 for their participation (see Sprouse 2011 for discussion of the reliability of AMT-collected judgment data). Participants were qualified to participate based on US-internal IP addresses and passing a native-speaker screening test, consisting of a short series of acceptability judgments (unrelated to the

current study) that prior testing had shown to effectively distinguish native and non-native speakers of English.

7.3.2.2 Procedure

Presentation used the IBEX Farm web-based, experimental presentation platform, developed and managed by Alex Drummond (www.spellout.net/ibexfarm/docs). Sentences were presented one-at-a-time centered on the screen. A 7-pt acceptability scale was presented as an array of numbered boxes below each sentence, with endpoints marked 'bad' (1), and 'good' (7). Participants were instructed to rate sentences on the scale, with a rating of 1' corresponding to bad', unacceptable' or doesn't make sense', and a rating of 7' corresponding to good', totally acceptable', or easy to understand'. Participants could record a response by either clicking a box, or pressing a number on their keyboard. Though participants were not explicitly instructed to attend to pronouns, practice items illustrated cases where intra-sentential coreference and binding failed.

7.3.2.3 Materials

24 experimental items were distributed across 4 lists according to a Latin Square, and were interspersed in a pseudo-randomized order among 64 fillers. This resulted in lists of 88 items per participant, composed of roughly 60% acceptable sentences. 18 ungrammatical fillers contained errors such as unlicensed NPIs, agreement errors, or were missing functional words such as *the*. 8 fillers contained QPs and pronouns that could not be bound by them (either due to number mismatch, or configurational constraints).

	But	When
Quantificational	4.17 (.25)	4.76 (.22)
Referential	5.69 (.19)	5.61 (.18)

Table 7.1: Experiment 6a: Acceptability Ratings (Test Items)

	Bad-Binding	Good-Binding	Ungrammatical Fillers
Fillers	4.28 (.17)	5.82 (.16)	3.05 (.11)

Table 7.2: Experiment 6a: Acceptability Ratings (Binding Control Conditions)

These sentences, referred to as *Bad-Binding* sentences below, were included to provide a base-line against which to test the acceptability of the *Quantificational-But* conditions. 4 unacceptable sentences featured an unheralded or infelicitous pronoun inside a *when* adverbial clause, to prevent participants from developing superficial tendencies to rate all sentences containing *when* as acceptable due to its frequency in the test-items. Similarly, 4 unacceptable sentences with the connectives *but* or *and* were included for the same reason. 30 acceptable fillers were also included, 6 of which contained acceptable examples of structurally-licensed variable-binding to use as a baseline to test participants' willingness to posit a bound reading of the pronoun when necessary.

7.3.2.4 Analysis

Statistical analysis used logistic-mixed effect regression. CONTINUATION and *Referentiality* were treated as fixed effects. Subject and item were treated as random effects.

7.3.2.5 Results

There was a reliable effect of REFERENTIALITY (β = -0.60, s.e. = .08, t = -7.71, p <.001), driven by higher scores in the *Referential* conditions than in the *Quantificational* conditions. A marginally significant effect of CONTINUATION was also observed (β = -0.13, s.e. = 0.08, t = -1.76, p <.10), due to higher acceptability scores in the *When* conditions. The REFERENTIALITY x CONTINUATION interaction was significant (β = -0.18, s.e. = .08, t = -2.35), with *Quantificational-But* conditions receiving the lowest acceptability scores overall. The pairwise difference between the two *Quantificational* conditions was significant (t = -2.86, p <.01). *Quantificational-When* sentences were rated reliably more acceptable than *Quantificational-But* sentences. No comparable pairwise difference was observed between *Referential* conditions (t < 1).

7.3.2.6 Discussion

The qualitative pattern of acceptability scores observed is consistent with the general notion that the pronoun in the *Quantificational-But* cannot be bound. The main effect of REFERENTIALITY reflects that processing simple coreference is generally easier than processing a quantificational binding dependency between a QP and pronoun. Reduced acceptability in the *Quantificational-But* sentence relative to its *When* counterpart suggests that the pronoun was either (i) interpreted as disjoint from the QP, or (ii) that the pronoun and QP were linked, but at a cost.

The relatively high acceptability scores of the *Quantificational-But* and *Bad-Binding* conditions relative to the ungrammatical fillers is unexpected. Interpretation of this unexpected acceptability might suggest that binding is indeed possible in these conditions, contrary to what is traditionally assumed. Alternatively, it is consistent with the possibility that accommodation of an antecedent-less pronoun in offline judgment studies may not always result in dramatic reductions in acceptability. The relatively high acceptability rating of the Bad-Binding condition is consistent with this possibility. It appears that even where binding is unavailable, due to the lack of a feature-matching noun phrase, participants rated the sentences relatively high. A third possibility for explaining the comparability of the scores in both quantificational conditions is that the main effect of REFERENTIALITY observed actually reflects the penalty for accommodation of an antecedent-less referential pronoun. This interpretation would require us to assume that pronouns in all Quantificational conditions were treated as unheralded, irrespective of c-command between QP and pronoun. This could be seen as a general avoidance of bound readings. This possibility is the least likely of the three discussed as it does not provide a rationale for the difference within the *Quantificational* level. In order to distinguish between these possibilities we conducted a test of participants interpretations of these sentences.

Distinguishing between these various possibilities requires probing participants' interpretations of these sentences.

7.3.3 Experiment 6b: Sentence Judgments

In order to test whether the relatively high acceptability ratings for the Quantificational-But condition in the study above reflects that participants do entertain a binding dependency between a QP and pronoun in the absence of c-command, contrary to the generalizations in the formal linguistics literature, we conducted a forced-choice interpretive task to measure participants consideration of bound-variable interpretations.

7.3.3.1 Participants

21 participants were recruited through the Amazon Mechanical Turk marketplace. Participants were paid $3.50 for their participation.

7.3.3.2 Materials

The interpretation task used the same target sentences as the acceptability rating task, but each sentence was combined with two paraphrases to create a triplet. The first sentence in each triplet was the test sentence drawn from the items from the 2x2 design above. Each test sentence was paired with two corresponding response sentences. Participants choice of the paraphrase that best matched their interpretation of the test sentence indicated their interpretation of the pronoun in the second clause of the target sentence. These response sentences were designed to be either consistent with a single-individual interpretation (response sentences marked *SI*), or a quantificational interpretation of the continuation (marked *BV*). Sentences consistent with a single-individual interpretation used an existential construction to assert the existence of an individual who performed the action described by the post-pronominal VP. Sentences consistent with a quantificational interpretation used a paraphrase of the relevant portion of the test sentence using the quantifier *every*, instead of *any*. Examples are given below.[8]

(21) EXAMPLE BUT SENTENCE

Kathi didn't think the janitor enjoyed performing his custodial duties, but he had to clean up the messes left after prom anyway.

1. There was someone who had to clean up after prom.
2. Every janitor had to clean up after prom.

(22) EXAMPLE WHEN SENTENCE

Kathi didn't think any janitor enjoyed his custodial duties when he had to clean up messes left after prom anyway.

1. There was someone who disliked having to clean up after prom.
2. Every janitor disliked having to clean up after prom.

Items were distributed across 4 lists in a Latin-square design and presented in a pseudo-randomized order. Order of response sentences was also randomized.

7.3.3.3 Procedure

On each trial participants saw a screen on which all three sentences in a test item were displayed, all center-aligned. Test sentences were presented above possible response sentences, which were presented as a numbered list in randomized order. After choosing a

[8] The semantics of *But* sentences differed from *When* sentences in terms of how the continuation modified the first clause. In *When* sentences the continuation introduces a temporal restriction on the action described by the main VP. The continuation in *But* sentences does not restrict the interpretation of the previous VP. Predicates in response sentences differed minimally based on the semantics of the test sentence to account for this difference.

sentence, participants were shown another screen which asked them to rate their confidence in the previous decision on a scale of 1 to 7, 1 corresponding to ‘low confidence’, 7 corresponding to ‘high confidence’. An array of numbered boxes was presented below the question *How confident are you in your response?*

Test sentences were interspersed randomly among 28 filler items, consistent either with single-individual, or quantificational interpretations. Filler sentences, as well as test sentences with referential gender-matching noun phrases were counted as having a correct response. Participants were encouraged to answer as accurately as possible, and received feedback on their accuracy throughout the experiment on both fillers and single-interpretation test sentences. If an inaccurate response was recorded, the phrase ‘Wrong.’ was presented on the screen briefly before the next trial. Sentences with quantificational noun phrases were not coded as having correct responses, so participants received no error message on these trials, regardless of their input.

7.3.3.4 Analysis

Data from four participants who scored lower than 70% accuracy on either the filler questions or the single-individual test sentences were excluded from analysis. Statistical analysis used logistic-mixed effect regression. CONTINUATION and REFERENTIALITY were treated as fixed effects. Subject and item were treated as random effects.

	But	*When*	*Filler*
Quantificational	.32	.90	.89
Referential	.08	.06	.22

Table 7.3: Experiment 6b: Proportion of Quantificational-Reading Responses on Test Conditions.

7.3.3.5 Results

The proportion of trials, by condition, on which participants chose the sentence corresponding to the quantificational interpretation of the continuation is shown in Table 7.3.

Logistic mixed-effects regression showed that a main effect of CONTINUATION was significant (β = -3.37, s.e. = 0.43, z = -7.84, p <.001). This was driven by an increased proportion of quantificational responses in the *Quantificational-When* condition condition in comparison to the *Quantificational-But* condition. A main effect of REFERENTIALITY was also significant (β = -2.21, s.e. = 0.51, z = -4.37, p <.001), driven by an increased proportion of quantificational responses in Quantificational conditions relative to Referential conditions. The CONTINUATION x REFERENTIALITY interaction was also significant (β = -3.05, s.e. = 0.74, z = -4.13, p <.001). This was due to an increased proportion of quantificational responses in the Quantificational-When condition relative to all other conditions.

A main effect of CONTINUATION was also observed in the pairwise comparison between *Quantificational* and *Single-Individual* filler conditions (β = -3.40, s.e. = 0.26, z = -12.96, p ¡.001). Filler-Quantificational sentences were more consistently given a quantificational interpretation than Filler-Single sentences.

7.3.3.6 Discussion

For filler conditions, participants reliably picked the correct interpretation. In *Referential* conditions, participants consistently chose the single-individual reading. This suggests that participants understood the task, and did not have difficulty with the existential construction used in single-individual responses.

The proportion of quantificational readings in the *Quantificational-When* condition was comparable to the proportion of quantificational readings chosen in *Quantificational* fillers. This indicates that participants chose to form quantificational binding dependencies for feature-matching pronouns in the sentence, rather than positing an extra-sentential referent for the pronoun.

Participants preferred to resolve the pronoun in *Quantificational-But* sentences as referring to an extra-sentential individual and not as a bound pronoun. On around 70% of trials participants favored the relatively costly strategy of coercing an unknown referent for resolution of the pronoun, rather than considering a grammatically illicit bound reading using the intra-sentential QP as the pronouns antecedent. Individual participant analyses show that most participants chose the quantificational interpretation for *Quantificational-But* trials at very low rates. 11 of the 17 participants chose the quantificational interpretation on 30% or fewer trials (average proportion 0.16). In contrast, the remaining 6 participants chose the quantificational response at much higher rates (range 0.50 1.00).

The findings demonstrate a clear effect of structure, but only in the *Quantificational* conditions, where the structural manipulation affects grammatical availability of a bound variable reading. Whether these results reflect a grammatical distinction or simply an interpretive preference cannot conclusively be determined because the task only probed participants interpretations. Regardless of the origin of the decision process, the results show that participants do make the distinction between pronouns that can and cannot be bound by a QP. These results also suggest that the relatively high ratings of *Quantificational-But* sentences in Experiment 6a do not reflect participants choosing a bound variable reading for the pronoun. The results are instead more consistent with an interpretation under which coercion of an extra-sentential referential antecedent for the pronoun in *Quantificational-But* conditions does result in drastic reductions in acceptability. Having validated the structural manipulations ability to affect the acceptability of quantificational phrases as antecedents, Experiment 6c tests whether real-time indices of retrieval reflect this distinction.

7.3.4 Experiment 6c. Eye-tracking while reading

Having established participants' offline sensitivity to the grammatical generalization that a QP's c-command relation in relation to a pronoun affects its eligibility as an antecedent, we tested whether this offline sensitivity maps transparently onto real-time behavior. In an eye-tracking study we investigated whether a feature-matching QP that does not c-command the pronoun is initially retrieved as a potential antecedent for a pronoun

7.3.4.1 Materials and Procedure

The 24 experimental item sets were distributed into 4 lists in a Latin Square design. Each list also contained 40 sentences from another experiment on pronoun resolution and 40 filler sentences for a total of 104 sentences per participant. The order of each list was pseudo-randomized such that no two experimental sentences were presented in succession. Sentences were presented in a 12-point fixed-width font (Courier), and each character was 9x16 pixels on the display. The resolution of the visual display was 1280 x 720 pixels on an LCD screen. The maximum number of characters allowed on a single line on the visual display was 142 characters, and all sentences in the experiment fit on one line.Eye movements were recorded using an Eyelink 1000 tower-mount eye-tracker, which sampled eye-movements at 1000Hz. Participants had binocular vision while movements were measured, but only the right eye was tracked. The tower was 32 inches from the visual display, giving participants approximately 5 characters per degree of visual angle.

Before beginning the experiment, participants were familiarized with the apparatus and given four practice trials. While seated, participants heads were immobilized using a chin rest and forehead restraint that was adjusted for comfort. Before the experiment, and whenever necessary throughout the experiment, the experimenter calibrated the eye-tracker with a 9-point display to ensure an accurate record of eye-movements across the screen. Participants began each experimental trial by fixating on a marker at the beginning of the sentence, triggering display of the entire test sentence. Participants terminated the presentation of sentences via button-press on a response pad, which triggered presentation of a yes/no comprehension question. Participants were allowed to take breaks at their discretion throughout the experiment. Following each break, participants were recalibrated to ensure accurate measurement.

7.3.4.2 Data Analysis

Test sentences were divided into 5 regions of interest, as indicated in Figure 7.4. The sentence-initial region was comprised of all words in the beginning of the sentence up to two words before the manipulated connective. The pre-critical region was comprised of those two pre-connective words. The critical pronoun region contained the manipulated connective and the critical pronoun. The connective was included in the critical pronoun region due to the tendency towards pronouns being skipped at high rates due to their relatively short length (Ehrlich & Rayner, 1983; Garrod, Freudenthal, & Boyle, 1994). The post-pronoun region contained the first three words of the continuation-internal verb-phrase. The remainder of each sentence was treated as a single region. I present data from four regions: pre-critical, critical pronoun, post-pronoun and final regions.

For each region of interest we report four different measures. These measures can be divided into two categories: early measures and late measures. Early measures reported here are First-pass reading time and right-bound reading time. First pass reading time is calculated by summing all fixations in a region of interest after participants first enter the region until the first saccade out of that region (either to the right or the left). Right-bound reading time is the sum of all fixations in a region beginning when the region is first entered from the left to when it is first exited to the right. Right-bound times includes fixations that occur if a participant makes a leftward regression, and then re-enters the region. First-pass times therefore form a subset of right-bound times. The late measures reported here are reread time and total time. Reread times are calculated by summing all fixations that occur in a region after a participant has first exited the region to the right. Total times sum over all fixations in a particular region of interest, including first pass reading time and any time spent rereading the region.

Statistical analysis was carried out using linear mixed-effect models (LMEMs) with maximal random-effects structures. Each model included sum-coded fixed effects of REFERENTIALITY (whether the feature-matching NP was *Referential*, or *Quantificational*), CONTINUATION (whether the critical pronoun was preceded by but, or when), and their interaction, as well as random intercepts for participants and items (Baayen, Davidson & Bates 2008), together with participant and item random slopes for all fixed effects (Barr, Levy, Scheepers & Tily 2013). In adopting a mixed-effect models analysis, missing values were left out from analysis, rather than being replaced as zeros. Analyses were also conducted on data in which missing observations contributed a value of 0ms to the cell mean. Analyses were also computed with missing values simply omitted. Effects did not change between the two analysis regimes.

A fixed effect was considered significant if its absolute t-value was greater than 2, which indicates that its 95% confidence interval does not include 0 (Gelman & Hill 2005). Reported coefficients reflect the magnitude of the difference between levels of a given factor in milliseconds. Reported coefficients whose absolute t-value is greater than 2 are significant at the $p < 0.05$ level.

Raw mean fixation times are presented in Table 7.5 below. A fixed effect was considered significant if its absolute t-value was greater than 2, which indicates that a

Kathi didn't think ...	
QUANT-BUT	any janitor enjoyed performing his/ custodial duties,/ but he/
QUANT-WHEN	any janitor enjoyed performing his/ custodial duties /when he/
REF-BUT	the janitor enjoyed performing his/ custodial duties,/ but he/
REF-WHEN	the janitor enjoyed performing his/ custodial duties /when he/
	... had to clean/ up the messes/ left after prom anyway./

Table 7.4: Experiment 6c: Regions for Items

given effect's 95% confidence interval does not include 0 (Gelman & Hill 2005). Reported coefficients whose absolute t-value is greater than 2 are significant at the $p<0.05$ level.

7.3.4.3 Results

Pre-critical Region. In the pre-critical region, a main effect of CONTINUATION was marginally significant in First-pass times (β = -27.0, s.e. = 14.4, t = -1.88 $p < .10$). This reflected longer reading times in When conditions relative to But conditions. No effects achieved significance in Right-bound times. The REFERENTIALITY x CONTINUATION interaction was significant (β =165.0, s.e. = 90.3, t = -1.82, $p < .10$), due to longer RTs in the Quantificational-When condition than in the other three conditions.

In total time, a main effect of REFERENTIALITY was marginally significant (β = 47.2, s.e. = 30.0 , t = 1.57) and a main effect of CONTINUATION was significant (β = -68.9, s.e. = 30.4, t = -2.26, $p < .05$). The interaction was significant, as well (β = -121.0 s.e. = 55.4, t = -2.19, $p < .05$). In total times the interaction was driven by elevated RTs in the Quantificational-When condition only. Pairwise comparison between the Quantificational conditions was significant (β = -128.80 s.e. = 47.01, t = -2.74, $p < .05$), but not between Referential conditions ($t < 1$). The interaction in total reading times in the pre-pronoun

	Pre-Pronoun	Critical Pronoun	Post-Pronoun	Wrap-up
First-Pass Time				
Quantificational-But	422 (16)	306 (12)	416 (17)	455 (18)
Quantificational-When	457 (18)	324 (11)	377 (14)	437 (16)
Refer-But	420 (16)	313 (11)	400 (16)	443 (19)
Refer-When	442 (19)	316 (12)	411 (15)	411 (16)
Right-bound Time				
Quantificational-But	545 (19)	333 (14)	475 (21)	505 (20)
Quantificational-When	570 (21)	357 (13)	431 (17)	492 (17)
Refer-But	557 (21)	350 (16)	454 (18)	493 (21)
Refer-When	538 (21)	343 (13)	472 (16)	456 (27)
Reread Time				
Quantificational-But	519 (33)	384 (28)	469 (34)	486 (35)
Quantificational-When	698 (39)	445 (29)	518 (36)	542 (35)
Refer-But	567 (49)	390 (28)	501 (34)	457 (28)
Refer-When	536 (34)	426 (27)	466 (25)	523 (37)
Total Time				
Quantificational-But	701 (28)	458 (19)	662 (28)	680 (29)
Quantificational-When	831 (38)	532 (22)	648 (29)	731 (29)
Refer-But	710 (28)	478 (21)	652 (27)	657 (27)
Refer-When	722 (28)	507 (21)	633 (24)	670 (27)

Table 7.5: Reading Time Measures, Experiment 6c. Standard Errors in Parentheses

region suggests that the parser distinguishes between the two Quantificational conditions.

Critical Pronoun Region. No significant main or interaction effects were observed in any measures at the critical pronoun region.

Post-Pronoun Region. At the post-pronoun region, there was a marginally significant REFERENTIALITY x CONTINUATION interaction ($\beta = 49.2$, s.e. = 28.4, t = 1.73, $p < .10$) in First-pass times. The interaction was driven by faster RTs in the Quantificational-When condition relative to the Quantificational-But condition. This pairwise difference was significant ($\beta = 43.5$, s.e. = 18.5, t = 2.36, $p < .05$). No corresponding pairwise difference was observed between Referential conditions ($t < 1$).

The REFERENTIALITY x CONTINUATION interaction was also marginally significant in Right-bound times ($\beta = 59.4$, s.e. = 30.6, t = 1.94, $p < .10$). As with the effects in First-pass times in the region RTs in the Quantificational-But condition are significantly higher in comparison to Quantificational-When (pairwise comparison: $\beta = 45.8$, s.e. = 21.3, t = 2.15, $p < .05$), whereas no differences are observed between the Referential conditions ($t < 1$).

7.3.4.4 Discussion

The study tested whether antecedent retrieval distinguishes between c-commanding feature-matching QPs and matching QPs that do not c-command a pronoun.

Participants rated sentences in which a pronoun was c-commanded by a feature-matching QP as more acceptable than sentences in which no c-command relation obtained. This preference for a c-commanding antecedent was not observed in minimally different

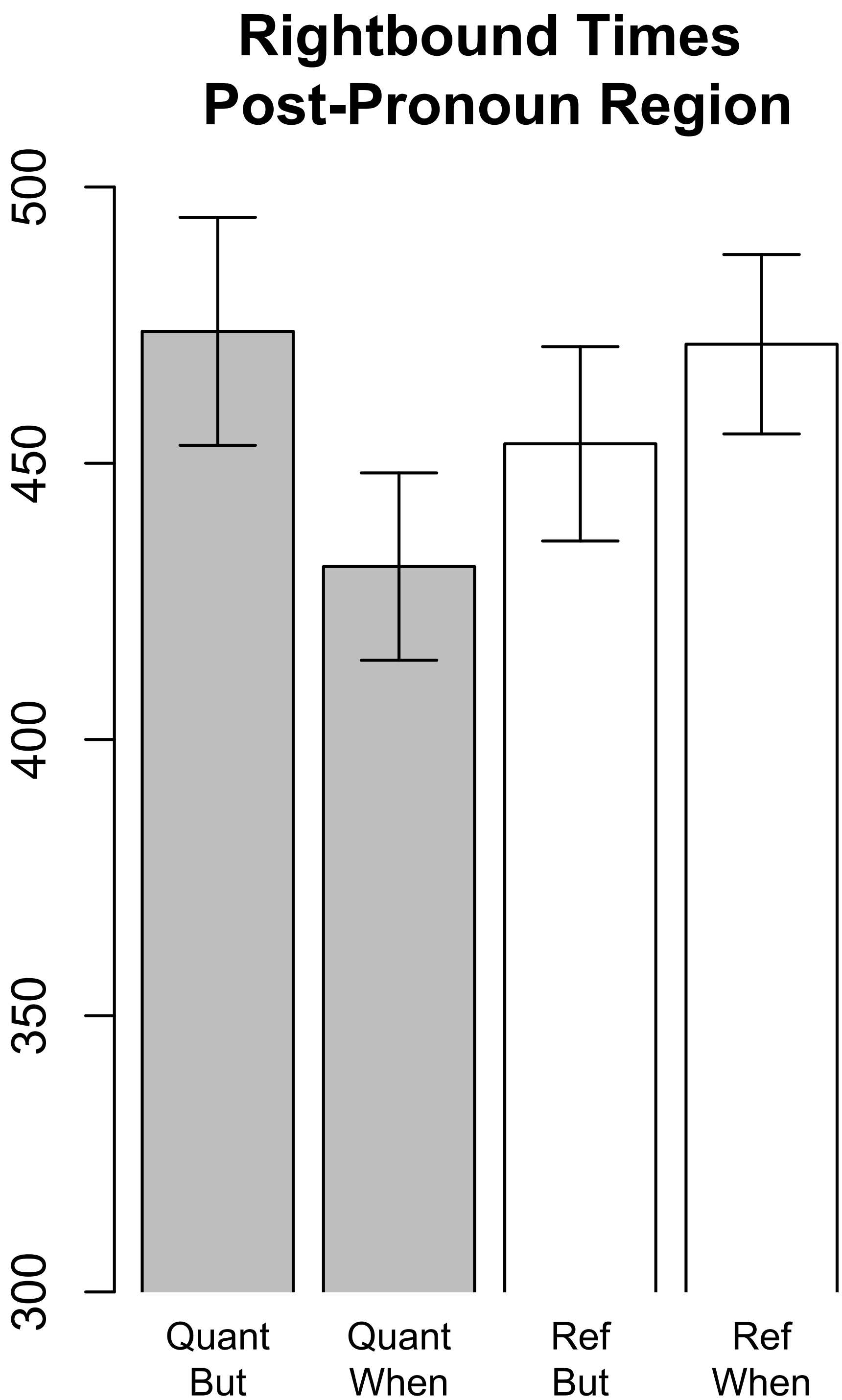


Figure 7.1: Right-bound Reading Times, Experiment 6c. Error bars represent standard errors.

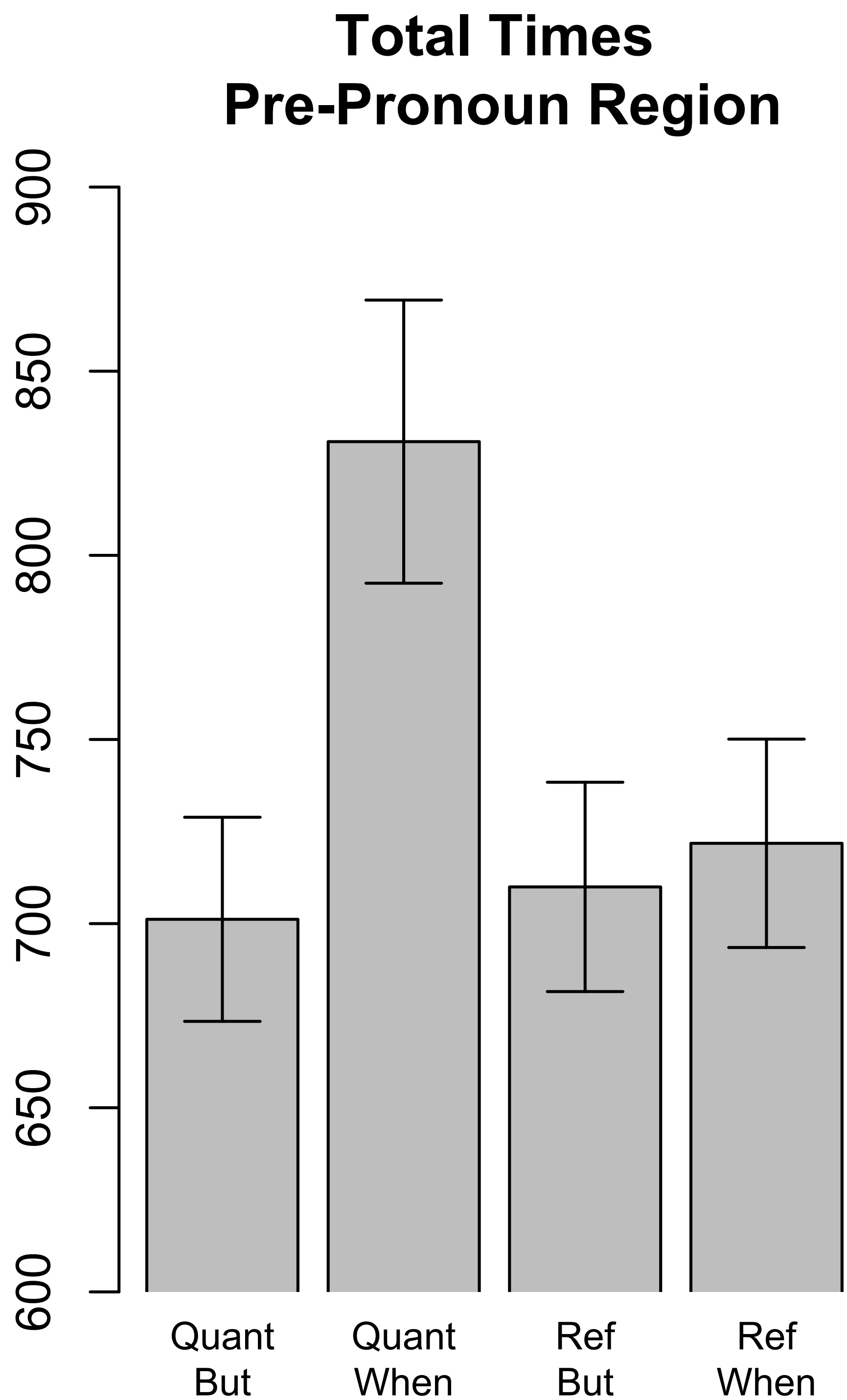


Figure 7.2: Total Reading Times, Experiment 6c. Error bars represent standard errors.

sentences in which the feature-matching phrase was a referential NP. In a second study participants reliably chose to interpret pronouns as bound by a QP only when the QP c-commanded the pronoun.

Eye-tracking measures showed that antecedent retrieval distinguished structurally-appropriate QPs from inappropriate QPs. This distinction was manifested in two patterns in the reading time measures. In early measures (First-pass and Right-bound time), the *Quantificational-When* condition was read significantly more quickly than its *Quantificational-But* counterpart in the region immediately following the critical pronoun. No pairwise difference was observed between the *Referential* conditions in this region. In later measures (Reread and Total times), a different pattern of effects emerged. Longer RTs were observed in the *Quantificational-When* condition in the pre-critical pronoun region relative to all other conditions.

These results support the contention that pronoun resolution distinguishes potential QP antecedents based on the structural relation they bear to the critical pronoun. The interaction effect indicating sensitivity to c-command was observed in early measures, immediately following the critical pronoun.

Participants display no difficulty in accessing a c-commanding quantificational antecedent (as evidenced by relatively short early measures in the *Quantificational-When* condition). Longer reading times in the *Quantificational-But* condition are consistent with the parser encountering difficulty at the pronoun because it must coerce a referent for the pronoun.

Longer reading times were observed in later measures in the *Quantificational-When* condition relative to the other conditions. I suggest that these effects should not be interpreted as indices of retrieval difficulty. Instead, the increased reading time in late measures may reflect the additional computational cost of processing a bound quantificational interpretation during the resolution stage following retrieval. Previous studies have shown that establishing a licit dependency between a quantificational antecedent and a pronoun results in longer reading times in comparison to the time required to establish a coreferential dependency (e.g. Carminati et al. 2002).

If the resolution of a pronoun with a referential antecedent results in shorter reading times in late measures than the resolution of a quantificationally bound pronoun, the pronoun in the *Quantificational-But* displays the RT hallmarks of a referential pronoun. With respect to Reread and Total times the *Quantificational-But* condition is comparable to the two *Referential* conditions. We conclude that the pronoun a bound variable reading is not considered for the pronoun in that condition. Rather, it is treated as referential. This aligns with the judgment patterns observed in Experiment 6b, where participants consistently chose to interpret the pronoun as a referential, and not a bound, pronoun.

The results of Experiment 6 establish that the c-command constraint exerts an early effect on retrieval. The study does not determine whether this information eliminates interference. It is not clear from these data whether the results observed because c-command information gates access to QPs or whether such information acts as one retrieval cue among many other cues (e.g. morphological feature-match) in determining the accessibility of an item to retrieval. Experiment 6 cannot distinguish between these two hypotheses because potential antecedents were identical in their feature composition except for their c-command information. Experiment 7 employs a standard interference-paradigm to test whether the morphological features of a structurally inappropriate QP interfere with antecedent retrieval.

7.4 Experiment 7: Bound Variable Pronoun Licensing II: Interference

Experiment 6 showed that c-commanding feature-matching QPs enjoy an advantage over non-c-commanding QPs in antecedent retrieval. Experiment 7 investigated whether this advantage reflects a categorical ban on the retrieval of non-c-commanding QPs, or whether antecedent retrieval is susceptible to partial-match interference from the morphological features of non-c-commanding QPs.

This section reports the results of an eye-tracking study that tested whether c-command information can be used to eliminate retrieval interference from feature-matching structurally inappropriate QPs.

7.4.1 Materials

The experimental items were made up of 36 item sets, each containing 6 conditions. 4 of the 6 conditions consisted of a 2x2 factorial design and the remaining 2 conditions were lexically similar, though structurally distinct, control conditions. The design of the factorial manipulation is first discussed, followed by discussion of the design of the control conditions.

7.4.1.1 Test Conditions

A critical pronoun appeared as the object of the matrix verb. Critical pronouns were either masculine or feminine (17 masculine, 19 feminine) and singular. Matrix subjects were plural, referential NPs modified by a RC. The head of the RC was related to a gap that was either the object of a preposition, or of the RC-internal verb. As critical pronouns were always singular, the matrix subject's plural marking rendered it ineligible as a potential antecedent for the pronoun.

(23) a. *Quantificational-Match*

The troop leaders [that **no girl scout** had respect for] had scolded *her* after the incident at scout camp.

b. *Quantificational-Mismatch*

The troop leaders [that **no boy scout** had respect for] had scolded *her* after the incident at scout camp.

c. *Referential-Match*

The troop leaders [that **the girl scout** had no respect for] had scolded *her* after the incident at scout camp.

d. *Referential-Mismatch*

The troop leaders [that **the boy scout** had no respect for] had scolded *her* after the incident at scout camp.

A potential antecedent for the critical pronoun was embedded as the subject of the RC attached to the matrix subject. As in Experiment 6, REFERENTIALITY manipulated whether the potential antecedent for the pronoun was *Quantificational*, or *Referential*. Unlike Experiment 6, the potential antecedent's structural position in relation to the pronoun was not manipulated: the potential antecedent, as the subject of the RC, never had scope over/c-commanded the critical pronoun across any test condition. When *Quantificational* the RC-internal potential antecedent bore the DE quantifier *no*. When *Quantificational* the RC-internal phrase was not a grammatical antecedent for the pronoun.[9]

The GENDERMATCH manipulation controlled the RC-internal potential antecedent's gender features. It could either *Match* or *Mismatch* the critical pronoun. Strongly gender-biased nouns, or inherently gendered nouns (e.g. widow, grandfather) were used.[10]

REFERENTIALITY crossed with GENDERMATCH determines the possibility of intra-sentential reference/valuation for the critical pronoun. In *Referential-Match* conditions, the pronoun can take the RC-internal matching NP as its antecedent, because coreference is allowed between two items in their relative position. In *Referential-Mismatch* conditions, there is no sentence-internal antecedent that matches the pronoun in features, so retrieval of potential antecedents will fail and difficulty ensue (Gerrig 1986; Osterhout & Mobley 1995; van Hoek 1997; Filik, Sanford & Leuthold 2008). When *Quantificational*, the RC-internal noun cannot grammatically antecede the pronoun, regardless of its gender, because (i) the fact that it does not scope over the pronoun precludes the possibility of it binding it, and (ii) coreference is not possible between QPs and pronouns.

[9]RCs are generally accepted to be *scope islands*, regardless of whether scope is taken to require c-command or otherwise. The only exception to this generalization come from so-called 'functional' readings of relative clauses (Cooper 1978, Lakoff 1970, Rodman 1976, Sharvit 1999, a.o.).

- The woman [that *every man* loves] is *his* mother.

These readings arise under very strict conditions, none of which are met in this study. First, the readings are generally thought to be restricted to copular or specificational constructions (cf. Heycock 1991, 1992; though see Sharvit 1999 for discussion of this generalization). Deviations from the restriction to identity sentences is only observed with the quantifiers *every*, and *each*, not with the quantifier *no*. Therefore, our use of *no* rules out the possibility of a functional reading of these relatives in which the quantifier might be thought to take scope over the pronoun.

[10]Many nouns were selected from the list of gender-biased referential nouns normed by Kennison & Trofe (2003). Other nouns were selected by intuition. See Sturt (2003), and van Gompel and Liversedge (2003) for discussion of using gender-biased nouns as a means of manipulating GENDERMATCH.

Unlike the previous experiment, this experiment provides a baseline against which the degree of interference exerted by structurally inappropriate QPs can be measured. If retrieval is able to use the relative position of the QP and the critical pronoun to constrain retrieval, we expect a REFERENTIALITY x GENDERMATCH interaction that follows the possibilities of pronoun interpretation determined by the grammar. The gender features of the *Quantificational* NPs should not exert any influence on the retrieval of antecedents for the pronoun, whereas the NP in *Referential-Match* condition will facilitate processing of the critical pronoun, relative to the *Referential-Mismatch* condition. If, on the other hand, only content features drive antecedent retrieval, we expect a main effect of *GenderMatch*: processing of the pronoun should be facilitated in both *Match* conditions. While the size of the effect does not have to be symmetric (as facilitatory interference effects often do not extinguish difficulty entirely), a pairwise effect of GENDERMATCH is expected between *Quantificational* conditions, with lower RTs in the *Quantificational-Match* condition.

7.4.1.2 Control Conditions

The test conditions pit the profile of coreference against that of potential illusory binding, rather than comparing grammatical binding with illusory licensing. As such, the test conditions do not provide an independent index of BVP processing.

We included two Control conditions to provide an experiment-internal baseline that showed the profile of successful variable binding under c-command. Control conditions were adapted from test items. For each test item, there existed a corresponding control. In control conditions a QP acted as the subject of an embedded propositional attitude, or manner-of-speech verb. The clausal complement of this verb contained a subject pronoun that the QP could either match or mismatch in gender. The predicate of the most deeply embedded clause was the passive of the corresponding predicate in the test conditions. Passives were used to avoid introducing an additional overt argument between QP and pronoun, to maximize similarity with test conditions.

(24) a. *Control-Match*

The troop leaders were sure **no girl scout** was afraid that *she* would be scolded after the incident at scout camp.

b. *Control-Match*

The troop leaders were sure **no boy scout** was afraid that *she* would be scolded after the incident at scout camp.

If the scope condition categorically restricts access to ungrammatical QPs, we expect the 3 conditions *Quantificational-Match*, *Quantificational-Mismatch*, and *Referential-Mismatch* to pattern similarly. The pronoun in each of these conditions should introduce difficulty, as it will lack a sentence-internal antecedent. In *Referential-Match* conditions, on the other hand, we would expect facilitation relative to these conditions because the pronoun could find a referential antecedent in the RC.

If the scope condition does not restrict access to the embedded QP, we expect interference from the gender-matching QP subject in *Quantificational-Match* condition relative to the *Quantificational-Mismatch* condition. Based on previous work, interference should manifest as relative facilitation in reading times in the *Quantificational-Match* condition in comparison to the *Quantificational-Mismatch* condition. That is, the experiment is design to find a measure of facilitative interference in RTs (Vasishth et al. 2008, Wagers et al. 2009, King et al. 2012). *Quantificational-Match* conditions should be read faster than the *Quantificational-Mismatch* conditions at the earliest point that retrieval effects are visible. The effect of GENDERMATCH does not need to be comparable in scale to the effect of GENDERMATCH expected between *Referential* conditions.

7.4.2 Experiment 7a: Acceptability Judgment Study

To first verify offline sensitivity to the structural manipulate, an online acceptability judgment study was run.

7.4.2.1 Participants

20 participants were recruited through Amazon Mechanical Turk (AMT) marketplace, and paid \$3.50 for their participation. Participant eligibility was determined in two ways. First, AMT automatically restricted access to the experiment to those with IP addresses corresponding to US-internal locations. Second, participants were required to take a Native speaker evaluation test commonly used in laboratory settings and score above a certain threshold.

7.4.2.2 Procedure

Presentation used the IBEX Farm web-based, experimental presentation platform, developed and managed by Alex Drummond (www.spellout.net/ibexfarm/docs). Sentences were presented one-at-a-time centered on the screen. Below each sentence was a series of 7 boxes corresponding to values on a 7 pt acceptability scale. Each box was marked with a number, 1 -7, in ascending order. Next to the left-hand side of the scale was a label 'bad', whereas the right-hand side of the scale bore the label 'good'. Participants could record a response by either clicking a box, or pressing a number on their keyboard.

Participants were instructed to rate sentences on the scale, with a rating of '1' corresponding to 'bad', 'unacceptable' or 'doesn't make sense', and a rating of '7' corresponding to 'good', 'totally acceptable' , or 'easy to understand'.

Though participants were not explicitly instructed to attend to pronouns, practice items illustrated cases where intrasentential coreference and binding failed. Explanations were given for ratings, as below.

(25) The duke believes no princess would say no to a pony if she were offered one.

. . .

Explanation: The last sentence was a bit more complicated, but most people would say it's good because it expresses a thought that Rudolph has about princesses in general.

(26) The baron thought no princess would come, and she didn't.

. . .

Explanation: That sentence was bad, or difficult to understand, because it went from talking about princesses in general to talking about some mystery person 'she' who hasn't been talked about before. It should therefore be rated low.

Condition	Average Rating	Standard Error
Quantificational-Match	2.35	.16
Quantificational-Mismatch	1.97	.15
Referential-Match	4.14	.20
Referential-MisMatch	2.18	.16
Control-Match	3.97	.20
Control-Mismatch	2.19	.14
Referential-Command	4.66	.23
Good-Binding	5.4	.15

Table 7.6: Experiment 7a: Acceptability Ratings for Items

7.4.2.3 Materials

30 of the 36 items used in the eye-tracking study were used in the judgment study. 4 sentences involving the same configuration as the control conditions of the test items, but with referential embedded subjects, were also included. These were referred to as 'Referential-Command' sentences. 6 additional sentences involving variable binding in various configurations were included in the list to provide an index of the general availability/acceptability of variable binding (Good-Bind Sentences). 58 filler sentences, all grammatical and of comparable complexity, were also included. This resulted in the presented sentences being roughly 70% acceptable.

7.4.2.4 Results

A main effect of REFERENTIALITY was significant (β = -1.00, s.e. = 0.148, t = -6.75, $p < .001$), due to higher average acceptability ratings for *Referential* sentences over *Quantificational* sentences. A significant effect of GENDERMATCH was also observed

(β = -1.16, s.e. = 0.148, t = -7.84, $p < .001$). *Match* conditions were rated, on average, more acceptable than *Mismatch* conditions. The REFERENTIALITY x GENDERMATCH interaction was also significant (β= 1.59, s.e. = 0.297, t = 5.36, $p < .001$). *Referential-Match* conditions received the highest mean acceptability rating, while all other conditions were rated relatively low. A pairwise difference between *Quantificational* conditions was also observed. *Quantificational-Match* conditions were rated slightly, but significantly, higher than *Quantificational-Mismatch* conditions (β = 0.37, s.e. = 0.176, t = -2.11, $p < .05$).

7.4.3 Discussion

The study sought to validate the grammatical generalization that non-c-commanding referential NPs are acceptable antecedents for a pronoun, but QPs that do not c-command a pronoun are not eligible antecedents.

Conditions that lack a feature-matching antecedent received low ratings. The condition in which a feature-matching referential antecedent was available to the pronoun exhibited the highest average acceptability. Similar acceptability was observed in the *Control-Match* condition, where a matching QP c-commanded a subject pronoun in the lower clause.

Filler conditions that assess the acceptability of the syntactic configuration used in control sentences, and the acceptability of uncontroversial cases of grammatical binding both receive high average acceptability scores. This confirms that the constructions used in the control conditions easily allow binding/coreference. The ratings are consistent with the ratings of the *Control-Match* conditions: participants found bound variable readings available and acceptable.

Conditions in which the antecedent was QUANTIFICATIONAL were rated, on average, less acceptable than their *Referential* counterparts. The condition in which the structurally inappropriate QP antecedent matches the pronoun in gender was rated slightly more acceptable than the condition in which the QP did not match the pronoun. This effect can be interpreted in one of two ways. First, the effect could indicate a lingering effect of retrieval interference. If interference was facilitative, it may have lead participants to occasionally judge the sentences as more acceptable. The second option is that the effect reflects not interference, but rather some minor facilitation of a later interpretive process. Perhaps the presence of a structurally inappropriate matching QP facilitates the coercion of an extra-sentential referent for the pronoun.

7.4.4 Experiment 7b: Eye-tracking while Reading

7.4.4.1 Participants

30 participants from the University of Maryland community participated in the study for pay, or course credit. Paid participants were compensated $10 for an hour of their time. All participants had normal, or corrected-to-normal vision, and were self-reported native speakers of English.

7.4.4.2 Procedure

The 36 experimental item sets were distributed into 6 lists in a Latin Square design. Each list contained 114 filler sentences for a total of 150 sentences per participant. The order of each list was pseudo-randomized. The maximum number of characters allowed on a single line on the visual display was 142 characters, and all sentences in the experiment fit on one line. The testing procedure and equipment were the same as those in Experiment 6c.

7.4.4.3 Analysis

Test sentences were divided into 6 regions of interest, as indicated below.

(27) TEST CONDITIONS

a. *Quantificational-Match*

The troop leaders that/ no girl scout/ had respect/ for had scol/ded her/ after the incident at scout camp./

b. *Quantificational-Mismatch*

The troop leaders that/ no boy scout/ had respect/ for had scol/ded her/ after the incident at scout camp./

c. *Referential-Match*

The troop leaders that/ the girl scout/ had no respect/ for had scol/ded her/ after the incident at scout camp./

d.*Referential-Mismatch*

The troop leaders that/ the boy scout/ had no respect/ for had scol/ded her/ after the incident at scout camp./

(28) CONTROL CONDITIONS

a.CONTROL-MATCH

The troop leaders were sure no girl scout was afraid that she would be scolded after the incident at scout camp.

b.CONTROL-MISMATCH

The troop leaders were sure no boy scout was afraid that she would be scolded after the incident at scout camp.

For the test conditions, the fifth region is the critical region because it contains the critical pronoun and the three characters preceding the pronoun. The three characters preceding the pronoun were included in the region to (i) account for possible parafoveal processing of the pronoun when reading the preceding context, and (ii) to increase the number of observations for the critical region. Small, high frequency words are often skipped in eye-tracking tasks (Rayner 1997, 1998; Rayner et al. 2004; Rayner, Ashby, Pollatsek, & Reichle 2004; Ashby, Rayner & Clifton 2005; Staub & Rayner 2007 a.o.), therefore limiting the region to just the pronoun would result in a large amount of data loss.

Reading times reported are those reported in Experiment 6c. Experimental fixed effects were the factors GENDERMATCH and REFERENTIALITY and their interaction.

Fixed effects were coded using simple difference sum coding (*Match* conditions were coded as -.5, *Mismatch* conditions as .5; *Referential* conditions -.5, *Quantificational* as .5). Reported coefficients reflect the magnitude of the difference between levels of a given factor in milliseconds. The model also has a maximal random effects structure with random intercepts for subject and items, as well as random slopes for the fixed effects and their interaction by participants and by items. A fixed effect was considered significant if its absolute t-value was greater than 2, which indicates that a given effect's 95% confidence interval does not include 0 (Gelman & Hill 2005). Reported coefficients whose absolute t/z-value is greater than 2 are significant at the $p < 0.05$ level.

Raw mean fixation times are presented in Tables 7.7 and 7.8 below.

7.4.4.4 Results

Reading Time Results

Test Conditions

Pre-critical Region. In the pre-critical region, no effects achieved significance (main effects $ts < 1$; interaction $t < 1.6$) in First-pass time, Right-bound, or Reread times. In total time a marginal main effect of GENDERMATCH was observed ($\beta = 56.4$, s.e. = 29.5, $t = 1.91$), driven by faster Total times in Match conditions.

Critical Pronoun Region. In First-pass times, the model reveals a significant main effect of REFERENTIALITY ($\beta = 31.6$, s.e. = 14.0, $t = 2.25$, $p < .05$) driven by increased

	Pre-Critical	Critical Pronoun	Post-Pronoun	Final
First-Pass Times				
Quant-Match	344 (13)	311 (13)	338 (27)	741 (38)
Quant-MisMatch	368 (16)	293 (13)	356 (24)	819 (46)
Ref-Match	364 (16)	254 (10)	294 (21)	848 (40)
Ref-MisMatch	358 (14)	285 (11)	318 (18)	803 (49)
Right-bound Times				
Quant-Match	399 (17)	342 (16)	403 (38)	1232 (61)
Quant-MisMatch	418 (19)	317 (14)	415 (32)	1383 (82)
Ref-Match	404 (17)	271 (12)	381 (39)	1280 (74)
Ref-MisMatch	396 (14)	330 (17)	418 (41)	1405 (70)
Reread Times				
Quant-Match	518 (34)	410 (44)	398 (32)	1005 (86)
Quant-MisMatch	502 (48)	365 (29)	497 (59)	1268 (115)
Ref-Match	536 (49)	319 (34)	506 (79)	1133 (137)
Ref-MisMatch	558 (43)	455 (48)	502 (59)	1121 (95)
Total Times				
Quant-Match	585 (29)	438 (24)	527 (36)	1232 (61)
Quant-MisMatch	631 (33)	453 (23)	527 (35)	1428 (88)
Ref-Match	601 (33)	378 (22)	488 (42)	1280 (76)
Ref-MisMatch	673 (34)	473 (27)	533 (38)	1405 (70)

Table 7.7: Reading Time Measures (Test Conditions), Experiment 7b

First-pass times in Quantificational conditions relative to Referential conditions. The REFERENTIALITY x GENDERMATCH interaction was marginally significant (β = -48.0, s.e. = 25.2, t= -1.91, $p < .05$), driven by lower RTs in the Referential-Match condition relative to the other three. Post-hoc pairwise comparisons revealed a significant effect of REFERENTIALITY between Match conditions (β = 54.0, s.e. = 15.5, t = 3.49, $p < .01$), driven by longer RTs in the Quantificational-Match condition in comparison to the Referential-Match condition. No corresponding effect was observed between Mismatch conditions ($t < 1$). No pairwise effect of GENDERMATCH was found between Quantificational conditions ($t < 1$).

Effects similar to those observed in First-pass times were observed in Right-bound times. The main effect of REFERENTIALITY was marginally significant (β = 31.7, s.e. = 17.9, t= 1.77). The REFERENTIALITY x GENDERMATCH interaction was significant (β = -78.5, s.e. = 32.8, t = -2.40, $p < .05$). Pairwise comparisons revealed a significant effect of REFERENTIALITY within Match conditions (β = 69.2, s.e. = 18.4, t = 3.76, $p < .01$), but not within Mismatch conditions ($t < 1$). Pairwise comparison between Quantificational conditions was not significant ($t < 1.2$), suggesting no reliable difference between the Quantificational-Match and Quantificational-Mismatch conditions.

In reread times, there were no significant main or interaction effects. In total times, there was a marginally significant main effect of GENDERMATCH (β = -51.9, s.e. = 26.1, t = 1.99). This effect was largely driven by a pairwise difference between Referential conditions (β = 85.0, s.e. = 28.7, t = -2.97, $p < .05$), with the Referential-Match condition read more quickly than the Referential-Mismatch condition. The pairwise difference between Quantificational conditions was not significant ($t < 1$).

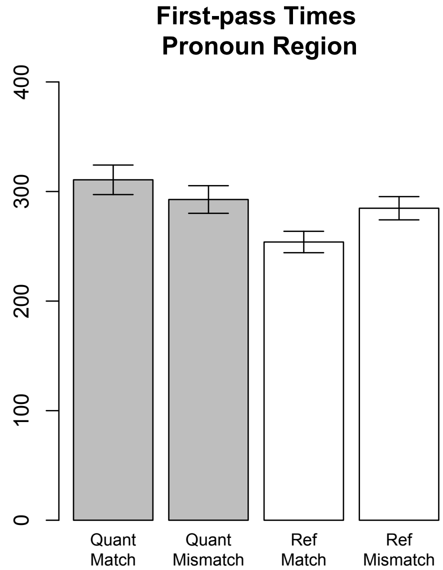


Figure 7.3: First-pass Reading Times, Experiment 7b. Error bars represent standard error.

Post-Pronoun Region. In the post-pronoun region there were no significant First-pass, Right-bound, Reread, or Total Time effects (ts $<$ 1).

Sentence-Final Region. In the sentence-final region, a marginal main effect of GENDERMATCH was found (β = 206.0 , s.e. = 117.0, t = 1.77, $p < .10$) in reread times. This effect was driven by lower RTs in Match conditions, relative to Mismatch conditions. The marginal effect of GENDERMATCH was also observed in Total time (β = 118.0, s.e. = 69.5, t = 1.70, $p < .10$), once again carried by lower RTs in GENDERMATCH conditions.

Control Conditions

First Pass Reading Times

No main effect of GENDERMATCH was observed in pre-critical, critical, or post-critical regions.

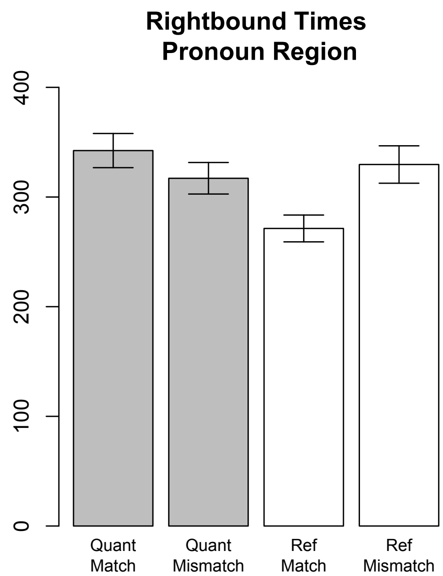


Figure 7.4: Right-bound Reading Times, Experiment 7b. Error bars represent standard error.

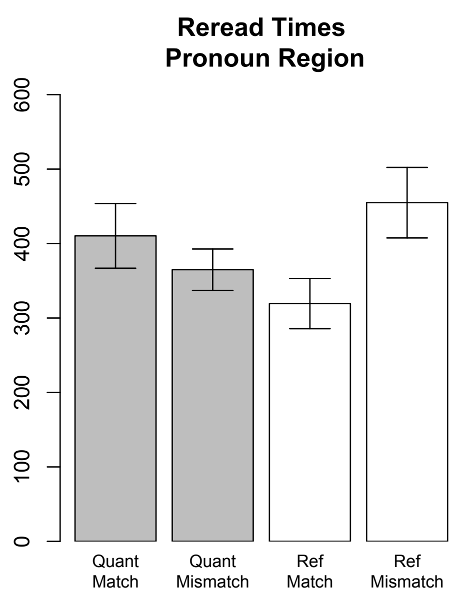


Figure 7.5: Reread Times, Experiment 7b. Error bars represent standard error.

	Pre-critical	Critical Pronoun	Post-Pronoun	Final
First-Pass Times				
Control-Match	775 (31)	309 (14)	250 (11)	471 (18)
Control-Mismatch	798 (36)	293 (13)	256 (9)	489 (20)
Right-bound Times				
Control-Match	924 (34)	326 (12)	262 (14)	533 (20)
Control-Mismatch	960 (43)	324 (13)	275 (11)	555 (21)
Reread Times				
Control-Match	1011 (73)	389 (29)	324 (26)	599 (39)
Control-Mismatch	1373 (164)	462 (45)	322 (29)	682 (44)
Total Times				
Control-Match	1447 (63)	483 (21)	363 (19)	834 (36)
Control-Mismatch	1772 (132)	532 (30)	400 (21)	931 (40)

Table 7.8: Experiment 7b: Reading Time Measures (Control Conditions)

Pre-critical Region. No effect of GENDERMATCH was observed in early measures (ts < 1). A significant effect of GENDERMATCH was observed in Reread times, with the Control-Mismatch condition taking longer to read than the Control-Match condition (β = 382.0, s.e.= 173.0, t= 2.21). This pairwise effect achieved significance in Total times, once again with Mismatch conditions taking longer to read than Match regions (β = 332.0, s.e.= 132.0, t= 2.52, $p < .05$).

Critical Pronoun Region. In the critical pronoun region no effects achieved significance (ts < 1.5).

Post-Pronoun Region. In the post-pronoun region no effects achieved significance (ts $<$ 1.5).

Pronoun+2 Region. Two words after the pronoun a significant effect of GENDERMATCH was observed in Total times (β = 91.8, s.e.= 43.9, t= 2.09), due to elevated RTs in Control-Mismatch conditions relative to Control-Match conditions. No other effects were found in other measures.

7.4.5 Discussion

Experiment 2 sought to determine whether grammatically inappropriate feature-matching QPs interfered with antecedent retrieval for a pronoun. The study contrasted the accessibility of a feature-matching referential NP embedded inside an RC with that of a QP in the same position. Conditions with corresponding non-matching NPs and QPs provided a comparison measurement of retrieval failure.

An additional pair of Control conditions provided an index of successful bound-variable pronoun licensing. An offline acceptability study verified that comprehenders do not consider a feature-matching QP as a grammatical antecedent for a pronoun that the QP c-commands.

In sentences with a matching referential antecedent for the pronoun, the pronoun and the immediately subsequent regions are read more quickly than in sentences where no matching antecedent is present. These results indicate that antecedent retrieval can rapidly access a referential matching NP embedded in a relative clause.

In sentences with a matching non-c-commanding quantificational antecedent for the pronoun, reading times at and after the pronoun do not differ from sentences in which the quantificational antecedent does not match the pronoun in gender. This lack of partial-match effects suggests that structurally inappropriate QPs do not interfere with antecdent retrieval.

Given the absence of observable interference in the eye-tracking times, the slightly raised acceptability rating of the *Quantificational-Match* test conditions relative to *Quantificational-Mismatch* should not be interpreted as the result of interference. They more likely reflect eased coercion of a referent for the unheralded pronoun. This interpretation is consistent with the late effects (in reread and total time) of GENDERMATCH. Here it does seem that the presence of a feature-matching, though inaccessible, QP does ease processing of the pronoun, broadly construed. The task of positing an extra-sentential referent for the unheralded pronoun in test sentences with a feature-matching QP is eased because the QP's reference set provides salient and relevant properties that can be predicated of the extra-sentential reference. This minimizes cost of accommodating a new referent into the discourse (as represented by the sentence).

The differential time-course of mismatch effects between test and control conditions does raise questions regarding the accessibility of quantificational NPs generally. A number of possible explanations exist for the absence of a GENDERMATCH effect in early measures in the Control conditions, all of which will discussed at greater length in the general discussion. Two general kinds of explanations exist. First, it is possible that GENDERMATCH effects are masked in earlier measures by increases in processing time associated with successful binding. The second explanation would take late effects as indicative of a general delay, or dispreference, for attempting bound variable readings for pronouns. This latter option seems unlikely given that successful binding in the previous experiment, and in other experiments (e.g. Cunnings et al. 2012), displays a time-course comparable to coreference resolution.

Regardless of the interpretation of the late effects, however, it is clear that the feature content of structurally inappropriate quantificational NPs does not interfere with antecedent retrieval for pronouns.

7.5 General Discussion

Results from Experiment 6 suggested that antecedent retrieval can distinguish between c-commanding and non-c-commanding QPs. Experiment 7's results indicated that non-c-commanding matching QPs are not accessed during with antecedent retrieval, although referential matching NPs in the same position are readily accessible.

Taken together the studies demonstrate that structurally inappropriate feature-matching QPs do not interfere with the retrieval of antecedents for a pronoun. Prima facie, these results pose a challenge for any retrieval system whose antecedent retrieval probes are comprised of morphological and local structural cues, but not relational information like MCCOMMAND. These results do not, necessarily, undermine the idea that the parser exploits cue-based memory access function that does not make reference to relations. As stated at the outset, an alignment between the extension of a grammatical constraint stated in terms of relational information, and the dependencies considered generated by a function that makes use of retrieval does not automatically entail that retrieval is relation-sensitive. The restriction comes from some part of the parsing model, but it need not be retrieval. In what follows I explore different methods for achieving or approximating the observed relational sensitivity without altering fundamental assumptions about the retrieval mechanism.

7.6 Accommodating Sensitivity to the Scope Constraint With Cue-Based Retrieval

The version of the scope condition stated at the outset of the chapter holds that a QP is a viable potential antecedent for a pronoun so long as the QP MCCOMMANDs the pronoun. Another way of stating this constraint is the following: A pronoun must either (i) sit on the same spine as the minimal clause containing the QP or (ii) be contained in some constituent dominated by that spine.

In this section I propose a feature-based account of the scope condition. I first consider treating the scope condition as a type of spine-mate condition, requiring match between a QP and its antecedent on a spine feature. I argue that this strategy faces two problems. First, it undergenerates licit BVP readings. A spine-mate condition only captures the first term in the disjunctive statement above; the account would not allow covaluation between a QP and a pronoun not directly connected to the QP's spine. The second problem is that imposing a spine-mate condition on retrieved items singles out only those antecedents for the pronoun that are MCCOMMAND the pronoun. Such a retrieval probe would not serve to identify referential NPs that do not MCCOMMAND the pronoun, but are nevertheless potential coreferential antecedents.

The second solution to the problem that I consider utilizes a dynamically-updated feature, ACTIVE, to encode whether a QP remains syntactically viable. ACTIVE is a binary feature. Similar to LOCAL, ACTIVE is subject to rewriting at the end of syntactically-defined domain. Unlike LOCAL, the domain of ACTIVE is not rigidly set. At the end of a spine, *S*, any QPs whose scope is restricted to *S* have their ACTIVE features rewritten to reflect that they should be treated as effectively inert for later retrieval. Importantly, it is shown that the ACTIVE account of the scope constraint does not face the same challenges as a spine-mate strategy. Modulo one or two technical complications, the ACTIVE account provides a robust reformulation of the unbounded, relational scope condition.

7.6.1 A Spine-based Retrieval Strategy Is Not Viable

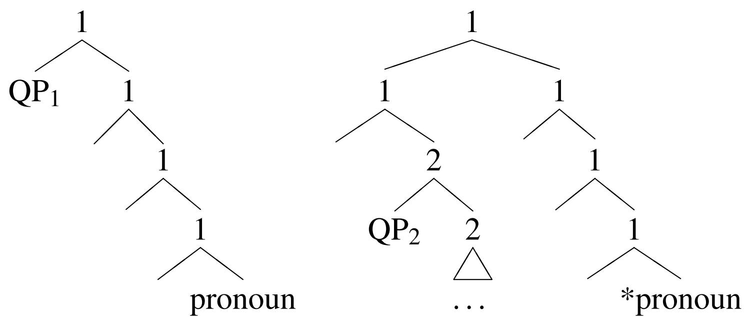
As noted before, use of a proxy feature to approximate the effects of a relational constraint under certain conditions is a strategy that can be used to explain the apparent insensitivity of reflexive licensing to similarity-based interference. One possible feature-based strategy would exploit a feature that tracks the command-path, or 'spine' of a QP. With such a feature a 'spine-mate' strategy could be implemented. Such a strategy was considered as a proxy for a general 'c-command feature' by Alcocer and Phillips (2012).

Under the strict c-command account of bound-variable licensing, a QP and the pronoun it binds are *often* daughters of the same 'spine' in the syntactic tree. It is possible to encode the closest spine dominating an NP as a feature of that NP. Suppose that every time a new spine was started, a counter was incremented, and all NPs on that spine were tagged with the current value of the counter.

(29) $$\textsc{spine}(x) = \begin{cases} \text{the current spine index} & \text{if x is on the current spine} \\ \textit{null} & \text{otherwise} \end{cases}$$

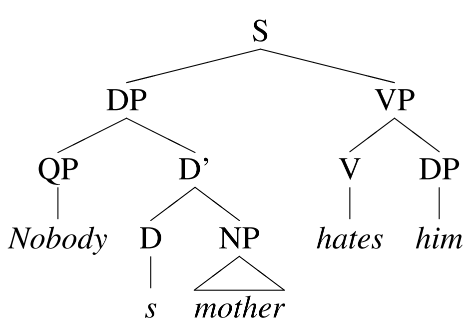
Under this encoding schema, QPs and pronouns in our experiments would mismatch on [spine]. For example, in Experiment 7b, the RC-internal QP would bear the feature [spine:2], but the pronoun would bear [spine:1], because it is an element on the main spine of the tree.

(30)



A strict spine-based account leaves cases of McCommand that go beyond surface c-command unaccounted for. We saw that a surface c-command constraint is not adequate to characterize all licit cases of BVP-licensing. Consider the case of binding out of possessors. In (31) *nobody* does not c-command *him* (and is therefore not its spine-mate). Yet, the QP may nevertheless bind the pronoun.[11]

(31) Nobody's mother hates him.



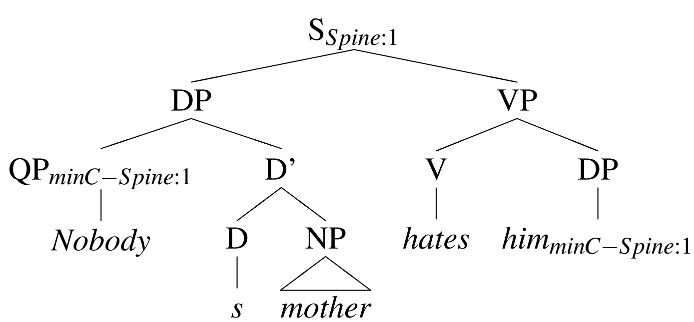
[11]Though no studies have, to my knowledge, compared the real-time licensing of these cases with the licensing of cases of variable binding under strict c-command, intuition suggests that there is little reason to suspect *nobody* is more difficult to retrieve than it would be if it were a c-commander of *him*.

Any spine-mate strategy used to encode the scope condition must loosen the criteria for spine feature assignment. MCCOMMAND is a relation that tracks the *minimal clause* containing the QP. Instead of a strict spine feature, a minimal clause-spine feature appears to be warranted.

(32) $$\text{MINC-SPINE}(x) = \begin{cases} \text{spine\# of x's minimal clause} & \text{if x is a QP} \\ \text{current spine\#} & \text{if x is a non-quantificational DP} \\ \textit{NULL} & \text{otherwise} \end{cases}$$

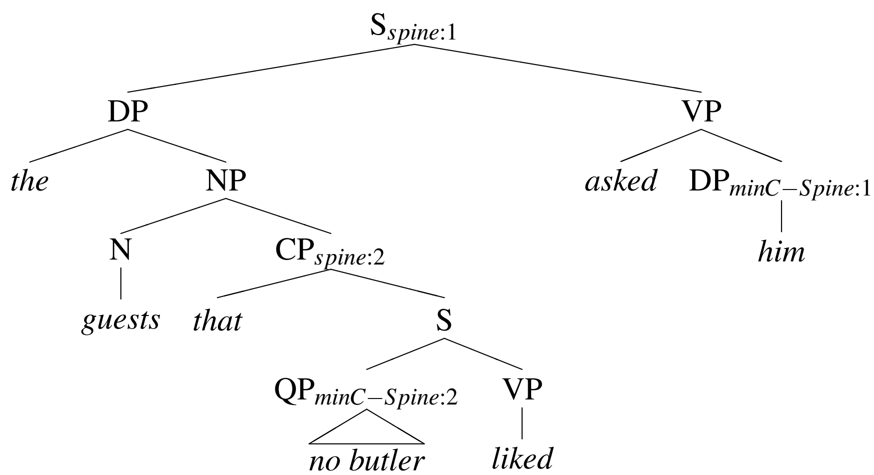
The scope condition under such an account would become a 'minC-spine-mate' constraint. The minC-spine approach to the scope condition achieves the right results in the core cases considered so far. Binding from a possessor position falls under the new account.

(31) Nobody$_{minC-Spine:1}$'s mother hates him$_{minC-Spine:1}$.



While binding out of a RC attached to a subject, for example, does not. The QP in the sentence below receives a MINC-SPINE: 2 feature because it bears the spine index of the embedded RC.

(33) The guests [that no butler $_{minC-Spine:2}$ liked] asked him$_{minC-Spine:1}$.

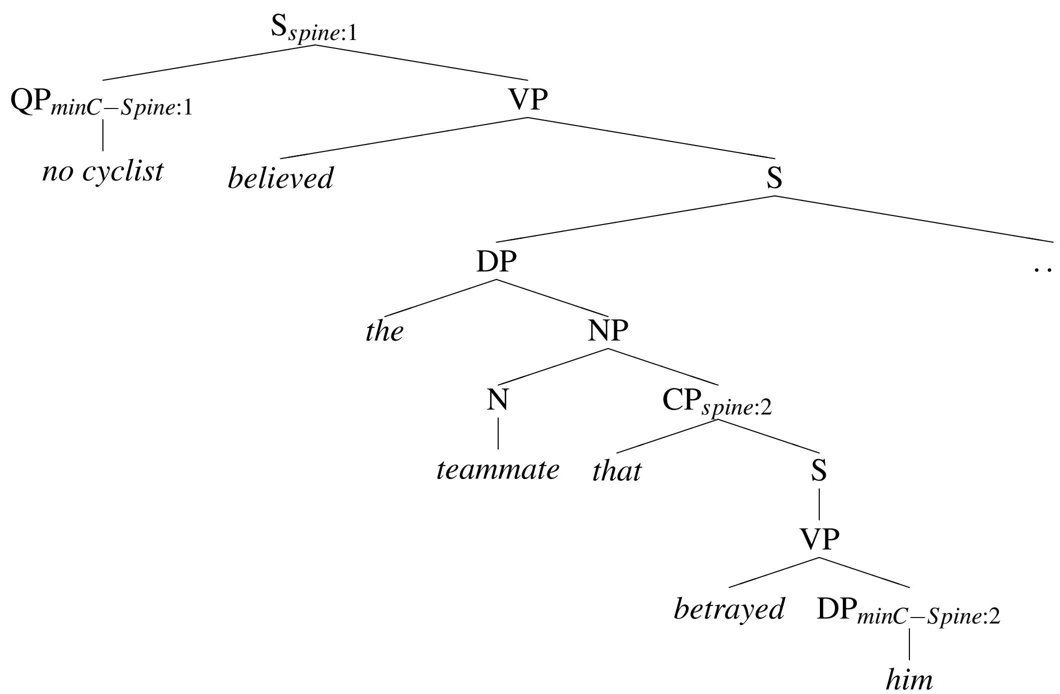


The minC-spine approach appears to have a minimum level of empirical purchase. Moreover, it allows the scope condition to be reformulated in a manner that does not require retrieval to appeal to relations between nodes in a tree. This is because an item is assigned a minC-spine feature independently of other nodes in the tree. Retrieval reduces to a simple feature-matching procedure.

Despite its initial appeal, there are good reasons to doubt that even a minC-spine-mate strategy is at work: even with modifications such a strategy strategy under-generates possible readings. As mentioned above, this is due to the fact that not all QP-pronoun pairs that can enter into a BVP relation are minC-spine-mates.

The primary case of under-generation is found with a pronoun that is not a minC-spine-mate of the QP, but is nevertheless MCCOMMANDed by it. In the example below, the pronoun *him* can be bound by the QP *no cyclist*, even though it is embedded inside a RC, viz. it lies on a distinct spine from the QP.

(34) *No cyclist* believed [the teammate [that betrayed *him*]] would dare show enter the race.



Because *him* and *no cyclist* do not share a minC-spine feature, *no cyclist* would not be retrieved as a potential antecedent.

The second problem that any strategy that attempts to include structural restrictions on BVP into a probe for retrieving a pronoun's potential antecedents is that such a probe would only contact a subset of all of the possible structural positions in which an antecedent might be found. Cues that restrict access to a c-command domain are only useful in retrieving quantificational NPs. Referential antecedents do not exhibit the same restrictions. If an effective cue set for restricting access to scope-taking potential antecedents were devised and deployed in all retrieval probes, retrieval would be forced to exclusively consider c-commanding/scoping NPs as potential antecedents, and would therefore fail to access potential referential antecedents in non-scoping positions.

If a spine-mate strategy were implemented by retrieval, two distinct retrieval probes would be required for every pronoun: one to retrieve potential binders, the other referential antecedents. The simplest assumption is that a single retrieval probe is used to contact all possible antecedents in a given context. If the same probe is used to retrieve potential referential and quantificational antecedents simultaneously, it should not contain any reference to structural constraints specific to one kind of antecedent.

7.6.2 ACTIVE-ly Restricting Retrieval

The strategy pursued in this section is to use a dynamically-updated feature to track whether an NP is accessible for retrieval. This feature is exclusively used to distinguish QPs that no longer take scope over the segment of the sentence currently being processed from all other chunks in memory. Unlike the minC-spine-mate strategy, the account does not require retrieval to use BVP-specific structural restrictions. That is, retrieval itself does not make even oblique reference to the scope condition. Instead, the feature 'pre-computes' the structural accessibility of an item before it is retrieved and encodes the result as a non-relational content feature on that item.

The account begins from the following observation of the distribution of accessible QPs: once the parser has left the scope domain of a QP, that QP should be rendered inert for subsequent retrieval. Suppose there is a feature that encodes whether a QP is a viable retrieval target at some point in the parse t. Call this feature t-ACTIVE. A chunk could bear t-ACTIVE:1 if it was still a possible target, and t-ACTIVE: 0 if it was inert. If such a feature existed, the scope condition could be simply encoded as below.

(35) SCOPE CONDITION (non-final):

Pronouns must have $t-$ACTIVE: 1 antecedents.

An illustration of the desired distribution of the t-ACTIVE feature for three sentences is given below. Because *no cyclist* is an eligible antecedent for the pronoun *him* in (36-a), it should bear t-ACTIVE:1. In (36-b) it should be inert for retrieval because it does not scope over the pronoun, but in (36-c) it should be t-ACTIVE:1 in order to reflect that it does, in fact, scope over the pronoun.

(36) a. *No cyclist*$_{t-Active:1}$ would admit the spectators hated *him*.

b. The spectators [that *no cyclist*$_{t-Active:0}$ saw] would admit the people hated *him*.

c. The spectators [that *no cyclist*$_{t-Active:1}$ would admit hated *him*] stood by the roadside.

The problem with such a feature is that it is inherently relational, due to the fact that it makes indexical reference to a particular stage, t, in the parse of the sentence. As such, it would be just as difficult to encode a t-ACTIVE feature as it would a c-command feature. In order for such a feature to be viable, it must be stripped of its inherent indexicality. Instead of a t-ACTIVE feature, we need an ACTIVE feature.

A QP's scope terminates at the end of its minC-spine. If a pronoun is encountered after the edge of the spine, the QP does not scope over it. This is the case in (36-b). However, if the pronoun precedes the edge, the QP does scope over it. It would appear that the edge of the QP's scope domain is the point relevant for determining whether a chunk bears ACTIVE: 0. If we suppose that all NPs are introduced as ACTIVE: 1, we must simply specify a procedure that selectively de-activates a QP at its respective domain edge to obtain the right results. This procedure can be described as follows: Upon encountering the end of a spine (if the parser must 'pop-up' a level), any QPs whose scope is limited to that spine must be marked ACTIVE: 0.

(37) UPDATE PROCEDURE TO RENDER NON-SCOPING QPS INERT TO LATER RETRIEVAL:

```
If the parser reaches end of a minC-spine S:

    While any QPs on S are ACTIVE: 1:
        ret = retrieve([+Quant,minC-spine:  S,ACTIVE:1]);
        ret[ACTIVE] = 0;
```

Note that although assignment of ACTIVE: 0 is, in some sense, structure-dependent, it is not relation-dependent. The domain edge acts to 'switch off' the ACTIVE feature, without reference to other items that the QP does, or does not, enter into a relation with. Once we have this procedure, the scope condition can be restated in terms of the ACTIVE feature.

(38) SCOPE CONDITION (final):

Pronouns must have ACTIVE: 1 antecedents.

A step-by-step illustration of how the system would restrict access to a feature-matching QP that no longer scopes over a pronoun is given below. When a QP on a spine is encountered, it is initially encoded just as any other syntactically-viable chunk would be - it is assigned a value of ACTIVE: 1. As long as the parser continues to encounter input on the same spine the QP will remain marked ACTIVE: 1. As an ACTIVE: 1-marked NP, the QP would be accessible to retrieval. The QP's accessibility ends at the right edge of the spine.

(39) The guests that [$_{spine:2}$ . . .

Recognize new spine

The guests that *no butler*$_{Active:1}$ . . .

Assign QP ACTIVE: 1

The guests that *no butler*$_{Active:1}$ liked . . .

Continue

The guests that *no butler*$_{Active:1}$ liked *asked* . . .

'asked' Signals End of Spine: 2

The guests that *no butler*$_{Active:0}$ liked asked . . .

Rewrite ACTIVE *on QP*

The guests that *no butler*$_{Active:0}$ liked asked *him* . . .

Retrieve ACTIVE: 1 antecedent, FAIL

The pronoun in the example above would trigger a retrieval for an ACTIVE: 1 antecedent. Assuming that the ACTIVE feature is weighted in such a way to restrict access to mismatching NPs,*no butler* not be retrieved.

7.6.2.1 Squaring ACTIVE with LOCAL

If we adopt the ACTIVE account of the Scope Condition, antecedents for pronouns must meet two criteria. They must:

i) be ACTIVE: 1

ii) be LOCAL: 0 (when the pronominal trigger is LOCAL: 1)

(40) illustrates why imposing both constraints is necessary. In (40), *him* matches both *the spectator* and *no cyclist* on morphological features. As it stands, to exclude both matching NPs, the retrieval triggered by the pronoun must specify that its antecedent be LOCAL: 0 and ACTIVE: 1. Exclusive use of either gating feature would predict interference from the NP marked with the other feature.

(40) The spectator$_{Local:1,Active:1}$ [that no cyclist$_{Active:0,Local:0}$ could stand] applauded *him*$_{Local:1}$.

Use of multiple gating features engenders complications with the preferential weighting system I adopted to limit partial-match interference. The problem arises from the interaction of multiple gating cues with the independent assumption cues contribute to activation individually (the individual contribution assumption, or ICA).

The ICA relates to the terms allowed in the function that determines a chunk's level of activation. The ICA is implicit in cue-combinatorics schemes such as the linear function assumed by Lewis & Vasishth (2005), in which each cue receives an individual term in the calculation of a chunk's activation. Suppose that the activation of a pronoun was determined according to the following equation (adapted from Lewis & Vasishth 2005).

(41) $A = w_{Local}S_{Local} + w_{Active}S_{Active}) + w_{gender}S_{gender} + w_{number}S_{number}$

Because each gating cue has its own individual term, it can contribute to a chunk's activation independently of the other cues. This renders the function subject to partial-match interference. Consider the activation values that would be assigned to the chunks *no cyclist* or *the spectator* from (40). If an item matches one gating cue, but not another, its activation will nevertheless be increased. One might consider attempting to nullify these effects through cue weighting, but regardless of how weights were distributed, the effects would persist. If LOCAL bore the weight mass, *the spectator* would be correctly excluded from retrieval, but we would expect retrieval interference from *no cyclist*, which matches the probes LOCAL: 0 specification. The opposite would be the case if ACTIVE were assigned the mass: *the spectator* would interfere. If the weights were more evenly distributed, partial-match interference would persist.

These examples illustrate that when multiple gating cues are required, it is not the sum of each feature's individual contribution that guarantees correct access, but rather the conjunctive contribution, or interaction, of the features. Chunks should only be considered if they match both gating cues.

I now consider two possible responses to the challenges presented by multiple gating cues. The first involves dispensing with the ICA. The second considers how the interrelationship between ACTIVE and LOCAL might obviate the need to use multiple gating cues.

If the ICA were weakened, or abandoned altogether, it would be possible to allow *interaction* terms into the activation functions. For example, a term that multiplicatively combined the strengths of two gating features would provide roughly the right results. A low strength of association between a chunk *i* and either of the gating features would result in a low value for the activation term.

(42) $A_i = w_{interaction}(S_{Local} * S_{Active}) + w_{gender}S_{gender} + w_{number}S_{number}$

This kind of cue-combinatorics scheme is a kind of hybrid system that incorporates both multiplicative and additive components. Though there is nothing, in principle, incoherent about such a function, those that strive for a completely linear cue-combinatorics might look askance at such a move. Moreover, invoking this mechanism for one specific case might appear *ad hoc*.

The second option I consider is that ACTIVE and LOCAL exist in a hierarchy, such that under certain circumstances specification of one feature presupposes a particular value on another. To be more specific, I propose that LOCAL: 1/0 marking implies ACTIVE: 1 marking. Therefore, probes that specify LOCAL: 1, for example, will only retrieve chunks that bear ACTIVE: 1. Thus, the scope condition can often be implicitly enforced, without specifying ACTIVE: 1 in the retrieval probe. I consider cases when this implicit specification works and when it does not. The upshot of this final segment is the following: with minor modifications to the LOCAL feature and slight elaboration of the cue generation procedure, a system that uses a single gating feature with a linear cue-combinatorics scheme can achieve the same results as a system that uses multiple features and interaction terms.

We saw above that when two gating cues independently contributed to a chunk's activation, partial-match interference seemed unavoidable. The previous account introduced an interaction term between the features in the cue-combinatorics to overcome the problems of cue independence. The present account explores placing the locus of interdependence in the encoding and update functions, rather than the cue-combinatorics.

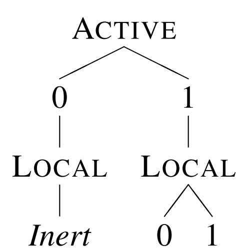
The procedure for updating ACTIVE discussed above assigns a QP the ACTIVE: 0 feature at the end of its scope domain. This renders it inert for the purposes of retrieval cued on ACTIVE: 1. We saw, however, that this does not prevent the same QP from being retrieved based on its LOCAL feature. This is not desired, nor does it accord with our intuition. A QP that is ACTIVE: 0 should not be accessible for the purposes of LOCAL-based retrieval at all. To render a chunk inert for the purposes of LOCAL retrieval, we must mark it with a LOCAL feature that will never match a retrieval probe. This entails that the values LOCAL: 1, nor LOCAL: 0 are eligible. Therefore, a third value for LOCAL is required. Call this value LOCAL: INERT. This elaboration yields the following possible values of LOCAL.

(43) Complete List of Values For Local

$$\textsc{Local}(x, \text{curr.Clause}) = \begin{cases} 1 & \text{x is an argument of curr.Clause's verb} \\ 0 & \text{if x is any other DP} \\ \text{Inert} & \text{if x is } \textsc{Active}\text{: 0} \\ \textit{NULL} & \text{if x is not a DP} \end{cases}$$

If the assignment of Local: Inert is tied to the assignment of Active: 0 to a chunk, then the Active update function bears responsibility for not only changing a QP's Active value, but also its Local value. Once we assume the existence of Local: Inert, the features fall into the hierarchy below.

(44)



This hierarchy produces the following result: If a retrieval probe specifies a value of 0 or 1 for Local, only Active: 1 chunks will be considered. Therefore, when retrieving the antecedent of a Local: 1 pronoun, only the cue Local: 0 need be used as a gating feature. This is shown to produce the correct results in our example sentences modeled after our experimental items.

(45) The guests that [$_{spine:2}$ . . .

Recognize new spine

The guests that *no butler*$_{Active:1,Local:1}$ . . .

Assign QP ACTIVE: 1

The guests that *no butler*$_{Active:1,Local:1}$ liked . . .

Continue

The guests that *no butler*$_{Active:1,Local:1}$ liked *asked* . . .

'asked' Signals End of Spine: 2

The guests that *no butler*$_{Active:0,Local:1}$ liked asked . . .

Rewrite ACTIVE *on QP*

The guests that *no butler*$_{Active:0,Local:Inert}$ liked asked . . .

Rewrite LOCAL *to Reflect* ACTIVE *Update*

The guests that *no butler*$_{Active:0,Local:Inert}$ liked asked *him* . . .

Retrieve LOCAL: 0 antecedent

Mismatch on LOCAL

Retrieval Fails.

So long as the probe specifies a definite value for the LOCAL feature, QPs marked LOCAL: INERT will not be accessed. Unfortunately, not all retrieval probes positively specify a value for LOCAL that can be used to block access to inappropriate QPs. Recall that pronouns bearing the LOCAL: 0 feature have probes underspecified for LOCAL, allowing retrieval of NPs regardless of their LOCAL value. If retrieval never exploited the ACTIVE feature, we would expect INERT-marked QPs to interfere with retrieval of antecedents of LOCAL: 0 pronouns. To be concrete, *no butler* would still remain accessible to *his* in the minimally different sentence.

(46) The guests that [no butler$_{Local:Inert,Active:0}$ liked] asked *his*$_{Local:0,Active:1}$ supervisor a question.

To block access to *no butler* as the antecedent for *his* in the sentence above, retrieval would need to use ACTIVE: 0 as a cue. Thus, our cue-generation rule for pronoun-triggered antecedent retrieval would have to be altered:

(47) ANTECEDENT RETRIEVAL FOR PRONOUNS:

When retrieving an antecedent for a pronoun P, use the following features:

- If P is LOCAL: 1 $\rightsquigarrow$ Probe [LOCAL: 0]
- If P is LOCAL: 0 $\rightsquigarrow$ Probe [ACTIVE: 1]

7.6.3 Possible Interference as a Result of Failure To Recognize Domain Edges

The account of scope sensitivity outlined above uses an ACTIVE update procedure that requires the parser to recognize when it has reached the edge of a spine (e.g. the edge of a relative clause). But how does the parser determine the boundaries of a RC? Natural language sentences do not come annotated with clause boundaries. Unambiguous evidence of the end of the clause can come from bottom-up evidence that a new clause has been reached (Frazier 1979; Frazier 1987). In the case of Experiment 7 in this chapter, unambiguous bottom-up evidence came in the form of the matrix verb. Low attachment of *asked* is not possible, which indicates that the parser should pop out of the embedded clause and resume parsing the matrix clause.

(48) The guests [that no maid had prepared lunch for] *asked* . . .

In all of the sentences above, evidence that the edge of the syntactic domain has been reached before the pronoun is encountered. This means that state-transition or LOCAL update have had time to apply. As such, the parser will be in a state to resist interference from the RC-internal QP. If a pronoun preceded the evidence inducing update or state-transition, the parser might not yet be in a state to resist. Thus, if it were possible to place a pronoun between the edge of a relative clause and the matrix verb of the higher clause, we might predict interference from the QP. English does not have pronouns that could be inserted in this position, but Spanish does, in the form of clitics.

(49) Los hombres [que ninguna mujer vió] *la*/*lo* patearon.
The men that no woman saw *her*/*him* kicked.
'The men that no woman saw kicked her.'

If retrieval of the antecedent for the pronoun is immediate, i.e. it is initiated before the parser has time to process the structural position of the clitic, we predict *no woman* to interfere with the clitic *la*, because they match in features. It is also possible, though, that recognition and syntactic integration of the clitic precedes retrieval. If this were the case, the clitic itself might provide enough evidence that the edge of the domain had been reached.

7.7 Conclusion

In this chapter I showed that feature-matching structurally inappropriate QPs do not interfere with the retrieval of antecedents for a pronoun. It appears that the parser exhibits retrospective sensitivity to the Scope Constraint on Bound Variable Pronoun licensing. This sensitivity does not entail that retrieval makes explicit reference to the relation MCCOMMAND or perform any relational computations, however. A non-relational feature, ACTIVE, was proposed as the basis of sensitivity to the scope condition. This feature was able to capture the relational nature of the constraint by pre-computing a QP's accessibility at a given point in time, and the state of the parser at that time. A separate incremental update procedure was responsible for determining the point at which a QP would be inert for subsequent retrieval, and rewriting the value of ACTIVE.

Chapter 8: Conclusion

8.1 Summary

The empirical contribution of the dissertation was to show that antecedent retrieval displays fine-grained sensitivity to the distinction between structurally-appropriate and inappropriate NPs, as determined by a variety of c-command constraints. Where observed, the sensitivity was also categorical (i.e. any evidence of retrieval interference from structurally-inappropriate NPs was entirely absent). The one constraint which it appears retrieval is not effective in implementing was *Weak Crossover* (Wasow 1972). Unlike the other constraints, *Weak Crossover* does not invoke c-command (or some comparable structural relation) in the determination of grammatical accessibility.

The strategies that underlie the retrieval account represent the rational strategy for approximating the extension of grammatical constraints stated in terms of c-command given the architectural constraints imposed by the nature of memory and the temporal

Dependency	Interference Found?	C-command Required in Description?
Hindi Reciprocals	N	Y
Strong Crossover	N	Y
Weak Crossover	Y	N
Bound Variables	N	Y

Table 8.1: Summary: Empirical Results

Dependency	Retrieval Cue	Dynamic Update Point
Hindi Reciprocals	[LOCAL:1]	At All Clause Edges
English Reflexives	[LOCAL:1]	At All Clause Edges
Principle B (Mostly)	[LOCAL:0]	At All Clause Edges
Strong Crossover	[LOCAL:0]	At All Clause Edges
Bound Variables	[ACTIVE: 1]	At All Spine Edges

Table 8.2: Summary: Gating Cues and Update Points for Various Dependencies

demands of incremental retrieval. Though the constraints to which retrieval displayed sensitivity were all constraints often stated in terms of c-command, I showed that feature-based re-codings of the constraints that did not require retrieval to make direct reference to c-command relations were possible. Sensitivity to the Binding Principles A, B, and C was explained with appeal to the feature LOCAL. The Scope Condition on bound variable pronouns (BVPs) was restated in terms of the feature LOCAL. Both features, ACTIVE and LOCAL were argued to be *dynamically updated* features. Their values were subject to change as a function of later syntactic structure. For each of the dependencies considered, a specific retrieval cue was responsible for distinguishing structurally-accessible from inaccessible NPs.

By re-coding the constraints in terms of dynamically-updated features I was able to reconcile the sensitivity of retrieval to apparently relational constraints with a direct-access cue-based memory retrieval mechanism.

8.2 Where's C-Command?

I assumed at the outset that c-command constraints were an integral component of the grammatical principles governing the distribution of anaphoric elements and their antecedents. A maximally transparent mapping between grammar and parser would predict c-command to have some counterpart or corresponding construct in our parsing system. The most transparent method for representing c-command relations would be to explicitly encode them as cues for retrieval. As we saw in Chapter 2, this is not a computationally efficient strategy. Alternative methods for implicitly encoding c-command that maintain a degree of transparency seem more plausible.

In prior parsing models, c-command has almost always been implicitly encoded and c-command constraints implemented as procedures on search. In the parser proposed by Berwick & Weinberg (1984), 'privileged c-commanders' – those NPs that fell along a command-path – were maintained in a stack. C-command relations were represented implicitly through the ordering of objects in the stack; those items that were pushed further down on the stack c-commanded those that were higher up. C-command-sensitive retrieval is implemented via a backwards iterative search through the stack.

C-command constraints cannot be procedurally encoded with a direct-access mechanism. Since we are using a rigid retrieval mechanism, encoding is left as the domain of transparency. One way to is to achieve transparency in encoding is to exhaustively specify c-command relations as item information. As was shown in Chpater 2, this is not a computationally efficient strategy. My proposal does not explicitly encode c-command relations, per se, as features on items. Rather, c-command relations are (approximately) represented through the 'spine' feature in my system and constraints that make reference to c-command are procedurally implemented in the functions that determine the encoding of features like LOCAL and ACTIVE and the dynamic updating procedures that perform the book-keeping functions of determining accessibility relative to a given step in the parse.

For each of the conditions discussed, the notion of 'spine' was invoked. Items bear a spine index and the encoding functions that determine the LOCAL and ACTIVE values of an individual item make reference to that item's spine feature. As discussed in Chapter 7, a 'spine-mate' condition can be used as an imperfect proxy for c-command. If two items are spine-mates, it is guaranteed that one c-commands the other. Using a spine feature to encode c-command is only a rough proxy because spine features do not encode item-to-item information. A system that enforces a condition of this kind cannot distinguish the direction of the c-command relation between the two domain-mates, which can make unfortunate predictions with respect to the licensing of certain grammatical elements.

An example of a problem that a 'spine-mate' strategy faces can be illustrated using the 'clause-mate' formulation of the constraint on reflexive licensing. If reflexives were only subject to the loose constraint that they be co-indexed with a feature-matching clause-mate, our grammar would have no way to distinguish between the two sentences below (e.g. Klima 1964).

(1) a. John$_i$ likes himself$_i$.

b. Himself$_i$ likes John$_i$.

In Berwick and Weinberg's stack architecture relative c-command was implicitly represented by relative order in the stack. Similarly, linear order is used as a proxy for relative c-command between two items that are marked accessible (either through specification of LOCAL or ACTIVE). This simplifying assumption leads to a relatively grammaticallyfaithful implementation, but there is at least one instance where complete coverage fails. In the sentence below, reliance on a 'precedence + spine/domain-mate' strategy (as implemented by my system) would predict *Shaggy* to be judged as a grammatical antecedent for the reflexive (contra the predictions of an account that relied on strict c-command).

(2) *To Shaggy, Thelma revealed himself to be the thief.

8.2.1 Is C-command Still Needed?

If it were the case that my account were able to cover the full range of facts that a c-command-based theory of binding covered, we would potentially have the beginning of an argument for the elimination of constraints stated in terms of c-command. Such an argument would proceed from the premise that all 'grammatical primitives' should be linked to 'parsing primitives', or from the premise that our grammatical ontology should follow the concepts furnished by the performance system insofar as possible. Since the retrieval-based account of the constraints makes reference to spine and precedence only (and does not need c-command), so the argument would go, the grammatical account should eschew reference to c-command.

This reduction, or elimination, appears unmotivated from the standpoint of retrieval, however, because there is a gap in the grammatical coverage of the retrieval mechanism. Sentences like (2) are clearly ungrammatical, but the retrieval account cannot make the appropriate distinctions to rule them out. The fact that they are ungrammatical suggests that an additional procedure - and additional information above and beyond precedence and domain-mate - is required to rule them out. C-command would appear to be the best candidate for this task, though the task could also be accomplished using something like an obliqueness hierarchy as proposed by Pollard and Sag (1992, 1994). If c-command information is still required to rule out those remaining ungrammatical items that pass through the initial filter of retrieval, it maintains its secure footing as a productive component of grammatical theory.[1]

8.3 Cue-combination (and its Discontents)

Following a suggestion in Van Dyke and McElree (2011), it was assumed throughout the thesis that the absence of partial-match interference from structurally-inappropriate items was the product of a system that assigns differential weights to linearly combined retrieval cues.Relying on a preferential weighting scheme to restrict access based on a particular feature is not without its complications. Weighted, linear cue-combinatorics trade probability of partial-match interference from one feature (set) for another. Take some feature set comprised of the features {A, B, C}. Preferentially weighting the activation of A decreases the influence of partial match on features B and C, but simultaneously increases susceptibility to interference from non-targeted items that match on A. In a world where A is absolutely diagnostic of the target image, this weighting is not a problem. On the other hand, in instances where multiple [+A] images abound, interference is predicted

[1]This does not rule out the possibility that c-command is not a grammatical primitive. Perhaps the fact that the relations conform to a c-command generalization does not reflect the relation's ontological primacy. Hornstein (2011) explores this view. In Hornstein's view, c-command generalizations can be descriptively correct without being explanatorily correct. According to Hornstein, c-command is not a constitutive property of of syntactically-mediated relations themselves, but instead a 'side-effect' of some other grammatical process responsible for establishing the dependency. Hornstein proposes that c-command is an emergent relation that is brought about by (i) the fact that all grammatical dependencies are mediated by movement, and (ii) movement operations obey extension. If movement obeys extension, a moved object's landing site will (almost) always target a position that c-commands its base position.

to be high. Pronoun resolution provides a concrete illustration of the problem. The proposed solution for avoiding interference from structurally-inaccessible QPs involved preferentially weighting the feature *Referential*. This reduced the contribution of gender and number features. However, it increases the likelihood of erroneously retrieving any referential NP, regardless of its gender features.

8.3.1 Reconciling the Power of the Account With Previous Work On Agreement

Suppose we accept that context-dependent cue-weighting strategies restrict the influence of interference-prone morphological features on retrieval. We saw how this could be used to achieve categorical effects. Though we saw its weaknesses above, preferential weighting is also a relatively powerful mechanism. It appears too powerful at first blush: it seems possible to provide a method for quashing partial-match interference even in cases where such interference is observed. Take a representative sampling of agreement attraction cases.

(3) a. The key to the cabinets *were* rusty from years of disuse.

b. The manager that despised the employees apparently were tired of working in the office.

In both of the above examples, interference from the plural-marked NP could be overwhelmed by preferentially weighting a cue diagnostic for the grammatical subject. Strongly weighting [Local: 1], [clause], or [subject] would heavily bias towards retrieval of the structurally-appropriate NP. Yet, we see interference effects in these constructions, suggesting that this strategy is not employed. Why would retrieval choose an error-prone strategy in one case (agreement), when it does not in other cases (antecedent retrieval)? I would like to claim that responsibility for this difference lies in the *retrieval trigger*. I suggest that what distinguishes the two retrieval types is that one is error-driven (agreement), and the other is not (antecedent retrieval).

Wagers, Lau and Phillips (2009) argue, on the basis of the *grammatical asymmetry*, that retrieval of the local subject for agreement-licensing that begets agreement attraction effects is entirely error-driven. Let's consider how an error signal influences the cue-generation process. The error signal arises from the mismatch between the agreement morphology *predicted* on the verb, and what is encountered. An error signal induces *uncertainty* in the veridicality of one's encoding. If one is uncertain about a previous structural analysis, relying heavily on structural cues that encode that analysis is foolish. To be less abstract - if the parser has reason to believe that it has mis-encoded the subject, then it has reason to believe that using [subject] as a cue will not retrieve the desired target. Structural cues are the most likely to be non-veridical, because they are constructed on-the-fly. Lexical or morphological cues, on the other hand, are less likely to be mis-encoded, as they are context-invariant. A plural noun bears a plural feature regardless of the environment that it is found in. Thus, under uncertainty, the parser does not weigh structural cues in such a manner as to reduce the chance of partial-match interference.

Anaphoric dependencies, and their associated retrievals, are not born of error signals. Anaphoric items are not predicted (unlike subject-verb agreement morphology), thus there exists no consistent expectation that their presence can violate. It may be the case that an anaphoric item lacks an appropriate antecedent in context, but a parser does not know immediately upon encountering the pronoun. Therefore, a parser should always 'have confidence' in the diagnosticity of structural cues in finding an antecedent.

A consequence of attributing agreement attraction effects in comprehension to mis-retrieval is that it undermines a unified explanation of agreement attraction effects in production and comprehension. Something else must be said to explain the occurrence of agreement attraction in online production. Following Gillespie & Pearlmutter (2011), I assume that the effects in production are a result of faulty encoding during utterance planning. This explanation captures the difference between comprehension and production in that a faulty-encoding explanation can handle the absence of the grammatical asymmetry in production.

8.4 Remaining Issues

8.4.1 Methods of Retrieval

In the thesis I proposed feature-based methods for implementing relational constraints while holding the model of retrieval constant. Both the LOCAL- and ACTIVE-based solutions presupposed retrieval of antecedents proceeds via direct-access cue-based retrieval. In chapter 2 evidence was presented motivating the use of a direct-access retrieval mechanism in the computation of other dependencies (e.g. filler-gap dependencies), but no direct evidence was offered to motivate using a direct-access mechanism for antecedent retrieval. Using a direct-access mechanism for antecedent retrieval has its appeal; for example, doing so would maintain a uniform account of retrieval mechanisms in linguistic processing. However, use of a direct-access mechanism in the resolution of one type of dependency does not automatically entail its general use across all dependency types. In the thesis, I offer no dispositive arguments against employing a serial-search strategy for the retrieval of a pronoun's antecedent. The empirical results presented in this thesis are, to a great extent, compatible with a search procedure that traverses a syntactically guided path to find antecedents for a pronoun.

Determining whether antecedent retrieval employs a serial-search or a direct-access mechanism can be done empirically. As was done by McElree et al. (2003) for filler-gap resolution, a study should test whether hierarchical distance between a QP and a pronoun affects the dynamics of retrieving that QP as an antecedent. An example of a design that would do this is given below. In the sentences, the position of the QP *every employee* is varied such that its hierarchical distance from the pronoun changes. In the first sentence *every employee* is distanced from the pronoun *him* by two clause boundaries. In the second sentence, only one clause boundary intervenes. If hierarchical distance influences retrieval time, longer times are expected in (4-a) than in (4-b) condition.[2]

(4) a. The executives believed that *every employee* had reported [that the secretaries said [the others hated *him*.

b. The executives believed that the secretaries read the report that *every employee* said the others hated *him*.

[2]The sentences differ in that (4-b) contains a complex NP, inside of which the antecedent and pronoun are embedded. The reason for including this complex NP is to effectively trap the QP in its surface position.

8.4.2 On the Origin of Features By Means of Grammatical Specification

My account of the Binding Principles online makes extensive use of the LOCAL feature. The feature is functionally motivated. Its application aids the partitioning of accessible from inaccessible NPs. However, one might wonder about its etiology. Two options present themselves:

i) It is innately specified, or

ii) it is developed over time, as a function of increased skill in language use.

Why might a feature be innately specified? One simple idea is to assume that a one-to-one correspondence holds between grammatical primitives and ‘parsing primitives’. An slightly different way of envisioning the connection would be that it is not merely a correspondence relation that links parsing and grammatical primitives, but rather an identity relation. On either view, features used by the encoding system would correspond directly to primitive constructs in our grammatical theories. In the next subsection I consider the consequences of assuming this kind of transparency exists between parser and grammar, but here I simply note that innate specification of the feature would be assumed under such an account.

Could a feature like LOCAL emerge as a result of skilled language use rather than innate specification? In the literature there is already a tradition that (at least tacitly) assumes that this may be possible. Lewis & Vasishth (2005) and Caplan & Waters (2013) both analogize the encoding and retrieval operations used in real-time sentence processing with ‘skilled’, or ‘expert’ memory performance.

A great deal of work has gone into studying task-specific memory encoding procedures that are developed and honed through repeated exposure to a task, so-called *expert chunking strategies*. Researchers have studied, among other topics, expert mnemonic strategies for encoding large amounts of information (e.g. recalling a number of random digits larger than the assumed capacity of working memory - Ericsson & Kintsch 1995), and the recall of game-related information, such as piece position, by expert chess players (Simon & Chase 1973; Simon & Gilmartin 1973; Gobet & Simon 1996; Gobet 1998). Memory of positions of chess pieces on a board is achieved through an implicit, automatized encoding of the input referred to as *perceptual chunking*, whereas deliberative, conscious strategies that allow control over the chunking process are referred to as *goal-oriented chunking* strategies (Gobet et al. 2001). I assume that the memory encoding operations in subconscious language processing form a kind of perceptual chunking regime.

Expert perceptual chunking strategies are encoding schemas that are developed a function of skill. An extant encoding vocabulary can be altered (made more efficient, added to, etc.) to better aid in recalling particular types of information relevant for the task. The results of Chase and Simon's (1973) study typify the beneficial effects of expert perceptual chunking strategies: expert chess players appear to be better at remembering the positions of chess pieces on a board better than lower skill players. Evidence for task-specific encoding benefits come from the fact experts show a disproportionate ability to recall a distribution of pieces that follows the rules of chess than a random distribution of the same pieces.

Suppose the expert advantage emerges over time as the creation of *novel* encoding strategies and features. If general purpose strategies exist for the creation of task-specific encoding features, they should surely be operative in language processing. Thus, it might be possible to assume that some or all of the higher-order features that are not transparently instantiated may be developed over time. This type of account makes some interesting predictions:

(i) Language users may differ in their parsing-specific feature vocabulary,

(ii) A 'non-expert' stage of language use should exist, at which point users have not had sufficient evidence to induce creation of features like LOCAL from their experience, and

(iii) The number of potential features is potentially unbounded, though functionally constrained.

As for point (iii), I have little to say. I know of no effective way to probe the bounds of idiosyncratic feature systems that individuals might have arrived at by differentially parsing latent information available in linguistic representations.

As for the the first two points, I believe current evidence speaks neither conclusively for, nor against, them. If we follow the analogy proposed, the expert strategy hypothesis predicts that children should, at some point, not have LOCAL in their repertoire of encoding features. This in turn would predict that they should exhibit difficulty in implementing Principle B in real-time tasks. Chien and Wexler (1990) argued that children are, under some circumstances, susceptible to considering pronominal coreference relations that violate Principle B. This could be taken as confirmatory evidence for the idea that the features responsible for implementing such constraints are developed over time as a function of language experience. However, Elbourne (2005b) argued that these findings were artifactual in origin and that children are able to successfully deploy Principle B. Conroy, Takahashi, Lidz & Phillips (2009) provided a more nuanced interpretation. The authors concluded, based on their own experiments and an extensive review of prior studies investigating Principle B application cross-linguistically, that children display robust knowledge of Principle B. However, the authors note that children's application of Principle B online appears less robust than application of other constraints (such as Principle C). The authors took this to indicate that children apply Principle B as a 'late filter' on generated representations, which would entail that the initial stages of pronoun resolution (i.e. retrieval) did, at times, have access to ungrammatical NPs. If children's performance differs from adults in this domain, we may have evidence that supports the idea of skill- or experience-contingent strategies.

Appendix: Items for Experiment 3

a. *Crossover* Sentences
b. *NoCrossover* Sentences

1. a. Jane asked which {janitor/lunch-lady} at the boarding school Donna had said that he already spoke with about the foodfight.
   b. Jane asked which {janitor/lunch-lady} at the boarding school had said that he already spoke with Jim about the foodfight.

2. a. Mary knew which {policewoman/policeman} at the scene Hailey had claimed that he quickly talked to after the boating accident.
   b. Mary knew which {policewoman/policeman} at the scene had claimed that he quickly talked to Martin after the boating accident.

3. a. Christina remembered which {mother/fireman} from the PTA Diane had claimed that he gladly worked with during the heatwave.
   b. Christina remembered which {mother/fireman} from the PTA had claimed that he gladly worked with John during the heatwave.

4. a. Olivia said which {midwife/doctor} in the crowded ward Kaitlyn had thought that he already spoke with about the diagnosis.
   b. Olivia said which {midwife/doctor} in the crowded ward had thought that he already spoke with Jeff about the diagnosis.

5. a. Jessica asked which {cowgirl/cowboy} from the new hires Theresa had admitted that he actually slept with on the ranch.
   b. Jessica asked which {cowgirl/cowboy} from the new hires had admitted that he actually slept with James on the ranch.

6. a. Ashley wondered which {prostitute/gangster} in the line-up Susanne had said that he definitely had seen in the alleyway.
   b. Ashley wondered which {prostitute/gangster} in the line-up had said that he definitely had seen Daniel in the alleyway.

7. a. Emily knew which {typist/judge} at the trial Eve had said that he quietly argued with outside the courtroom.
   b. Emily knew which {typist/judge} at the trial had said that he quietly argued with Robert outside the courtroom.

8. a. Sarah remembered which {nun/monk} from the neighboring church Nora had claimed that he quietly talked to at the pancake breakfast.
   b. Sarah remembered which {nun/monk} from the neighboring church had claimed that he quietly talked to Sean at the pancake breakfast.

9. a. Megan said which {secretary/engineer} from the drafting division Christine had claimed that he actually agreed with about the revised blueprints.
   b. Megan said which {secretary/engineer} from the drafting division had claimed that he actually agreed with Thomas about the revised blueprints.

10. a. Hannah couldn't decide which {girl/boy} scout in the cabin Bridget had implied that he frequently fought with during the trip.
    b. Hannah couldn't decide which {girl/boy} scout in the cabin had implied that he frequently fought with Pedro during the trip.

11. a. Lauren couldn't decide which {nurse/private} in the regiment Catherine had admitted that he actually learned from during boot camp.
    b. Lauren couldn't decide which {nurse/private} in the regiment had admitted that he actually learned from Paul during boot camp.

12. a. Michael asked which {bouncer/stripper} from the sleazy club Omar had said that she secretly called up after the night's end.
    b. Michael asked which {bouncer/stripper} from the sleazy club had said that she secretly called up Martina after the night's end.

13. a. Christopher wondered which {uncle/aunt} at the family reunion Felix had said that she frequently called on after the pregnancy.
    b. Christopher wondered which {uncle/aunt} at the family reunion Felix had said that she frequently called on Roberta after the pregnancy.

14. a. Matthew knew which {designer/supermodel} at fashion week. Jorge had claimed that she accidentally ran into behind the runway.
    b. Matthew knew which {designer/supermodel} at fashion week. had claimed that she accidentally ran into Kristen behind the runway.

15. a. Joshua remembered which {prince/princess} from the distant kingdom Brian had claimed that she successfully cast out from the court.
    b. Joshua remembered which {prince/princess} from the distant kingdom had claimed that she successfully cast out Alice from the court.

16. a. Jacob said which {baritone/soprano} in the exclusive choir Brad had thought that she possibly sang with many years ago.
    b. Jacob said which {baritone/soprano} in the exclusive choir had thought that she possibly sang with Allison many years ago.

17. a. Nick asked which {frat boy/sorority girl} at the Greek mixer Joseph had thought that she unintentionally bumped into during pledge week.

b. Nick asked which {frat boy/sorority girl} at the Greek mixer had thought that she unintentionally bumped into Mandy during pledge week.

18. a. Andrew wondered which {dentist/hygienist} in the big office Peter had said that she often spoke to about proper brushing.

    b. Andrew wondered which {dentist/hygienist} in the big office had said that she often spoke to Felicia about proper brushing.

19. a. Daniel knew which {waiter/waitress} from the greasy diner George had said that she already worked with on Tuesday night.

    b. Daniel knew which {waiter/waitress} from the greasy diner had said that she already worked with Donna on Tuesday night.

20. a. Tyler remembered which {doctor/nurse} from the emergency room James had claimed that she immediately talked to after the seizure.

    b. Tyler remembered which {doctor/nurse} from the emergency room had claimed that she immediately talked to Eileen after the seizure.

21. a. Joseph said which {executive/secretary} at the ad agency Howard had claimed that she obviously had assisted with the typing.

    b. Joseph said which {executive/secretary} at the ad agency had claimed that she obviously had assisted Hailey with the typing.

22. a. Brandon couldn't decide which {priest/nun} Paul had believed that she actually visited with during the holidays.

    b. Brandon couldn't decide which {priest/nun} had believed that she actually visited with Ellen during the holidays.

23. a. David couldn't decide which {boy/girl scout} Jeff had believed that she clearly had outsold during the fundraiser.

    b. David couldn't decide which {boy/girl scout} had believed that she clearly had outsold Maxine during the fundraiser.

24. a. Lily wondered which {seamstress/tailor} from the studio Alice had said that he recently called on to get black thread.

    b. Lily wondered which {seamstress/tailor} from the studio had said that he recently called on Edward to get black thread.

Appendix: Items for Experiment 4

1. Jane asked which {maintenance man/lunch lady} {it appeared / had said } that {he /Donna} might have already spoken with regarding the food-fight in the cafeteria.
2. Lily wondered {which tailor/seamstress} {it seemed/ had said} that {he /Alice} could have recently called on for sewing assistance before the wedding.
3. Mary knew which {policeman/policewoman} {it appeared/had admitted} that {he /Hailey} should have immediately talked to after the accident on the highway.
4. Christina remembered {which fireman/mother} {it seemed/had admitted} that {he /Diane} should have immediately reported to after the catastrophe in the city.
5. Olivia said which {doctor/midwife} {it appeared/had thought} that {he /Kaitlyn} might have already spoken with about the prognosis for good recovery.
6. Jessica asked which {cowboy/cowgirl} {it seemed/had thought} that {he /Theresa} should have actually fought with during the cattle-drive on the ranch.
7. Ashley wondered which {gangster/prostitute} {it appeared/had said} that {he /Susanne} should have probably gone after in the alleyway during the sting.
8. Emily knew which {judge/typist} {it seemed/had said} that {he /Eve} must have actually disagreed with about the ruling by the court.
9. Sarah remembered which {monk/nun} {it appeared/had confessed} that {he /Nora} might have privately talked to in the rectory after the mass.
10. Megan said which {engineer/secretary} {it seemed/had admitted} that {he /Christine} might have actually agreed with about the blueprints for the skyscraper.
11. Hannah couldn't remember which {boy/girl} scout {it appeared/had implied} that {he /Bridget} would have gone fishing with on troop trips in the summers.
12. Lauren couldn't remember which {soldier/nurse} {it seemed/had disclosed} that {he /Catherine} might have secretly conspired with to steal provisions from the infirmary.
13. Michael asked which {stripper/bouncer} {it appeared/had said} that {she /Omar} should have immediately looked for outside the club after last call.

14. Christopher wondered which {aunt/uncle} {it appeared/had said} that {she /Felix} could have easily corresponded with according to staff at the retirement_home.

15. Matthew knew which {nanny/boy} {it seemed/had claimed} that {she /Jorge} would get stuck waiting for during afternoon dismissal outside the school.

16. Joshua remembered which {princess/prince} {it appeared/had said} that {she /Brian} would have discreetly taken aside during the ball to discuss rumors.

17. Jacob found out which {maid/butler} {it seemed/had insisted} that {she /Brad} would have gladly spied on during the lunch-break for the investigators.

18. Nick asked which {sorority girl/frat boy} {it appeared/had thought} that {she /Mandy} would have definitely bumped into at the mixer during welcome week.

19. Andrew wondered which {receptionist/dentist} {it seemed/had said} that {she /Peter} could have already spoken to about the benefits of dental insurance.

20. Daniel knew which {waitress/waiter} {it appeared/had complained} that {she /Tina} must have already worked with during dinner rush on Tuesday night.

21. Tyler remembered which {nurse/doctor} {it seemed/had believed that} {she /Eileen} should have immediately talked to after the outbreak in the ward.

22. Joseph said which {secretary/executive} {it appeared/had conceded} that {she /Howard} should have frequently consulted with with the typing of the reports.

23. Brandon couldn't recall which {nun/priest} {it seemed/had thought} that {she /Paul} might have actually talked with about the interpretation of the scripture.

24. David couldn't recall which {girl/boy} scout {it appeared/had believed} that {she /Jeff} could have actually worked with during the fundraiser for new uniforms.

25. Mary asked which {knight/baroness} {it appeared/had declared} that {he /Evelyn} would have valiantly fought for during the joust at the fair.

26. Sam described which {bride/groom} {it seemed/had insisted} that {she /Andrew} would have never waited for before the dress-fitting at the store.

27. Eliza questioned which {bully/girl} {it seemed/had said} that {he /Harry} might have been hiding from behind the slide on the playground.

28. Sean asked which {ballerina/boxer} {it appeared/had thought} that {she /Samantha} could have actually tied with in the exhibition at the gym.

29. Rosie knew which {linebacker/cheerleader} {it appeared/had said} that {he /Shelly} would have rather spoken with for the article in the newspaper.

30. Adam recalled which {actress/actor} {it seemed/had implied} that {she /Ellen} would have rather studied under in the workshops at the conservatory.

31. Angela mentioned which {professor/kindergarten teacher} it {seemed/had} claimed that {he /Melanie} would have formally written up for teaching evolution to the class.

32. Benjamin described which {salesgirl/salesman} {it appeared/had claimed} that {she /Richard} would have angrily yelled at for the faulty merchandise in stock.

33. Ingrid reported which {emperor/empress} {it appeared / had implied} that {he /Isabel} might have secretly conspired with to plan imprisoning many local revolutionaries.

34. Allan forgot which {sorceress/sorcerer} {it seemed/had admitted} that {she /Kevin} would have gladly fought against during the battle in the novel.

35. Leah didn't know which {farmer/milkmaid} {it seemed/had said} that {he /Carolyn} might have accidentally ripped off at the market on Sunday afternoon.

36. Jack reported which {policewoman/policeman} {it appeared/had admitted} that {she /Eric} might have unintentionally shot at in the basement during the firefight.

Appendix: Items for Experiment 5

1. Jane asked which {maintenance man/lunch lady} {it appeared / had said } that {his supervisor/Donna} might have already spoken with regarding the food-fight in the cafeteria.
2. Lily wondered {which tailor/seamstress} {it seemed/ had said} that {his apprentice/Alice} could have recently called on for sewing assistance before the wedding.
3. Mary knew which {policeman/policewoman} {it appeared/had admitted} that {his partner/Hailey} should have immediately talked to after the accident on the highway.
4. Christina remembered {which fireman/mother} {it seemed/had admitted} that {his buddy/Diane} should have immediately reported to after the catastrophe in the city.
5. Olivia said which {doctor/midwife} {it appeared/had thought} that {his associate/Kaitlyn} might have already spoken with about the prognosis for good recovery.
6. Jessica asked which {cowboy/cowgirl} {it seemed/had thought} that {his employer/Theresa} should have actually fought with during the cattle-drive on the ranch.
7. Ashley wondered which {gangster/prostitute} {it appeared/had said} that {his rival/Susanne} should have probably gone after in the alleyway during the sting.
8. Emily knew which {judge/typist} {it seemed/had said} that {his colleague/Eve} must have actually disagreed with about the ruling by the court.
9. Sarah remembered which {monk/nun} {it appeared/had confessed} that {his abbott/Nora} might have privately talked to in the rectory after the mass.
10. Megan said which {engineer/secretary} {it seemed/had admitted} that {his manager/Christine} might have actually agreed with about the blueprints for the skyscraper.
11. Hannah couldn't remember which {boy/girl} scout {it appeared/had implied} that {his leader/Bridget} would have gone fishing with on troop trips in the summers.
12. Lauren couldn't remember which {soldier/nurse} {it seemed/had disclosed} that {his sergeant/Catherine} might have secretly conspired with to steal provisions from the infirmary.

13. Michael asked which {stripper/bouncer} {it appeared/had said} that {her admirer/Omar} should have immediately looked for outside the club after last call.

14. Christopher wondered which {aunt/uncle} {it appeared/had said} that {her nephew/Felix} could have easily corresponded with according to staff at the retirement_home.

15. Matthew knew which {nanny/boy} {it seemed/had claimed} that her {employer/Jorge} would get stuck waiting for during afternoon dismissal outside the school.

16. Joshua remembered which {princess/prince} {it appeared/had said} that {her advisor/Brian} would have discreetly taken aside during the ball to discuss rumors.

17. Jacob found out which {maid/butler} {it seemed/had insisted} that {her supervisor/Brad} would have gladly spied on during the lunch-break for the investigators.

18. Nick asked which {sorority girl/frat boy} {it appeared/had thought} that {her sisters/Mandy} would have definitely bumped into at the mixer during welcome week.

19. Andrew wondered which {receptionist/dentist} {it seemed/had said} that {her manager/Peter} could have already spoken to about the benefits of dental insurance.

20. Daniel knew which {waitress/waiter} {it appeared/had complained} that {her friend/Tina} must have already worked with during dinner rush on Tuesday night.

21. Tyler remembered which {nurse/doctor} {it seemed/had believed that} {her patient/Eileen} should have immediately talked to after the outbreak in the ward.

22. Joseph said which {secretary/executive} {it appeared/had conceded} that {her boss/Howard} should have frequently consulted with with the typing of the reports.

23. Brandon couldn't recall which {nun/priest} {it seemed/had thought} that {her priest/Paul} might have actually talked with about the interpretation of the scripture.

24. David couldn't recall which {girl/boy} scout {it appeared/had believed} that {her mother/Jeff} could have actually worked with during the fundraiser for new uniforms.

25. Mary asked which {knight/baroness} {it appeared/had declared} that {his compatriots/Evelyn} would have valiantly fought for during the joust at the fair.

26. Sam described which {bride/groom} {it seemed/had insisted} that {her girlfriends/Andrew} would have never waited for before the dress-fitting at the store.

27. Eliza questioned which {bully/girl} {it seemed/had said} that {his victim/Harry} might have been hiding from behind the slide on the playground.

28. Sean asked which {ballerina/boxer} {it appeared/had thought} that {her rival/Samantha} could have actually tied with in the exhibition at the gym.

29. Rosie knew which {linebacker/cheerleader} {it appeared/had said} that {his interviewer/Shelly} would have rather spoken with for the article in the newspaper.

30. Adam recalled which {actress/actor} {it seemed/had implied} that {her replacement/Ellen} would have rather studied under in the workshops at the conservatory.

31. Angela mentioned which {professor/kindergarten teacher} it {seemed/had} claimed that {his student/Melanie} would have formally written up for teaching evolution to the class.

32. Benjamin described which {salesgirl/salesman} {it appeared/had claimed} that {her client/Richard} would have angrily yelled at for the faulty merchandise in stock.

33. Ingrid reported which {emperor/empress} {it appeared / had implied} that {his advisors/Isabel} might have secretly conspired with to plan imprisoning many local revolutionaries.

34. Allan forgot which {sorceress/sorcerer} {it seemed/had admitted} that {her enemies/Kevin} would have gladly fought against during the battle in the novel.

35. Leah didn't know which {farmer/milkmaid} {it seemed/had said} that {his helpers/Carolyn} might have accidentally ripped off at the market on Sunday afternoon.

36. Jack reported which {policewoman/policeman} {it appeared/had admitted} that {her back-up/Eric} might have unintentionally shot at in the basement during the firefight.

Appendix: Items for Experiment 6

1. Kathy didn't think {any/the} janitor enjoyed performing his custodial duties{, but/when} he had to clean up the messes left after prom anyway.
2. Molly doesn't think {any/the} fan likes watching his favorite team lose{, but/when} he feels the need to watch games against better teams still.
3. Anne didn't believe {any/the} professor agreed with his freshmen students{, but/when} he was forced to listen to them during office hours anyway.
4. Margaret didn't believe {any/the} lawyer was comfortable violating his moral code{, but/when} he had to defend guilty parties that paid for counsel.
5. Evelyn doesn't think {any/the} doctor likes meeting his patients' families{, but/when} he has to invite them in to discuss bad test results.
6. Catherine didn't expect {any/the} priest to endorse his parishioners' sins{, but/when} he gave them forgiveness during confession after Sunday mass.
7. Mary doesn't think {any/the} detective likes revealing his informants' identities{, but/when} he is forced to make them known during criminal trials.
8. Bridget didn't think {any/the} judge doubted his earlier rulings{, but/when} he was asked to discuss them critically with first-year students.
9. Ruby didn't expect {any/the} chef to share his secret recipe{, but/when} he was forced to give inspectors a list of ingredients.
10. Jill didn't expect {any/the} boy scout to underplay his wild adventures{, but/when} he was asked to tell younger scouts an inspiring story.
11. Kristin doesn't believe {any/the} bartender enjoys refusing drinks to his regular patrons{, but/when} he has to stop making them cocktails when they're drunk.
12. Jenna doesn't think {any/the} celebrity wants the media scrutinizing his private life{, but/when} he has to attend counseling or enter a rehab program.
13. Robert doesn't expect {any/the} milkmaid to agree to medicating her organic cows{, but/when} she has to comply with FDA regulations that require drugs.
14. Harry didn't expect {any/the} teacher to ignore her gifted students{, but/when} she was obligated to spend more time with failing students.

15. Arnold doesn't want {any/the} girl scout offering discounts on here fundraiser cookies{, but/when} she has to find ways to lure in cheap
16. Ron doesn't think {any/the} maid stole from her employer's home{, but/when} she was tempted to in order to settle personal debts.
17. Herbert doesn't believe {any/the} feminist enjoys defending her personal beliefs{, but/when} she has to attend parties and talk to conservatives.
18. Brian didn't think {any/the} stewardess should let them search her carry-on bags{, but/when} she had to wait with passengers at airport security anyway.
19. Philip doesn't want {any/the} secretary to talk about her superiors' email{, but/when} she finds juicy work-related and personal gossip inside it.
20. Daniel didn't believe {any/the} ballerina complained about her swollen feet{, but/when} she was forced to stand on tiptoe for hours during rehearsal.
21. Alex didn't think {any/the} diva enjoyed taking orders from her personal entourage{, but/when} she had to surrender control of her busy schedule on tour.
22. George didn't want {any/the} babysitter to invite friends to her client's houses{, but/when} she was allowed to have one stop by after bedtime.
23. Ralph doesn't think {any/the} waitress enjoys old people at her best table{, but/when} she is stuck working the early-bird dinner shift on Tuesdays.
24. Timothy didn't expect {any/the} nun to be embarrassed about her outdated outfit{, but/when} she had to wear it during Sunday prayer group meetings.

Appendix: Items for Experiment 7

a. Test Conditions
b. Control Conditions

1. a. The guests that {no/the} butler/maid prepared lunch for had asked him about the picnic on the lawn.
   b. The guests said that no butler/maid thought that he would be asked about the picnic on the lawn.

2. a. The counselors that {no/the} father/mother considered very responsible had reassured him before the field-trip to the canyon.
   b. The counselors worried that no father/mother thought that he could be convinced about the field trip to the canyon.

3. a. The jesters that {no/the} lord/lady found very funny had entertained him after the feast in the hall.
   b. The jesters worried that no lord/lady would say that he had been entertained after the feast in the hall.

4. a. The teachers that {no/the} schoolboy/schoolgirl cared about impressing had praised him after the competition in the gym.
   b. The teachers knew that no schoolboy/schoolgirl thought that he would be praised after the competition in the gym.

5. a. The reporters that {no/the} linebacker/cheerleader could ever avoid had hounded him after the game for an interview.
   b. The reporters hoped that no linebacker/cheerleader said that he had been hounded after the game for an interview.

6. a. The administrators that {no/the} doctor/nurse had agreed with had reprimanded him after the outbreak in the hospital.
   b. The administrators hoped that no doctor/nurse would admit that he had been reprimanded after the outbreak in the hospital.

7. a. The mercenaries that {no/the} blacksmith/barmaid wanted to serve had threatened him during the brawl in the tavern.
   b. The mercenaries hoped that no blacksmith/barmaid would admit that he had been threatened during the brawl in the tavern.

8. a. The workers that {no/the} farmboy/farmgirl had difficulty understanding had helped him with the crops during the harvest.
   b. The workers worried that no farmboy/farmgirl would admit that he had been helped with the crops during the harvest.

9. a. The monks that {no/the} priest/nun would discuss scripture with had consulted him while gathering passages for the service.
   b. The monks said that no priest/nun would reveal that he had been consulted while gathering passages for the service.

10. a. The board members that {no/the} executive/secretary cared about pleasing had criticized him during the meeting with the clients.
    b. The board members thought that no executive/secretary would admit that he had been criticized during the meeting with the clients.

11. a. The ranchers that {no/the} cowboy/cowgirl got along with had hired him for the summer to tend sheep.
    b. The ranchers were sure that no cowboy/cowgirl thought that he would be hired for the summer to tend sheep.

12. a. The gangsters that {no/the} policeman/policewoman had under surveillance had spotted him outside the hideout during the stake-out.
    b. The gangsters were confident no policeman/policewoman knew that he had been spotted outside the hideout during the stake-out.

13. a. The activists that {no/the} feminist/chauvinist would speak out against had criticized her during the protest outside the courthouse.
    b. The activists hoped that no feminist/chauvinist believed that she had been criticized during the protest outside the courthouse.

14. a. The delegates that {no/the} congresswoman/congressman thought about double-crossing had supported her during the nominations at the convention.
    b. The delegates were sure no congresswoman/congressman doubted that she would be supported during the nominations at the convention.

15. a. The officers that {no/the} prostitute/pimp provided information to had protected her after the shooting in the neighborhood.
    b. The officers feared that no prostitute/pimp would testify that she had been protected after the shooting in the neighborhood.

16. a. The businessmen that {no/the} stripper/bouncer liked having around had harassed her after closing time outside the bar.
    b. The businessmen argued that no stripper/bouncer thought that she had been harassed after closing time outside the bar.

17. a. The troopleaders that {no/the} girl/boy scout had respect for had scolded her after the incident at scout camp.

b. The troopleaders were sure no girl/boy scout was afraid that she would be scolded after the incident at scout camp.

18. a. The lords that {no/the} maiden/knight swore allegiance to had banned her from the tournament at the castle.
    b. The lords said no maiden/knight could believe that she had been banned from the tournament at the castle.

19. a. The designers that {no/the} seamstress/tailor enjoyed working with had rehired her for the show during fashion week.
    b. The designers thought no seamstress/tailor believed that she would be rehired for the show during fashion week.

20. a. The guards that {no/the} stewardess/pilot paid attention to had frisked her outside the gate before the flight.
    b. The guards were sure no stewardess/pilot would complain that she had been frisked outside the gate before the flight.

21. a. The servers that {no/the} waitress/waiter could deal with had ignored her during their break in the kitchen.
    b. The servers knew that no waitress/waiter would mind that she had been ignored during their break in the kitchen.

22. a. The waiters that {no/the} hostess/host liked working with had cheated her out of tips during their shifts.
    b. The waiters thought no hostess/host would suspect that she had been cheated out of tips during their shifts.

23. a. The contest judges that {no/the} bikini model/bodybuilder dared to second-guess had dismissed her from the contest during the floorshow.
    b. The contest judges thought no bikini model/bodybuilder would complain that she had been dismissed from the stage during the floorshow.

24. a. The editors that {no/the} poetess/poet sent manuscripts to had rejected her for the anthology of local writers.
    b. The editors said no poetess/poet could believe that she had been rejected for the anthology of local writers.

25. a. The cousins that {no/the} uncle/aunt kept in touch with had invited him to the reunion at the lakehouse.
    b. The cousins worried that no uncle/aunt thought that he would be invited to the reunion at the lakehouse.

26. a. The critics that {no/the} actor/actress expected good reviews from had praised him in the reviews at the festival.
    b. The critics knew that no actor/actress would that believe that he had been praised in the reviews at the festival.

27. a. The villagers that {no/the} witch/wizard cast any spells on had persecuted her during the trials after the plague.
    b. The villagers knew that no witch/wizard would suspect that she would be persecuted during the trials after the plague.

28. a. The homeowners that {no/the} cleaning lady/plumber wanted to work for had paid her below minimum wage three months late.
    b. The homeowners claimed that no cleaning lady/plumber could say that she had been paid below minimum wage three months late.

29. a. The inspectors that {no/the} landlady/landlord was completely honest with had fined her for unhealthy conditions in the buildings.
    b. The inspectors suspected that no landlady/landlord would care that she had been fined for unhealthy conditions in the buildings.

30. a. The caterers that {no/the} bride/groom planned to hire had served her stale pineapple cake as a sample.
    b. The caterers were relieved that no bride/groom discovered that she had been served stale pineapple cake as a sample.

31. a. The linebackers that {no/the} quarterback/cheerleader shouted encouraging words to had surprised him with the skillful block.
    b. The linebackers were proud that no quarterback/cheerleader admitted he had been surprised by the skillful block.

32. a. The sorcerors that {no/the} fairy/warrior wanted to fight against had lured her into a trap during the battle.
    b. The sorcerors knew that no fairy/warrior would admit she had been lured into a trap during the battle.

33. a. The skeptics that {no/the} god/goddess expected a sacrifice from had asked him for more rain during the drought.
    b. The skeptics believed that no god/goddess would notice that he had been asked for more rain during the drought.

34. a. The ambassadors that {no/the} emperor/empress had much respect for had snubbed her at the summit before the Olympics.
    b. The ambassadors worried that no empress/emperor would remember she had been snubbed at the summit before the Olympics.

35. a. The bankers that {no/the} heir/heiress was brave enough to sue had deceived him with fraudulent statements on his account.
    b. The bankers believed that no heir/heiress would testify she had been deceived with fraudulent statements on his account.

36. a. The orderlies that {no/the} widow/grandfather said pleasant things about had bathed her after physical therapy in the pool.

b. The orderlies couldn't believe that no widow/grandfather would admit she had been bathed after physical therapy in the pool.

References

Aho, A. V., & Ullman, J. D. (1972). *The theory of parsing, translation, and compiling*. Prentice-Hall, Inc.

Alcocer, P. (2011). There is more than one way to search working memory. Ms. University of Maryland, College Park.

Alcocer, P. & Phillips, C. (2012). Using Relational Syntactic Constraints in Content-Addressable Memory Architectures for Sentence Parsing. Ms. University of Maryland, College Park.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology*, 5, 451-474.

Anderson, J. R. (1989). A rational analysis of human memory. In Roediger III, H., and Craik, F., (eds.) Varieties of Memory and Consciousness: Essays in Honor of Endel Tulving: 195-210. Erlbaum, Hillsdale, NJ.

Anderson, J. R. (1990). The Adaptive Character of Thought. Erlbaum, Hillsdale, NJ.

Anderson, J. R., & Lebiere, C. (1998). The atomic components of thought. Erlbaum, Mahwah, NJ.

Anderson, J. R. & Milson, R. (1989). Human Memory: An Adaptive Perspective. Psychological Review, 96, 703-719.

Anderson, M.C. & Neely, J.H. (1996). Interference and inhibition in memory retrieval. E.L. Bjork & R.A. Bjork (Eds.), *Memory. Handbook of Cognition and Perception (2nd ed.).* San Diego: Academic Press, pp. 237-313.

Anderson, J. R. & Reder, L. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General*, 128, 186-197.

Anderson, J. R., & Schooler, L. (1991). Reflections of the environment in memory. Psychological Science, 2, 396-408.

Anderssen, J. (2011). Quantification, misc. Unpublished Doctoral Dissertation. University of Massachusetts, Amherst.

Aoshima, S., Phillips, C., & Weinberg, A. (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language*, 51, 23-54.

Ashby, J., Rayner, K., & Clifton, C. (2005). Eye movements of highly skilled and average readers: Differential effects of frequency and predictability. *Quarterly Journal of Experimental Psychology*, 58A, 1065-1086.

Avrutin, S. (1994). Psycholinguistic investigations in the theory of reference. Doctoral dissertation, Massachusetts Institute of Technology.

Baayen, H. (2008). Analyzing Linguistic Data: A Practical Introduction to Statistics Using R. Cambridge University Press, Cambridge, UK.

Baayen, H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390-412.

Badecker, W. A feature principle for partial agreement. *Lingua*, 117, 1541-1565.

Badecker, W., & Straub K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 28, 748-769.

Barker, Chris. (2002). Continuations and the nature of quantification. *Natural Language Semantics* 10: 211-242.

Barker, Chris, and Geoffrey K. Pullum. (1990). A theory of command relations. *Linguistics and Philosophy* 13:1-34.

Barker, Chris, and Chung-chieh Shan. (2008). Donkey anaphora as in-scope binding. Semantics and Pragmatics 1:1-42.

Barwise, J & Cooper, R.H. (1981). Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4.

Beghelli, F. & Stowell, T. (1997). Distributivity and negation: The syntax of 'each' and 'every'. In Anna Szabolcsi (ed.), *Ways of Scope Taking*. Dordrecht: Kluwer, pp. 71107.

Berwick, R., & Weinberg, A. (1984). The Grammatical Basis of Linguistic Performance. MIT Press, Cambridge.

Bhatt, R. (2003) Locality in correlatives. *Natural Language and Linguistic Theory*, 21, 485-541.

Bhatt, R & Dayal, V. (2007). Rightward scrambling as rightward remnant movement. *Linguistic Inquiry*, 38, 287-301.

Bock, K. & Miller, C. (1991). Broken Agreement. *Cognitive Psychology*, 23, 45-93.

Bock, K., Nicol, J., & Cutting, J. C. (1999). The ties that bind: Creating number agreement in speech. *Journal of Memory and Language*, 40, 330-346.

Bock, K., Eberhard, K., & Cutting, J. C. (2004). Producing number agreement: How pronouns equal verbs. *Journal of Memory and Language*, 51, 251-278.

Bock., K., Butterfield, S., Cutler, A., Cutting, J. C., Eberhard, K., & Humphreys, K. (2006). Number agreement in British and American English: Disagreeing to agree collectively. Language, 82, 64-113.

Boeckx, C. & Hornstein, N. (2003). Reply to "Control is not movement", *Linguistic Inquiry*, 34,269-280

Boeckx, C. & Hornstein, N. (2004). Movement under control. *Linguistic Inquiry*, 35, 431-452.

Boeckx, C., Hornstein, N. & Nunes, J. (2010) Control as movement. Cambridge University Press: Cambridge, UK.

Boland, J. E. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton (eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond*, Psychology Press, pp. 51-76.

Bruening, B., (2001). QR obeys superiority: ACD and frozen scope. *Linguistic Inquiry* 32: 233-273.

Büring, D. (2005). *Binding Theory*. Cambridge University Press, Cambridge.

Caplan, D., & Waters, G. (1998). Verbal working memory and sentence comprehension. *Brain and Behavioral Sciences*, 22, 77-126.

Chao, W. and P. Sells. (1983). On the interpretation of resumptive pronouns. *NELS 13*, 47-61.

Chomsky, N. (1955/1975). *Logical structure of linguistic theory*. Springer.

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

Chomsky, N. (1973). Conditions on Transformations. In Kiparsky, P., & Peters, S. (eds), A Festschrift for Morris Halle. Mouton, The Hague.

Chomsky, N. (1977). *Essays on Form and Interpretation*. Elsevier, North-Holland.

Chomsky, N. (1981).*Lectures on government and binding*. Mouton de Gruyter: Berlin.

Chomsky, N. (1986). *Knowledge of language: its nature, origins and use*. New York.

Chomsky, N., & Lasnik, H. 1993. The theory of principles and parameters. In Jacobs, J., Von Stechow, A., & Sterneveld, W., (eds.) *Syntax: An International Handbook of Contemporary Research, Vol. 1*. Mouton de Gruyter, Berlin.

Chomsky, N. (2000). Minimalist Inquiries: The Framework. In Martin, R., Michaels, D., & Uriagereka, J. (eds), *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*. MIT Press, Cambridge, MA.

Chow, W-Y., Lewis, W., Lee, S., & Phillips, C. (2011). Immediate structural constraints on pronoun antecedent retrieval. Poster presented at the 24nd Annual Meeting of the CUNY Conference on Human Sentence Processing, Stanford, CA: March 24-26.

Chung, Sandra, William Ladusaw, and James McCloskey. 1995. Sluicing and logical form. *Natural Language Semantics* 3:239-282.

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review*, 3, 37-60.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica*, 11, 11-39.

Clifton, C., Kennison, S. M., & Albrecht, J. E. (1997). Reading the words *Her, His, Him:* Implications for parsing principles based on frequency and on structure. *Journal of Memory and Language*, 36, 276-292.

Conroy, A., Takahashi, E., Lidz, J. & Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry*, 40(3): 446-486.

Corbett, A. T., & Chang, F. R. (1983). Pronoun disambiguation: Accessing potential antecedents. *Memory & Cognition*, 11, 283-294.

Cowan, N. (1995). *Attention and Memory: An Integrated Framework*. New York: Oxford UP.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *Behavioral and Brain Sciences* 24, 87-185.

Cooper, R. (1978). 'Variable binding and relative clauses.' in F. Guenthner & S.J. Schmidt (eds.) *Formal Semantics and Pragmatics for Natural Language*, D. Reidel Publishing Company, Dordrecht, pp. 131-169.

Cowart, W., & Cairns, H. S. (1987). Evidence for an anaphoric mechanism within syntactic processing: some reference relations defy semantic and pragmatic constraints. *Memory & Cognition*, 15, 318-331.

Crain, S. and McKee, C. (1985). The acquisition of structural restrictions on anaphora. *Proceedings of NELS 15*. Amherst, Mass.: GLSA.

Demirdache, H. (1991). Resumptive Chains in Restrictive Relatives, Appositives and Dislocation Structures. Ph.D. Dissertation, MIT.

Den Dikken, M. (2001). Pluringulars, pronouns, and quirky agreement. *The Linguistic Review*, 18, 19-41.

Dillon, B., Chow, W.-Y., Wagers, M., Guo, T., Liu, F., & Phillips, C. (submitted). The structure-sensitivity of search: Evidence from Mandarin Chinese. Ms. University of Massachusetts.

Dillon, B., Mishler, A., Slogett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*

Dosher, B. A. (1984). Discriminating pre-experimental (semantic) information from learned (episodic) associations: A speed-accuracy study. *Cognitive Psychology*, 16, 519-555.

Dosher, B. A., & Rosedale, G. (1991). Judgments of semantic and episodic relatedness: Common time-course and failure of segregation. *Journal of Memory and Language*, 30, 125-160.

Drenhaus, H., Frisch, S., & Saddy, D. (2005). Processing negative polarity items: When negation comes through the backdoor. In: Kepser, S. and Reis, M. (eds.), *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*. Mouton de Gruyter, Berlin.

Drummond, A. & Kush, D. (To appear). Reanalysis as raising to object. *Syntax*.

Drummond, A., Kush, D. & Hornstein, N. Minimalist construal: Two approaches to A and B. In C. Boeckx (ed.), *The Oxford Handbook of Linguistic Minimalism*, Oxford University Press: Oxford, pp. 396-426.

Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and*

Verbal Behavior, 22, 75-87.

Eich, J.E. (1980). The cue-dependent nature of state-dependent retrieval. *Memory & Cognition*, 8, 157-173.

Eich, J.M. (1982). A composite holographic associative recall model. *Psychological Review*, 89, 627-61.

Elbourne, P. (2005a). *Situations and individuals*. Cambridge, MA: MIT Press.

Elbourne, P. (2005b). On the acquisition of Principle B. *Linguistic Inquiry*, 36:3, 333-365.

Engdahl, E. (1983). Parasitic gaps. *Linguistics and Philosophy*, 6, 5-34.

Ericsson, K.A. and Kintsch, W. (1995). Long-term working memory. *Psychol. Rev*. 102, 211-245.

Evans, G. (1977). Pronouns, quantifiers, and relative clauses (I). *Canadian Journal of Philosophy*, 7(3): 467-536.

Evans, G. (1980). Pronouns. *Linguistic Inquiry*, 11: 337-362.

Ferreira, M. (2009). Null subjects and finite control in Brazilian Portuguese. In J. Nunes (ed.), *Minimalist essays on Brazilian Portuguese syntax*. John Benjamins: Amsterdam, pp. 17-49.

Fiengo, R., & May, R. (1994). Indices and identity. Cambridge, Mass. ; London: MIT Press.

Filik R, Sanford A.J. & Leuthold H. (2008). Processing pronouns without antecedents: evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 20:7, 1315-1326.

Fodor, J.A., Bever, T., & Garrett, J. (1974). *The psychology of language: an introduction to psycholinguistics and generative grammar*. McGraw-Hill, New York.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry*, 9, 427-473.

Fodor, J. D., & Sag, I. A. (1982). Referential and quantificational indefinites. *Linguistics and Philosophy*, 5, 355-398.

Foley, C., Nuñez del Prado, Z., Barbier, I., & Lust, B. (1997). Operator-variable binding in the initial stage: An argument from English VP-ellipsis. *Cornell Working Papers in Linguistics*, 15, 119.

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Availability versus accessibility. *Journal of Memory and Language*, 56, 357-383.

Fox, D. (2000). *Economy and semantic interpretation*. Cambridge, Mass.: MIT Press.

Frank, R, and Vijay-Shanker, K. 2001. Primitive c-command. *Syntax* 4:164-204.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: a new two-stage parsing model. *Cognition*, 6, 291-325.

Frazier, L., Clifton, C., & Randall, J. (1983). Filling gaps: decision principles and structure in sentence comprehension. *Cognition*, 13, 187-222.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. Language and Cognitive Processes, 4, 93-126.

Frazier, L., & Clifton, C. (1996). Construal. MIT Press, Cambridge.

Frazier, L., & Clifton, C. (1997). Construal: Overview, motivation and some new evidence. *Journal of Psycholinguistic Research*, 26, 277-295.

Frazier, L. & Clifton, C. (2005). The syntax-discourse divide: processing ellipsis. *Syntax* 8:2, 121-174.

Frazier, L., & Clifton Jr, C. (2000). On bound variable interpretations: The LF-only hypothesis. *Journal of Psycholinguistic Research*, 29, 125-140.

Frazier, L., Flores-dArcais, G. (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language*, 28, 331-344.

Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, 33, 39-68.

Gawron, J.M., & Peters, S. (1990). *Anaphora and quantification in situation semantics*. Stanford, CA: CSLI Publications.

Geach, P. (1962) Reference and Generality. Cornell University Press, Ithaca, NY.

Gelman, A., & Hill, J. (2005). Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge, UK.

Gerrig, R. J. (1986). Process models and pragmatics. In N.E. Sharkey (ed.), *Advances in cognitive science*. Ellis Horwood: Chichester, England, pp. 23-42.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68, 1-76.

Gigerenzer, Gerd & Selten, R. (2002). *Bounded Rationality*. Cambridge: MIT Press.

Gillespie, M. & Pearlmutter, N.J. (2011). Hierarchy and scope of planning in subject-verb agreement production. *Cognition*, 118, 377-397.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. Psychological Review, 91, 1-65.

Gobet, F. (1998) Expert memory: a comparison of four theories. *Cognition*, 66, 115-152.

Gobet, F. and Simon, H.A. (1996) Recall of rapidly presented random chess positions is a function of skill. *Psychonomic Bulletin and Review* 3, 159-163.

Goldstein, D.G. & Gigerenzer, G. (2002). Models of ecological rationality: The recognition heuristic. *Psychological Review*, 109:1, 75-90.

Gordon, P.C., Grosz, B.J. & Gilliom, L.A. (1993). Pronouns, names and the centering of attention in discourse. *Cognitive Science*, 17(3):311-348
Gordon, P. C., & Hendrick, R. (1997). Intuitive knowledge of linguistic co-reference. *Cognition*, 62, 325-370.

Gordon, P.C., Hendrick, R., & Johnson, M. (2001). Memory interference during sentence-processing. Psychological Science, 13, 425-430.

Gordon, P.C., Hendrick, R., & Levine, W. (2002). Memory-load interference in syntactic processing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 27, 1411-1423.

Gordon, P.C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51, 97-114.

Gordon, P.C., Hendrick, R., Johnson, M., & Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 32, 1304-1321.

Greene, S.B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 266-283.

Grimshaw, J. (1991). Extended projection. Ms., Brandeis University, Waltham, Mass.

Grosz, B.J., Joshi, A.K., & Weinstein, S. 1995. Centering: a framework for modeling the

local coherence of discourse. *Computational Linguistics*, 21(2):203-225.
Guasti, M. T., & Chierchia, G. (1999/2000). Backward versus forward anaphora: Reconstruction in child grammar. *Language Acquisition*, 8, 129-170.

Guo, F. Foley, C., Chien, Y.-C., Chiang, C.-P., & Lust, B. (1996). Operator-variable binding in the initial state: A cross-linguistic study of VP-ellipsis structures in Chinese and English. *Cahiers de Linguistique-Asie Orientale*, 25, 334.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. Trends in Cognitive Sciences, 11(6), 236-242.

Hankamer, J., & Sag, I. (1976). Deep and surface anaphora. *Linguistic Inquiry*,7, 391-426
Harley, Heidi. (2003). Possession and the double object construction. In *Linguistic variation yearbook 2*, ed. by Pierre Pica and Johan Rooryck, 31-70. Amsterdam: John Benjamins.

Harris, C. L., & Bates, E. A. (2002). Clausal backgrounding and pronominal reference: a functionalist approach to c-command. Language and Cognitive Processes, 17,237-269.

Hart, P.E., Nilsson, N. J. & Raphael, B. (1968). A formal basis for the heuristic determination of minimum cost paths. IEEE Transactions on Systems Science and Cybernetics SSC4 4 (2): 100-107

Häussler, J. & Bader, M. (2009). Agreement checking and number attraction in sentence comprehension: Insights from German relative clauses. *Travaux du cercle linguistique de Prague 7*.

Heim, Irene. (1982). The semantics of definite and indefinite noun phrases. Doctoral dissertation, University of Massachusetts, Amherst.

Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Malden MA: Blackwell.

Hellan, L. 1988. *Anaphora in Norwegian and the theory of grammar.* Dordrecht: Foris.

Hendriks, H. (1987). Type change in semantics: the scope of quantification and coordination. In E. Klein & J. van Benthem (Eds.), Categories, Polymorphism and Unification: Centre for Cognitive Science University of Edinburgh Institute for Language Logic and Information University of Amsterdam.

Hintzman, D.L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments & Computers*. 16, 96-101.

Hintzman, D.L. (1988). Judgments of frequency and recognition memory in a multiple-trace model. *Psychological Review*, 95, 528-551.

Heycock, C. (1991). Layers of predication: The non-lexical syntax of clauses. Doctoral dissertation, University of Pennsylvania.

Heycock, C. (1992). Layers of predication and the syntax of the copula. *Belgian Journal of Linguistics* 7, 95-123

Holmberg, A., Nayudu, A., & Sheehan, M. (2009). Three partial null-subject languages: a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, 63:1, 59-97.

Heycock, C & Kroch, A. (1999). Pseudocleft connectivity: Implications for the LF interface. *Linguistic Inquiry* 30.3: 365-397.

Hornstein, N. (1995). *Logical form: From GB to Minimalism*. Blackwell: Cambridge, MA.

Hornstein, N. (2001). Move! A minimalist theory of construal. Blackwell: Cambridge, MA.

Hornstein, N. (2009). *A theory of syntax.* Cambridge University Press: Cambridge.

Hornstein, N., & Witkos, J. (2003). Yet another approach to existential constructions. Grammar in Focus: Festschrift for Christer Platzack, 167-184.

Israel, M. (1998). Ever: Polysemy and polarity sensitivity. *Linguistic Notes from La Jolla* 19, 29-45.

Jackendoff, R. S. (1977). *X-Syntax: A Study of Phrase Structure*. : MIT Press: Cambridge, MA.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness (Vol. 6)*. Harvard University Press.

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes and in reading comprehension. *Journal of Experimental Psychology: General*, 3,228-238.

Karlsson, F. (2007). Constraints on multiple center-embedding of clauses. *Journal of Linguistics* 43:2, 365-392.

Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., & Phillips, C. (2007). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language*, 56, 384-409.

Kazanina, N., & Phillips, C. (2001). Coreference in child Russian: Distinguishing syntactic and discourse constraints. *Proceedings of the 25th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press

Kazanina, N. & Phillips, C. (2010). Multiple routes to pronoun resolution in Russian cataphora. *Quarterly Journal of Experimental Psychology*, 63, 371-400.

Kehler, Andrew. (2002). *Coherence in discourse*. Stanford, Calif.: CSLI Publications.

Keller, F. and Alexopolou, A. (2005). A Crosslinguistic, Experimental Study of Resumptive Pronouns and that-Trace Effects Bruno G. Bara, Lawrence Barsalou, and Monica Bucciarelli, eds., *Proceedings of the 27th Annual Conference of the Cognitive Science Society*, 1120-1125.

Kaiser, E., Runner, J., Sussman, R., & Tanenhaus, M. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112, 55-80.

Kayne, R.S. Connectedness. *Linguistic Inquiry*, 14: 223-248.

Kennsion, S. (2003). Comprehending the pronouns her, him and his: Implications for theories of referential processing. *Journal of Memory and Language*, 49, 335-352.

Kennsion, S., & Trofe, J. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, 32, 355- 378.

Kimball, J. & Aissen, J. (1971). I think, you think, he think. *Linguistic Inquiry*, 2, 241- 246.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, 2, 15-47.

King, J., Andrews, C. & Wagers, M.W. (2012). Do reflexives always find a good antecedent for themselves? Poster presented at the 25th Annual Meeting of the CUNY Conference on Human Sentence Processing, New York, NY.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.

Klima, E. (1964). Negation in English. In Jerry A. Fodor & J. J. Katz, (eds.), The Structure of Language. Prentice Hall, New Jersey.

Kluender, R., & Kutas, M. (1993). Subjacency as a Processing Phenomenon. Language and Cognitive Processes, 8, 573-633.

Knuth, D. (1965/1997). *The Art of Computer Programming. Volume 1: Fundamental Algorithms. 3rd ed.* Addison-Wesley: Reading, MA.

Kohonen, T. (1980). Content-addressable memories. Springer Verlag, Berlin.

Koopman, H., & Sportiche, D. (1991). The position of subjects. Lingua, 85, 211-258.

Koornneef, A.W. (2008). *Eye-catching Anaphora*. Utrecht: LOT International Dissertation Series.

Koornneef, A. W., Wijnen, F., & Reuland, E. (2006). Towards a modular approach to anaphor resolution. In *Ambiguity in Anaphora Workshop Proceedings*, 65-72.

Kratzer, A. (2009). Making a pronoun: fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, 40, 187-237.

Kroch, Anthony. (1981). On the role of resumptive pronouns in amnestying island violations. In *Proceedings of the Seventeenth Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, University of Chicago, Chicago, IL.

Kuno, S. (1987). Functional syntax: anaphora, discourse and empathy. University of Chicago Press, Chicago.

Lakoff, G. (1970). 'Repartee'. *Foundations of Language* 6, 389 - 422.

Langacker, R. (1969). On pronominalization and the chain of command. In Modern studies in English, ed. by David Reibel & Sanford Schane, 160-186. Englewood Cliffs, NJ: Prentice-Hall.
Larson, R. K., (1985). Quantifying into NP. Manuscript, MIT.

Lasnik, H. (1976). Remarks on coreference. Linguistic Analysis, 2, 1-22.

Ladusaw, W. A. (1979). Negative Polarity as Inherent Scope. Unpublished Doctoral Dissertation. University of Texas at Austin.

Leddon, E., & Lidz, J. (2005). Reconstruction in child language. *Proceedings of the 30th Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press

Lees, R. B., & Klima, E. S. (1963). Rules for English pronominalization. *Language*, 39, 17-28.

Lewis, R. (1996). Interference in short-term memory: the magical number two (or three) in Sentence Processing. *Journal of Psycholinguistic Research*, 25,93-115.

Lewis, R., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science, 29(3), 375-419
Lewis, R., Vasishth, S., & Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. Trends in Cognitive Science, 10, 447-454.

Lidz, J., & Idsardi, W. (1999). Chains and phono-logical form. University of Pennsylvania

Working Papers in Lingusitics, 5, 119-125.

Lindström, P. (1966). First order predicate logic with generalized quantifiers. Theoria, 32, 186-195.

Lust, B., & Clifford, T. (1986). The 3-D study: Effects of depth, distance and directionality on children's acquisition of anaphora. In Lust, B. (Ed.), *Studies in the acquisition of anaphora. (Volume 1)*. Defining the constraints. Dordrecht, Holland: Reidel Press, (pp. 203-244).

Marcus, M. P. (1980). A theory of syntactic recognition for natural language. Cambridge, Mass.: MIT Press.

Martin, A. E., & McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb phrase ellipsis. *Journal of Memory and Language*, 58, 879-906.

Martin, A. E., & McElree, B. (2009). Memory operations that support language comprehension: Evidence from verb-phrase ellipsis. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 35, 1231-1239.

May, R. (1977). The Grammar of Quantification. Unpublished Doctoral Dissertation, MIT, Cambridge, MA.

May, R. (1985). Logical form : its structure and derivation. Cambridge, Mass.: MIT Press.

McCloskey, J. (2006). Resumption. In M. Everaert and H. van Riemsdijk (eds.) *The Blackwell Companion to Syntax*. Blackwell Publishing: Oxford, pp. 94-117.

McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: A time-course analysis.Journal of Memory and Language, 32, 536-571.

McElree, B. (1998). Attended and non-attended states in working memory: Accessing categorized structures. Journal of Memory & Language, 38, 225-252.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29, 111-123.

McElree, B., & Dosher, B. (1989). Serial position and set size in short-term memory: time course of recognition. *Journal of Experimental Psychology: General*, 18, 346-373.

McElree, B., & Dosher, B. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, 122, 291-315.

McElree, B., Foraker, S. & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48, 67-91.

McElree, B., & Griffith, T. (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. Learning, Memory, 24, 432-460.

McKinnon, R. & Osterhout, L. (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. Language and Cognitive Processes, 11(5), 495-524.

Miller, G., & Chomsky, N. (1963). Finitary models of language users. In Luce, R.D., Bush, R.R., & Galanter, E. (eds), Handbook of Mathematical Psychology, Volume II. John Wiley, New York.

Merchant, Jason. 2001. The syntax of silence: Sluicing, islands, and the theory of ellipsis. Oxford and New York: Oxford University Press.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. CSLI Publications: Stanford.

Montague, R. (1974). The proper treatment of quantification in ordinary English. In R. Thomason (Ed.), Formal Philosophy: selected papers of Richard Montague, 247-270. New Haven: Yale University Press.
Mostowski, A. (1957). On a generalization of quantifiers. *Fundamenta Mathematicae*, 44, 12-36.

Nairne, J.S. (1990). A feature model of immediate memory. *Memory & Cognition*, 18, 251- 269.

Nairne, J.S. (2002). Remembering over the short-term: The case against the standard model. *Annual Review of Psychology*, 53, 53-81.

Nederhof, M-J. 1993. Generalized left-corner parsing. In *Proceedings of the sixth conference on European chapter of the Association for Computational Linguistics: (EACL 93)*, 305-314.

Newell, A. (1973). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, pp. 463-526.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Nicol, J. (1988). Coreference processing during sentence comprehension. Doctoral dissertation, MIT.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18, 5-19.

Nieuwland, M., Otten, M. & van Berkum, J.J. A. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of Cognitive Neuroscience* 19:2, 228-236.

Nunes, J. (2008) Inherent case as a licensing condition for A-movement: the case of hyper raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics*, 7, 83-108.

Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, 31, 785-806.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition*, 25, 273-285.

Parker, D., & Phillips, C. (2011). Illusory negative polarity item licensing is selective. Poster presented at the 24nd Annual Meeting of the CUNY Conference on Human Sentence Processing, Stanford, CA: March 24-26.

Parker, D., MacMillan, G. & Phillips, C. (2013). Illusory NPI licensing: Now you see it, now you don't. Poster presented at the 26th annual CUNY human sentence processing conference. University of South Carolina

.

Partee, B. (1986). Noun phrase interpretation and type-shifting principles. In J. Groenendijk, D. de Jongh & M. Stokhof (Eds.), Studies in discourse representation theory and the theory of generalized quantifiers (pp. 115-143). Dordrecht: Foris.

Partee, B., & Rooth, M. (1983). Generalized conjunction and type ambiguity. In R. Buerle, C. Schwarze & A. von Stechow (Eds.), Meaning, use and the interpretation of language, 361-383. Berlin: Walter de Gruyter.

Pearlmutter, N., & MacDonald, M. (1995). Individual differences and probabilistic constraints in syntactic ambiguity resolution. *Journal of Memory and Language*, 34, 521-542.

Pearlmutter, N., Garnsey, S., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427-456.

Pesetsky, David. 1995. *Zero Syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.

Phillips, C. 2006. The real-time status of island phenomena. *Language*, 82, 4.

Phillips, C., Wagers, M. W., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (ed.), *Experiments at the Interfaces, Syntax & Semantics*, vol. 37. Bingley, UK: Emerald Publications, pp. 153-186.

Pickering, M. J., Barton, S., & Shillcock, R. (1994). Unbounded dependencies, island constraints and processing complexity. In C. Clifton, L. Frazier, &K. Rayner (Eds.), Perspectives onsentence processing (pp. 199-224). London: Lawrence Erlbaum.

Poesio, M. & Zucchi, A. (1992). On telescoping. In Chris Barker & David Dowty, (eds.), Proceedings of the Second Conference on Semantics and Linguistic Theory (SALT II), 347-366. Ohio State University Working Papers in Linguistics 40, Columbus, Ohio.

Pollard, C., & Sag, I. (1992). Anaphors in English and the scope of binding theory. *Linguistic Inquiry*, 23, 261-303.

Pollatsek, A., Juhasz, B.J., Reichle, E.D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements during reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Peformance*, 34, 726-750.

Postal, P. (1971). *Crossover Phenomena*. New York: Holt, Rinehart and Winston.

Prince, E. 1990. Syntax and discourse: A look at resumptive pronouns. *Berkeley Linguistic Society: Proceedings of the 16th Annual Meeting*, 482-497.

Raajimakers, J. G. W., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, 88, 93-134.

Ratcliff, R. (1980). A note on modeling accumulation of information when the rate of accumulation changes over time. *Journal of Mathematical Psychology*, 21, 178- 184.

Ratcliff, R. (1983). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510-532.

Ratcliff, R. & McKoon, G. (1995). Sequential effects in lexical decision: Tests of compound-cue retrieval theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 21:S, 1380-1388.

Rayner, K. (1997). Understanding eye movements in reading. *Scientific Studies of Reading*, 1, 317-339.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.

Rayner, K., Ashby, J., Pollatsek, A., & Reichle, E.D. (2004). The effects of word frequency and predictability on eye movements in reading: Implications for the E-Z Reader model. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 720-732.

Reinhart, T. (1976). The Syntactic Domain of Anaphora. Doctoral dissertation, Massachusetts Institute of Technology.

Reinhart, T. (1983). Anaphora and semantic interpretation. London: Croom Helm.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24, 657-720.

Riemsdijk, Henk van, & Williams, E. (1986). *Introduction to the theory of grammar*. Cambridge, Mass.: MIT Press.

Reuland, E. (2001). Primitives of binding. *Linguistic Inquiry*, 32, 439-492.

Roberts, C. (1987). Modal Subordination, Anaphora, and Distributivity. Doctoral Dissertation. University of Massachusetts, Amherst. GLSA, Amherst.

Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12(6): 683-721.

Rodman, R. (1976). 'Scope phenomena, movement transformations and relative clauses.' In B. Partee (ed.) *Montague Grammar*, Academic Press, New York.

Rohde, H. (2008). Coherence-driven effects in sentence and discourse processing. Unpublished Doctoral Dissertation. University of California, San Diego.

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

Ross, John R. (1969). Guess who? in CLS 5: Papers from the fifth regional meeting of the Chicago Linguistic Society, eds. Robert Binnick, Alice Davison, Georgia Green, and Jerry Morgan, 252-286. Chicago, Illinois: Chicago Linguistic Society.

Runner, J. T., Sussman, R. S. & Tanenhaus, M. K. (2006). Assigning Reference to Reflexives and Pronouns in Picture Noun Phrases. *Cognitive Science* 30, 1-49.

Safir, Ken. (2004a). *The syntax of anaphora*. Oxford: Oxford University Press.

Safir, Ken. (2004b). *The syntax of (in)dependence*. Cambridge, MA: MIT Press.

Sag, I. (1976). Deletion and logical form. Doctoral dissertation, Massachusetts Institute of Technology.

Sag, I., & Hankamer, J. (1984). Toward a theory of anaphoric processing. *Linguistics and Philosophy*, 7, 325-345.

Sells, P. (1985). Restrictive and non-restrictive modification. Technical report, CSLI Stanford.
Sells, P. (1987). Aspects of logophoricity. *Linguistic Inquiry*, 18, 445-479.

Shapiro, L., & Hestvik, A. (1995). On-line comprehension of VP-Ellipsis: Syntactic reconstruction and semantic influence. *Journal of Psycholinguistic Research*, 24, 517532.

Shan, C. & Barker, B. (2006). Explaining crossover and superiority as left-to-right evaluation. *Linguistics and Philosophy* 29: 91-134.

Sharvit, Y. (1999). Functional relative clauses. *Linguistics and Philosophy*, 22, 447-478.

Simon, H. (1957). A Behavioral Model of Rational Choice. In *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. Wiley: New York.

Simon, H.A. and Chase, W.G. (1973) Skill in chess. *American Scientist*, 61, 393-403.

Simon, H.A. and Gilmartin, K.J. (1973) A simulation of memory for chess positions. *Cognitive Psychology* 5, 29-46.

Sipser, M. (2006). Introduction to the Theory of Computation. Course Technology Inc.

Sprouse, J., Fukuda, S., Ono, H., & Kluender, R. (2011). Reverse island effects and the backward search for a licensor in multiple wh-questions. *Syntax*, 14, 179-203.

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, 114, 447-454.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (ed.), *The Oxford Handbook of Psycholinguistics*. Oxford, UK: Oxford University Press, pp. 327-342.

Staub, A., White, S.J., Drieghe, D., Hollway, E.C., & Rayner, K. (2010). Distributional effects of word frequency on eye fixation durations. *Journal of Experimental Psychology: Human Perception and Performance*, 36, 1280-1293.

Steedman, M. (2000). *The syntactic process*. Cambridge, Mass.: MIT Press.

Sternberg, S. (1966). High speed scanning in human memory. *Science*, 153, 652-654.

Sternberg, S. (1969). Memory scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57, 421-457.

Sternberg, S. (1975). Memory-scanning: New findings and current controversies. Quarterly *Journal of Experimental Psychology*, 27, 1-32.

Stowe, L. (1986). Evidence for online gap creation. *Language and Cognitive Processes*, 1, 227-245.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542-562.

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries. *Psychological Review*, 95, 15.

Treisman, A., & Souther, J. (1985). Search asymmetry: a diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, 114, 285.

Townsend, D. & Bever, T. (2001). *Sentence comprehension: the integration of habits and rules*. MIT Press, Cambridge.

Traxler, M. J., & Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: an eye-tracking study. *Journal of Memory and Language*, 35, 454-475.

Trommershäuser, J., Körding, K. & Landy, M.S., eds. (2011). *Sensory Cue Integration*. New York: Oxford University Press.

Tunstall, S. L. (1998). The interpretation of quantifiers: Semantics and processing. Unpublished Doctoral dissertation, University of Massachusetts, Amherst.

Uriagereka, J. (1999). Multiple Spell-out. In Epstein, S.D. & Hornstein, N. (eds.), *Working Minimalism*. Cambridge, MA: MIT Press.

Van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 33, 407-430.

Van Dyke, J. A. & McElree, B. (2006). Retrieval interference in sentence processing. *Journal of Memory and Language*, 55, 157-166.

Van Dyke, J. A., & McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*.

Van Gompel, R.P.G., & Liversedge, S.P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 128-139.

van Hoek, K. (1997). *Anaphora and conceptual structure*. Chicago: University of Chicago Press.

Vasishth, S., Brssow, S., Lewis, R. & Drenhaus, H. (2008). Processing Polarity: How the ungrammatical intrudes on the grammatical. Cognitive Science, 32, 685-712.

Vigliocco, G., & Nicol J. (1998). Separating hierarchical relations and word order in language production: is proximity concord syntactic or linear? *Cognition*, 68, 13-29.

Wagers, M. (2008). The structure of memory meets memory for structure in linguistic comprehension. Doctoral dissertation, University of Maryland, College Park.

Wagers, M., & McElree, B. (2009). Focal attention and the timing of memory retrieval in language comprehension. Talk given at the Architectures and Mechanisms for Language Processing Conference, September, Barcelona.

Wagers, M., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: representations and processes. *Journal of Memory and Language*, 61, 206-237.

Wanner, E., & Maratsos, M. (1978). An ATN approach to comprehension. In Halle, M., Bresnan, J., & Miller, G.A. (eds.), Linguistic Theory and Psychological Reality. MIT Press, Cambridge.

Wasow, T. (1972). Anaphoric relations in English. Doctoral dissertation, Massachusetts Institute of Technology.

Williams, E. S. (1977). Discourse & Logical Form. *Linguistic Inquiry*, 8(1), 101-139.

Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. Acta Psychologica, 41, 67-85.

Wickelgren, W. A., Corbett, A. T., & Dosher, B. A. (1980). Priming and retrieval from short-term memory: A speed-accuracy trade-off analysis. *Journal of Verbal Learning and Verbal Behavior*, 19, 387-404.

Xiang, M., Dillon, B. & Phillips, C. (2009). Illusory licensing across dependency types:ERP evidence. *Brain and Language*, 108, 40-55.

Xiang, M., Grove, J. & Giannakidou, A. 2011. Interference licensing of NPIs: Pragmatic reasoning and individual differences. Poster presented at the 24th annual meeting of the CUNY Conference of Human Sentence Processing, March, Stanford.
Yoshida, M. (2006). Constraints and mechanisms in long-distance dependency formation. Unpublished doctoral dissertation, University of Maryland, College Park.

Zribi-Hertz, A. (1989a). Anaphor Binding and Narrative Point of View English Reflexive Pronouns in Sentence and Discourse. *Language* 65(4): 695-727.

Zribi-Hertz, A. (1989b). Emphatic or reflexive? On the endophoric character of French luimeme and similar complex pronouns. *Journal of Linguistics*, 31, 333-374.

Zukowski, A. and Larsen, J. (2004). The production of sentences that we fill their gaps. Poster presented at the CUNY Sentence Processing Conference, April (University of

Maryland).