WORKING MEMORY IN SENTENCE COMPREHENSION: PROCESSING HINDI CENTER EMBEDDINGS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By
Shravan Vasishth

* * * * *

The Ohio State University
2002

Dissertation Committee:
Professor Shari Speer, Adviser
Professor Richard Lewis, Adviser
Professor John Josephson
Professor Keith Johnson
Professor Chris Brew

Approved by

Adviser
Department of Linguistics

Adviser
Department of Psychology
University of Michigan
Copyright by
Shravan Vasishth
2002
ABSTRACT

I used acceptability rating tasks and self-paced reading studies to investigate three working memory-based theories’ predictions regarding sentence processing difficulty in Hindi center-embedding constructions: Hawkins’ Early Immediate Constituents (EIC), Gibson’s Discourse Locality Theory (DLT), and Lewis’ Retrieval Interference Theory (RIT). Two main issues were investigated: (a) the effect of definiteness marking on direct objects; and (b) the effect of increasing head-dependent distance.

First, definite-marked direct objects were found to be harder to process than bare (indefinite) direct objects, contra EIC, and contra DLT. I argue that, due to discourse constraints, indefinites are harder to process when they are in subject position, whereas definites are harder to process in the direct-object position.

Second, regarding distance between heads and dependents, distance was manipulated in two distinct ways: (a) by fronting indirect objects, and by fronting direct objects in center embeddings like ‘siitaa-ne hari-ko kitaab khariid-neko kahaa’, “Sita told Hari to buy a book” ; and (b) by inserting an adverb between the final NP and the innermost verb in canonical order center embeddings.

One finding was that if distance is increased between heads and dependents by reordering the dependents ((a) above), processing becomes more difficult, as predicted by EIC and DLT, and contra RIT. However, processing is, suprisingly, easier when distance is increased between the head and its dependents by inserting an adverb between them ((b) above). This goes against EIC, DLT, and RIT’s predictions. I explain these results as follows: fronting indirect or direct objects renders them more
similar to subjects (since fronted objects are in a typical subject position), causing increased similarity-based interference between the actual subject and the fronted object; by contrast, the easier processing due to adverb insertion occurs because the adverb strengthens the activation level of the current hypothesis (in working memory) regarding the sentence completion.

In sum, EIC, DLT, and RIT are only partly able to correctly characterize important cross-linguistic aspects of human sentence parsing. This incomplete coverage of the empirical results motivates a new, more general model of human sentence parsing that correctly accounts for reading-time and acceptability rating data from four languages.
To my parents
ACKNOWLEDGMENTS

Most people know more as they grow older.
I give all that the cold shoulder.

I spent my second quarter-century
Losing what I learnt at university

And refusing to take in what has happened since.
Now I know none of the names in the public prints,

And am starting to give offence by forgetting faces
And swearing I’ve never been in certain places.


My greatest debt is to my advisers, Richard Lewis and Shari Speer. Most of the work presented here is a result of their detailed feedback at every stage. I also thank my other committee members, Keith Johnson, John Josephson, and Chris Brew, for their many insightful and helpful comments. The following people had a powerful impact on this research, and in many cases altered the entire course of the investigation: Mary Beckman, Martin Jansche, Neal Johnson, Brian Joseph, Geert-Jan M. Kruijff, Edson Miyamoto, Mineharu Nakayama, and Scott Schwenter. Many others (too many to list) also contributed a great deal. My thanks to them all.

Five people in particular have had a lasting and significant impact on my university career in general. Gautam Sengupta (Linguistics, University of Hyderabad), and Michael Wescoat (formerly at Osaka University) influenced my early years as a graduate student through their outstanding courses. Later, at Ohio State, Tim Carlson’s extraordinary lectures on mathematical logic and proof theory, and
the many discussions with Scott Schwenter (Spanish and Portuguese), transformed my academic life. Also at Ohio State, Martin Jansche patiently educated me on linguistic, mathematical, and computational matters.

The faculty and graduate students at the Jawaharlal Nehru University, New Delhi, deserve special mention, especially Professor Ayesha Kidwai, Aarti Venkataraman, Professor R. S. Gupta, and Professor Sushama Jain. They provided the space and assistance for running the experiments described here. Much other logistical support was provided by my parents, Satish and Saroj Vasishth, and my wife, Andrea Vasishth. Without all these people the field work would have been impossible. Andrea helped considerably with editing draft versions, and transcribed the stimuli that appear in Appendix A; she also critically evaluated every aspect of this research project. Thanks also go to Andrea for putting up with my daily academic excesses and long mental absences.

I thank Tom Santner and Sumithra Mandrekar, Consultants at the Statistical Consulting Service, Ohio State University, for providing detailed guidance on the proper use of linear mixed-effects models and multiple linear regression. Thanks also to Mary Beckman, Martin Jansche, and John Josephson, for many useful and educative discussions about statistical matters.

Finally, my family, along with the team of renal transplant doctors and surgeons at the All India Institute of Medical Sciences, New Delhi, in particular, Dr. S. N. Mehta, Dr. S. C. Tiwari, Dr. S. C. Dash, Dr. S. K. Agarwal, and Dr. S. Saxena, conspired to give me another chance at life back in 1985. Without this, of course, nothing would be possible.

This research was funded by very generous travel and research grants from the Department of Linguistics, the Department of Computer and Information Science, and the Center for Cognitive Science, The Ohio State University. I gratefully acknowledge this funding.
VITA

March 20, 1964  Born in New Delhi, India
1983          High school certificate
              St. Columba’s School, New Delhi
1986-89      BA (Honours), Japanese
              Jawaharlal Nehru University, New Delhi
1986-89      Diplôme de langue française
              Alliance Française, New Delhi
1989-90      Diploma, Japanese Studies Program
              Osaka University of Foreign Studies, Japan
1990-92      Technical translator (Japanese to English)
              Hara Kenzo Patent Law Firm, Osaka, Japan
1992-94      MA, Linguistics
              Jawaharlal Nehru University, New Delhi
1995-96      Research student
              Faculty of Language and Culture, Osaka, Japan
1996-97      Graduate Teaching Assistant
              Faculty of Language and Culture, Osaka, Japan
1997-99, 2001-02 Graduate Research Associate
              Department of Linguistics, The Ohio State University
March 22, 2002 MS, Computer and Information Science
              The Ohio State University
PUBLICATIONS

Research Publications


**FIELDS OF STUDY**

Major Field: Linguistics
Areas of Specialization: Computational and Experimental Psycholinguistics
# TABLE OF CONTENTS

Abstract .......................................................... ii
Acknowledgments .................................................. v
Vita ................................................................. vii
List of Tables ..................................................... xiii
List of Figures ..................................................... xviii
List of Algorithms ............................................... xxiii

Chapters:

Preface ............................................................. 1

1 Introduction ..................................................... 2
  1.1 Self-center embeddings and human sentence processing .... 4
  1.2 Self-center embeddings in Hindi .......................... 9
    1.2.1 Some syntactic properties of self-center embeddings ... 9
    1.2.2 Semantics/pragmatics: Differential Object Marking in Hindi . 15
  1.3 Case-marked objects are specific, but conversationally implicate definiteness ........................................... 30
  1.4 Bare singular NPs really are indefinites .................... 31
    1.4.1 Dayal’s arguments against treating bare singular NPs as indefinites ............................................. 32
    1.4.2 Alternative explanations for the Hindi facts ........... 33

2 Three models of sentence processing ........................ 38
  2.1 Hawkins’ Early Immediate Constituents (1994), (1998) ..... 41
  2.2 Gibson’s Dependency Locality Theory (2000) ............... 51
  2.3 Lewis’ Retrieval Interference Theory (2001, 2002) ....... 53
4.2 Experiment 5: Indirect-object fronting ........................................ 112
  4.2.1 Method ................................................................. 113
  4.2.2 Results: The effect of definiteness marking (canonical word order) 115
  4.2.3 Results: The effect of definiteness marking (fronted indirect objects) ......................................................... 118
  4.2.4 Results: The effect of increasing distance between a verb and its arguments (bare direct object) .............................. 121
  4.2.5 Results: The effect of increasing distance between a verb and its arguments (definite-marked direct object) .................. 123
  4.2.6 Discussion ............................................................... 126
4.3 Experiment 6: Direct-object fronting ................................. 129
  4.3.1 Method ................................................................. 129
  4.3.2 Results: The effect of definiteness marking (canonical order) ........ 131
  4.3.3 Results: The effect of definiteness marking (direct object fronted) 134
  4.3.4 Results: The effect of direct-object fronting (bare direct object) 137
  4.3.5 Results: The effect of direct-object fronting (definite-marked direct object) ......................................................... 139
  4.3.6 Discussion ............................................................... 141
5 Increasing distance can facilitiate processing: Experiment 7 ........ 144
  5.1 Method ................................................................. 145
  5.1.1 Subjects ............................................................... 145
  5.1.2 Materials and Procedure ........................................... 145
  5.2 Results: Adverb insertion in single embeddings ...................... 147
  5.3 Results: Adverb insertion in double embeddings ..................... 149
  5.4 Discussion ............................................................... 151
6 Word length, position, and reading time: Myths and misunderstandings . 156
  6.1 Hindi orthography ....................................................... 157
  6.2 Does word length affect reading time in Hindi? ...................... 160
6.2.1 Computing word length for Hindi .......................... 160
6.2.2 Reading time and word length – results from Experiment 2 .......................... 160
6.2.3 Correcting for word length ................................ 167
6.3 Does position affect reading time? ................................. 168

7 Concluding remarks, and a new sentence processing model ............................. 181
7.1 Multiple constraints in language processing: A summing up ............................. 181
7.2 A model of sentence processing based on abductive inference .......................... 187
7.2.1 The underlying competence grammar $G$ ................................ 188
7.2.2 The set $H$ of hypothesis formers ................................ 189
7.2.3 Some definitions .......................................................... 191
7.2.4 The algorithm .............................................................. 194
7.2.5 The complexity metric .................................................... 196
7.2.6 Computational properties of the abductive parser ................................. 200
7.2.7 Empirical coverage of AIM .............................................. 203
7.2.8 The connection between the abduction-based model and the complexity metric; and integrating AIM, DLT, and RIT .......................... 217

APPENDIX A: Stimuli for the experiments ........................................ 219
A.1 Experiment 1 ................................................................. 220
A.2 Experiment 2 ................................................................. 222
A.3 Experiment 3 ................................................................. 224
A.4 Experiment 4 ................................................................. 231
A.5 Experiments 5 and 6 ....................................................... 232
A.6 Experiment 7 ................................................................. 233

Index ................................................................. 237

References ................................................................. 240
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Inflections on infinitivals and nominals</td>
<td>11</td>
</tr>
<tr>
<td>1.2</td>
<td>The percentage (of all occurrences) of indefinites versus definites occurring as subjects and as objects. Note here that Givón’s claim is not that indefinites occur more frequently than definites in object position, but that “the accusative object position in English is a prime environment in which referential nouns are introduced into the discourse for the first time, as indefinites.”</td>
<td>21</td>
</tr>
<tr>
<td>1.3</td>
<td>The number of indefinite versus definite determiners occurring immediately post-verbally, and the percentages of each type. This table is consistent with Hopper and Thompson’s claim that definites are the ‘norm’ for direct objects, for the reasons discussed in the text.</td>
<td>22</td>
</tr>
<tr>
<td>1.4</td>
<td>The total number of indefinite versus definite determiners occurring in the corpus. There are nearly twice as many definites as indefinites. This shows that Table 1.3 does not necessarily support Hopper and Thompson’s claim: the larger number of definites immediately post-verbally might be simply because there are more definites overall, and not for the reasons Hopper and Thompson provide.</td>
<td>23</td>
</tr>
<tr>
<td>1.5</td>
<td>Percentages of subjects and direct objects correctly found by heuristic from samples of 100 sentences of each type</td>
<td>25</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>1.6</td>
<td>Total numbers of indefinites and definites (second column) in the North American News corpus; the number (third column) and percentage (fourth column) of indefinites/definites occurring immediately after a verb; and the number (fifth column) and percentage (sixth column) appearing sentence-initially.</td>
<td></td>
</tr>
<tr>
<td>3.1</td>
<td>Expt. 2 (Conditions A vs. C); contrast analysis for NP3</td>
<td></td>
</tr>
<tr>
<td>3.2</td>
<td>Expt. 2 (Conditions A vs. C); ANOVA based on a linear mixed-effects model for NP3 (cf. page 168)</td>
<td></td>
</tr>
<tr>
<td>3.3</td>
<td>Expt. 2 (Conditions A vs. C); contrast analysis at V2</td>
<td></td>
</tr>
<tr>
<td>3.4</td>
<td>Expt. 2 (Conditions B vs. D); contrast analysis for NP4</td>
<td></td>
</tr>
<tr>
<td>3.5</td>
<td>Expt. 2 (Conditions B vs. D); ANOVA based on a linear mixed-effects model for NP4</td>
<td></td>
</tr>
<tr>
<td>3.6</td>
<td>Expt. 2 (Conditions B vs. D); contrast analysis at V3</td>
<td></td>
</tr>
<tr>
<td>3.7</td>
<td>Expt. 3 (Condition C vs. D); contrast analysis for NP4</td>
<td></td>
</tr>
<tr>
<td>3.8</td>
<td>Expt. 3 (Condition C vs. D); ANOVA based on linear mixed effects model for NP4</td>
<td></td>
</tr>
<tr>
<td>3.9</td>
<td>Expt. 3 (Condition C vs. D); contrast analysis for V2</td>
<td></td>
</tr>
<tr>
<td>3.10</td>
<td>Expt. 3 (Conditions B vs. G); contrast analysis for NP4</td>
<td></td>
</tr>
<tr>
<td>3.11</td>
<td>Expt. 3 (Conditions B vs. G); ANOVA based on linear mixed-effects model for NP4</td>
<td></td>
</tr>
<tr>
<td>3.12</td>
<td>Expt. 3 (Conditions B vs. G); contrast analysis for V2</td>
<td></td>
</tr>
<tr>
<td>3.13</td>
<td>Expt. 3; contrast analysis for NP3</td>
<td></td>
</tr>
<tr>
<td>3.14</td>
<td>Expt. 3 (Condition D vs. G); contrast analysis for NP4</td>
<td></td>
</tr>
<tr>
<td>3.15</td>
<td>Expt. 3 (Condition D vs. G); contrast analysis for V2</td>
<td></td>
</tr>
<tr>
<td>3.16</td>
<td>Expt. 3 (Condition B vs. C); contrast analysis for NP2</td>
<td></td>
</tr>
<tr>
<td>3.17</td>
<td>Expt. 3 (Condition B vs. C); contrast analysis for NP3</td>
<td></td>
</tr>
<tr>
<td>3.18</td>
<td>Expt. 3 (Condition B vs. C); contrast analysis for NP4</td>
<td></td>
</tr>
<tr>
<td>3.19</td>
<td>Expt. 3 (Condition B vs. C); contrast analysis for V2</td>
<td></td>
</tr>
<tr>
<td>3.20</td>
<td>Expt. 3 (Condition E vs. G); contrast analysis for NP2</td>
<td>100</td>
</tr>
<tr>
<td>3.21</td>
<td>Expt. 3 (Condition E vs. G); contrast analysis for NP4</td>
<td>100</td>
</tr>
<tr>
<td>3.22</td>
<td>Expt. 3 (Condition E vs. G); contrast analysis for V2</td>
<td>100</td>
</tr>
<tr>
<td>3.23</td>
<td>Expt. 3 (Conditions A vs. B); contrast analysis for NP3</td>
<td>102</td>
</tr>
<tr>
<td>3.24</td>
<td>Expt. 3 (Conditions A vs. B); contrast analysis for V2</td>
<td>102</td>
</tr>
<tr>
<td>3.25</td>
<td>Expt. 3 (Conditions A vs. B); contrast analysis for V1</td>
<td>102</td>
</tr>
<tr>
<td>3.26</td>
<td>Summary of predictions and results in Experiment 3. The tick mark (✔) indicates that the prediction was correct, and the cross (✘) indicates that the prediction was incorrect.</td>
<td>103</td>
</tr>
<tr>
<td>4.1</td>
<td>Expt. 5 (Conditions A vs. C); contrast analysis for NP3</td>
<td>116</td>
</tr>
<tr>
<td>4.2</td>
<td>Expt. 5 (Conditions A vs. C); ANOVA based on linear mixed-effects model for NP3</td>
<td>116</td>
</tr>
<tr>
<td>4.3</td>
<td>Expt. 5 (Conditions A vs. C); contrast analysis for V2</td>
<td>116</td>
</tr>
<tr>
<td>4.4</td>
<td>Expt. 5 (Conditions B vs. D); contrast analysis for NP3</td>
<td>119</td>
</tr>
<tr>
<td>4.5</td>
<td>Expt. 5 (Conditions B vs. D); ANOVA based on linear mixed-effects model for NP3</td>
<td>119</td>
</tr>
<tr>
<td>4.6</td>
<td>Expt. 5 (Conditions B vs. D); contrast analysis for V2</td>
<td>119</td>
</tr>
<tr>
<td>4.7</td>
<td>Expt. 5 (Conditions A vs. B); contrast analysis for first position NP</td>
<td>122</td>
</tr>
<tr>
<td>4.8</td>
<td>Expt. 5 (Conditions A vs. B); contrast analysis for second position NP</td>
<td>122</td>
</tr>
<tr>
<td>4.9</td>
<td>Expt. 5 (Conditions A vs. B); contrast analysis for V2</td>
<td>122</td>
</tr>
<tr>
<td>4.10</td>
<td>Expt. 5 (Conditions C vs. D); contrast analysis for second position NP</td>
<td>124</td>
</tr>
<tr>
<td>4.11</td>
<td>Expt. 5 (Conditions C vs. D); contrast analysis for third position NP</td>
<td>124</td>
</tr>
<tr>
<td>4.12</td>
<td>Expt. 5 (Conditions C vs. D); contrast analysis for V2</td>
<td>124</td>
</tr>
<tr>
<td>4.13</td>
<td>Expt. 5 (Conditions C vs. D); contrast analysis for V1</td>
<td>124</td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>4.14</td>
<td>Expt. 6 (Conditions A vs. C); contrast analysis for NP1</td>
<td>132</td>
</tr>
<tr>
<td>4.15</td>
<td>Expt. 6 (Conditions A vs. C); contrast analysis for NP3</td>
<td>132</td>
</tr>
<tr>
<td>4.16</td>
<td>Expt. 6 (Conditions A vs. C); ANOVA based on linear mixed-effects model for NP3</td>
<td>132</td>
</tr>
<tr>
<td>4.17</td>
<td>Expt. 6 (Conditions A vs. C); contrast analysis for V2</td>
<td>132</td>
</tr>
<tr>
<td>4.18</td>
<td>Expt. 6 (Conditions A vs. C); contrast analysis for V1</td>
<td>132</td>
</tr>
<tr>
<td>4.19</td>
<td>Expt. 6 (Conditions B vs. D); contrast analysis for direct-object NP in first position</td>
<td>135</td>
</tr>
<tr>
<td>4.20</td>
<td>Expt. 6 (Conditions B vs. D); ANOVA based on linear mixed-effects model</td>
<td>135</td>
</tr>
<tr>
<td>4.21</td>
<td>Expt. 6 (Conditions B vs. D); contrast analysis for V2</td>
<td>135</td>
</tr>
<tr>
<td>4.22</td>
<td>Expt. 6 (Conditions A vs. B); contrast analysis for V2</td>
<td>138</td>
</tr>
<tr>
<td>4.23</td>
<td>Expt. 6 (Conditions C vs. D); contrast analysis for V2</td>
<td>140</td>
</tr>
<tr>
<td>5.1</td>
<td>Expt. 7 (Conditions A vs. B); contrast analysis for NP3</td>
<td>148</td>
</tr>
<tr>
<td>5.2</td>
<td>Expt. 7 (Conditions A vs. B); contrast analysis for V2</td>
<td>148</td>
</tr>
<tr>
<td>5.3</td>
<td>Expt. 7 (Conditions C vs. D); contrast analysis for NP4</td>
<td>150</td>
</tr>
<tr>
<td>5.4</td>
<td>Expt. 7 (Conditions C vs. D); contrast analysis for V3</td>
<td>150</td>
</tr>
<tr>
<td>6.1</td>
<td>Expt. 2, ANOVA of word length (wl), using stimulus words in Conditions A and C and filler words in position three, as a predictor of reading time</td>
<td>163</td>
</tr>
<tr>
<td>6.2</td>
<td>Expt. 2, ANOVA of word length (wl), using stimulus words in Conditions B and D and filler words in position four, as a predictor of reading time</td>
<td>164</td>
</tr>
<tr>
<td>6.3</td>
<td>Expt. 2, ANOVA of word length (wl), using all words in all positions, as a predictor of reading time</td>
<td>165</td>
</tr>
<tr>
<td>6.4</td>
<td>Raw reading time correlation coefficients (a) when position is held constant; and (b) when words in all positions are considered</td>
<td>166</td>
</tr>
</tbody>
</table>
6.5 t- and p-values for each successive position in the sentence (82a));
   Multiple R-squared: 0.1357, Adjusted R-squared (penalizing for
   higher p): 0.1207. .................................................. 175

6.6 t- and p-values for each successive position in the sentence (82b));
   Multiple R-squared: 0.0582, Adjusted R-squared (penalizing for
   higher p): 0.04204. ................................................... 176

6.7 t- and p-values for each successive position in the sentence (82c));
   Multiple R-squared: 0.02471, Adjusted R-squared (penalizing for
   higher p): 0.007976. ................................................... 177

7.1 Summary of conclusions. Each cell indicates the prediction of each
   theory regarding a particular issue; a tick mark (✔) stands for an
   affirmative answer to questions like “Does only syntax matter?”,
   while a cross (✘) stands for an answer in the negative. A boxed
   tick mark or cross indicates that the Hindi results suggest that
   the answer is either not exactly correct, or is wrong. An up-
   arrow (↑) indicates that a slowdown in reading time is predicted
   (at the innermost verb), while a down-arrow (↓) indicates that a
   speedup is predicted. A question mark indicates that the theory is
   underspecified regarding that particular issue. .......................... 183
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Syntactic tree for Example (3)</td>
<td>9</td>
</tr>
<tr>
<td>1.2</td>
<td>Mean percentage of indefinite and definite determiners appearing sentence-initially and immediately following a verb (North American News corpus). Also see Table 1.6.</td>
<td>26</td>
</tr>
<tr>
<td>1.3</td>
<td>Percentage of determiners appearing sentence-initially and immediately following a verb (North American News corpus). Also see Table 1.6.</td>
<td>27</td>
</tr>
<tr>
<td>2.1</td>
<td>The CRD of the VP in (35a)</td>
<td>42</td>
</tr>
<tr>
<td>2.2</td>
<td>The CRD of the VP in (35b)</td>
<td>43</td>
</tr>
<tr>
<td>2.3</td>
<td>Illustration of DLT’s storage cost computation</td>
<td>51</td>
</tr>
<tr>
<td>2.4</td>
<td>Schematic illustration of interference; the solid lines collectively indicate a strong match, and the broken lines show interference effects.</td>
<td>61</td>
</tr>
<tr>
<td>3.1</td>
<td>Experiment 1 results, with 95% confidence intervals. A rating of 1 represents the acceptability judgement “completely unacceptable”, and a rating of 7 represents “completely acceptable”</td>
<td>67</td>
</tr>
<tr>
<td>3.2</td>
<td>Illustration of presentation on the computer screen</td>
<td>72</td>
</tr>
<tr>
<td>3.3</td>
<td>Expt. 2, single embeddings, bare versus case-marked final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals</td>
<td>77</td>
</tr>
<tr>
<td>3.4</td>
<td>Expt. 2, single embeddings, bare versus case-marked final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals</td>
<td>77</td>
</tr>
</tbody>
</table>
3.5 Expt. 2, double embeddings, bare versus case-marked final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals ........................................... 79
3.6 Expt. 2, double embeddings, bare versus case-marked final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals 79
3.7 Expt. 3 (Conditions C vs. D); raw RTs, with 95% confidence intervals ................................................................. 90
3.8 Expt. 3 (Conditions C vs. D); residual RTs, with 95% confidence intervals ................................................................. 90
3.9 Expt. 3 (Conditions B vs. G); raw RTs, with 95% confidence intervals ................................................................. 93
3.10 Expt. 3 (Conditions B vs. G); residual RTs, with 95% confidence intervals ................................................................. 93
3.11 Expt. 3 (Condition D vs. G); raw RTs, with 95% confidence intervals ................................................................. 95
3.12 Expt. 3 (Condition B vs. C); raw RTs, with 95% confidence intervals ................................................................. 98
3.13 Expt. 3 (Condition E vs. G); raw RTs, with 95% confidence intervals ................................................................. 100
3.14 Expt. 3 (Conditions A vs. B); raw RTs, with 95% confidence intervals ................................................................. 102
4.1 Experiment 4 results, with 95% confidence intervals. A rating of 1 represents the acceptability judgement “completely unacceptable”, and a rating of 7 represents “completely acceptable.” ........ 110
4.2 Expt. 5, canonical order with case marking absent or present on final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals ................................................................. 117
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3</td>
<td>Expt. 5, canonical order with case marking absent or present on final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals</td>
<td>117</td>
</tr>
<tr>
<td>4.4</td>
<td>Expt. 5, IO fronted, case marking absent or present on final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals</td>
<td>120</td>
</tr>
<tr>
<td>4.5</td>
<td>Expt. 5, IO fronted, case marking absent or present on final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals</td>
<td>120</td>
</tr>
<tr>
<td>4.6</td>
<td>Expt. 5, canonical versus IO-fronted order, bare final NP (Conditions A vs. B); raw RTs, with 95% confidence intervals</td>
<td>122</td>
</tr>
<tr>
<td>4.7</td>
<td>Expt. 5, canonical versus fronted IO, case-marked final NP (Conditions C vs. D); raw RTs, with 95% confidence intervals</td>
<td>125</td>
</tr>
<tr>
<td>4.8</td>
<td>Expt. 6, canonical order, case marking absent or present on final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals</td>
<td>133</td>
</tr>
<tr>
<td>4.9</td>
<td>Expt. 6, canonical order, case marking absent or present on final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals</td>
<td>133</td>
</tr>
<tr>
<td>4.10</td>
<td>Expt. 6, DO fronted, case marking absent or present on final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals</td>
<td>136</td>
</tr>
<tr>
<td>4.11</td>
<td>Expt. 6, DO fronted, case marking absent or present on final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals</td>
<td>136</td>
</tr>
<tr>
<td>4.12</td>
<td>Expt. 6, canonical versus fronted DO, bare final NP (Conditions A vs. B); raw RTs with 95% confidence intervals</td>
<td>138</td>
</tr>
<tr>
<td>4.13</td>
<td>Expt. 6, canonical versus fronted DO, case-marked final NP (Conditions C vs. D); raw RTs in single embeddings, with 95% confidence intervals</td>
<td>140</td>
</tr>
<tr>
<td>5.1</td>
<td>Expt. 7 (Conditions A vs. B); raw RTs, with 95% confidence intervals</td>
<td>148</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5.2</td>
<td>Expt. 7 (Conditions C vs. D); raw RTs, with 95% confidence intervals</td>
<td>150</td>
</tr>
<tr>
<td>6.1</td>
<td>A partial inventory of Devanagari characters</td>
<td>159</td>
</tr>
<tr>
<td>6.2</td>
<td>Expt. 2; plot of RT against word length (using words from Conditions A, C, and from fillers in position three)</td>
<td>163</td>
</tr>
<tr>
<td>6.3</td>
<td>Expt. 2; plot of RT against word length (using words from Conditions B, D, and from fillers in position four)</td>
<td>164</td>
</tr>
<tr>
<td>6.4</td>
<td>Expt. 2; plot of RT against word length (all words, ignoring position)</td>
<td>165</td>
</tr>
<tr>
<td>6.5</td>
<td>Expt. 2; raw RTs by position in single versus double embeddings, non-case-marked final NP</td>
<td>172</td>
</tr>
<tr>
<td>6.6</td>
<td>Expt. 2; raw RTs by position in single versus double embeddings, case-marked final NP</td>
<td>172</td>
</tr>
<tr>
<td>6.7</td>
<td>Expt. 7; single embeddings with adverb present versus absent</td>
<td>173</td>
</tr>
<tr>
<td>6.8</td>
<td>Expt. 7; double embeddings with adverb present versus absent</td>
<td>173</td>
</tr>
<tr>
<td>6.9</td>
<td>Tukey box-and-whiskers plot for sentence (82a)</td>
<td>175</td>
</tr>
<tr>
<td>6.10</td>
<td>Tukey box-and-whiskers plot for sentence (82b)</td>
<td>176</td>
</tr>
<tr>
<td>6.11</td>
<td>Tukey box-and-whiskers plot for sentence (82c)</td>
<td>177</td>
</tr>
<tr>
<td>6.12</td>
<td>Expt. 2; plot of RT against position</td>
<td>179</td>
</tr>
<tr>
<td>6.13</td>
<td>Expt. 3; plot of RT against position</td>
<td>179</td>
</tr>
<tr>
<td>6.14</td>
<td>Expt. 5; plot of RT against position</td>
<td>180</td>
</tr>
<tr>
<td>6.15</td>
<td>Expt. 6; plot of RT against position</td>
<td>180</td>
</tr>
<tr>
<td>7.1</td>
<td>Tree representation of schemas</td>
<td>190</td>
</tr>
<tr>
<td>7.2</td>
<td>Complete derivation history for (87a); total cost = 48</td>
<td>207</td>
</tr>
<tr>
<td>7.3</td>
<td>Complete derivation history for (87b); total cost = 39</td>
<td>208</td>
</tr>
<tr>
<td>7.4</td>
<td>Complete derivation for (88a); total cost = 38</td>
<td>209</td>
</tr>
<tr>
<td>7.5</td>
<td>Complete derivation for (88b); total cost = 25</td>
<td>209</td>
</tr>
<tr>
<td>7.6</td>
<td>Derivation history for (89a); total cost = 38</td>
<td>210</td>
</tr>
<tr>
<td>Section</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.7</td>
<td>Complete derivation history for (89b); total cost = 48</td>
<td>211</td>
</tr>
<tr>
<td>7.8</td>
<td>Derivation for (90)</td>
<td>212</td>
</tr>
<tr>
<td>7.9</td>
<td>Derivation for (91a); total cost = 50</td>
<td>213</td>
</tr>
<tr>
<td>7.10</td>
<td>Derivation for (91b); total cost = 28</td>
<td>214</td>
</tr>
<tr>
<td>7.11</td>
<td>Derivation for Dutch crossed embedding (92a); total cost = 41</td>
<td>215</td>
</tr>
<tr>
<td>7.12</td>
<td>Derivation for German nested embedding (92b); total cost = 41</td>
<td>215</td>
</tr>
<tr>
<td>7.13</td>
<td>Derivation for Hindi examples (93a,b)</td>
<td>216</td>
</tr>
</tbody>
</table>
LIST OF ALGORITHMS

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1    Main loop of the parser</td>
<td>195</td>
</tr>
<tr>
<td>2    The main function <strong>Process</strong>(S)</td>
<td>195</td>
</tr>
</tbody>
</table>
PREFACE

The frontispiece needs an explanation: it is a scanning electron microscope image of an ant. It adorns this dissertation because, during the period that I was running experiments in New Delhi, hundreds (perhaps thousands) of ants infiltrated one of the laptops being used for conducting the experiments. The ant image is intended to commemorate this now-amusing, but then-harrowing, moment of my field work.

I also provide a potentially useful bug report here. If ants happen to invade your laptop, do the following. Take three plastic shopping bags. Place the laptop inside one and tightly wrap the bag around it and seal it with duct tape. Now tightly wrap the wrapped-up laptop in another plastic bag, and then a third one; each time, seal the bag with duct tape. Leave the laptop alone for three or four hours, and then quickly unwrap and throw away the plastic bags. The multiple wrapping is essential since the ants (New Delhi ants anyway) will have bitten their way through the bags and will be at Level Three in those few hours. Then repeat the procedure with new plastic bags until all the ants are dead – death by asphyxiation.
CHAPTER 1

INTRODUCTION

Strange to know nothing, never to be sure
Of what is true or right or real,
But forced to qualify or so I feel,
Or Well, it does seem so:
Someone must know.
Strange to be ignorant of the way things work:
Their skill at finding what they need,
Their sense of shape, and the punctual spread of seed,
And willingness to change;
Yes, it is strange,
Even to wear such knowledge – for our flesh
Surrounds us with its own decisions –
And yet spend all our life on imprecisions,
That when we start to die
Have no idea why.

Ignorance, Philip Larkin, 1955.

One goal common to human sentence processing theories is to develop a cross-linguistically applicable account of human parsing processes. There is much empirical evidence consistent with such theories, based on experiments involving diverse languages such as English, Japanese, Korean, Dutch, and German. However, processing facts for many other languages are yet to be determined, and it is still unknown how such cross-linguistic theories fare when their predictions are applied to the other languages. Hindi\(^1\) is a particularly appropriate language for this purpose;

\(^1\)Hindi, also known as Urdu, or Hindi-Urdu, is an Indo-Aryan language spoken primarily in South Asia; it has about 424 million speakers in India (source: 1991 Census of India, www.censusindia.net),
although much is known about Hindi syntax, semantics, pragmatics, etc., currently there exists almost no experimentally grounded work on aspects of Hindi sentence processing. As discussed later in this chapter, Hindi also has certain interesting properties which make it a useful test case for evaluating current processing models.

In this research, I present a series of experiments that primarily investigate phrase-by-phrase processing difficulty associated with Hindi self-center embedding constructions (these are defined later in this chapter). The results from these experiments are used to evaluate three models of human sentence processing: Hawkins' Early Immediate Constituents (Hawkins, 1994), (Hawkins, 1998), Gibson's Discourse Locality Theory (Gibson, 2000), and Lewis' Retrieval Interference Theory (Lewis & Nakayama, 2001), (Lewis, 2002). As this dissertation will demonstrate, a rather complex and perhaps unexpected picture emerges of the cross-linguistically applicable constraints on human language.

This introduction is structured as follows. Section 1.1 defines self-center embeddings (SCEs) and places them in the context of existing research on human sentence processing. Section 1.2 discusses some relevant facts about Hindi syntax and semantics. Section 1.3 briefly discusses specificity/definiteness marking in Hindi, and Section 1.4 examines whether bare singular object nominals are indefinites. All this forms the background for the discussion in subsequent chapters.

---

and about 10 million in Pakistan (source: www.sil.org). Although Urdu and Hindi use different scripts, and even though political pressures have resulted in an increasingly divergent lexicon, the syntax is essentially indistinguishable. Accordingly, the present research is equally relevant to Urdu as to Hindi, with the caveat that a different picture may emerge of word length effects (see Chapter 6) in Urdu.

2Exceptions are (Vaid & Pandit, 1991) on sentence processing in normal and aphasic Hindi speakers, and (Vaid & Gupta, to appear) on aspects of Hindi word-recognition.

3This dissertation is not directly concerned with the garden-path sentence — sentences such as The horse raced past the barn fell — which has been the staple diet of sentence processing research in the English-speaking world. Accordingly, I will not discuss the considerable literature on this subject.
1.1 Self-center embeddings and human sentence processing

Self-center-embedding constructions (hereafter, SCEs) are grammatical structures in which a constituent occurs medially within a larger instance of the same kind of syntactic category. Examples from English are sentences like (1a), and the naturally occurring examples (1b), and (1c); see (Sampson, 2001, 15,19,18). In each example the embedded clauses are enclosed in square brackets.

(1)  a. Don’t you find [that sentences [that people you know produce] are easier to understand]?
   
   b. The odds [that your theory will be in fact right, and the general thing [that everybody’s working on] will be wrong,] are low.
   
   c. Your report today [that any Tory constituency party [failing [to deselect its MP], should he not vote in disobedience with a prime ministerial diktat,] might itself be disbanded], shows with certainty that ... an ‘elective dictatorship’ is now with us.

Contrary to a commonly held belief, SCEs occur suprisingly frequently in language (Roeck et al., 1982), particularly in head-final languages. They have been the subject of much psycholinguistic research. E.g., for Dutch, see (Dickey & Vonk, 1997), (Kaan & Vasić, 2000); for Dutch and German, see (Bach, Brown, & Marslen-Wilson, 1986); and for Japanese, see (Lewis & Nakayama, 2001), (Babylonyshev & Gibson, 1999), (Uehara & Bradley, 1996).

The reason that SCEs have attracted attention in sentence processing research is that such constructions (especially in head-final languages) necessarily overload working memory,\(^4\) which may be defined as follows.

\(^4\)Cognitive Psychology textbooks such as (Ashcraft, 1994, 145-146) consider short-term memory (STM) and working memory as referring to distinct things. For Ashcraft, STM refers to input and storage of information, whereas working memory refers more to the retrieval and use of stored items in the service of some task. This distinction is certainly useful when focusing on specific aspects of
Working memory is those mechanisms or processes that are involved in the control, regulation, and active maintenance of task-relevant information in the service of complex cognition, including novel as well as familiar, skilled tasks. It consists of a set of processes and mechanisms and is not a fixed “place” or “box” in the cognitive architecture. It is not a completely unitary system in the sense that it involves multiple representational codes and/or different subsystems. Its capacity limits reflect multiple factors and may even be an emergent property of the multiple processes and mechanisms involved. Working memory is closely linked to [long-term memory (LTM)], and its contents consist primarily of LTM representations, but can also extend to LTM memory representations that are closely linked to activated retrieval cues, hence can be quickly reactivated.

(Miyake & Shah, 1999, 450)

It is easy to see that working memory is taxed during the processing of SCEs, particularly in head-final languages. Consider the examples in (2) from Dutch, German, Japanese, and Hindi.

(2)  a. (dat) Aad Jantje de leerares de knikkers liet helpen opruimen
    that Aad Jantje the teacher the marbles let help collect
    ‘(that) Aad let Jantje help the teacher collect the marbles.’

    b. (dass) die Männer haben Hans die Pferde füttern lehren
    that the men have Hans the horses feed teach
    ‘(that) the men have taught Hans to feed the horses.’
c. Keiko-ga Tadashi-ga Kenji-o kiraida-to omotteiru
   Keiko-nom Tadashi-nom Kenji-acc hates-comp
   ‘Keiko thinks that Tadashi hates Kenji.’

d. Sitaa-ne Hari-ko kitaab khariid-neko kahaa
   Sita-erg Hari-dat book buy-inf said
   ‘Sita told Hari to buy a/the book.’

During the course of real-time processing of such sentences, certain memory processes must necessarily occur: pre-theoretically, the noun phrases (NPs) must somehow be temporarily encoded and stored in working memory until one or more verbs necessitate the NPs’ retrieval and subsequent integration with the verb(s).\(^5\)

The above observation about storage and retrieval/integration processes during sentence processing, along with the fact that working memory is resource bounded (Miller, 1956), suggests that inherent constraints on working memory are likely to be the source of processing difficulty\(^6\) (Yngve, 1960).\(^7\)

\(^5\)I am not suggesting that encoding/storage is a processing mode reserved for nouns and retrieval/integration for verbs. Encoding/storage of NPs involves retrieval/integration processes as well, but the nature of these processes is different in comparison to verbs. Similarly, verbs must also be encoded and stored. There is, however, a distinction between verb- and noun-processing: in grammatical sentences nouns are arguments or adjuncts of the verb and are, in this sense, dependent on the verb. Pre-theoretically, it follows that nouns must be integrated with the verbs in a manner that is quite distinct from the case where only nouns have been seen successively: in the latter case, any retrieval/integration processes occurring at the nouns may involve “retrieval” of predictions about sentence endings, and integration of all the nouns seen so far with existing predictions about sentence structure. For convenience, I will refer to all such noun-phrase based retrieval/integration as encoding/storage.

\(^6\)This is perhaps not so obvious. In principle, there could be an explanation for processing constraints on center embeddings that does not rely on such questions like memory constraints. Christiansen and MacDonald (1996) provide such an explanation. Theirs is a connectionist account of center embeddings, where the observed processing constraints are an emergent property of the network. This is certainly a very interesting and promising approach. However, it is not clear yet whether their results, based on strings generated from the specification of an artificial language, are able to provide fine-grained accounts of the kinds of processing difficulty I address here. Also see (Gordon, Hendrick, & Johnson, 2001, 1421-1422) for a similar discussion of the (at-present) limited nature of such connectionist accounts.

\(^7\)Cf. (Sampson, 2001), which provides a compelling alternative explanation, other than the resource boundedness of memory, for Yngve’s original account for the fact that left-branching constructions are less preferred than right-branching ones in English.
This characterization raises several questions: what exactly is stored and encoded, and how? What factors affect retrieval and integration? What has a greater impact on memory overload, retrieval or integration? The answers to these questions go beyond sentence processing; they are fundamental to understanding human attention and have wide-ranging applications in any area concerned with the consequences of cognitive overload in humans, such as aphasiology, attention-related disorders, the design of time-critical systems, multi-task performance, pedagogy, foreign language learning, etc.

Simple assumptions about human working memory, for example, that NPs are stored in a stack-like buffer and are assembled sequentially with incoming verbs (Joshi, 1990), (Rambow & Joshi, 1994), cannot be the way sentence processing proceeds; see (Vasishth, 2002a) for a detailed discussion of the empirical problems with Joshi’s embedded pushdown automaton model of sentence processing. However, research in cognitive psychology has provided a rich body of results regarding working memory that suggest answers to some of the basic questions. One factor is likely to be its general resource-bounded nature (see, e.g., (Lewis, 1996), (Cowan, 2001)).

Other plausible candidates are phonological similarity (Baddeley, 1966) (see also (Baddeley & Hitch, 1974), (Baddeley, Thompson, & Buchanan, 1975) (Gathercole & Baddeley, 1993), (Baddeley, 1992)), decay of stored items (Brown, 1958), (Peterson & Peterson, 1959) and interference (Waugh & Norman, 1965), specifically proactive interference (Müller & Pilzecker, 1900) and retroactive interference (Keppel & Underwood, 1962) among stored items.\(^8\)

\(^8\)The resource-boundedness assumption is not uncontroversial; see, for example, the commentary appended to (Cowan, 2001), and (Richardson, 1996, 122-124). It is, however, orthogonal to the present discussion whether resource boundedness per se constrains working memory, or whether it is a side-effect of some other fact about working memory or cognitive architecture.

\(^9\)In a recall task, proactive interference occurs when the retrieval of an item suffers from interference by an item or items seen earlier than the item to be retrieved. Retroactive interference is the opposite: the retrieval of an item suffers from interference from items that were seen after the to-be-recalled item.
Lewis’ Retrieval Interference Theory directly relies on research on interference (Lewis, 1996), and proposes similarity-based interference during retrieval as being the main factor affecting processing difficulty (cf. (Stabler, 1994)). Hawkins and Gibson define other constraints that are also derived (although less directly) from the psychology literature and are ultimately related to working memory. In Hawkins’ EIC theory, the main factor is the number of words the perceiver needs to see/hear in order to recognize a phrasal category: the more the number of words per phrase, the harder a sentence is to process. This translates to a complexity metric that yields specific predictions about particular sentence structures across languages; the complexity metric can be regarded as quantifying the strain on a resource-bounded working memory. By contrast, Gibson’s model (DLT) quantifies the constraints on a limited working memory in terms of (inter alia) the number of new discourse referents introduced so far.

Because it overloads working memory, the self-center embedding construction is an important test case for these working-memory based models. Existing research on center embeddings in several languages lends a great deal of support to Lewis and Gibson’s models (Hawkins’ EIC fares less well with the center embedding facts; EIC’s predictions for center embeddings and the corresponding center embedding facts are discussed in Chapter 2). However, it is an open question whether the predictions of these models are correct for other languages which have not yet been studied.

Before the experiments are presented, it is necessary to spell out the assumptions regarding the syntactic and semantic properties of Hindi self-center embeddings. This is discussed in the next section.
1.2 Self-center embeddings in Hindi

1.2.1 Some syntactic properties of self-center embeddings

An important property of Hindi self-center embeddings is that these are control constructions (Bickel & Yadava, 2000). That is, the structure of a double embedding like (3) is as shown in Figure 1.1 (single embeddings have a similar structure, but with one less level of embedding).

(3) Sita-ne Hari-ko [Ravi-ko [kitaab khariid-ne-ko] bolne-ko] kahaa
Sita-erg Hari-dat Ravi-dat book buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy a/the book.’

![Syntactic tree for Example (3)](image)

Figure 1.1: Syntactic tree for Example (3)

Here, the indirect object of a clause at a given level (matrix or embedded) obligatorily controls a PRO in subject position in the clause embedded within it. This fact becomes relevant in Lewis’ model. It is of course also possible to have obligatory subject control in such sentences, as the following example demonstrates:
(4) Siitaa-ne [PROi Hari-se kitaab le-ne-kaa] nirṇaya-kiyaa
   Sita-erg PRO Hari-ABL book take-inf decided
   ‘Sita decided to take the a/book from Hari.’

That is, the PRO in subject position can in principle be coindexed with a subject or
a non-subject in the higher clause, depending on the properties of the matrix verb.

Embedded infinitivals: nominal or verbal?

One open question relating to embedded infinitivals in self-center embeddings is
that these infinitivals may be nominal in nature and not verbs at all.10 Notice
that the embedded infinitivals have ko-marking, which is phonologically identical
to the dative/accusative case marker. In fact, Masica (1993, 321-325), and Butt
(1993), among others, have argued that ko-marked infinitivals are simply nominals
with ordinary case marking. In sum, the claim is as follows:

   “... the constituent headed by the infinitive not only has the distribution
   of an NP, it can take case markers and undergo some further morphological
   processes that only apply to NPs. The entire infinitival ‘clause’ must
   therefore be analyzed as an NP.”

   (Butt, 1993, 52)

Butt summarizes the evidence that this may be the case. First, in an infinitival,
ko can be replaced by clear postpositions, like keliye, “for”:

(5) a. Siitaa-ne Hari-ko [kitaab khariid-ne]-ko kahaa
   Sita-erg Hari-DAT book buy-inf-ko told
   ‘Sita told Hari to buy a/the book.’

   b. Siitaa-ne Hari-ko [kitaab khariid-ne]-keliye kahaa
   Sita-erg Hari-DAT book buy-inf-for told
   ‘Sita told Hari to buy a/the book.’

---

10 The remainder of this sub-section is partly based on joint research with Brian Joseph (Vasishtth & Joseph, 2002).
Second, the inflections on the infinitival are similar to nominal inflections.
Consider *laṛkaa*, “boy”, and *piinaa*, “to drink” in Table 1.1.

<table>
<thead>
<tr>
<th>Nom. Sing.</th>
<th>Oblique</th>
<th>Dat./Acc.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masc. Noun</td>
<td>laṛkaa</td>
<td>laṛke</td>
</tr>
<tr>
<td>Inf.</td>
<td>piinaa</td>
<td>piine</td>
</tr>
</tbody>
</table>

Table 1.1: Inflections on infinitivals and nominals

Finally, coordination of two NPs versus two infinitivals is identical:

(6) a. Siitaa-ne [Ravi *aur* Hari]-ko dekhaa
   Sita-erg Ravi and Hari-ko saw
   ‘Sita saw Ravi and Hari.’

   b. Siitaa-ne Hari-ko [[kitaab khariid-ne] *aur* [akhbaar becne]]-ko kahaa
      Sita-erg Hari-dat book buy-inf and newspaper sell-inf-ko told
      ‘Sita told Hari to buy the book and sell the newspaper.’

On the other hand, Mohanan (1994, 13-14), Bickel and Yadava (2000), and others suggest that such infinitivals are verbal. Although they present no arguments supporting this view, there is some justification for treating the infinitival *ko* as distinct from the case marker *ko*. First, there is some non-uniformity in the behavior of the two kinds of *ko*. Consider the ordering possibilities of the focus particle *bhii*, ‘also, even’, in (7) versus (8).

(7) **Indirect Object ko**

   a. Riitaa Siitaa-*ko-bhii* akhbaar de-gii
      Rita Sita-ko-even newspaper give-fut
      ‘Rita will give even Sita (a/the) newspaper.’

   b. Riitaa Siitaa-*bhii-ko* akhbaar de-gii
      Rita Sita-even-ko newspaper give-fut
      ‘Rita will give even Sita (a/the) newspaper.’

11
(8) **Infinitival marker ko**

a. Kalikaa-ne Siimaa-ko kitaab khariidhe-**ko-bhii** bolaa
   Kalika-erg Seema-dat book buy-inf-even told
   ‘Kalika told Seema to even buy a/the book.’

b. ?? Kalikaa-ne Siimaa-ko kitaab khariidhe-**bhii-ko** bolaa
   Kalikaa-erg Seema-dat book buy-even-ko told
   ‘Kalika told Seema to even buy a/the book.’

In (7) both orders, **bhii-ko** and **ko-bhii** are permitted, but in (8) the preferred order is **ko-bhii**. This suggests that once the **ko** is affixed to the infinitival, it does not have the same reordering properties with elements like **bhii** as a noun. Another example of this asymmetry is with the particle **hii**, ‘only’:

(9) **Indirect Object**

a. Riitäa Siitaa-**ko-hii** akhbaar de-gii
   Rita Sita-ko-only newspaper give-fut
   ‘Rita will give only Sita (a/the) newspaper.’

b. Riitäa Siitaa-**hii-ko** akhbaar de-gii
   Rita Sita-only-ko newspaper give-fut
   ‘Rita will give only Sita (a/the) newspaper.’

(10) **Infinitival marking**

a. Kalikaa-ne Siimaa-ko kitaab khariidhe-**ko-hii** bolaa
   Kalika-erg Seema-dat book buy-inf-only told
   ‘Kalika told Seema to only buy a/the book.’

b. ?? Kalikaa-ne Siimaa-ko kitaab khariidhe-**hii-ko** bolaa
   Kalika-erg Seema-dat book buy-only-ko told
   ‘Kalika told Seema to only buy a/the book.’

---

11These facts are based on my own acceptability judgements, and are therefore open to empirical investigation. The two question marks before (8b) merely indicate that the sentence sounds very odd, if not outright ungrammatical.
Second, (Masica, 1993, 232-235) develops a functional, historically derived typology of affixes in Indo-Aryan languages. Under this typology, the Hindi dative marker *ko* is a Layer II affix, whereas the infinitival marker *ko* is a Layer III affix. The definitions of Layers I, II and III are as follows (Masica, 1993, 232).

**Layer I:** Affixes inherited from Old and Middle Indo-Aryan that attach directly to the base, with morphophonemic adjustments which are occasionally complex. An example from Hindi is the oblique marker *e* in *laḍke* (see Table 1.1).

**Layer II:** Affixes that are (a) attached to the base indirectly, through the mediation of a Layer I element; and/or (b) invariant for all nouns and the same for both numbers. Morphophonemic variation occurs in Layer II as well, but the situation is simpler than in Layer I (Masica, 1993, 232). A Hindi example is the dative marker *ko* in *laḍke-ko* (see Table 1.1).

**Layer III:** Affixes that (a) are potentially mediated by a Layer II element, most often a genitive (e.g., Hindi *laḍk-e ke saath*, boy-oblique gen with, ‘with the boy’); (b) lack morphophonemic variants, and may be longer than one syllable, and usually retains a fairly transparent connection with an independent word, of which it is generally a Layer I case form; and (c) are semantically more specific.

Distributionally, the *ko* marker in infinitivals patterns with clear Layer III purpose clause markers like *keliye* in (5b): it is mediated by a Layer II element (*-e*), lacks morphophonemic variants, and (most importantly) its meaning as an infinitival suffix corresponds closely to the purpose clause marker *keliye*. In other words, instead of the argument that infinitival plus *ko* is nominal because *ko* can be replaced by clear postpositions, one could equally say that this is an argument *against* treating the infinitival as a nominal, since the clear postposition *keliye* is functionally (and historically) a very different element compared to the Level II dative marker
ko, and the infinitival ko distributionally behaves like this Layer III suffix. In fact, Experiments 3 and 7 suggest that the infinitivals are actually verbal in nature. This is discussed in detail on pages 105 and 155.

However, the question of whether the infinitival is verbal or nominal is predicated on the assumption that entities must necessarily have a uniform set of properties. This assumption is merely a reductionist convenience, and is certainly a useful means for modeling complex facts in cases where no serious empirical consequences follow from oversimplification. However, the morpheme ko need not be any one mutually exclusive entity; it could be a nominal affix in a certain context, and a purpose clause marker in another context. Such entities, which display unity-in-diversity, are well-documented cross-linguistically and have been termed CONSTELLATIONS (Janda & Joseph, 1999). The constellation view is articulated in terms of two constructs:

(11) a. **THE CONSTELLATION:** A group of elements which share at least one characteristic property of form but are distinguished by individual idiosyncrasies — of both form and function — that prevent their being collapsed with one another.

b. **META-TEMPLATE:** A meta-level redundancy statement, which ranges over all relevant candidates and equates instances of a particular formal configuration that meet certain criteria of uniting properties.

Diversity is characterized here by the Constellation, and unity by the Meta-template. This approach is not merely of meta-theoretical relevance, but has real empirical consequences. For example, in the above misleading debate about the nature of infinitivals, one might conclude that infinitivals are nominals based on Butt’s and others’ evidence. But this leaves unexplained the results of Experiment 3 and 7, as discussed further on.
A final point to note about Hindi is that in general word order is fairly free (although it is of course subject to certain constraints; see, e.g., (Gambhir, 1981), (Mahajan, 1990), (Kidwai, 2000), (Vasishth, 2002b)). As Experiments 4-6 will demonstrate, the important point for present purposes is that the indirect object or the direct object in example (3) on page 3 can be fronted without rendering the sentence ungrammatical.

Another interesting semantic/pragmatic property of Hindi is differential object marking; this property is exploited in Experiments 1-6, and is discussed next.

1.2.2 Semantics/pragmatics: Differential Object Marking in Hindi

Differential object marking (DOM) refers to the presence of case marking on objects, based on certain pragmatic and/or semantic constraints (Comrie, 1979), (Bosson, 1985, 1991), (Aissen, 2000). Many languages have this property, and Hindi is one of them, as example (12) shows:

(12)  a. Sita kîtaab paḍh-rahii hai
    Sita book reading is
    ‘Sita is reading a/the book.’

   b. Sita kîtaab-ko paḍh-rahii hai
    Sita book-acc reading is
    ‘Sita is reading *a/the book.’

In (12a) the bare direct object can be construed as specific/definite or indefinite, depending on the preceding context. However, with ko marking (12b) the object can only be specific/definite.12

Aissen, following the existing literature on typology (see, e.g., (Comrie, 1979, 1989)), characterizes the cross-linguistically observed constraints on DOM as being

---

12I will gloss the specific/definite marker ko as acc(usative), but this is only for convenience; nothing hinges on considering it an accusative marker, rather than a dative case marker (Masica, 1993, 365), which is what the specific/definite-marker -ko is etymologically. Henceforth, for convenience, I will refer to the specific/definite-marker -ko as the definite marker -ko.
a function of markedness of objects.\textsuperscript{13} Markedness can be defined as follows: Given the well-known (see, e.g., (Croft, 1988), (Comrie, 1989)) relational, animacy, and definiteness scales shown in (13), a noun phrase that is a subject or object is marked if its position on the relational scale (high or low, assuming the ordering implied by “>”) is not identical with (or aligned with) its position on the animacy and/or definiteness scales. Exactly which scales are involved is language-dependent. Henceforth, I will call an object that is marked in this sense a MARKED OBJECT; I distinguish this from DEFINITE- or CASE-MARKED OBJECT, which refers to the object along with the relevant morphological marking.

(13) \begin{itemize}
  \item[a.] \textbf{Relational Scale:} Subject > Object
  \item[b.] \textbf{Animacy Scale:} Human > Animate > Inanimate
  \item[c.] \textbf{Definiteness Scale:}
    Pronoun > Name > Definite > Indefinite Specific > NonSpecific
\end{itemize}

For example, a human direct object is marked compared to an inanimate direct object. In Hindi, this correlates with obligatory case marking on human objects, whereas inanimate objects have optional case marking, as shown in (14):

(14) \begin{itemize}
  \item[a.] Sita\textsuperscript{a} Ravi\textsuperscript{b}(-ko) piti\textsuperscript{c}-rahii hai
    \begin{align*}
    \text{Sita} & \quad \text{Ravi-acc} \quad \text{beating is} \\
    & \quad \text{‘Sita is beating up Ravi.’}
    \end{align*}
  \item[b.] Sita\textsuperscript{a} kitaab\textsuperscript{b}(-ko) pa\textsuperscript{c}dh-rahii hai
    \begin{align*}
    \text{Sita} & \quad \text{book-acc} \quad \text{reading is} \\
    & \quad \text{‘Sita is reading a/the book.’}
    \end{align*}
\end{itemize}

\textsuperscript{13}Aissen casts her presentation in terms of harmonic alignment (Prince & Smolensky, 1993), from Optimality Theory, but nothing hinges on using that approach for the present purposes; accordingly, I present the relevant insights due to the typology literature as descriptive generalizations.
Similarly, a definite human direct object (e.g., one referred to by a proper name) is more marked than a (possibly) indefinite human direct object; in Hindi this correlates with obligatory case marking for definite human direct objects, but optional case marking for indefinite human direct objects:\footnote{\cite{mohan1995correlates} regards the non-case marked version of the human object in (15b) as being incorporated with the verb, but see \cite{wesc02} for the irreconcilable problems associated with treating these as simply a case of noun incorporation.}

\begin{enumerate}
  \item Sitaa Ravi\(^*\)(-ko) piit-rahii hai
      Sita Ravi-acc beating is
      ‘Sita is beating up Ravi.’
  \item Sitaa bacca/bacce(-ko) sambhaal-rahii hai
      Sita child(-acc) looking-after is
      ‘Sita is looking after a child/the child(ren).’
\end{enumerate}

To give another example, among indefinite inanimate direct objects, specific objects are correctly predicted to be more marked than nonspecific ones: in Hindi, such specific objects can be case marked, but the nonspecific ones are not. Consider the three sentences in (16). Example (16a) is a possible preceding-context sentence for (16b,c). The important contrast is between (16b) and (16c). The bare direct object in (16b) is construed as specific/definite or indefinite depending on the context: if there is a context like (16a), the direct object is construed as specific/definite, but if there is no preceding context the direct object is construed as indefinite. By contrast, the case-marked direct object in (16c) is necessarily construed as specific/definite (or marked) irrespective of preceding context.

\begin{enumerate}
  \item mez par ek kitaab pa\d{i}thi
      table on one book lying was
      ‘A book was lying on the table.’
  \item Sitaa-ne kitaab \uthaayii
      Sita-erg book picked-up
      ‘Sita picked up a/the book.’
\end{enumerate}
c. Sītāa-ne kitaab-ko uṭhaayaa
   Sīta-erg book-ko picked-up
   ‘Sita picked up the book.’

In other words, if (16a) is the utterance preceding (16b), the bare object NP is
cstrued as specific/definite, but without (16a) as preceding context, *kitaab* in (16b)
is construed as indefinite (there is a principled reason for this, as discussed below).
By contrast, if (16a) precedes (16c), the *ko* marking occurs optionally with *kitaab*
in (16c), but if (16c) is uttered out of context, the *ko*-marked NP *kitaab* must be
cstrued as specific/definite.

Thus, the notion of markedness leads to the generalization that, cross-
linguistically, marked objects tend to be case marked, and the greater the degree of
markedness the greater the tendency for obligatory case marking. This generalization
can, in a sense, be considered an explanation. However, if there were a principled
account of why inanimate, indefinite direct objects and human, definite subjects are
canonical, that would be the real explanation for the observed canonicity of subjects
and objects.

In the remainder of this sub-section, I primarily address this question: Why do
marked objects need case marking? This issue significantly affects the interpretation
of the experiment results in this dissertation, as discussed in later chapters, but
for the moment I would like to consider the question for its own sake. I begin
by summarizing the various explanations for DOM, and then present a preliminary
corpus investigation that evaluates some of the claims made in the literature.
Why do marked objects need to be case marked?

Comrie (1975, 1979) offers the explanation, echoed in (Bossong, 1991), that case marking may serve to distinguish subjects from objects. The more marked an object, the less distinguishable it is from a subject and, presumably in order to facilitate comprehension, there has to be some way to distinguish the subject from the object. Case marking serves this purpose.

Comrie’s confusability explanation is debatable and has been challenged in the literature. For example, Hopper and Thompson (1980, 291) argue that definite/referential direct objects NPs in fact constitute the unmarked or ‘normal’ case, and that “the tendency to mark just definite/animate [objects] reflects purer object-ness of such [objects], and simultaneously marks the higher Transitivity of the clause as a whole” (Hopper & Thompson, 1980, 291). In order to understand Hopper and Thompson’s argument, it is necessary to briefly examine three key concepts in their theory: (a) Transitivity; (b) Individuation; and (c) Grounding. I will not present the details of Hopper and Thompson’s general theory here, but only the essential ideas that are relevant to the present discussion.

Transitivity is a scale, just like the definiteness and animacy scales, along which a sentence as a whole can be characterized as more (or less) transitive (Hopper & Thompson, 1980, 252-253). The second component is Individuation. This is defined by Hopper and Thompson as the extent to which a patient is distinct from an agent, and the extent to which it is distinct from its own background (defined below). Specifically, an NP is highly individuated if it is proper, human/animate, concrete, singular, count, and referential/definite. An NP is non-individuated to the extent that it has the properties of being common, inanimate, abstract, plural, mass, and non-referential (in other words, “individuated” and “non-individuated” are not boolean categories). Individuation is a key factor which indicates that an NP is a canonical object; Comrie’s notion of confusability is assumed to be a secondary factor, and
dependent upon the degree of individuation. Finally, they posit FOREGROUND and BACKGROUND as primitive theoretical notions that determine (among other things) what counts as a canonical direct object. These notions are intended to characterize regularities in narratives as opposed to isolated sentences, but presumably apply even in the case of single sentences. Background is “that part of a discourse which does not immediately and crucially contribute to the speaker’s goal, but which merely assists, amplifies, or comments on it”; foreground is “the material which supplies the main points of the discourse” (Hopper & Thompson, 1980, 280). The relevance of all this for markedness in direct objects is that, if a sentence is foregrounded, “there is a marked tendency for [objects] to be individuated, i.e., to have properties associated with referentiality/definiteness and animacy” (Hopper & Thompson, 1980, 291). Furthermore, as mentioned above, Hopper and Thompson explain case marking on direct objects as follows: “the tendency to mark just definite/animate [objects] reflects purer object-ness of such [objects], and simultaneously marks the higher Transitivity of the clause as a whole” (Hopper & Thompson, 1980, 291).\(^\text{15}\)

Hopper and Thompson’s conclusion is intended to provide an argument against Givón’s observation (1978a) that if there is an indefinite NP in the sentence, it is more likely to be the object than the subject, from which it follows that objects tend to be indefinite compared to subjects. Givón’s argument was that, for discourse reasons, “the accusative position is the ‘most indefinite’ of all major arguments of the verb, in contrast with the subject and dative which are overwhelmingly definite” (Givón, 1978a, 306) (also see (Givón, 1978b, 72-73)). Although Givón does not specify what exactly these discourse reasons are, one plausible reason is the need for nouns to conform to the grammatical-relation scales discussed above. That is, direct objects are canonically inanimate indefinites, whereas subjects are canonically human/animate definites. Marked (specific/definite) objects are case marked in order

\(^{15}\)They do not explain what “purer” might mean in this context.
to mark their noncanonicity. Givón’s formulation turns out to be important for the experiment results presented in subsequent chapters; therefore, it is important to look at the evidence for his claims versus Hopper and Thompson’s.

His observation that direct object positions typically introduce new referents is motivated by corpus evidence; he conducted a hand count of short texts, such as the front page news of *Los Angeles Times* (Givón, 1978b, 72-73). His counts showed that (a) “on the average about 50 percent of accusative object nouns are INDEFINITE,” and (b) English indefinites are more likely to occur in object position than in subject position. Some of his key results are summarized in Table 1.2.

<table>
<thead>
<tr>
<th>Determiners</th>
<th>%age in subject position</th>
<th>%age in object position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indefinites</td>
<td>9%</td>
<td>44%</td>
</tr>
<tr>
<td>Definites</td>
<td>91%</td>
<td>56%</td>
</tr>
</tbody>
</table>

Table 1.2: The percentage (of all occurrences) of indefinites versus definites occurring as subjects and as objects. Note here that Givón’s claim is not that indefinites occur more frequently than definites in object position, but that “the accusative object position in English is a prime environment in which referential nouns are introduced into the discourse for the first time, as indefinites.”

However, Hopper and Thompson cast doubts on Givón’s corpus results: “Givón’s statistics may be questionable, since these statistics did not consider the crucially important distinction between foregrounding and backgrounding” (Hopper & Thompson, 1980, 291). Although Hopper and Thompson may be correct that definites, rather than indefinites, are more likely to be objects, they do not actually address any details of Givón’s corpus data. Consequently, it is difficult to decide whether the argument against Givón’s account has any empirical basis. In any case, for Hopper and Thompson, the important point is that there is no evidence that ‘normal’ objects are indefinite, and that objects (in foregrounded contexts) are canonically definite.
Hopper and Thompson’s arguments against Givón’s have been assumed in the literature to mean that “the view that objects are typically indefinite is mistaken” (Dahl, 1997), (Schwenter & Silva, 2002). However, I argue below that this conclusion is not necessarily valid.

In order to test Hopper and Thompson’s conclusion that objects are typically definites, I investigated a large (139,316,701 tokens, 939,082 types) part-of-speech tagged English text corpus (North American News, available at the Linguistic Data Consortium, www.ldc.upenn.edu). The results of the search appear to bear out Hopper and Thompson’ claim that ‘normal’ objects are definites (Table 1.3). However, this is not a conclusive result, since in the corpus definites in general occur twice as often as indefinites (Table 1.4). If we normalize the counts so that the total number of indefinites and definites occurring in the corpus are assumed to be equal, the number of definite objects would be only 770,012 (and not 1,551,752). That is, indefinites would occur 60% of the time rather than 43% of the time, as is actually the case (Table 1.3). This suggests that if one takes into account the fact that definites in general occur more often than indefinites in the corpus, indefinites are the ‘norm’ in object position, contra Hopper and Thompson’s claims.\(^{16}\)

<table>
<thead>
<tr>
<th>Determiners</th>
<th>No. post-verbally</th>
<th>%age of Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indefinites: a, an, some, any</td>
<td>1,191,257</td>
<td>43%</td>
</tr>
<tr>
<td>Definites: the, this, those, these</td>
<td>1,551,752</td>
<td>57%</td>
</tr>
<tr>
<td>Total:</td>
<td>2,743,009</td>
<td></td>
</tr>
</tbody>
</table>

Table 1.3: The number of indefinite versus definite determiners occurring immediately post-verbally, and the percentages of each type. This table is consistent with Hopper and Thompson’s claim that definites are the ‘norm’ for direct objects, for the reasons discussed in the text.

In addition, there is an important point that has become obscured in Hopper and Thompson’s critique of Comrie and Givón. Their claim is that direct objects occur

\(^{16}\)I show below that this result is not necessarily inconsistent with their main argument.
Table 1.4: The total number of indefinite versus definite determiners occurring in the corpus. There are nearly twice as many definites as indefinites. This shows that Table 1.3 does not necessarily support Hopper and Thompson’s claim: the larger number of definites immediately post-verbally might be simply because there are more definites overall, and not for the reasons Hopper and Thompson provide.

as definites more often than indefinites, whereas Givón’s claim was that indefinites occur more often in direct object position than in subject position. Although Hopper and Thompson are aware of this difference in claims (Hopper & Thompson, 1980, 291), they speculate that even Givón’s claim may not hold up once their notions of foregrounding and backgrounding are considered.

However, apart from the fact that the greater number of definites than indefinites overall (Table 1.4) renders their hypothesis difficult to test,\(^\text{17}\) Givón’s claims may in fact be consistent with Hopper and Thompson’s: the relational, definiteness, and animacy scales discussed earlier may also impose their own canonicity to subjects and objects, independent of Hopper and Thompson’s parameters, such as the Transitivity scale, foregrounding etc. In this connection, notice also that Hopper and Thompson claim that definite objects will be the norm in \textit{foregrounded} clauses; but they never indicate whether foregrounded clauses dominate in language, and if so, to what extent. In the extreme case, foregrounded clauses may be in such a small minority in a text corpus that the dominating background clauses may result in indefinite objects occurring more frequently overall than definite objects. Thus, Hopper and Thompson, as well as Givón, could both be right about their various claims. Accordingly, in the

\(^{17}\)The normalization discussed above is not a conclusive argument, but merely points out the confound in Hopper and Thompson’s claims.
remainder of this section, I explore Givón’s hypothesis using the same corpus (North American News) that was used for investigating Hopper and Thompson’s claims.

A corpus investigation of Givón’s hypothesis

Givón’s hypothesis was that indefinites occur more often in direct object position than in subjects. I tested this hypothesis using the following heuristic. For estimating the rate of occurrence of definites and indefinites in subject position, I used the percentage of definite (indefinite) determiners in sentence-initial position, with respect to the total number of definites (respectively, indefinites) in the corpus. Similarly, the percentage immediately following a verb was used as a rough estimate of the rate of occurrence of indefinites and definites as direct objects.

Clearly, the above heuristic does not guarantee that all subjects and direct objects will be found. In order to quantify the accuracy of the heuristic in finding subjects and direct objects, the following sampling procedure was followed. A native speaker of English checked 100 example sentences in the output and found that in ten of these, the indefinite a in sentence-initial position found by the heuristic was an adverb such as A few hours later. Seven other sentences in this set of 100 sentences began as shown in (17) below.

(17) A mother of two, Kennedy-Powell is a member of a family that is to the law profession what the Wallendas are to trapezes.

Among these seven sentences, the subject (which was non-sentence-initial) was actually a definite noun. That is, the heuristic used was overestimating the number of indefinites in subject position. In sum, 83 clear indefinite subjects were found by the heuristic in the 100 sample sentences. For the definite determiner the, in another sample of 100 sentences, 23 were not subjects of the sentence. Thus, the heuristic found 77 clear definite subjects in the 100-sentence sample.
In the immediately post-verbal position, from a new sample of 100 sentences, the indefinite determiner *a* appeared in 98 sentences as part of a direct object; the other two sentences were instances of the string *if a ...*, where *if* had been tagged as a verb. For the determiner *the*, out of a sample of 100 sentences, the distribution was as follows. In two sentences the NP was a subject of an embedded clause ... *but also said the testimony Thursday will force Simpson’s camp to produce an explanation ...*; four sentences involved phrases like *at the moment*, where *at* was tagged as a verb; there was one instance of *if the* and one of *and the*, where both *if* and *and* were tagged as verbs; and two other sentences involved a noun being tagged as a verb, e.g., the noun *services* was tagged as a verb in ... *from each of the major uniformed services: the army, navy, air force, marines and coast guard*. Thus, among the 100 sample sentences, 98 indefinite direct objects were found correctly by the heuristic. For definites, in the 100 sample sentences 90 were actual direct objects.

These results, summarized in Table 1.5, demonstrate the limitations of the heuristic I have used, and suggest further refinements, which I leave for future research.

<table>
<thead>
<tr>
<th>Determiner</th>
<th>Sentence-initial subjects</th>
<th>Post-verbal direct objects</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>a</em></td>
<td>83%</td>
<td>98%</td>
</tr>
<tr>
<td><em>the</em></td>
<td>77%</td>
<td>90%</td>
</tr>
</tbody>
</table>

Table 1.5: Percentages of subjects and direct objects correctly found by heuristic from samples of 100 sentences of each type

The results of the search are summarized in Figures 1.2, 1.3, and Table 1.6. As Figure 1.2 shows, on average a larger percentage of indefinites than definites appear immediately following the verb, and on average a larger percentage of definites than indefinites appear in sentence-initial position. The percentages differentiated by

---

18In this simple search, I did not include the word *that*, which occurs 1,191,645 times in the corpus, since it is often mistagged as a determiner when it is a complementizer.
determiner are shown in Figure 1.3; here, it may appear that Givón’s hypothesis is wrong, since a larger percentage of each definite appears in direct object position than in sentence-initial position. However, this is probably a result of (a) the search being constrained to sentence-initial position rather than subject position (as discussed earlier, the heuristic does not successfully identify all subjects); and (b) the fact that definites in general occur approximately twice as often as indefinites (see Table 1.4).

Figure 1.2: Mean percentage of indefinite and definite determiners appearing sentence-initially and immediately following a verb (North American News corpus). Also see Table 1.6.
Figure 1.3: Percentage of determiners appearing sentence-initially and immediately following a verb (North American News corpus). Also see Table 1.6.
<table>
<thead>
<tr>
<th>Determiner</th>
<th>Total</th>
<th>Post-verbally</th>
<th>%age post-verbally</th>
<th>Sentence-initially</th>
<th>%age as sentence initial</th>
</tr>
</thead>
<tbody>
<tr>
<td>a, an</td>
<td>3,050,362</td>
<td>1,100,924</td>
<td>36%</td>
<td>133,365</td>
<td>4%</td>
</tr>
<tr>
<td>some</td>
<td>167,518</td>
<td>52,566</td>
<td>31%</td>
<td>25,730</td>
<td>15%</td>
</tr>
<tr>
<td>any</td>
<td>93,778</td>
<td>37,767</td>
<td>40%</td>
<td>2,104</td>
<td>2%</td>
</tr>
<tr>
<td>the</td>
<td>6,612,389</td>
<td>1,446,878</td>
<td>22%</td>
<td>677,805</td>
<td>10%</td>
</tr>
<tr>
<td>this</td>
<td>336,272</td>
<td>75,848</td>
<td>23%</td>
<td>54,001</td>
<td>16%</td>
</tr>
<tr>
<td>those</td>
<td>86,258</td>
<td>15,922</td>
<td>19%</td>
<td>8,506</td>
<td>10%</td>
</tr>
<tr>
<td>these</td>
<td>64,510</td>
<td>13,104</td>
<td>20%</td>
<td>10,213</td>
<td>16%</td>
</tr>
<tr>
<td><strong>Mean %age:</strong> 36%</td>
<td><strong>Mean %age:</strong> 7%</td>
<td><strong>Mean %age:</strong> 21%</td>
<td><strong>Mean %age:</strong> 13%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1.6: Total numbers of indefinites and definites (second column) in the North American News corpus; the number (third column) and percentage (fourth column) of indefinites/definites occurring immediately after a verb; and the number (fifth column) and percentage (sixth column) appearing sentence-initially.
Summary of DOM discussion

To summarize the discussion on DOM and its causes, Comrie argued that case marking of direct objects occurs in order to distinguish objects from subjects. Hopper and Thompson maintain that Comrie’s observation falls out of a more general phenomenon, Transitivity: in foregrounded contexts, direct objects are canonically \emph{definite}, not indefinite, as claimed by Givón. I have argued that Givón’s and Hopper and Thompson’s arguments may be mutually compatible, and that preliminary corpus evidence suggests that Givón’s generalization could in fact be correct: indefinites may occur more frequently (when considered as a percentage of the overall occurrence of indefinites) in direct object position than in subject position. Moreover, indefinites may also occur more frequently than definites in direct object position. This would not necessarily falsify Hopper and Thompson’s arguments, since they restrict their claims about definite objects to foregrounded clauses, and it is currently unknown whether foregrounded clauses dominate in discourse or not. If they do not dominate, Hopper and Thompson’s arguments could be valid even if indefinites occur more frequently than definites in direct object position. This is an empirical question which awaits further research, but I hope to have demonstrated that Givón’s hypothesis need not be abandoned, and that we need not give up the generalization (as suggested in recent work (Dahl, 1997), (Schwenter & Silva, 2002)) that indefinite direct objects are ‘canonical’ or ‘normal’.
1.3 Case-marked objects are specific, but conversationally implicate definiteness

Interpreting an NP in Hindi as specific/definite is mainly a function of discourse conditions, not of the formal shape of the NP. This is clear from the fact that case marking is optional for specific/definite NPs: in (12a) the direct object is indefinite if the sentence appears out of context, but can be specific/definite when there is an appropriate referent in a preceding sentence in the discourse.

The existing literature on direct object case marking occasionally refers to it as definiteness marking (Masica, 1986), (Hopper & Thompson, 1980) and occasionally as specificity marking (Singh, 1994). I show below that Hindi case-marked objects are specificity markers, not definiteness markers; definiteness is conversationally implicated in case-marked NPs which have no indefinite determiner.

It is clear that -ko is used optionally to mark specificity and not definiteness, since ko-marking is possible with indefinites:

(18) Ravi-ne (kisii) ek kitaab-ko utthayaa
    Ravi-erg (some) one book-acc picked-up
    ‘Ravi picked up a (certain) book (indefinite specific).’

If -ko were marking definiteness, then it would be odd for it to occur with indefinite determiners such as ek. However, in the absence of the indefinite determiner, -ko marking on the object results in an implicature that the NP is definite. To see this, consider (19a). Presented out of context, the case-marked object is construed as definite. However, if the sentence were continued as in (19b), the case marked object is now construed as indefinite. That is, in (19a) there is a conversational implicature that the NP is definite, and this implicature is cancelled when more information becomes available, as in (19b).
(19)  
   a. kitaab-ko uṭhaaoo
      book-acc pick-up
      ‘Pick up the book.’

   b. kitaab-ko uṭhaaoo, kisii-bhii ek kitaab-ko
      book-acc pick-up any one book-acc
      ‘(Pick up a book,) any one book.’

Such a cancellation would be impossible if -ko were a marker of definiteness: contrast this situation with the English example below.

(20)  
      Pick up the book, #just any (one) book.

Here, the definiteness implied by the first NP cannot be cancelled.

1.4  Bare singular NPs really are indefinites

The properties of unmarked object NPs are considered next. Throughout this dissertation I will treat the bare singular direct object NP as an indefinite. This view goes against (Dayal, 1999), which argues that they are not simply indefinites, but are rather kind-denoting terms. I show in this section that there is no evidence for the more complex picture of the bare singular NP as presented by Dayal.

Dayal points out that Hindi bare NPs have three possible readings: generic, definite, and indefinite (Verma, 1971), (Porterfield & Srivastava, 1988), (Dayal, 1992), (Mohanand, 1994), (Dayal, 1999), as the following examples from (Dayal, 1992) show:

(21)  
   a. kutta aam jaanvar hai
       Dog common animal is
       ‘The dog is a common animal.’

   b. Ravi ek ladkii-se milaa. ladkii bahut acchi thi
       Ravi one girl-with met. girl very nice was
       ‘Ravi met a girl. The girl was very nice.’

   c. Anu kitaab padh-rahii hai
       Anu book reading is
       ‘Anu is reading a book/books.’
Dayal (1992) has argued that bare NPs are ambiguous between definite and kind-denoting terms. Any existential readings associated with the bare NPs are dependent on their kind reference; thus, there is no independent indefinite reading of the bare singular.

I show below that the evidence for such a view is not convincing, and that the bare NP does in fact have an independent indefinite reading. A wider consequence of this result is that Dayal’s Carlsonian treatment of bare NPs as uniformly kind-denoting (Carlson, 1977) is no longer well-motivated, since the evidence from Hindi is completely consistent with the simpler Discourse Representation Theory view (Kratzer, 1995) according to which bare NPs are treated uniformly as indefinites. I first present the four arguments that Dayal provides for her conclusion, and then I present alternative explanations for Dayal’s data. If these explanations are valid, there is no need to abandon the view that bare NPs are in fact indefinites.

1.4.1 Dayal’s arguments against treating bare singular NPs as indefinites

Dayal presents four pieces of evidence from Hindi favoring the Carlsonian view of bare NPs. First, the indefinite reading is available when the NP is an object, but not when it is a subject:

\[(22) \quad \text{a. Anu kitaab pa\dh-rahii hai} \\
\text{'Anu is reading a book.'} \]

\[\text{b. la\dki\i kha\dli thii} \\
\text{'A girl was standing.'} \]

Second, sentential negation blocks the wide scope reading of the bare NP, whereas this reading is available with NPs which have determiners like \(ek\), ‘one’/‘a’, and \(koi\), ‘some’:
(23) a. Anu kitaab nahī paḍhe-gii
   Anu book not read-will
   ‘Anu won’t read any book.’/*There is a book Anu won’t read.’

   b. Anu ek/koi kitaab nahī paḍhe-gii
   Anu one/some book not read-will
   ‘Anu won’t read any book.’/*There is a book Anu won’t read.’

Third, the bare NP takes scope under adverbs, whereas the indefinite-marked NP can take wider scope (resulting in the pragmatically odd reading in (24b)).

(24) a. Anu puure din macchlii pakaḍtii-rahii
   Anu all day fish catching-was
   ‘Anu kept catching fish all day.’

   b. Anu puure din ek macchlii pakaḍtii-rahii
   Anu all day fish catching-was
   ‘Anu kept catching a fish all day.’

Finally, according to Dayal, anaphoric reference is not allowed with bare NPs:

(25) Ravi kitaab, paḍh-rahaa hai. voṣi bahnut acchii hai.
   Ravi book reading is. it very good is
   ‘Ravi is reading a/the book. It is very good.’

It is not clear that any of these arguments rule out bare NPs as having non-derived indefinite interpretations. This is discussed in the next subsection.

1.4.2 Alternative explanations for the Hindi facts

Regarding the first contrast, although it is true that bare NPs in subject position do not allow an indefinite reading, this appears to be not due to their being subjects, but rather due to their being in the sentence-initial position. In (26a), an adjunct in sentence-initial position allows the indefinite reading, and in (26b), the subject of an embedded clause, gaḍlii, ‘car’, can similarly have either the definite or indefinite reading. Regarding (26a,b), note also that for the first bare NP in each case (sadalak, ‘street’, and dukaan, ‘shop’ respectively), the most accessible reading is as a definite.
(26) a. sadak par ladkii khaḍii thii
   street on girl standing was
   ‘A/the girl was standing on *a/the street.’

   b. jab mai dukaaں-se baahar-niklaa, to mai-ne dekhaa ki dukaa
   when I shop-from came-out, then me-er saw that shop
   ke-saanme gaaḍii khaḍii thii
   of-in-front car standing was
   ‘When I came out of *a/the shop, I saw a/the car standing in outside.’

This suggests that in the absence of any preceding context the first bare noun
phrase in a sentence tends to be construed as definite, and that this should be true
no matter whether the NP is a subject or not. Now, it is often observed (see, e.g.,
(Mohanan, 1994)) that fronting a bare NP object also results in an unambiguously
definite reading:

(27) a. anu gaaḍii laaye-gii
   anu car bring-will
   ‘Anu will bring a/the car.’

   b. gaaḍii anu laaye-gii
   car anu bring-will
   ‘Anu will bring the car.’

But this is not an entirely correct generalization because, as in the case of subject
bare NPs, placing an adjunct in sentence-initial position results in a possible indefinite
reading of the fronted bare NP, and the adjunct’s NP is again only interpretable as
definite:

(28) shaadi-mें gaaḍii Anu laaye-gii
    wedding-in car Anu bring-will
    ‘Anu will bring a/the car to *a/the wedding.’

Regarding the claim that negation blocks the wide scope reading of the bare NP,
there is some other property of the bare singular NP, other than its co-occurrence
with negation, that is responsible for blocking this wide scope: the Hindi equivalent
of the ‘There is a . . . ’ construction (29a) doesn’t allow bare NPs anyway.
(29)  a. ek kitaab hai jo Anu paḍhe-gii
   a book is that Anu read-will
   ‘There is a book that Anu will read.’

   b. * kitaab hai jo Anu paḍhe-gii
      book is that Anu read-will
      ‘*There is a book that Anu will read.’

This general prohibition on bare NPs may be due to the fact that sentence-initial bare NPs always have a definite interpretation.

Moreover, the above constrast appears to have nothing to do with negation, since the same effect is seen in the absence of negation:

(30)  a. Anu kitaab paḍhe-gii
      Anu book read-will
      ‘Anu will read a/the book.’/*There is a book Anu will read.’

      b. Anu ek/koi kitaab paḍhe-gii
         Anu one/some book read-will
         ‘Anu will read a/some/*the book.’/*There is a book Anu will read.’

Turning next to the scope-restricting properties of adverbs, Dayal’s example (24) does not appear to be generally true. It is likely that the phrase ‘fish-catching’, macchīī pakaṁnaa, in (24a) is lexicalized towards a preferentially generic reading, due to the fact that catching fish is a common pastime or hobby. Contrast this with the case where the bare NP is a less common activity, such as ship-repairing, kishtīī ṭhiik karnaa. Now, the bare NP still admits a generic reading, but the indefinite reading is also available:

(31)  a. Anu puure din kishtīī ṭhiik-kartii-rahii
      Anu all day boat repairing-kept
      ‘Anu kept fixing a/the boat (or boats) all day.’

      b. Anu puure din ek kishtīī ṭhiik-kartii-rahii
         Anu all day one boat repairing-kept
         ‘Anu kept fixing a boat all day.’
Finally, regarding the constraint on anaphoric reference (25), there is a more plausible explanation for why the pronoun cannot refer to the bare NP direct object, and this has nothing to do with the properties of the bare NP per se. Prasad and Strube (2000), working in the Centering Theory framework, have shown through corpus research that in Hindi the grammatical function of NPs determines anaphoric reference: pronouns tend to refer to NPs with grammatical functions that are higher in the hierarchy shown in (32).

(32) subject > direct object > indirect object

Prasad and Strube investigated an annotated corpus which had 149 utterance pairs of sentences, where each pair had an initial sentence containing a subject and a direct object, and a following sentence containing a pronoun. They found that in 144 of the 149 cases (i.e., 96% of the time), the subject was realized as the pronoun in the second utterance. Similarly, of 22 pairs of sentences in which the initial sentence had a direct object and an indirect object, all the pronouns referred to the direct object.

This suggests that the constraint on anaphoric reference in (25) is due to a grammaticalized preference for subject anaphora over object anaphora. When a direct object bare NP and an indirect object can in principle be coindexed with an anaphor, the direct object is the appropriate antecedent for the pronoun.

(33) Anu Hari-ko kitaab, dikhaa-rahii hai. yaha, Anu-kaa sabse-puraanaa
Anu Hari-dat book showing is. this Anu’s most-first
prakaashana thaa
publication was
‘Anu is showing a book to Hari. This was her first-ever publication.’

Moreover, although Dayal implies that an indefinite-marked NP should allow pronominal reference, this does not appear to be so; see (34b). The unacceptable coindexing in (34b) is correctly predicted by Prasad and Strube’s analysis.
(34) a. Ravi kitaab\textsubscript{i} pa\dhy\textsubscript{a} rahaa hai. vo\textsubscript{s} bahut acchii hai.
Ravi book reading is. it very good is
‘Ravi is reading a/the book. It is very good.’

b. Ravi ek kitaab\textsubscript{i} pa\dhy\textsubscript{a} rahaa hai. vo\textsubscript{s} bahut acchii hai.
Ravi one book reading is. it very good is
‘Ravi is reading a book. It is very good.’

The above alternative explanations suggest that there is no reason to consider Hindi bare singular nominals as being anything other than simple indefinites in the DRT-sense.

The rest of the dissertation is structured as follows. Chapter 2 describes the three models whose predictions I will investigate. Chapters 3 to 5 present the results of seven experiments involving Hindi self-center embeddings. Chapter 6 discusses the effect of word length and position on reading time. Chapter 7 concludes the dissertation, and also presents a new model of human sentence processing based on abductive inference (originally described in (Vasishth, 2002a)). Appendix A provides all the stimuli in Devanagari script.
CHAPTER 2

THREE MODELS OF SENTENCE PROCESSING

The explanation goes like this, in daylight:
To be ambitious is to fall in love
With a particular life you haven’t got
And (since love picks your opposite) won’t achieve.
That’s clear as day. But come back at night,
You’ll hear a curious counter-whispering:
Success, it says, you’ve scored a great success.
Your wish has flowered, you’ve dodged the dirty feeding,
Clean past it now at hardly any price —
Just some pretence about the other thing.

From Success Story, Philip Larkin, 1957.

In this chapter, I discuss details of the three sentence processing models whose predictions I investigate. These are: Hawkins’ Early Immediate Constituents, Gibson’s Discourse Locality Theory, and Lewis’ Retrieval Interference Theory. I chose these models because they make precise, fine-grained, cross-linguistically applicable claims, and have empirically wide coverage.

It is surprisingly difficult to find theories of sentence processing that satisfy these (not unreasonable) criteria. By way of illustration, consider a very promising line of research in the modeling of sentence processing: probabilistic parsing. Recent contributions in this area have been by Jurafsky (1996), Narayanan and Jurafsky (1998), Brants and Crocker (2000), and Hale (2001), among others. Although the details vary, the central idea is that sentence processing difficulty can be accounted
for in terms of probabilities of syntactic structures (using, e.g., stochastic context-free grammars (Booth, 1969)). However, this is more of a promissory note at the moment than a well-developed research direction. To quote Jurafsky (to appear): “The role of probabilities in non-lexical syntactic structure, while assumed in most probabilistic models, rests on very little psychological evidence . . . As for models, it is clear that probabilistic models of linguistic processing are still in their infancy. Most models only include a very small number of probabilistic factors and make wildly unjustified assumptions about conditional independence.” Jurafsky goes on to point out the achievements of probabilistic models: “Probabilistic models do a good job of selecting the preferred interpretation of ambiguous input, and are starting to make headway in predicting the time-course of this disambiguation process.” However, even this claim (about predicting the time course of processing) may be too strong. For example, consider a recent model developed by Hale (2001) and cited by Jurafsky as “an important new contribution” in that it provides “much more fine-grained predictions about parsing difficulty and hence reading time.” The Hale model accounts for two kinds of ambiguous sentence structures in English: (a) the standard reduced-relative garden path sentence, *The horse raced past the barn fell*, which is harder to process than its unreduced version, *The horse that was raced past the barn fell* (Bever, 1970); and (b) the subject-object relative clause asymmetry, i.e., the fact that subject relatives like *The man who saw you saw me* are more difficult to process than object relatives like *The man who you saw saw me* (see, e.g., (Gibson, 1998)). Although the predictions of Hale’s model are correct in terms of overall processing difficulty, the prediction for the subject-object relatives asymmetry is incorrect when one considers moment-by-moment difficulty: Gordon et al. (2001) have shown that the hardest-to-process regions are at the verbs following the embedded subject, not at the embedded subject itself. One might, of course, ascribe this observed slowdown

---

1Stochastic context-free grammars associates each rule in a context-free grammar with the conditional probability that the left-hand side expands to the right-hand side (Jurafsky, to appear).
at the verb to “spillover”, or delayed manifestation of processing difficulty, from
the embedded subject. But appealing to spillover in such cases raises a second
problem: alternative explanations exist in terms of other, so-called “rationalist”,
sentence processing theories for facts such as those presented by Gordon et al. (2001).
For example, Gordon et al. account for the subject-object relative clause asymmetry
in terms of similarity-based interference. Yet, probabilistic parsing-based accounts
such as Hale’s make no attempt to rule out such alternative and independently
motivated explanations of processing difficulty. All this – the absence of any real
empirical coverage, and the inability to place probabilistic components of parsing
in the larger picture of cognitive processes – makes it considerably more difficult to
compare probabilistic parsing-based explanations with existing wide-coverage models.
Sentence parsing almost certainly involves probabilistic components, but the relative
role of probabilistic parsing compared to other components of parsing is yet to be
determined.

In sum, it is clear that approaches such as probabilistic parsing are likely
to provide new insights into human sentence parsing processes, and will raise the
standards of formal and mathematical rigor in the field. But these approaches are
not yet robust enough to be considered full-fledged theories of sentence processing.

The above discussion regarding probabilistic parsing in sentence processing
research is intended to illustrate the difficulties involved in modeling human sentence
processing, and to justify the choice of the three models I compare in this dissertation.
In comparison to probabilistic parsing-based approaches, the models discussed below
make precise predictions which can account for a wide variety of facts about different
languages.2

---

2AsJurafskypointsout(personalcommunication),probabilistic parsing-based approaches also
have the problem that their predictions for different languages and constructions are not readily
available or derivable, unless a text corpus is available for determining, e.g., the probabilities
associated with context-free grammar rules. Once such corpus-based information becomes more
widely available for different languages, it will become possible to evaluate such approaches empirically.
In the remainder of this chapter I present the details of each of the three models.


Although Hawkins (1994; 1998) presents EIC as a theory of human sentence processing, its central claim is that constraints on processing (such as constraints on extraposition, scrambling, etc.) determine the grammars of different languages. Competence grammar is an “evolutionary movement” towards an ideal specified by the theory. Under this view, the grammar only contains general principles of linguistic structure (such as the structure of verb phrases); there are no innate and parameterized universals for word order in the competence grammar; rather, performance constraints become “conventionalized” (grammaticalized) to determine preferences for optimal and sub-optimal word orders.

An appealing feature of this processing-driven view of grammar is that it offers to provide true explanations of grammatical phenomena, as opposed to merely recoding some observed fact as an innately present principle of language. To take one of Hawkins’ examples (Hawkins, 1994, 257), when verbs (Vs) modify prepositional phrases (PPs), cross-linguistically the preferred order is one where the V and head of the PP are adjacent (that is, the orders [V [P NP]] and [[NP P] V] are preferred over [V [NP P]] and [[P NP] V]). A typical syntactic analysis would be to engage in what I will call the fact-as-principle fallacy. This amounts to reasserting as a universal principle something that we have just observed. For example, one could assert that languages are parameterized to be head-initial or head-final, and that any deviances from the head-parameter are “marked” and therefore dispreferred. Although such a reformulation of the facts has the reassuring feel of an explanation, in reality not much is gained by way of explanatory force. The EIC provides a different account

\[\text{Of course, there is nothing wrong with limiting oneself to reorganizing facts systematically, but to call it an explanation is clearly incorrect.}\]
whereby the preferred orders are a result of the grammaticalization of immediate-constituent-to-word ratios (these are discussed below in detail).

I present the theory with the help of some motivating examples; the discussion below is based on (Hawkins, 1998). I begin with some definitions. Hawkins defines a complexity metric which relies on (i) the notion of Constituent Recognition Domain\(^4\) (CRD); (ii) the principle of Early Immediate Constituents (EIC); (iii) the Mother Node Construction principle; and (iv) the parsing axiom of Constructibility. I first informally motivate each of the concepts, and then provide the precise definitions.

Informally, the CRD of a given phrase (such as a VP) is the set of words that must be scanned before that phrase and its immediate constituents can be recognized. To illustrate, consider (35a,b) and the corresponding Figures 2.1 and 2.2.

![Diagram of CRD](image)

Figure 2.1: The CRD of the VP in (35a)

\(^4\)Given a syntactic tree for any natural language sentence, a Constituent is any part of the sentence that occurs under a node in the tree, and a node \(A\) is an Immediate Constituent of a node \(B\) if \(B\) immediately dominates \(A\). Any node \(B\) immediately dominates another node \(A\) if no other node intervenes between \(A\) and \(B\).
(35)   a. He \(v_P\) \([v\ \text{donated}]\ \[_{PP\ \text{to charity}}\] \([N_P\ \text{the beautiful desk dating from the early Victorian period}]\).

   b. He \(v_P\) \([v\ \text{donated}]\ \[_{NP\ \text{the beautiful desk dating from the early Victorian period}}\] \[_{PP\ \text{to charity}}\].

In (35a), the VP’s CRD consists of the four words that need to be scanned before the VP is recognized, and in (35b) the VP’s CRD consists of eleven words; see Figures 2.1 and 2.2.

The CRD, a set of words, is the basis for making parsing decisions, and parsing difficulty is quantified by the ratio of immediate constituents (ICs) to non-ICs in the CRD. A less accurate but simpler version of this is to simply take the IC-to-word ratio (Hawkins, 1994, 74-77). For example, in the CRD of the VP in (35a) there are three ICs (V, PP, NP) and 4 words, so the IC-to-word ratio is 3/4. The VP’s CRD in (35b) also has 3 ICs, but it now has 11 words, so its IC-to-word ratio is 3/11. The overall complexity of a sentence is taken to be the mean of the complexities of each phrase in the sentence. These ratios quantify the relative differences in parsing difficulty: the higher the ratio, the easier the processing. In other words, the fewer the words per CRD, the easier the processing. This preference for fewer words per CRD relates to constraints on human working memory, which prefers buffering fewer words rather than more in order to recognize a CRD. To quote Hawkins: “Smaller CRDs contain
fewer elements that must be held in working memory and that are relevant to a given parsing decision, and this reduces the number of phrasal nodes whose structure must be computed simultaneously” (Hawkins, 1998, 732).

Formally, the CRD is defined as follows:

**CONSTITUENT RECOGNITION DOMAIN (CRD):**

The CRD for a phrasal mother node M consists of the set of terminal and non-terminal nodes that must be parsed in order to recognize M and all ICs of M, proceeding from the terminal node in the parse string that constructs the first IC on the left, to the terminal node that constructs the last IC on the right, and including all intervening terminal nodes and the non-terminal nodes that they construct.

The main goal in parsing is thus to recognize constituent structure. The claim that higher ratios lead to easier processing is codified as the principle of Early Immediate Constituents (EIC).

**EARLY IMMEDIATE CONSTITUENTS (EIC):**

The human parser prefers linear orders that maximize the IC-to-non-IC (or IC-to-word) ratios of Constituent Recognition Domains. Orders with the most optimal ratios will be preferred over their non-optimal counterparts in the unmarked case; orders with non-optimal ratios will be more or equally preferred in direct proportion to the magnitude of their ratios. For finer distinctions, ratios can be measured left-to-right.

Apart from recognizing a phrase M based on the CRD of M, the parser also has to build the corresponding node for M in the tree. The next principle, Mother Node Construction, stipulates that this node is built as soon as a word appears that uniquely determines the existence of M. For example, a verb uniquely determines a VP; therefore, seeing a verb is necessary and sufficient information for the parser to build a VP.
MOTHER NODE CONSTRUCTION (MNC):

In the left-to-right parsing of a sentence, if any word\textsuperscript{5} of syntactic category C uniquely determines a phrasal mother node M, in accordance with the PS rules of the grammar, then M is immediately constructed over C.

The parse succeeds if a phrasal node M has been built in accordance with MNC. Hawkins states this as an axiom:

AXIOM OF CONSTRUCTIBILITY:

For each phrasal node M there will be at least one word of category C dominated by M that can construct M on each occasion of use.

The above discussion means that the model essentially uses a head-driven parser: a head of a phrase must be encountered before the corresponding phrasal node is built. For example, a verb must be seen before the subtree rooted at a verb phrase node is constructed. Before ending this section, I briefly discuss two issues that arise in Hawkins’ model. One relates to constraints on the kinds of center embeddings permissible in natural language, and the other with the role of pragmatics in sentence processing.

Turning first to the issue of center embeddings, an important question for Hawkins is: how can we account for the fact that only certain kinds of center embedding (e.g., center-embedded NPs versus sentential complements) occur in some languages? His answer is to posit an implicational hierarchy of non-optimal basic orders of increasing complexity (Hawkins, 1994, 102):

\textsuperscript{5}This includes bound morphemes, such as suffixes marking infinitival verbs (Hawkins, 1994, 359-379), (Hawkins, 1998, 733).
EIC Grammatical Hierarchy Prediction

Given: an IC position P center-embedded in a CRD for a phrasal node D (i.e., there is non-null material to both the left and right of P within D);
a set of alternative categories \{C\} that could occur in position P according to the independently motivated and unordered PS-rules of the grammar;

Then: if a category \( C_i \) from \{C\} produces low EIC ratios for D and is grammatical, then all the other categories from \{C\} that produce improved ratios will also be grammatical.

Hawkins illustrates this as follows (Hawkins, 1994, 319-320). Consider the following pair of sentences:

(36) a. \(*_S \text{ Did } [\overline{S} \text{ that John failed his exam}] [_{VP} \text{ surprise Mary}]]\)?
b. \(_S \text{ Did } [_{NP} \text{ this fact}] [_{VP} \text{ surprise Mary}]]\)?

According to Hawkins, sentential complements cannot be center-embedded (see (36a)), but NPs can (see (36b)). From these facts, he derives a hierarchy:

(37) \( S : NP > \overline{S} \) (">") = “more center-embeddable than”

In his words, “if a language permits center embedding of \( \overline{S} \), it will permit the center embedding of NP. Japanese and Korean permit both . . .; English permits a center-embedded NP only” (Hawkins, 1994, 320).

The hierarchy is based on the EIC (shorter constituents are more center embeddable than longer ones), and can explain why certain kinds of embeddings are allowed and others are not. However, there are several problems with this account of center embeddings.
First, Hawkins’ model provides no explanation for why languages like Japanese allow center embedding of $\overline{S}$ in the first place; this is just assumed to be a fact of the grammar of Japanese. Hawkins (1994, 8-12) argues that there cannot be any processing explanation, due to the differing degrees of availability of center embeddings in different languages. His argument rests on the assumption that center embedding of $\overline{S}$ does not occur in English (Hawkins, 1994, 8,12). However, this assumption is demonstrably false. Consider two of the examples from (1), repeated below:

(38)  
   a. $[s \text{ The odds } [\overline{S} \text{ that your theory will be in fact right, and the general thing } \overline{S} \text{ that everybody’s working on} \text{ will be wrong}] \text{ are low}.\]$
   
   b. $[s \text{ Your report today } [\overline{S} \text{ that any Tory constituency party } \overline{S} \text{ failing to deselect its MP, should he not vote in disobedience with a prime ministerial diktat,} \text{ might itself be disbanded}, \text{ shows with certainty that } \ldots \text{ an ‘elective dictatorship’ is now with us}.\]$
   
Both of these satisfy the operational definition of center embedding in Hawkins’ EIC Grammatical Hierarchy Prediction above: “an IC position P center-embedded in a CRD for a phrasal node D.” Yet both contain not one but two center-embedded $\overline{S}$’s, which directly contradicts Hawkins’ prediction and his assumption that English center embeddings are ungrammatical.

Second, Hawkins’ model cannot account for acceptability differences in Dutch versus German center embeddings. Bach et al. (1986) showed that Dutch center embeddings were easier to process for native Dutch speakers than German ones are for native German speakers. Subjects were asked to give an acceptability rating to each sentence, and were also tested on comprehension. The sentences were matched across German and Dutch, and varied from simple, easy-to-understand sentences, to center embeddings with three levels of embedding. The key finding was that, compared to the German subjects evaluating German embeddings, Dutch subjects
rated the Dutch embeddings to be better, and performed better on the comprehension test. Examples of Dutch and German SCEs are shown in (39). The Dutch SCEs are called “crossed dependencies” because of the fact that the dependencies between the verbs and the subjects form crossing chains, and the German SCEs are called “nested dependencies” since each embedded dependency is nested within another dependency.

(39) a. Jan Piet Marie zag laten zwemmen
   Jan Piet Marie saw make swim
   ‘Jan saw Piet make Marie swim.’
   NP1 NP2 NP3 V1 V2 V3

b. … dass Hans Peter Marie Schwimmen lassen sah
   … that Hans Peter Marie swim make saw
   ‘… that Hans saw Peter make Marie swim.’
   NP1 NP2 NP3 V3 V2 V1

The EIC incorrectly predicts that German center embeddings should be as acceptable as Dutch ones. In the Dutch case, at V1 the ratio is 3/4 since four words are seen before the three constituents of the VP1 headed by V1 are constructed. At this point during the parse, the tree structure will contain nodes NP1, V1, and a third center-embedded node VPx whose existence is implied by V1 (“said”). This last node has no terminal (leaf) since the embedded verb has not been seen yet. Then V2 is seen. By MNC, a VP2 node is constructed over it, and this is presumably unified with the node VPx. The ratio for VP2 is again 3/4. The same process is repeated for V3, and the ratio there is also 3/4. The average ratio is thus 3/4 for the Dutch sentence. Now consider the German example: for V3 the ratio is 1, for V2, 3/4; and for V1 it is 3/6 (i.e., 1/2). The average is 3/4, the same as the Dutch example.

---

6 A common objection to the Bach et al. paper is that one cannot really compare two different languages directly. The main point of the Bach et al. paper was that one can in fact do this. There is clearly a possible confound in such a study, but to the extent that this cross-linguistic approach is valid, the Dutch-German contrast is relevant to sentence processing questions.
Due to these reasons, EIC and the implicational hierarchy may not be able to account for the (non-)acceptability of center embeddings in different languages, or of different types within the same language.

I turn now to Hawkins’ claims about the role of pragmatics in human sentence parsing. Hawkins’ claim is that parsing decisions are based solely on the principle of EIC; pragmatic factors such as given-ness or new-ness play no role per se; in this sense it is a syntax-only theory (Hawkins, 1998, 751-759). In Hawkins’ own words: “[T]he selection of alternative orderings in [free word order] languages is driven by syntactic processing, and not by pragmatic information status and the “givenness”, “newness”, and “predictability” of entities that are referred to in discourse, as is widely believed” (Hawkins, 1998, 751-752). According to Hawkins, effects ascribed previously to “givenness” or “newness” can be explained by the fact that shorter constituents refer to more referentially given and definite entities, and longer ones to newer and more indefinite entities. Hawkins’ main claim here is that pragmatic theories cast in terms of information structure are not in themselves explanatory but are epiphenomenal; they derive from “a more primitive generalization, EIC, and of an independent correlation between constituent length and pragmatic status” (Hawkins, 1998, 756).

This is an interesting claim because it goes against the observation that semantic information is used early for parsing decisions. For example, Steedman’s Principle of Parsimony (Steedman, 2000, 238) states that “the analysis whose interpretation carries fewest unsatisfied but accommodatable presuppositions or consistent entailments will be preferred.” Given that there is much evidence (see, e.g., (Altmann & Steedman, 1988), (Kaiser & Trueswell, 2002)) supporting theories such as Steedman’s, Hawkins’ claim about pragmatic effects being epiphenomenal is problematic. Moreover, there already exists evidence that Hawkins’ epiphenomenal approach to pragmatic effects is empirically inadequate; see, e.g., (Kurz, 2000).
To summarize, Hawkins’ view is that the processing principle of Early Immediate Constituents can account for both real-time processing facts and word order preferences cross-linguistically; the grammar of a language is a result of the grammaticalization of performance constraints. The theory is certainly very impressive in its empirical coverage: it can account for diverse word-order facts from an impressive array of languages (Hawkins, 1998, 734-741): English, Hungarian, Romanian, Polish, German, Greek, Japanese, and Korean, to name a few. However, it also has several empirical problems; it cannot account for center-embedding facts, and it is inconsistent with the finding that semantic information is used early in parsing. There is other evidence against Hawkins’ model as well; this is discussed in the concluding chapter (see page 182).

EIC makes very clear predictions for Hindi center embeddings. First, a lower immediate-constituent-to-word ratio should result in increased processing difficulty at the stage where recognition of the head takes place (in the Hindi experiments, one such region is the innermost verb in examples like (3) on page 9). Such a lowering of the ratio could be due either to reordering arguments so that more words must be seen before the head, or due to the insertion of non-arguments, such as adverbs, between a head and its dependents. Second, since pragmatic effects do not have any independent status, the prediction is that certain words (such as definites) that require more presuppositions than others should be no harder to process than other words (indefinites) which require fewer presuppositions. In other words, EIC makes exactly the opposite prediction compared to theories such as Steedman’s Principle of Parsimony.
2.2 Gibson’s Dependency Locality Theory (2000)

The Dependency Locality Theory (DLT) aims to account for processing difficulty in both ambiguous (garden-path) structures and unambiguous ones, such as center embeddings. The discussion in this section is based on (Gibson, 2000).

The DLT assumes that during the course of sentence parsing, computational resources in working memory are needed for two aspects, STORAGE of the structure built up thus far; and INTEGRATION of the current word into the structure built up thus far. Based on these two components, the DLT defines a cost metric which predicts relative processing difficulty. There is a STORAGE COST, measured in MEMORY UNITS (MUs), and an INTEGRATION COST, measured in ENERGY UNITS (EUs). I describe these two components next.

Storage cost is computed as follows: 1 MU is associated with each syntactic head required to complete the current input as a grammatical sentence. For example, as shown in Figure 2.3, for the first word, *The*, a noun and a verb are needed to complete the sentence grammatically, so the storage cost is 2 MUs. At the second word, *reporter*, only a verb is needed, so the storage cost is 1 MU. For the next word, *disliked*, one NP is needed, so the cost is 1 MU. The cost for the next word, *the*, is also 1 MU, since only a noun is needed to complete the input string as a grammatical sentence. The last word incurs no cost since it can complete the sentence.

<table>
<thead>
<tr>
<th>Input words:</th>
<th>The</th>
<th>reporter</th>
<th>disliked</th>
<th>the</th>
<th>editor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntactic heads needed:</td>
<td>Noun, Verb</td>
<td>Verb</td>
<td>NP</td>
<td>Noun</td>
<td>-</td>
</tr>
<tr>
<td>Storage cost (MUs):</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 2.3: Illustration of DLT’s storage cost computation

The cost of integrating a current word $A$ with an existing structure $S$ depends on the LOCALITY, i.e., the distance, between them. More precisely, the cost depends on the complexity of the computations that took place between $A$ and $S$. These
computations are assumed to be linearly related to the number of discourse referents (DRs) introduced between the two items. Thus, the cost of integrating $A$ with $S$ is quantified in terms of the number of new intervening DRs.

Apart from storage cost and integration cost, the other factors that are assumed to affect comprehension difficulty are:

1. The frequency of the lexical item being integrated: The lower the frequency, the greater the processing difficulty.

2. The contextual plausibility of the resulting structure: The less plausible the final structure, the greater the processing difficulty.

3. The discourse complexity of the final structure: Nonfocused entities and elements introduced using definite descriptions are less accessible (and are therefore harder to process) than focused entities. Least accessible of all are elements introduced using indefinite NPs. I will call this the DISCOURSE COMPLEXITY ASSUMPTION.

4. If the current word is not compatible with the highest ranked structure built so far, there is reanalysis difficulty.

The storage and integration costs together provide a measure of INTEGRATION TIME, which gives the complexity at any given point, as well as overall complexity of

---

7 Gibson defines discourse referents as follows: “A discourse referent is an entity that has a spatiotemporal location so that it can later be referred to with an anaphoric expression, such as a pronoun for NPs, or tense on a verb for events” (Gibson, 2000, 103). It is important to note that Gibson does not regard discourse referents as the only measure of distance: “Processing other kinds of elements [other than new discourse referents] also probably causes measurable increments in structural integration cost” (Gibson, 2000, 107). However, the DLT has nothing more to say about other measures of distance.

8 “Elements new to the discourse, which are usually introduced using indefinite NPs, require the most resources because they must be constructed in the discourse model ... [I]t is assumed that processing the head noun of an NP that refers to a new discourse object consumes substantial resources, and processing the head verb of a VP that refers to a new discourse event (also a discourse referent) consumes substantial resources, but processing other words does not consume substantial resources in the discourse processing component of structure building.” (Gibson, 2000, 103). Also see (Warren, 2001, 33-44).
a sentence. Gibson assumes the following regarding these two costs: (i) Integrations and storage access the same pool of resources; (ii) Resources have fixed capacity; (iii) Each predicted syntactic head takes up a fixed quantity of resources; and (iv) The overall acceptability of a sentence depends on the maximal integration time complexity experienced during the parsing process.

In sum, sentence processing according to the DLT is constrained by limited resources in working memory. This limit on resources is quantified in terms of integration and storage costs of linguistic elements as they are processed in real time. There is a clear distinction between constraints on competence and on performance, and in this respect Gibson's model differs from Hawkins'. What is impressive about this model is the wide array of of cross-linguistic data that it can account for (see, e.g., (Gibson, 1998), (Babyonyshev & Gibson, 1999), (Warren & Gibson, 1999), (Nakatani, Babyonyshev, & Gibson, 2000)).

The DLT raises interesting questions for Hindi sentence processing. For example, the Discourse Complexity Assumption (see page 52) predicts that indefinites will be harder to process than definites. Moreover, the distance-based metric predicts that arguments that are reordered so as to be further away from the verb selecting for them will result in greater integration difficulty at the verb. Finally, DLT predicts that increasing the distance between a verb and its arguments by inserting an element not introducing a discourse reference (such as an adverb) will either have no effect on integration cost, or will result in greater integration difficulty.

2.3 Lewis' Retrieval Interference Theory (2001, 2002)

I present two distinct versions of the theory; I will call the first RIT-01, and the second RIT-02. My presentation of RIT-01 relies mainly on (Lewis, 1998), (Lewis & Nakayama, 2001). The newer version is RIT-02 and its presentation here is based on discussions with Richard Lewis during January-May 2002.
2.3.1 Retrieval Interference Theory, 2001 (RIT-01)

This model treats parsing as a working memory task that is affected primarily by interference. This characterization is based on standard approaches in psychology; to quote Gordon et al. (in press), in cognitive psychology, memory interference is assumed to “emerge in a process where retrieval cues are matched against representations of targets in memory.” The standard assumption is that “the facility with which the target is retrieved is an increasing function of the cue-to-target strength for the to-be-retrieved trace and a decreasing function of the sum of the cue-to-trace strengths for all other traces stored in memory (e.g., (Gillund & Shiffrin, 1984), (Hintzmann, 1986)). When a distractor item is similar to the target, there is an increase in cue-to-trace strength for the distractor, which interferes with retrieval of the target.”

In RIT-01, these constraints on retrieval are the primary determinants of processing difficulty. In the context of center-embedding constructions, the main idea in RIT-01 is that, during real-time processing, the retrieval at a verb of an NP that is currently stored in working memory is affected by interference, specifically, RETROACTIVE INTERFERECE (RI) and PROACTIVE INTERFERENCE (PI) from other items in working memory.

Pro- and retroactive interference are defined as follows. Proactive interference (PI) occurs when the retrieval of an item causes it to suffer from interference by an item or items preceding the item to be retrieved. Retroactive interference (RI) is the opposite: the retrieval of an item causes it to suffer from interference from items that follow that item. There is a great deal of evidence in the psychology literature for PI and RI in intertrial list recall (Estes, 1988). It is an assumption of RIT-01 that PI and RI occur within a list of NPs (see (Humphreys & Tehan, 1998), which provides independent evidence for proactive interference within trials). Moreover, in the description below, it is assumed that PI and RI make an equal contribution, but
note that this may not necessarily be so; see, e.g., (Melton & von Lackum, 1941),
which systematically compared RI and PI and found that RI had a more substantial
effect than PI. The model’s behavior is illustrated next.

Consider a schematic representation of a sentence where the sequence of symbols
$\phi_1\phi_2\ldots\phi_nX\rho_1\rho_2\ldots\rho_mY$ represents a string of NPs followed by a verb $Y$.

\begin{equation}
\phi_1\phi_2\ldots\phi_nX\rho_1\rho_2\ldots\rho_mY
\end{equation}

If the verb $Y$ is the current word being processed, and a syntactic relation needs
to be established between a constituent projected from $Y$ and a constituent headed by
a prior word $X$ (a noun), the total amount of interference at $Y$ depends on the number
of similar items intervening between $X$ and $Y$ (RI) and the number of similar items
preceding $X$ (PI). ‘Similarity’ here crucially refers to syntactic similarity, which is
determined by the structural role to be assigned to $X$ and which occurs in a syntactic
working memory.\(^9\) For example, if $X$ is to be assigned the surface-structural position
of subject, then RI occurs due to all $\rho_1\rho_2\ldots\rho_m$ which could also fill subject positions,
and PI occurs due to all $\phi_1\phi_2\ldots\phi_n$ which could also fill subject positions.

For ease of exposition, simple numerical values are assigned to each component
of processing cost: if there are two elements causing RI, then the amount of RI is 2
units, and so on. In Lewis’ actual computational implementation (a parser written in
the ACT-R environment (Anderson & Lebiere, 1998)), the costs are not necessarily
simple integer values. Following (Gibson, 1998), the maximum complexity at any
given point in the processing in the sentence is taken as the measure of the processing
complexity of the sentence.

\(^9\)Lewis (1999) follows Caplan and Waters (1999) in assuming that there is a specialization in
working memory of language processing, and that a separate sentence interpretation resource is
used for sentence processing. In this connection, Gordon et al. (in press) provide important new
evidence inconsistent with the specialization view. This dissertation also provides evidence against
the specialization view; see page 154 for a detailed discussion of the Gordon et al. (in press) results.
The predictions for Japanese self-center embeddings illustrate the model’s operation. The pattern in (41a) is predicted to be easier to process than (41b).

(41)  

a. NP1-ga NP2-ni [NP3-ga NP4-o V2] V1

Announced-ga zyuumin-ni keikan-ga daigakusei-o
Announcer-nom people-dat policeman-nom college-student-acc
sirabeta-to hoososita
investigated-comp announced
‘The announcer announced to the people that the policeman investigated
the college student.’

b. NP1-ga NP2-ni [NP3-ga NP4-ga V2] V1

Announced-ga zyuumin-ni keikan-ga daigakusei-ga
Announcer-nom people-dat policeman-nom college-student-nom
sukida-to hoososita
likes-comp announced
‘The announcer announced to the people that the policeman likes the
college student.’

In example (41a), for expository purposes I consider the case where V2 is currently being processed. Its subject, NP3, is retrieved; potential candidates for proactive interference are NP1 and NP2, and for retroactive interference, NP4. However, among NP1 and NP2, only NP1 is a potential subject; this similarity to NP3 results in one unit of PI. Further, since NP4 is not a potential subject, there is no RI. By contrast, in (41b), at V2 there is one unit of PI and one unit of RI, since NP1 and NP4 are now both potential subjects and therefore similar to NP3.

Note here that in (41b), NP4 is an object (Kuno, 1994, 81), and yet is considered a potential subject. If these ga-marked nominals really are objects, the model must assume that superficial similarity\(^{10}\) results in increased similarity-based interference.

\(^{10}\)I use this term instead of “phonological similarity” since in this particular case the issue is morphophonemic identity of case markers. The term phonological similarity has much broader connotations in the cognitive psychology literature (Baddeley, 1966) and its use in this context may lead to misunderstanding.
This is at odds with the assumption in RIT-01 (page 55) that structural (syntactic) similarity is the main source of similarity-based interference. This problem is resolved in RIT-02 (see pages 62, 183), where feature bundle similarity rather than similarity of syntactic structure alone is the source of interference effects.

Lewis and Nakayama (2001) also posit another kind of similarity-based interference independent of pro- and retroactive interference: **positional similarity.** This component of the theory states that “placing otherwise syntactically indiscriminable NPs adjacent to one another will increase processing difficulty, just in those cases where the correct attachment of the NPs depends on their relative positions” (2001, 95). This predicts that in a structure like (42a) there will be less interference due to positional similarity of the two nominative (ga) marked NPs. By contrast, in (42b), the adjacency of the two ga-marked NPs will result in greater positional similarity-based interference during retrieval at the verbs.

(42) a. NP1-ga NP2-ni [NP3-ga V2] V1

Yamada-ga Uehara-ni Fukuzawa-ga asondeiru-to renrakusita
Yamada-nom Uehara-dat Fukuzawa-nom playing-comp notified
‘Yamada notified Uehara that Fukuzawa was playing.’

b. NP1-ga NP2-ga [NP3-o V2] V1

Haisya-ga daitooryoo-ga tuuyaku-o yonda-to oboeteita
dentist-nom president-nom interpreter-acc called-comp remembered
‘The dentist remembered that the President called the interpreter.’

Lewis’ positional similarity-based interference hypothesis predicts that processing becomes easier when NPs are reordered so that the distance between similar NPs is increased. For example, if the two nominative marked NPs in (43a) are moved away from each other as in (43b), positional similiarity-based interference should decrease during retrieval at the innermost verb, resulting in less processing difficulty.\(^{11}\)

\(^{11}\)Lewis (2001) also cites Uehara’s research on Japanese, where it was found that increasing distance between NPs (in canonical order sentences) by inserting an adverb between them results in greater overall acceptability of the sentence.
(43)  a. NP1-ga [NP2-ga NP3-o V2] V1

Keikan-ga ryoosin-ga kodomo-o sagasu-to kangaeta
policeman-nom parents-nom child-acc look-for-comp thought
‘The policeman thought that the parents would look for the child.’

b. NP1-ga [NP2-o NP3-ga V2] V1

Keikan-ga kodomo-o ryoosin-ga sagasu-to kangaeta
policeman-nom child-acc parents-nom look-for-comp thought
‘The policeman thought that the parents would look for the child.’

Thus, the main claim is that PI and RI, along with positional similarity, determine processing complexity. There is much evidence from English (Lewis, 1998), Japanese (Lewis & Nakayama, 2001) and Korean (Uehara & Bradley, 1996) that is consistent with this version of the RIT.

A final feature of the model is its assumptions about **FOCUS OF ATTENTION**. Focus of attention is assumed to include the current item and the item immediately preceding the current item (cf. Cowan 1995, 2001), and items in the focus of attention are assumed to be immune to any kind of interference at all. For example, when the innermost verb is being processed in (44a), the final NP, *onna-no-ko-ga*, ‘girl-nom’, is in the focus of attention. However, in (44b) the presence of the adverb *asa rokuji kara*, ‘from 6AM onwards’, results in the final NP no longer being in the focus of attention when the innermost verb is being processed. This means in (44b) the final NP experiences interference, but in (44a) the final NP does not. This predicts that in sentence pairs like (44), processing will be more difficult at the innermost verb when the adverb is present, in contrast to the case where there is no intervening adverb.

(44)  a. Ani-ga sensei-ni onna-no-ko-ga asondeiru-to renrakusita
ever brother-nom teacher-dat girl-nom playing-that notified
‘My older brother notified the teacher that a girl was playing.’
b. Ani-ga sensei-ni onna-no-ko-ga asa rokuji kara elder brother-nom teacher-dat girl-nom morning 6-o’clock from asondeiru-to renrakusita playing-that notified
‘My older brother notified the teacher that a girl was playing since 6AM.’

The above assumptions about focus of attention are based on results from short-term memory research such as (Humphreys & Tehan, 1998), which shows that PI and RI occur only when an item is out of the focus of attention.\(^\text{12}\)

An important question for RIT is: does superficial similarity matter? Uehara and Bradley (1996) found no evidence for this in Korean. Korean has the interesting property that the formal shape of nominative case marking varies between \(i\) and \(ka\), depending on whether the noun that it modifies ends in a consonant or a vowel, respectively. This allows adjacent, nominative case-marked NPs to have nominative case markers with the same or different formal shape. In other words, Korean allows a simple manipulation of phonological shape while holding the nominative case marking on the NPs constant. In a processability rating experiment, Uehara and Bradley did not find a significant difference between processability of adjacent, nominative case-marked NPs with phonologically identical case markers versus adjacent, nominative NPs with case markers having differing phonological shape. This was, however, a null result. I show later on (see Experiment 3, Section 3.3.8, page 103) that there is some evidence for morphophonemnic similarity causing similarity-based interference.

This concludes the first version of Lewis’ Retrieval Interference Theory. I present next the newer version, RIT-02.

\(^{12}\)Lewis (personal communication) assumes that positional similarity based-interference also occurs only when an item is out of the focus of attention.
2.3.2 Retrieval Interference Theory, 2002 (RIT-02)

As in RIT-01, in this version attachments in real-time parsing are construed as retrievals from working memory and are subject to similarity-based interference. However, in RIT-02, no distinction is made between proactive and retroactive interference; these, along with positional similarity, are subsumed under interference. In addition, the model includes a parser with a predictive or top-down component, which has rather different empirical consequences compared to RIT-01.\(^\text{13}\) The model is described next.

As items (e.g., nouns, verbs) are processed in real time, they can generate predictions and make retrieval requests. The process of making a retrieval request based on some cues is referred to as setting retrieval cues. Retrieval cues are set when an item that is currently being processed requests the retrieval of either a previously generated prediction or a lexically realized item that has already been seen but is now out of the focus of attention. The item that has been retrieved, along with the current item being processed, are assumed to be in the focus of attention. In contrast to RIT-01, in RIT-02, being in the focus of attention does not render an item immune to interference.

Each retrieval request results in interference between different retrieval cues active in working memory, depending on the degree of similarity of the cues. Interference can be illustrated schematically as follows.

In Figure 2.4, the letters in parentheses (\{x t\}) together represent a collection of retrieval cues consisting of syntactic features “x” and “t”. These retrieval cues are assumed to have been set by some item that is currently in the focus of attention. For example, a verb looking for a subject could set a retrieval cue for a nominative

\(^{13}\)This predictive component was partly a result of Lewis and my working together on human parsing. As a result, this predictive component has many similarities to the abductive parser that I developed in my Master’s thesis (Vasishth, 2002a). My own abductive parser is described in the concluding chapter.
case-marked subject, where “nominative case-marked subject” can be characterized in terms of a feature bundle. The items currently in working memory are the items \([x \ j \ l \ m], [o \ x \ m \ k], [r \ s \ v], [x \ t \ o]\). These could be feature bundles representing nouns that have already been seen, for example. After the retrieval cues are set by the item currently in the focus of attention, these items in working memory become cues that could in principle match the cues set by the retrieval request, and they compete with each other for the match. The broken lines in Figure 2.4 represent the fact that interference is being caused by items that match the retrieval cues to a lesser extent than the item \([x \ t \ o]\) in working memory. The match with the cues is stronger in the case of \([x \ t \ o]\), and is denoted by the unbroken lines.

There are two kinds of retrieval requests. One involves an attempted match with a predicted structure, and the other an attempted match with a lexically realized structure. An example of a predicted structure is the case where a nominative NP has been seen; here, a sentence-level node \(S\) is predicted. If a verb appears next, the lexical entry for the verb makes a retrieval request, that is, it looks for an \(S\) (the predicted structure) in working memory, and attempts to match its own \(S\)-features with that of the predicted \(S\). An example of the latter kind of retrieval request (that is, an attempted match with a lexically realized structure) is the case where a nominative NP has been seen. The NP makes a prediction for an \(S\) node as above, but if a PP
appears next instead of a verb, the PP makes a retrieval request for an NP. The NP is a lexically realized structure, in contrast to the predicted S in the first case.

There are four main differences between RIT-01 and RIT-02:

(a) RIT-02 does not assume that the current item and the immediately preceding one are in the focus of attention, only the current item, plus some other item that has been retrieved.

(b) In RIT-02, being in the focus of attention does not render an item immune from interference.

(c) RIT-02 assumes that retrieval requests are feature bundles, not just structural positions (including at least major category, subcategorization features, case, agreement features).

(d) RIT-02 assumes predictive parsing and retrieval; matching occurs against predictions as well as against lexically-realized structure.

RIT-02’s predictive parsing has the following consequence. Recall that in RIT-01, at the innermost verb in (45), the final NP is assumed to suffer interference since it is out of the focus of attention, compared to the case where no adverb is present (in the latter case, the final NP would be immune to the effects of interference, since it is in the focus of attention).

(45) Ani-ga sensei-ni onna-no-ko-ga asa rokuji kara
     elder brother-nom teacher-dat girl-nom morning 6AM from
     asondeiru-to renrakusita
     playing-that notified
     ‘My older brother notified the teacher that a girl has been playing since 6AM.’

By contrast, in RIT-02, a verb is predicted by the NPs seen so far and each prediction has an activation level associated with it. When the adverb is seen in a
sentence like (45), it is integrated with this prediction,\textsuperscript{14} and this has the consequence that the base-level activation of the predicted verb increases. Therefore, when the verb is seen the retrieval of the prediction is easier (compared to the case where no adverb was present) due to its higher activation level. This predicts that the presence of an adverb between the last NP and the first verb will result in \textit{easier} processing at the innermost verb compared to the case where no adverb is present. This prediction is interesting because it is the opposite to that of RIT-01, and of Gibson’s and Hawkins’ models. RIT-02’s prediction is tested in Experiment 7 and, as discussed in Chapter 5, turns out to be the correct one.

The predictions for Hindi of the models are as follows. RIT-01 predicts that increasing positional similarity, or increasing the distance between a verb and its dependents (by inserting an adverb as discussed above) should both result in greater processing difficulty at the verb. Conversely, decreasing positional similarity between certain arguments which are similar and adjacent to each other (by reordering them so that they are no longer adjacent) should result in less processing difficulty at the innermost verb, where retrieval takes place. Note that it is important here that similarity is defined as superficial similarity (morphophonemic identity of case markers). Figure 1.1 on page 9 shows the assumed structure: the definite-marked direct object and the indirect object(s) are assumed to be more confusible with each other if they are adjacent to each other. RIT-02 also predicts increasing difficulty due to similarity, but the predictive component implies that processing cost will be lower in the case where an adverb is inserted between a head and its argument.

This concludes the discussion of the three models. I turn next to the experiments themselves. In subsequent chapters, when I refer to “integration”, I use this term in the sense of Gibson’s model and also to refer to Hawkins’ notion of head recognition. The term “retrieval” will be used exclusively in the sense of Lewis’ models.

\textsuperscript{14}This means that the parse tree is restructured so that the adverb now modifies the predicted verb.
CHAPTER 3

THE EFFECT OF SIMILAR CASE MARKING:
EXPERIMENTS 1-3

‘Interesting, but futile,’ said his diary,
Where day by day his movements were recorded
And nothing but his loves received inquiry;
He knew, of course, no actions were rewarded,
There were no prizes: though the eye could see
Wide beauty in motion or a pause,
It need expect no lasting salary
Beyond the bowels’ momentary applause.

A writer, Philip Larkin, 1941

Early Immediate Constituents (EIC), Discourse Locality Theory (DLT), and Retrieval Interference Theory (RIT) make clear predictions about processing difficulties associated with definite versus indefinite direct object NPs in Hindi (see pages 50, 53, 63). EIC predicts that definiteness marking should have no effect on processing difficulty, since pragmatics plays no role in processing per se. DLT predicts that indefinites will be easier to process than definites, because indefinites introduce a new discourse referent whereas definites presuppose an old/familiar one; DLT thus predicts an increase in storage cost but no change in integration cost. RIT predicts that there will be greater processing difficulty in retrieving a dative marked NP at a verb when a definite-marked direct object NP (as opposed to a bare NP) intervenes between the dative NP and the verb. This is due to increased similarity (see page 57) of such NPs with superficially similar case marking.
The three experiments presented here investigate these predictions. In addition, Experiment 3 also investigates Lewis’ assumptions about interference and similarity. Specifically, Experiment 3 considers whether multiple occurrences of NPs with the postposition or case marker, se, ‘from’, can result in interference effects, and whether (non)-adjacency of ke-marked NPs results in increased similarity-based interference.

3.1 Experiment 1

3.1.1 Method and Procedure

Subjects

Nineteen of the subjects were Hindi-speaking graduate students at the Ohio State University (OSU), and were paid 5 US Dollars each for completing the questionnaire; the remaining thirty-four were undergraduate and graduate students at Jawaharlal Nehru University (JNU), New Delhi, India, and were paid 80 Indian Rupees each (equivalent approximately to 1.7 US Dollars in August 2000). The subjects were from a variety of majors (Computer Science, Statistics, Journalism, Engineering, Economics, Foreign Languages, Hindi, International Studies, etc.).

Materials

Experiment 1 had a 2 \times 2 factorial design, the two factors being level of embedding (single or double; compare (46a,b) and (46c,d)), and absence or presence of case marking on the final NP (compare (46a,c) and (46b,d)).
(46) a. Siitaa-ne Hari-ko [kitaab khariid-neko] kahaa
    Sita-erg Hari-dat book buy-inf  told
    ‘Sita told Hari to buy a book.’

    b. Siitaa-ne Hari-ko [kitaab-ko khariid-neko] kahaa
    Sita-erg Hari-dat book-acc buy-inf  told
    ‘Sita told Hari to buy the book.’

    Sita-erg Hari-dat Ravi-dat book buy-inf tell-inf told
    ‘Sita told Hari to tell Ravi to buy a book.’

    Sita-erg Hari-dat Ravi-dat book-acc buy-inf tell-inf told
    ‘Sita told Hari to tell Ravi to buy the book.’

In the test sentences, all but the final NPs were proper names; the final NP was always an inanimate common noun, such as ‘book’ or ‘letter’.

Procedure

This was a paper questionnaire where subjects were asked to rate each sentence on a scale from 1 (completely unacceptable) to 7 (completely acceptable). Four lists were prepared in a counterbalanced, Latin Square design, and 32 fillers were inserted between 16 target sentences in pseudorandomized order. The fillers consisted of eight examples of four syntactic structures: relative clauses, medial gapping constructions (Ross, 1970), simple declaratives, and sentences with that-clauses (all the stimuli are presented in Appendix A, and the fillers are available on request).

3.1.2 Results

A repeated measures analysis of variance (ANOVA) was done for subject (F1) and item (F2) means, with level of embedding and presence or absence of case marking on the final NP as the within-subject factors, and group as the between-subjects factor.
There was a main effect of level of embedding (F1(1,51) = 226.58, p < 0.00001; F2(1,15) = 134.97, p < 0.00001), a main effect of case marking (F1(1,51) = 210.57, p < 0.00001; F2(1,15) = 45.40, p < 0.00001), and a significant interaction (F1(1,51) = 33.62, p < 0.00001; F2(1,15) = 24.18, p = 0.0004). No group effect was found for the by-subjects analysis (F1(1,51) = 1.21, p = 0.3155), but there was a group effect in the by-items (F2(1,15) = 16.86, p = 0.0001). A contrast analysis showed that case marking on the final NP resulted in significantly lower acceptability; this was true for both single embeddings (F1(1,51) = 133.9390, p < 0.00001; F2(1,15) = 46.7984, p < 0.00001) and double embeddings (F1(1,51) = 29.4578, p < 0.00001; F2(1,15) = 10.1498, p = 0.007835). The results are summarized graphically in Figure 3.1.

![Graph showing acceptability rating vs. case marking on final NP](image)

Figure 3.1: Experiment 1 results, with 95% confidence intervals. A rating of 1 represents the acceptability judgment “completely unacceptable”, and a rating of 7 represents “completely acceptable”.

67
3.1.3 Discussion

This experiment showed that increasing the amount of embedding results in reduced acceptability: (46c,d) were less acceptable than (46a,b). This is correctly predicted by the three models, but since it is not a surprising result, I will say nothing more about this. More interesting is the result that case marking on the final NP results in reduced acceptability: (46b), (46d) were less acceptable than (46a), (46c) respectively. RIT-02 predicts this but EIC and DLT do not.

Regarding RIT, if similarity-based interference among *ko*-marked NPs is the source of processing difficulty at the innermost verb, then, compared to single embeddings, in double embeddings case marking the final NP should have a greater effect on the overall acceptability rating. This is because in the double embedding construction there are three NPs with *ko* marking, whereas in single embeddings there are only two. In other words, Lewis’ model predicts an interaction between case marking and level of embedding. This is the correct prediction.

The question then arises: is the reduced acceptability due to differences in processing difficulty at the nouns or due to differences in retrieval/integration difficulty at the innermost verb? Over the course of this and the next chapter I will show that the answer is: processing difficulty occurs both at nouns and at verbs. I also show that although the number of new discourse referents could be a factor affecting processing difficulty, storage cost at nouns is affected to a greater extent by whether the definite or indefinite noun phrase is in subject or direct object position (see the discussion concerning Givón’s hypothesis on page 20).

In order to explore center embeddings at the level of moment-by-moment processing difficulty, a self-paced reading experiment was carried out using the same stimuli. This experiment is described next.
3.2 Experiment 2

As discussed in Chapter 1, Hindi has differential object marking: inanimate direct objects optionally take accusative case marking to mark specificity/definiteness. As shown in (47), without any preceding discourse context the direct object *kitaab*, ‘book’, can be indefinite or definite, depending on the absence or presence of case marking.¹

(47) a. Siitaa-ne Hari-ko [kitaab khariid-neko] kahaa
   Sita-erg Hari-dat book buy-inf told
   ‘Sita told Hari to buy a book.’

b. Siitaa-ne Hari-ko [kitaab-ko khariid-neko] kahaa
   Sita-erg Hari-dat book-acc buy-inf told
   ‘Sita told Hari to buy the book.’

If, as Experiment 1 suggests, adding the specificity/definiteness marker on an NP like *kitaab*, ‘book’ (47b) results in increased processing difficulty compared to the indefinite (47a), this would be problematic for Hawkins’ and Gibson’s models. This is because in Hawkins’ theory the immediate-constituent-to-word ratio remains unchanged in both sentences, and this ratio is the sole determiner of processing complexity. Gibson’s DLT predicts that indefinites will result in greater processing difficulty than definite descriptions, due to the fact that a completely new, unpresupposed discourse referent must be introduced in the indefinite case, whereas in the case of definites the discourse referent is presupposed. In contrast to DLT’s predictions in terms of storage effects at nouns, Lewis’ 2001 model (Lewis & Nakayama, 2001) predicts that there will be an increase in processing difficulty (longer reading time) during retrieval at the innermost verb, due to increased

¹Recall that the specificity marking is optional: the bare NP can also be specific, depending on the preceding context. The important thing to note for current purposes is that without any preceding discourse context, the default interpretation of a bare object NP is indefinite (see pages 31-37 for details), and definite in the case-marked version (see page 30). Accordingly, I will gloss the bare NP as an indefinite, but the reader should keep in mind that this is only correct for the null-context sentences used in the experiments.
similarity between the similarly (ko) case-marked NPs. The present experiment further investigates EIC’s, DLT’s, and the 2001 version of RIT’s predictions regarding the effect of definiteness marking on reading time at the final NP and the innermost verb.

3.2.1 Method

Subjects

Fifty undergraduate and graduate students at Jawaharlal Nehru University, New Delhi participated in the experiment. The subjects were from a variety of majors (Economics, Foreign Languages, Hindi, International Studies, etc.). Each subject was paid Rupees 80 (equivalent approximately to 1.7 US dollars in August 2000) for participating. None of the subjects had participated in the previous experiment.

Materials

Experiment 2 had a 2 × 2 factorial design; as in Experiment 1, the two factors were level of embedding (single or double; compare (48a,b) and (48c,d)), and absence or presence of case marking on the final NP (see (48a) vs. (48c), and (48b) vs. (48d)).

(48) a. A: Single embedding, bare final NP

Siitaa-ne Hari-ko  [kitaab khariid-neko] kahaa
Sita-erg Hari-dat book buy-inf told
‘Sita told Hari to buy a book.’

b. B: Double embedding, bare final NP

Siitaa-ne Hari-ko  [Ravi-ko kitaab khariid-neko bol-neko] kahaa
Sita-erg Hari-dat Ravi-dat book buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy a book.’
c. **Single embedding, definite-marked final NP**

Sita-ne Hari-ko [kitaab-ko xariid-neko] kahaa
Sita-erg Hari-dat book-acc buy-inf told
‘Sita told Hari to buy the book.’

d. **Double embedding, definite-marked final NP**

Sita-erg Hari-dat Ravi-dat book-acc buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book.’

In each of the target sentences, all but the final NPs were proper names; the final NP was always an inanimate common noun, such as ‘book’ or ‘letter’. Four lists were prepared in a counterbalanced, Latin Square design, and 32 fillers were inserted between 16 target sentences in pseudorandomized order. The fillers consisted of eight examples of four syntactic structures: relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses (all the stimuli are in Appendix A). This was a non-cumulative self-paced moving window reading task (Just, Carpenter, & Woolley, 1982); the procedure is described next.

**Procedure**

A G3 laptop Macintosh running PsyScope (Cohen, MacWhinney, Flatt, & Provost, 1993) was used to present the materials to subjects. The task was to press the space key in order to see each successive phrase; each time the key was pressed, the previous phrase would disappear and the next phrase would appear. Reading time (in milliseconds) was taken as a measure of relative momentary processing difficulty. A yes/no comprehension question was presented after each sentence; these were meant to ensure that subjects were attending to the sentences. Subjects with less than 70% correct responses were not included in the data analysis; six subjects were excluded as a result. The computer screen presentation was as described below with the aid of Figure 3.2.
Figure 3.2: Illustration of presentation on the computer screen

Each line in Figure 3.2 shows a particular stage during a trial. At the beginning of a trial, the screen shows the uppermost line in Figure 3.2: a series of dotted lines separated by spaces, with each line corresponding to a phrase in the sentence that will eventually be seen in that trial. Then the subject presses the space bar. The screen now displays the second line in the figure; at this point, only the first phrase is visible. The subject then presses the space bar again, and the next phrase appears and the previous one disappears (see the third line in Figure 3.2). This procedure is repeated until the end of the sentence (marked by the period, a vertical line) is reached. After the entire sentence is seen, a yes/no question is shown, to which the subject responds by pressing the frontslash (“/”) key for “yes”, or the “z” key for “no”. After the subject responds, the screen reverts to displaying a line similar to the uppermost line in Figure 3.2, and the procedure is repeated for the next sentence (this could be a filler or critical sentence). Each subject practised the task with eight sample sentences before the actual experiment commenced.
The segmentation of the regions presented as a unit is as shown below. The segmentation is similar for each condition: each NP and verb constitutes a separate region. In example (49), each vertical line marks a region boundary.

(49) Raam-ne | Ravi-ko | kitaab | khariid-neko | kahaa
      Ram-erg  Ravi-dat  book  buy-inf  told
‘Ram told Ravi to buy a book.’

3.2.2 Results

Factoring out the effect of word length: A brief note on statistical methodology

In order to compare the reading times for definite-marked versus bare NPs, a common assumption is that it is necessary to factor out any effect of word length on reading time (Trueswell, Tanenhaus, & Garnsey, 1994). A standard method for factoring out word length effects is to do an analysis of variance (ANOVA) of residuals rather than of raw reading times. The procedure is as follows. Residual reading time are calculated for each region by subtracting from raw reading times each subject’s predicted reading time for regions with the same number of characters; this in turn is calculated from a linear regression equation across all of a subject’s sentences in the experiment (Ferreira & Cliffton, 1986; Trueswell et al., 1994).²

However, as Maxwell et al. (1985) point out, doing an ANOVA of residuals results in a model in which the parameter estimates used are not least-squares estimates of that model, resulting in the test statistic not being appropriately distributed. Accordingly, I used the linear mixed-effects (Pinheiro & Bates, 2000) to factor out any word-length effects.³

---

²See Chapter 6, Section 6.1 for a discussion of the Devanagari script used for Hindi and some relevant facts about character processing in Hindi, based on (Vaid & Gupta, to appear).

³In the linear mixed-effects model, subjects were the random effect and word length and case marking times were fixed effects. I acknowledge here the technical assistance provided by Professor Tom Santer and Sumithra Mandrekar (Graduate Student), both Consultants at the Statistical Consulting Service, Ohio State University.
Interestingly, this assumption, that word length affects reading time, contradicts a body of work in cognitive psychology (see, e.g., (Johnson, 1977), (Johnson, 1981), (Johnson, Turner-Lyga, & Pettergrew, 1986), (Johnson, 1991), and the references cited there). Indeed, in the experiments presented in this dissertation, there is no evidence that word length has a significant effect on reading time. I discuss this issue in Chapter 6.

In the results for the self-paced reading experiments presented in this dissertation, whenever word length is a potential confound I present results from ANOVAs by both raw reading times and by residuals, as well as an ANOVA using the linear mixed-effects model. I present the ANOVA results for all three methods for completeness; the misleading nature of the ANOVA based on residuals is discussed in Chapter 6 (see page 167) in the context of word length effects in reading time studies. When word length is not a potential confound, I present the results of an ANOVA based only on raw reading times. For the raw and residual reading times, I present both by-subject and by-item analyses, but for the linear mixed-effects model I present only the by-subject analyses. This is because, as discussed above, word length does not appear to have a significant effect on reading time, and the results based on raw reading times suffice for the issues of interest in this dissertation. However, the results of the by-subjects analyses for the linear mixed-effects model are presented because they do serve to demonstrate the absence of any effect of word length (this is discussed in detail in Chapter 6).
The effect on reading time of case marking on the direct object

One set of regions of interest was the third NP in single embeddings and the fourth NP in double embeddings; this direct object NP was either unmarked or marked for definiteness. Since this was a planned comparison, a contrast analysis was carried out using ANOVA. Another set of regions of interest was the innermost verb in the single and double embeddings. Two separate repeated measures ANOVAs were done for subject (F1) and item (F2) means of reading time (RT), one for single embeddings and another for double embeddings.

The statistical results for both the NP and innermost verb’s reading times are shown in Tables 3.1 to 3.6, and are summarized graphically in Figures 3.3 to 3.6. For single embeddings, a contrast analysis involving the third, direct-object NP showed that RT was significantly longer when the NP was case marked both for raw RTs and for residual RTs. In the ANOVA based on the linear mixed-effects model, the effect of case marking is also statistically significant. Although there is a significant effect of word length, this was due to a single outlier; once that was removed, the main effect goes away (see page 168 for details). Finally, the RTs at the innermost verb are significantly longer in single embeddings if the final NP is case marked (Table 3.3). Identical results were obtained for the fourth NP and the innermost verb in double embeddings; ANOVAs by both raw and (for the final NP) residual RTs yield significant results (see Tables 3.4-3.6 and Figures 3.5 and 3.6).
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,43)</th>
<th>p-value</th>
<th>F2(1,15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>15.12</td>
<td>0.0004</td>
<td>26.97</td>
<td>0.0002</td>
</tr>
<tr>
<td>Residuals</td>
<td>6.94</td>
<td>0.0120</td>
<td>9.93</td>
<td>0.0083</td>
</tr>
</tbody>
</table>

Table 3.1: Expt. 2 (Conditions A vs. C); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,42)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP3</td>
<td>11.24</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Word length of NP3</td>
<td>4.45</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 3.2: Expt. 2 (Conditions A vs. C); ANOVA based on a linear mixed-effects model for NP3 (cf. page 168)

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,43)</th>
<th>p-value</th>
<th>F2(1,15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>6.56</td>
<td>0.0143</td>
<td>4.83</td>
<td>0.0484</td>
</tr>
</tbody>
</table>

Table 3.3: Expt. 2 (Conditions A vs. C); contrast analysis at V2
Figure 3.3: Expt. 2, single embeddings, bare versus case-marked final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals

Figure 3.4: Expt. 2, single embeddings, bare versus case-marked final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,43)</th>
<th>p-value</th>
<th>F2(1,15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>12.62</td>
<td>0.0010</td>
<td>11.89</td>
<td>0.0048</td>
</tr>
<tr>
<td>Residuals</td>
<td>5.16</td>
<td>0.0285</td>
<td>9.19</td>
<td>0.0104</td>
</tr>
</tbody>
</table>

Table 3.4: Expt. 2 (Conditions B vs. D); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,42)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP4</td>
<td>12.25</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Word length of NP4</td>
<td>1.94</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 3.5: Expt. 2 (Conditions B vs. D); ANOVA based on a linear mixed-effects model for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,43)</th>
<th>p-value</th>
<th>F2(1,15)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>7.89</td>
<td>0.0077</td>
<td>7.52</td>
<td>0.0178</td>
</tr>
</tbody>
</table>

Table 3.6: Expt. 2 (Conditions B vs. D); contrast analysis at V3
Figure 3.5: Expt. 2, double embeddings, bare versus case-marked final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals

Figure 3.6: Expt. 2, double embeddings, bare versus case-marked final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals
3.2.3 Discussion

The results suggest that definiteness marking on the direct object results in increased processing difficulty compared to indefinite (bare) direct objects, both at the final NP and at the innermost verb. There are different implications of this result for the various theories.

In the case of Hawkins’ model, it is clear that EIC alone cannot account for processing difficulty in the case of Hindi. Pragmatic factors, such as the referential properties of NPs, play an important, early role in determining overall parsing difficulty (Altmann & Steedman, 1988), (Steedman, 2000). This implies that the syntax-only view advocated by Hawkins should be revised.

Regarding DLT, the model, based on experiment results from English (Warren, 2001, 33-44), assumes that indefinites result in greater processing difficulty than definites. In Hindi the opposite situation holds: definite NPs cause greater processing difficulty than indefinites. There is, however, a principled reason for this. In Warren’s experiment, comprehension difficulty ratings were obtained from subjects for several sentence types, two of which were as follows (the sentences were presented without any preceding context): (a) sentences with indefinites in the subject position of an embedded restricted relative clause, as in *The reporter who the senator who a professor met attacked disliked the editor*, versus (b) sentences with definite NPs, as in *The reporter who the senator who the professor met attacked disliked the editor*. The sentences with indefinites were found to be slightly harder to comprehend than those with the definite NP, although the difference did not reach significance. The key difference between Warren’s and my experiment is that in the Hindi experiments the bare or definite NP is invariably a direct object, not a subject. This is the source of the difference between the English and Hindi results.\(^4\)

\(^4\)Another possibility is that the Hindi indefinites are canonically inanimate (Dahl, 1997); one way to test that possibility would be to compare (non-human) animate and inanimate NPs in direct object position. I leave the exploration of this possibility for future research.
Recall the argument (page 20) due to Givón (1978a, 306) that “the accusative position is the ‘most indefinite’ of all major arguments of the verb, in contrast with the subject and dative which are overwhelmingly definite.” If indefiniteness is the default for direct objects, and definiteness the default for subjects, then it is not surprising that (English) definites in subject position are easier to process than indefinites (Warren’s result), and that (Hindi) indefinites in direct object position are easier to process than definites (the result in Experiment 2). Givón’s observation not only explains the asymmetry in Warren’s and my results, but also illustrates another important point. Patterns in language are often attributed to frequency effects, and the strong claim is made that frequency explains particular phenomena (e.g., (Jurafsky, 2002)). Although there is no doubt that frequency has an effect on processing ease, there are occasionally deeper reasons for the frequency effect itself. In the present case, it is not the case that in the direct object position indefinites are easier to process than definites because indefinites are more frequent, but because, for the reasons discussed earlier, indefinites are less “marked” than definites in object position. Of course, one could argue that this discourse-based explanation is manifested as a difference in actually observed frequencies (e.g., in a text corpus) and therefore counts as an explanation. However, the observed frequency differences in such cases are a consequence of an underlying principled explanation, and in that sense it is odd to claim that frequency “explains” the definite-indefinite effect. It is reasonable to assume that the underlying reason for the difference between indefinites and definites results in frequency differences and that this frequency is the immediate causal factor for the observed differences; but this is not the same thing as claiming that frequency per se is responsible for the phenomenon.

Turning next to Lewis’ model, it appears to make the correct prediction for the observed slowdown at the final, definite NP, but the model suggests a different explanation for the results than increased discourse load: the slowdown at the final NP may be due to increased similarity-based interference resulting from similar case
marking on adjacent NPs. Similarity-based interference also correctly predicts the observed slowdown at the innermost verb, that is, during retrieval at the verb.\footnote{It is possible that the slowdown at the verb in the case-marked condition may simply be a result of spillover from processing difficulty at the immediately preceding NP. This question is resolved in subsequent experiments.}

Recall that in Experiment 1 an interaction was found between case marking on the final NP and level of embedding (pages 67, 68). This was argued to provide support for Lewis' model. Similarity-based interference would predict a similar interaction at the innermost verb's RT; however, no such interaction was found (F1(1,43) = 1.57, p = 0.2172; F2(1,15) = 0.12, p = 0.7336).

The next experiment (Experiment 3) is designed to test whether similarity-based interference is responsible for the processing difficulty observed in the previous experiments.

### 3.3 Experiment 3

This experiment examines Lewis' similarity-based interference hypothesis from several related but different directions. First, if definiteness marking \textit{per se}, and not similarity due to identical case marking, is responsible for a longer reading time (RT) at the final NP, the case-marked NP4 in sentence type D should have a longer RT compared to NP4 in C (see the examples in (50)). This is because in C vs. D the only \textit{kó}-marked NP is the definite object, and therefore any difficulty in processing at the final NP in D cannot be due to similar case marking; it must be due to the inherent difficulty of processing the definite NP. Moreover, unlike in Experiment 2, no slowdown should be observed at the innermost verb, since there is no change in similarity-based interference in the two conditions.
(50) a. C. ne-se-se-0
   Sita-ne | Hari-se | Ravi-se | kitaab | le-neko | kahaa
   Sita-erg Hari-abl Ravi-abl book give-inf told
   ‘Sita told Hari to take a book from Ravi.’

b. D. ne-se-se-ko
   Sita-ne | Hari-se | Ravi-se | kitaab-ko | le-neko | kahaa
   Sita-erg Hari-abl Ravi-abl book-acc give-inf told
   ‘Sita told Hari to take the book from Ravi.’

Second, if processing difficulty is due to similarity-based interference and not definiteness marking per se, then one would expect (a) no significant difference in RT at the final NP in G compared to B (examples like (51)), since in G any interference effects should appear only during retrieval at the innermost verb; (b) a significant difference at the innermost verb’s RT in conditions G versus B.

(51) a. B. ne-ko-se-0
   Sita-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
   Sita-erg Hari-dat Ravi-abl book take-inf told
   ‘Sita told Hari to take a book from Ravi.’

b. G. ne-ko-se-ko
   Sita-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa
   Sita-erg Hari-dat Ravi-abl book-acc take-inf told
   ‘Sita told Hari to take the book from Ravi.’
Third, if se-marked NPs do not suffer interference effects but only ko-marked ones do, similarity-based interference predicts a slowdown at the innermost verb in condition G versus D in examples like (52).

(52)  a. D. ne-se-se-ko

Siitaa-ne | Hari-se | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

b. G. ne-ko-se-ko

Siitaa-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

Fourth, a slowdown is predicted at the innermost verb in condition D versus G (i.e., the opposite of the third prediction above) if the following hold true: (a) similarity-based interference affects any set of similarly case-marked NPs in the same way, and not just ko-marked ones; and (b) if adjacent se-marked NPs suffer greater interference than non-adjacent ko-marked NPs.

Fifth, if similarity-based interference affects se-marked NPs, then, in examples like (53), there should be a significantly longer RT at the innermost verb in C (compared to B).

(53)  a. B. ne-ko-se-0

Siitaa-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

b. C. ne-se-se-0

Siitaa-ne | Hari-se | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’
Sixth, greater positional similarity-based interference is predicted if ko-marked NPs are adjacent. That is, a longer RT is expected at the innermost verb in E compared to G in examples like (54).

(54) a. E. **ne-se-ko-ko**

Sitaa-ne | Hari-se | Ravi-ko | kitaab-ko | de-neko | kahaa
Sita-erg Hari-abl Ravi-dat book-acc give-inf told
‘Sita told Hari to give the book to Ravi.’

b. G. **ne-ko-se-ko**

Sitaa-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

Finally, similarity-based interference predicts a slower RT at the innermost verb in condition A versus B in examples like (55).

(55) a. A. **ne-ko-ko-0**

Sitaa-ne | Hari-ko | Ravi-ko | kitaab | de-neko | kahaa
Sita-erg Hari-dat Ravi-dat book give-inf told
‘Sita told Hari to give a book to Ravi.’

b. B. **ne-ko-se-0**

Sitaa-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

3.3.1 Method

Subjects

Eighty⁶ undergraduate and graduate students at Jawaharlal Nehru University, New Delhi participated in the experiment. The subjects were from a variety of majors

---

⁶The data from twenty subjects was not used in the data analysis because the percentage of correct responses to the yes/no comprehension questions for these subjects fell below 70%.
(Economics, Foreign Languages, Hindi, International Studies, etc.). Each subject was paid Rupees 100 (equivalent approximately to two US dollars in September 2001) for participating. None of the subjects had participated in previous experiments.

Materials and Procedure

In each of the test sentences, all but the final NPs were proper names; the final NP was always an inanimate common noun, such as 'book' or 'letter'. Six lists were prepared in a counterbalanced, Latin Square design, and 62 fillers were inserted between 36 target sentences in pseudorandomized order. The fillers consisted of various syntactic structures, such as relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses (all the stimuli are given in Appendix A). This was a non-cumulative self-paced moving window reading task (Just et al., 1982); the procedure is as described on pages 71 to 71 for Experiment 2. Examples of the stimuli are repeated below for convenience.
(56)  

a. A. **ne-ko-ko-o**

Sita-ne | Hari-ko | Ravi-ko | kitaab | de-neko | kahaa
Sita-erg Hari-dat Ravi-dat book give-inf told
‘Sita told Hari to give a book to Ravi.’

b. B. **ne-ko-se-o**

Sita-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

c. C. **ne-se-se-o**

Sita-ne | Hari-se | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

d. D. **ne-se-se-ko**

Sita-ne | Hari-se | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

e. E. **ne-se-ko-ko**

Sita-ne | Hari-se | Ravi-ko | kitaab-ko | de-neko | kahaa
Sita-erg Hari-abl Ravi-dat book-acc give-inf told
‘Sita told Hari to give the book to Ravi.’

f. G. **ne-ko-se-ko**

Sita-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’
3.3.2 Results: The effect of definite -ko with no other ko-marked NPs present

The main regions of interest were the fourth NP and the innermost verb; the fourth NP was either unmarked or marked for definiteness (see (57)).

(57) a. C. ne-se-se-0

Siitaa-ne | Hari-se | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

b. D. ne-se-se-ko

Siitaa-ne | Hari-se | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-abl Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

For each of the two regions, two sets of contrast analyses were carried out (as discussed on page 73). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT). The relevant statistical results are as shown in Tables 3.7 to 3.9, and are summarized graphically in Figures 3.7 and 3.8. The contrast analysis for the final, direct-object NP showed that RT was significantly longer when the NP was case marked. The linear mixed-effects model gave the same result, but no main effect of word length was found. The contrast analysis for the innermost verb (V2) showed no effect of case marking on the direct object.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>20.61</td>
<td>0.0000</td>
<td>11.20</td>
<td>0.0022</td>
</tr>
<tr>
<td>Residuals</td>
<td>3.11</td>
<td>0.0837</td>
<td>1.96</td>
<td>0.1722</td>
</tr>
</tbody>
</table>

Table 3.7: Expt. 3 (Condition C vs. D); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,58)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP4</td>
<td>13.53427</td>
<td>0.0003</td>
</tr>
<tr>
<td>Word length of NP4</td>
<td>0.04369</td>
<td>0.8345</td>
</tr>
</tbody>
</table>

Table 3.8: Expt. 3 (Condition C vs. D); ANOVA based on linear mixed effects model for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.01</td>
<td>0.9229</td>
<td>0.01</td>
<td>0.9208</td>
</tr>
</tbody>
</table>

Table 3.9: Expt. 3 (Condition C vs. D); contrast analysis for V2
Figure 3.7: Expt. 3 (Conditions C vs. D); raw RTs, with 95% confidence intervals

Figure 3.8: Expt. 3 (Conditions C vs. D); residual RTs, with 95% confidence intervals
3.3.3 Results: The effect of definite -ko with a non-adjacent ko-marked indirect object present

The main regions of interest were the fourth NP and the innermost verb; the fourth NP was either unmarked or marked for definiteness (see (58)).

(58)  a. B. ne-ko-se-0

Siitaa-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

b. G. ne-ko-se-ko

Siitaa-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa
Sita-erg Hari-dat Ravi-abl book-acc take-inf told
‘Sita told Hari to take the book from Ravi.’

For each of the two regions, two sets of contrast analyses were carried out (as discussed on page 73). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 3.10 to 3.12, and are summarized graphically in Figures 3.9 and 3.10. The contrast analysis for the final, direct-object NP showed that RT was significantly longer when the NP was case marked. The linear mixed-effects model gave the same result, but no main effect of word length was found. The contrast analysis for the innermost verb showed a near-significant effect of direct object case marking on the RT of the verb, but only in the by-subjects analysis.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>18.39</td>
<td>0.0001</td>
<td>19.84</td>
<td>0.0001</td>
</tr>
<tr>
<td>Residuals</td>
<td>4.22</td>
<td>0.0449</td>
<td>4.78</td>
<td>0.0367</td>
</tr>
</tbody>
</table>

Table 3.10: Expt. 3 (Conditions B vs. G); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,58)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP4</td>
<td>15.65</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Word length of NP4</td>
<td>0.29</td>
<td>0.59</td>
</tr>
</tbody>
</table>

Table 3.11: Expt. 3 (Conditions B vs. G); ANOVA based on linear mixed-effects model for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>3.61</td>
<td>0.0627</td>
<td>1.63</td>
<td>0.2112</td>
</tr>
</tbody>
</table>

Table 3.12: Expt. 3 (Conditions B vs. G); contrast analysis for V2
Figure 3.9: Expt. 3 (Conditions B vs. G); raw RTs, with 95% confidence intervals

Figure 3.10: Expt. 3 (Conditions B vs. G); residual RTs, with 95% confidence intervals

93
3.3.4 Results: The effect of adjacent *se*-marked NPs versus non-adjacent *ko*-marked NPs

The main regions of interest were the third and fourth NPs and the innermost verb; the fourth NP was marked for definiteness in both conditions (see (59)), and is a potential region of interest due to the possibility of spillover effects from the third NP.

(59)  

(a) D. **ne-se-se-ko**

Siitaa-ne | Hari-se | Ravi-se | kitaab-ko | le-neko | kahaa  
Sita-erg Hari-abl Ravi-abl book-acc give-inf told  
‘Sita told Hari to take the book from Ravi.’

(b) G. **ne-ko-se-ko**

Siitaa-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa  
Sita-erg Hari-dat Ravi-abl book-acc take-inf told  
‘Sita told Hari to take the book from Ravi.’

For each of the two regions, a contrast analysis was carried out using raw reading times. Each involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT). The statistical results are as shown in Tables 3.13 to 3.15, and are summarized graphically in Figure 3.11. The contrast analyses for the third NP and the final, direct-object NP showed no significant differences in RT for the two conditions. However, the contrast analysis for the innermost verb showed a near-significant effect on RT of the innermost verb; i.e., RT was longer at the verb when there were two non-adjacent *ko*-marked NPs preceding the verb.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.08</td>
<td>0.7851</td>
<td>0.07</td>
<td>0.7979</td>
</tr>
</tbody>
</table>

Table 3.13: Expt. 3; contrast analysis for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.53</td>
<td>0.4680</td>
<td>0.29</td>
<td>0.5939</td>
</tr>
</tbody>
</table>

Table 3.14: Expt. 3 (Condition D vs. G); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>3.37</td>
<td>0.0720</td>
<td>4.12</td>
<td>0.0513</td>
</tr>
</tbody>
</table>

Table 3.15: Expt. 3 (Condition D vs. G); contrast analysis for V2

Figure 3.11: Expt. 3 (Condition D vs. G); raw RTs, with 95% confidence intervals
3.3.5 Results: The effect of multiple se-marking

The main regions of interest were the third and fourth NPs and the innermost verb; the fourth NP was not marked for definiteness in the two conditions (see (60)), but is a region of interest due to possible spillover effects from NP3.

(60) a. B. ne-ko-se-0

Sita-na-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg  Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

b. C. ne-se-se-0

Sita-na-ne | Hari-se | Ravi-se | kitaab | le-neko | kahaa
Sita-erg  Hari-abl Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

For each of the two regions, a contrast analysis was carried out using raw reading times. Each involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT). The statistical results are as shown in Tables 3.16 to 3.19,\(^7\) and are summarized graphically in Figure 3.12. The contrast analyses for the third NP and the innermost verb showed no significant differences in RT for the two regions.

\(^7\)Results for the second NP is also included in the tables because Figure 3.12 seems to suggest that there may be a significant difference at this regions. This, however, turns out not to be the case, as Table 3.16 shows.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>2.67</td>
<td>0.1078</td>
<td>2.54</td>
<td>0.1214</td>
</tr>
</tbody>
</table>

Table 3.16: Expt. 3 (Condition B vs. C); contrast analysis for NP2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.11</td>
<td>0.7446</td>
<td>0.12</td>
<td>0.7300</td>
</tr>
</tbody>
</table>

Table 3.17: Expt. 3 (Condition B vs. C); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>1.48</td>
<td>0.2284</td>
<td>2.24</td>
<td>0.1452</td>
</tr>
</tbody>
</table>

Table 3.18: Expt. 3 (Condition B vs. C); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.11</td>
<td>0.7454</td>
<td>0.04</td>
<td>0.8462</td>
</tr>
</tbody>
</table>

Table 3.19: Expt. 3 (Condition B vs. C); contrast analysis for V2
Figure 3.12: Expt. 3 (Condition B vs. C); raw RTs, with 95% confidence intervals
3.3.6 Results: The effect of adjacency on positional similarity

The main regions of interest were the fourth NP and the innermost verb; the fourth NP was marked for definiteness in both the conditions.

(61)  a. E. ne-se-ko-ko

Siitaa-ne | Hari-se | Ravi-ko | kitaab-ko | de-neko | kahaa  
Sita-erg Hari-abl Ravi-dat book-acc take-inf told

‘Sita told Hari to take the book from Ravi.’

b. G. ne-ko-se-ko

Siitaa-ne | Hari-ko | Ravi-se | kitaab-ko | le-neko | kahaa  
Sita-erg Hari-dat Ravi-abl book-acc take-inf told

‘Sita told Hari to take the book from Ravi.’

For each of the two regions, a contrast analysis was carried out using raw reading times. Each involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 3.20 to 3.22, and are summarized graphically in Figure 3.13. The contrast analyses for the fourth NP and the innermost verb showed no significant differences in RT for the two conditions.

---

8Results for the second NP are also included in the tables because Figure 3.13 seems to suggest that there may be a significant difference at the second NP. This, however, turns out not to be the case, as Table 3.20 shows.
Table 3.20: Expt. 3 (Condition E vs. G); contrast analysis for NP2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.78</td>
<td>0.3800</td>
<td>1.55</td>
<td>0.2229</td>
</tr>
</tbody>
</table>

Table 3.21: Expt. 3 (Condition E vs. G); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.08</td>
<td>0.7823</td>
<td>0.06</td>
<td>0.8008</td>
</tr>
</tbody>
</table>

Table 3.22: Expt. 3 (Condition E vs. G); contrast analysis for V2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.27</td>
<td>0.6028</td>
<td>0.22</td>
<td>0.6392</td>
</tr>
</tbody>
</table>

Figure 3.13: Expt. 3 (Condition E vs. G); raw RTs, with 95% confidence intervals

100
3.3.7 Results: The effect of adjacent, dative case-marked NPs

The main regions of interest were the third NP and the innermost verb; the fourth NP was not marked for definiteness in the two conditions (see (62)).

(62) a. A. ne-ko-ko-0

Siitaa-ne | Hari-ko | Ravi-ko | kitaab | de-neko | kahaa
Sita-erg  Hari-dat Ravi-dat book give-inf told
‘Sita told Hari to give a book to Ravi.’

b. B. ne-ko-se-0

Siitaa-ne | Hari-ko | Ravi-se | kitaab | le-neko | kahaa
Sita-erg  Hari-dat Ravi-abl book take-inf told
‘Sita told Hari to take a book from Ravi.’

For each of the two regions, a contrast analyses was carried out using raw reading times. Each involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 3.23 to 3.25, and are summarized graphically in Figure 3.14. The contrast analyses for the third NP and the innermost verb showed no significant differences in RT for the two conditions. However, in the by-subjects analysis a significantly longer RT is observed at the final verb in Condition A, compared to the RT of the final verb in Condition B.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.74</td>
<td>0.3947</td>
<td>0.71</td>
<td>0.4058</td>
</tr>
</tbody>
</table>

Table 3.23: Expt. 3 (Conditions A vs. B); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.01</td>
<td>0.9336</td>
<td>&lt; 0.001</td>
<td>0.9528</td>
</tr>
</tbody>
</table>

Table 3.24: Expt. 3 (Conditions A vs. B); contrast analysis for V2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,59)</th>
<th>p-value</th>
<th>F2(1,35)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>5.17</td>
<td>0.0270</td>
<td>2.08</td>
<td>0.1597</td>
</tr>
</tbody>
</table>

Table 3.25: Expt. 3 (Conditions A vs. B); contrast analysis for V1

Figure 3.14: Expt. 3 (Conditions A vs. B); raw RTs, with 95% confidence intervals
3.3.8 Discussion

Before discussing the results, it may be useful to summarize the main findings. For each set of results, I present the main contrasts of interest, the predictions, and the results in Table 3.26.

<table>
<thead>
<tr>
<th>Section</th>
<th>Contrasts</th>
<th>RIT’s Predictions</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.2</td>
<td>C: ne-se-se-0 vs. D: ne-se-se-ko</td>
<td>No slowdown at NP4 in D</td>
<td>✗</td>
</tr>
<tr>
<td>3.3.3</td>
<td>B: ne-ko-se-0 vs. G: ne-ko-se-ko</td>
<td>(a) No slowdown at NP4 in G</td>
<td>✗</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(b) Slowdown at verb in G</td>
<td>✓</td>
</tr>
<tr>
<td>3.3.4</td>
<td>D: ne-se-se-ko vs. G: ne-ko-se-ko</td>
<td>Slowdown at verb in G</td>
<td>✓</td>
</tr>
<tr>
<td>3.3.5</td>
<td>B: ne-ko-se-0 vs. C: ne-se-se-0</td>
<td>Slowdown at verb in C</td>
<td>✗</td>
</tr>
<tr>
<td>3.3.6</td>
<td>E: ne-se-ko vs. G: ne-ko-se-ko</td>
<td>Slowdown at verb in E</td>
<td>✗</td>
</tr>
<tr>
<td>3.3.7</td>
<td>A: ne-ko-ko-0 vs. B: ne-ko-se-0</td>
<td>Slowdown at verb in A</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 3.26: Summary of predictions and results in Experiment 3. The tick mark (✓) indicates that the prediction was correct, and the cross (✗) indicates that the prediction was incorrect.

During the processing of NPs, definiteness marking *per se* seems to be the source of processing difficulty (Sections 3.3.2, 3.3.3); this is a discourse effect. Although this discourse effect is not inconsistent with Lewis’ RIT, his model does not have an explicit account for the observed effect. For Hawkins’ model, the result implies that pragmatic and semantic information does indeed play an early role in parsing, and such effects cannot be subsumed under the EIC as epiphenomenal and derivable
from other, more basic principles of parsing. In the case of Gibson’s DLT, this result serves to replicate those in Experiments 1 and 2 and confirms that Hindi indefinite direct objects are in fact harder to process than definite direct objects, as Givón’s hypothesis predicts. In subsequent chapters, I shall use the term **the definiteness effect** to refer to processing difficulty at a definite NP.

The other results mainly affect Lewis’ RIT-01, and in particular to his similarity-based interference hypothesis. Recall that RIT-01 assumes that if an NP is out of the focus of attention, there are two main sources of similarity-based interference: (a) presence of other NPs similar to the NP that is currently in working memory and that is being retrieved at a verb, and (b) adjacency of syntactically indiscriminable NPs, if the correct attachment of the NPs to a verb depends on their relative positions.

The main conclusions from Experiment 3 regarding RIT-01 are as follows.

- Similarity-based interference could be argued to occur when two NPs have ko-marking: there was a near-significant slowdown at the innermost verb when the NP-sequence was *ne-ko-se-ko* as opposed to *ne-ko-se-0* or *ne-se-se-ko* (Sections 3.3.3 and 3.3.4), and a significant slowdown at the final verb when the NP-sequence was *ne-ko-ko-0* versus *ne-ko-se-0* (Section 3.3.7).

- No evidence was found that se-marked NPs experience similarity-based interference: the difference in reading time at the innermost verb in the sequence *ne-ko-se-0* versus *ne-se-se-0* did not reach significance (Section 3.3.5).

- No evidence was found for positional confusability causing interference: the difference in reading time at the innermost verb in the sequence *ne-se-ko-ko* versus *ne-ko-se-ko* did not reach significance (Section 3.3.6).

The null results are inconclusive; therefore, one cannot conclude that positional similarity plays no role in similarity-based interference, or that multiple instances of se-marked NPs do not experience interference but ko-marked NPs do. Some evidence
was found for similarity-based interference, but this was limited to ko-marked NPs. Thus, the present experiment appears to provide some limited support for Lewis’ similarity-based interference hypothesis. Note that Hawkins’ and Gibson’s models alone cannot account for the facts that have an explanation in terms of similarity-based interference.

Finally, recall the issue (see pages 10-14) of whether infinitivals are nominals or verbal in nature. This experiment also shows that the infinitival, ko-marked verb in this experiment is not encoded in a manner similar to the preceding nouns. As the nouns are successively seen, they show monotonically nondecreasing reading time, whereas the verbs show a statistically significant speedup. This is consistent with Gibson’s, Hawkins’ and Lewis’ assumptions that verb-related integration/retrieval processes, rather than encoding of an NP, is taking place at the infinitival verb. Of course, one may argue that the semantic features of the nominal infinitive are causing the speedup at the infinitival. This is possible; but Experiment 7 decides the issue, since the results there are only consistent with the hypothesis that an adverb following the final NP predicts a verb (see page 155).

The next chapter explores the distance hypothesis in its various manifestations in the three models.
CHAPTER 4

DISTANCE EFFECTS OR SIMILARITY-BASED
INTERFERENCE?: EXPERIMENTS 4-6

Hours giving evidence
Or birth, advance
On death equally slowly.
And saying so to some
Means nothing; others it leaves
Nothing to be said.  

Nothing to be said, Larkin, 1961

I will refer to as the DISTANCE HYPOTHESIS any claim that increasing distance between certain arguments or between dependents and heads results in increased or decreased processing difficulty. Early Immediate Constituents (EIC) and Discourse Locality Theory (DLT) predict an increase in processing difficulty if certain arguments are displaced from their canonical position, but each does so for different reasons (see pages 50, 53). Interestingly, Retrieval Interference Theory (RIT-01) predicts improved processability if certain arguments are displaced with respect to their canonical order.

To illustrate the issues, first consider a single center embedding in its canonical order (63a) versus the situations where the indirect object (63b) or the direct object is fronted (63c).

(63)  a. sitaa-ne hari-ko kitaab khariid-neko kahaa
      Sita-erg Hari-dat book buy-inf told
      ‘Sita told Hari to buy a book.’
b. hari-ko siitaa-ne kitaab khariid-neko kahaa
Hari-dat Sita-erg book buy-inf told
‘Sita told Hari to buy a book.’

c. kitaab siitaa-ne hari-ko khariid-neko kahaa
book Sita-erg Hari-dat buy-inf told
‘Sita told Hari to buy a book.’

EIC predicts that increasing distance (63b,c) between dependents and heads will result in increased processing difficulty at the innermost verb, since more words need to be processed before the head of the embedded clause (the innermost verb) is recognized. That is, reading time (RT) at the innermost verb is predicted to be longer in the direct-object and indirect-object fronting cases than in the canonical case, but no difference in RT is expected at the verb between the fronted indirect-object versus fronted direct-object cases.

The DLT makes similar predictions, but the reason here is that the integration cost increases because the number of new discourse referents between the head and dependents increases in the noncanonical orders. Moreover, the integration cost at the innermost verb in the fronted direct-object sentence (63c) is higher than in the fronted indirect-object sentences (63b): in direct-object fronting, there are two new discourse referents between the fronted NP and the verb (as opposed to zero new discourse referents in the canonical case, a total increase of two Energy Units of integration cost), whereas in indirect-object fronting there are two new discourse referents intervening between the fronted indirect object and the verb (as opposed to one new discourse referent in the canonical case, a total increase of one Energy Unit). That is, RT at the innermost verb is predicted to be longer in the fronted cases compared to the canonical case, and longer in the fronted direct-object sentence compared to the fronted indirect-object sentence.

Lewis’ RIT-01 makes rather different predictions. If superficially similarly case-marked NPs are affected by positional similarity-based interference, then (64a) should
be harder to process than (64b,c). In other words, fronting the ko-marked NP should result in shorter RT at the innermost verb, compared to canonical order sentences.

(64)  a. siitaa-ne hari-ko kitaab-ko khariid-neko kahaa
    Sita-erg Hari-dat book-acc buy-inf told
    ‘Sita told Hari to buy the book.’

    b. hari-ko siitaa-ne kitaab-ko khariid-neko kahaa
    Hari-dat Sita-erg book-acc buy-inf told
    ‘Sita told Hari to buy the book.’

    c. kitaab-ko siitaa-ne hari-ko khariid-neko kahaa
    book-acc Sita-erg Hari-dat buy-inf told
    ‘Sita told Hari to buy the book.’

It is well-known that sentences with non-canonical order (presented out of context) are less acceptable in languages like English, German, Finnish, and, Hungarian, (see (Hyönä & Hujanen, 1997)). One would therefore expect non-canonical order sentences (with no preceding context) to involve more processing cost, contra RIT-01’s predictions. Lewis’ theory implicitly assumes that any decrease in acceptability due to non-canonical order will be smaller than the improvement in acceptability due to reduced similarity-based interference. Although this assumption is never stated by Lewis, this is the only way to make sense of the results presented there (and elsewhere; e.g., see (Miyamoto & Takahashi, 2001)) from Japanese, in which sentences with non-canonical order were rated to be more acceptable, or otherwise found to be easier to process, than those with canonical order.

Experiment 4 is an acceptability rating study that seeks to determine whether fronting the indirect object results in reduced acceptability compared to canonical order. This experiment leads to Experiment 5 (involving indirect object fronting) and Experiment 6 (involving direct object fronting). These two are reading time studies which investigate the predictions of the three models. Experiments 5 and 6 also attempt to replicate the definiteness effect.
4.1 Experiment 4

4.1.1 Method

Subjects

Sixty-four subjects participated in the experiment. They were undergraduate and graduate students at Jawaharlal Nehru University, New Delhi, India, and were paid 80 Indian Rupees each (equivalent approximately to 1.7 US Dollars in August 2000). None of the subjects had participated in the preceding experiments.

Materials

Experiment 4 has a $2 \times 2$ factorial design: one factor was absence or presence of case-marked final NPs (compare (65a,c) and (65a,c)) and the other factor was indirect object in canonical or fronted position (compare (65a,b) and (65c,d)).

(65)

a. siitaa-ne hari-ko kitaab khariid-neko kahaa
   Sita-erg Hari-dat book buy-inf told
   ‘Sita told Hari to buy a book.’

b. siitaa-ne hari-ko kitaab-ko khariid-neko kahaa
   Sita-erg Hari-dat book-acc buy-inf told
   ‘Sita told Hari to buy the book.’

c. hari-ko siitaa-ne kitaab khariid-neko kahaa
   Hari-dat Sita-erg book buy-inf told
   ‘Sita told Hari to buy a book.’

d. hari-ko siitaa-ne kitaab-ko khariid-neko kahaa
   Hari-dat Sita-erg book-acc buy-inf told
   ‘It was Hari who Sita told to buy the book.’

In the test sentences, all but the final NPs were proper names; the final NP was always an inanimate common noun, such as ‘book’ or ‘letter’.
Procedure

This was a paper questionnaire where subjects were asked to rate each sentence on a scale from 1 (completely unacceptable) to 7 (completely acceptable). Four lists were prepared in a counterbalanced, Latin Square design, and 32 fillers were inserted between 16 target sentences in pseudorandomized order. The fillers consisted of eight examples of four syntactic structures: relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses (all the stimuli are presented in Appendix A).

4.1.2 Results

![Graph](image)

Figure 4.1: Experiment 4 results, with 95% confidence intervals. A rating of 1 represents the acceptability judgement “completely unacceptable”, and a rating of 7 represents “completely acceptable.”

No main effect was found for order (F1(1,63) = 3.13, p = 0.0818; F2(1,15) = 2.21, p = 0.1631), but there was a main effect of case marking on the final NP
(F1(1,63) = 35.47, p < 0.00001; F2(1,15) = 155.78, p < 0.00001), and there was an interaction (F1(1,63) = 6.14, p = 0.0160; F2(1,15) = 3.94, p = 0.0703). No group effect was found in the by-subjects or by-items analysis (F1(1, 63) = 1.39, p = 0.2536; F2(1,15) = 3.94, p = 0.0703)). A contrast analysis showed that case marking on the final NP resulted in significantly lower acceptability; this was true for both canonical order sentences (F1(1,63) = 127.063, p < 0.00001; F2(1,15) = 81.1126, p < 0.00001) and fronted indirect-object sentences (F1(1,63) = 51.2496, p < 0.00001; F2(1,15) = 64.563, p < 0.00001). The results are summarized graphically in Figure 4.1.

4.1.3 Discussion

There are four conclusions to be drawn from this experiment. First, the definiteness effect was replicated in this experiment; this provides further support for the conclusions in the previous chapter. Second, with a bare direct-object NP, fronting the indirect object results in lower acceptability compared to the canonical-order sentence. This is consistent with the predictions of both Hawkins’ and Gibson’s models. Third, with a definite-marked direct object NP, no difference in acceptability was found between fronted indirect-object sentences compared to canonical order sentences; this is inconsistent with Lewis’ positional similarity hypothesis. However, there is some indirect evidence that is consistent with Lewis’ hypothesis. This comes from the interaction between the factors. This shows that the decrease in acceptability due to case marking is significantly greater in the canonical sentences than in the fronted indirect object ones. This could mean that a reduced amount of interference due to positional similarity occurs in the fronted indirect object sentences: RIT-01 predicts that adding a definiteness marker when there is no other adjacent ko-marked NP results in less processing difficulty than when there is such a similar NP present.

However, the interaction may equally well be due to the fact that, when no preceding context is provided, sentences with fronted indirect objects are worse
per se than canonical order sentences. Therefore, one could equally argue that no conclusive evidence was found for positional similarity in this experiment. Of course, this does not mean that this constitutes evidence against positional similarity: as pointed out earlier (page 108), the claim implicit in Lewis’ model is that the adverse effect on acceptability of reordering constituents is smaller than the improvement in acceptability due to reduced positional similarity. If anything regarding this question can be concluded from this experiment, it would be that the adverse effect of reordering constituents (without any preceding discourse context) is greater than the facilitation of processing due to reduced positional similarity-based interference. This is hardly a surprising result: the strong, adverse effect on acceptability of noncanonical order sentences in Hindi has been observed in the past in informal pilot studies (Gambhir, 1981). Future studies based on this experiment will involve the inclusion of appropriate discourse context for each sentence in order to resolve the relative roles of discourse-appropriateness of noncanonical order sentences versus positional similarity.

Having established that there is indeed some difference in acceptability between canonical and noncanonical orders, the next two experiments investigate, the reading time predictions of the three models.

4.2 Experiment 5: Indirect-object fronting

This experiment investigates reading time differences at the innermost verb resulting from increased distance between a verb (the innermost verb) and its dependents by fronting the indirect object. Although acceptability judgements are critical for initially determining the relative overall difficulty of syntactic structures (Experiment 4), the specific claims of current parsing models are more precisely tested in a task that reflects real-time processing demands. Experiment 5 addresses these issues (as does Experiment 6).
Experiment 5 is designed to determine whether reading time (RT) at the innermost verb will be faster or slower in canonical word-order sentences compared to sentences with fronted IOs. Hawkins’ EIC and Gibson’s DLT predict that RT will be faster in the canonical case for sentences with bare or definite-marked NPs, since the distance between the head and the dependent arguments increases when the IO is fronted. By contrast, Lewis’ Retrieval Interference Theory predicts that RT will be longer in the canonical case for sentences with definite-marked objects: in the fronted IO version, the similarly case-marked NPs (the fronted IO and the definite-marked NP) are no longer adjacent, and so (positional) similarity-based interference should decrease.

As discussed earlier, these models do not explicitly assume that prior discourse context will affect processing difficulty in noncanonical order sentences: the empirical results discussed by Hawkins, Gibson, and Lewis supporting their respective models involve canonical and noncanonical order sentences without any prior discourse context. However, research has shown that prior discourse context has a significant effect on processing difficulty. Accordingly, as before, I interpret Hawkins’, Gibson’s, and Lewis’ assumptions to be that they consider any effects of increasing distance (in Hawkins’ and Gibson’s case) or decreasing positional similarity (in Lewis’ case) to outweigh any effect of discourse context.¹

4.2.1 Method

Subjects

Sixty² undergraduate and graduate students at Jawaharlal Nehru University, New Delhi participated in the experiment. The subjects were from a variety of majors

¹I am making a more generous assumption for Hawkins than is really justified, since Hawkins explicitly rules out any effect of discourse as being relevant to processing difficulty.

²The data from eight subjects was excluded because these subjects had fewer than 70% correct answers to the yes/no comprehension questions.
(Economics, Foreign Languages, Hindi, International Studies, etc.). Each subject was paid Rupees 100 (equivalent approximately to two US dollars in September 2001) for participating. None of the subjects had participated in any of the previous experiments.

Materials and Procedure

Four lists were constructed with a counterbalanced design as in earlier experiments, with 44 fillers arranged in pseudorandomized order between 32 target sentences (i.e., six items per condition). The four conditions are illustrated below.

\[ 66 \]

a. A. **Canonical, no case marking on final NP**

\[
\begin{align*}
\text{Riinaa-ne} & | \text{Siitaa-ko} | \text{kitaab} | \text{khariid-neko} | \text{kahaa} \\
\text{Rina-erg} & | \text{Sita-dat} & \text{book} & \text{buy-inf} & \text{told} \\
\text{‘Rina told Sita to buy a book.’} 
\end{align*}
\]

b. B. **IO Fronted, no case marking on final NP**

\[
\begin{align*}
\text{Siitaa-ko} & | \text{Riinaa-ne} | \text{kitaab} | \text{khariid-neko} | \text{kahaa} \\
\text{Sita-dat} & | \text{Rina-erg} & \text{book} & \text{buy-inf} & \text{told} \\
\text{‘It was Sita who Rina told to buy a book.’} 
\end{align*}
\]

c. C. **Canonical, case marking on final NP**

\[
\begin{align*}
\text{Riinaa-ne} & | \text{Siitaa-ko} | \text{kitaab-ko} | \text{khariid-neko} | \text{kahaa} \\
\text{Rina-erg} & | \text{Sita-dat} & \text{book-acc} & \text{buy-inf} & \text{told} \\
\text{‘Rina told Sita to buy the book.’} 
\end{align*}
\]

d. D. **IO Fronted, case marking on final NP**

\[
\begin{align*}
\text{Siitaa-ko} & | \text{Riinaa-ne} | \text{kitaab-ko} | \text{khariid-neko} | \text{kahaa} \\
\text{Sita-dat} & | \text{Rina-erg} & \text{book-acc} & \text{buy-inf} & \text{told} \\
\text{‘It was Sita who Rinaa told to buy the book.’} 
\end{align*}
\]

As in earlier experiments, all but the direct object NP were proper names; the direct object NP was always an inanimate noun. The fillers consisted of various
syntactic structures, such as: relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses (all the stimuli are in Appendix A). This was a self-paced moving window reading task (Just et al., 1982); the procedure was as described on pages 71 to 71 for Experiment 2.

4.2.2 Results: The effect of definiteness marking (canonical word order)

The main regions of interest were the third NP and the innermost verb; the third NP was either unmarked or marked for definiteness (see (67)).

(67)  

a. A. Canonical, no case marking on final NP

Riinaa-ne | Siitaak | kitaab | khariid-neko | kahaa
Rina-erg  Sita-dat  book  buy-inf  told
‘Rina told Sita to buy a book.’

b. C. Canonical, case marking on final NP

Riinaa-ne | Siitaak | kitaab-ko | khariid-neko | kahaa
Rina-erg  Sita-dat  book-acc  buy-inf  told
‘Rina told Sita to buy the book.’

For each of the two regions, two sets of contrast analyses were carried out (one using raw reading times and another using residuals). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 4.1 to 4.3, and are summarized graphically in Figures 4.2 and 4.3. The contrast analysis for the final, direct-object NP showed that the RT was significantly longer when the NP was case marked in the case of raw mean RTs; the difference approaches significance in the by-subjects analysis of the residuals mean RTs. The linear mixed-effects model analysis also showed a main effect of case marking on the final NP; no main effect of word length was found. The contrast analysis for the innermost verb showed a significant increase in RT when the final NP was case marked.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>19.89</td>
<td>0.0000</td>
<td>14.04</td>
<td>0.0013</td>
</tr>
<tr>
<td>Residuals</td>
<td>3.16</td>
<td>0.0817</td>
<td>2.73</td>
<td>0.1139</td>
</tr>
</tbody>
</table>

Table 4.1: Expt. 5 (Conditions A vs. C): contrast analysis for NP3

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP3</td>
<td>18.01528</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word length of NP3</td>
<td>1.64959</td>
<td>0.1995</td>
</tr>
</tbody>
</table>

Table 4.2: Expt. 5 (Conditions A vs. C); ANOVA based on linear mixed effects model for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>20.36</td>
<td>0.0000</td>
<td>14.69</td>
<td>0.0010</td>
</tr>
</tbody>
</table>

Table 4.3: Expt. 5 (Conditions A vs. C); contrast analysis for V2
Figure 4.2: Expt. 5, canonical order with case marking absent or present on final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals

Figure 4.3: Expt. 5, canonical order with case marking absent or present on final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals
4.2.3 Results: The effect of definiteness marking (fronted indirect objects)

The main regions of interest were the third NP and the innermost verb; the third NP was either unmarked or marked for definiteness (see (68)).

(68) a. B. **IO Fronted, no case marking on final NP**

Siitaa-ko | Riinaa-ne | kitaab | khariid-neko | kahaa
Sita-dat Rina-erg book buy-inf told
‘It was Sita who Rina told to buy a book.’

b. D. **IO Fronted, case marking on final NP**

Siitaa-ko | Riinaa-ne | kitaab-ko | khariid-neko | kahaa
Sita-dat Rina-erg book-acc buy-inf told
‘It was Sita who Rinaa told to buy the book.’

For each of the two regions, two sets of contrast analyses were carried out (one using raw reading times and another using residuals). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 4.4 to 4.6, and are summarized graphically in Figures 4.4 and 4.5. The contrast analysis for the final, direct-object NP showed that the RT was significantly longer when the NP was case marked both in the case of raw and residual mean RTs. The linear mixed-effects model analysis also showed a main effect of case marking on the final NP; no main effect of word length was found. The contrast analysis for the innermost verb showed no significant increase in RT when the final NP was case marked.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>32.33</td>
<td>0.0000</td>
<td>23.78</td>
<td>0.0001</td>
</tr>
<tr>
<td>Residuals</td>
<td>12.81</td>
<td>0.0008</td>
<td>8.46</td>
<td>0.0087</td>
</tr>
</tbody>
</table>

Table 4.4: Expt. 5 (Conditions B vs. D); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,50)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP3</td>
<td>34.98582</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Word length of NP3</td>
<td>1.61788</td>
<td>0.2039</td>
</tr>
</tbody>
</table>

Table 4.5: Expt. 5 (Conditions B vs. D); ANOVA based on linear mixed-effects model for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>1.94</td>
<td>0.1705</td>
<td>1.88</td>
<td>0.1852</td>
</tr>
<tr>
<td>Residuals</td>
<td>2.08</td>
<td>0.1554</td>
<td>1.89</td>
<td>0.1847</td>
</tr>
</tbody>
</table>

Table 4.6: Expt. 5 (Conditions B vs. D); contrast analysis for V2
Figure 4.4: Expt. 5, IO fronted, case marking absent or present on final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals

Figure 4.5: Expt. 5, IO fronted, case marking absent or present on final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals
4.2.4 Results: The effect of increasing distance between a verb and its arguments (bare direct object)

The main region of interest was the innermost verb; the third NP was unmarked for definiteness in both conditions (see (69)).

(69) a. A. Canonical, no case marking on final NP
    Riinaa-ne | Siitaa-ko | kitaab | khariid-neko | kahaa
    Rina-erg  Sita-dat  book  buy-inf  told
    ‘Rina told Sita to buy a book.’

b. B. IO Fronted, no case marking on final NP
    Siitaa-ko | Riinaa-ne | kitaab | khariid-neko | kahaa
    Sita-dat  Rina-erg  book  buy-inf  told
    ‘It was Sita who Rina told to buy a book.’

Two sets of contrast analyses were carried out (one using raw reading times and another using residuals). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 4.7 to 4.9, and are summarized graphically in Figure 4.6. The contrast analysis for the innermost verb showed a significant increase in RT when the indirect-object NP was fronted. In addition, Figure 4.6 suggests that the NP in second position in the fronted IO sentence has a longer RT than in the canonical order sentence. Accordingly, one set of contrast analyses was performed for this position. These showed that RT at the second NP was significantly longer when the indirect object NP was fronted. Moreover, a contrast analysis for the NP in first position also showed a significant increase in RT in the fronted IO case.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>6.67</td>
<td>0.0129</td>
<td>7.08</td>
<td>0.0150</td>
</tr>
</tbody>
</table>

Table 4.7: Expt. 5 (Conditions A vs. B); contrast analysis for first position NP

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>7.38</td>
<td>0.0092</td>
<td>8.75</td>
<td>0.0078</td>
</tr>
</tbody>
</table>

Table 4.8: Expt. 5 (Conditions A vs. B); contrast analysis for second position NP

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>4.59</td>
<td>0.0373</td>
<td>6.11</td>
<td>0.0226</td>
</tr>
</tbody>
</table>

Table 4.9: Expt. 5 (Conditions A vs. B); contrast analysis for V2

Figure 4.6: Expt. 5, canonical versus IO-fronted order, bare final NP (Conditions A vs. B); raw RTs, with 95% confidence intervals
4.2.5 Results: The effect of increasing distance between a verb and its arguments (definite-marked direct object)

The main region of interest was the innermost verb; the third NP was unmarked for definiteness in both conditions (see (70)).

(70) a. C. Canonical, case marking on final NP

Rinaa-ne | Siitaa-ko | kitaab-ko | khariid-neko | kahaa
Rina-erg Sita-dat book-acc buy-inf told
‘Rina told Sita to buy the book.’

b. D. IO Fronted, case marking on final NP

Siitaa-ko | Rinaa-ne | kitaab-ko | khariid-neko | kahaa
Sita-dat Rina-erg book-acc buy-inf told
‘It was Sita who Rinaa told to buy the book.’

Two sets of contrast analyses were carried out (one using raw reading times and another using residuals). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 4.10 to 4.13, and are summarized graphically in Figure 4.7. The results indicate that the RT at the second NP was significantly longer when the indirect object NP was fronted. The RT at the third position NP approaches significance, but only in the by-subjects analysis. No difference in RT was found at the innermost verb (V2). Finally, the contrast analysis for the final verb (V1) shows a significant increase in RT in the fronted IO case, but only in the by-subjects analysis.
Table 4.10: Expt. 5 (Conditions C vs. D); contrast analysis for second position NP

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>6.58</td>
<td>0.0135</td>
<td>5.48</td>
<td>0.0297</td>
</tr>
</tbody>
</table>

Table 4.11: Expt. 5 (Conditions C vs. D); contrast analysis for third position NP

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>3.83</td>
<td>0.0562</td>
<td>2.80</td>
<td>0.1096</td>
</tr>
</tbody>
</table>

Table 4.12: Expt. 5 (Conditions C vs. D); contrast analysis for V2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>0.10</td>
<td>0.7581</td>
<td>0.06</td>
<td>0.8037</td>
</tr>
</tbody>
</table>

Table 4.13: Expt. 5 (Conditions C vs. D); contrast analysis for V1

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>4.95</td>
<td>0.0308</td>
<td>2.24</td>
<td>0.1503</td>
</tr>
</tbody>
</table>
Figure 4.7: Expt. 5, canonical versus fronted IO, case-marked final NP (Conditions C vs. D); raw RTs, with 95% confidence intervals
4.2.6 Discussion

This experiment replicates the definiteness effect discussed in the previous chapter. A longer RT is observed at the definite-marked NP in both canonical and fronted IO sentences (see Section 4.2.2 on page 115, and Section 4.2.3 on page 118). However, when the final NP is manipulated for case marking, a significantly longer RT at the innermost verb (V2) is observed only in the canonical order sentences (where the two ko-marked NPs are adjacent), and not in the indirect-object fronted sentences (where the two ko-marked NPs are non-adjacent).

Hawkins’ and Gibson’s models can account for neither the definiteness effect nor the different results for V2. By contrast, although Lewis’ RIT cannot account for the definiteness effect, his positional similarity hypothesis can explain the slowdown at the verb: if the bare final NP in fronted IO sentences (i.e., sentences with NPs occurring in the order NP2-ko NP1-ne NP3) receives definiteness marking (i.e., if the NPs are now NP2-ko NP1-ne NP3-ko), Lewis’ hypothesis predicts that there will be less interference due to positional similarity when retrieving the fronted indirect object, since it is now further away from the definite-marked NP. On the other hand, when the NPs are in canonical order and the direct object is bare (i.e., the NPs are now NP1-ne NP2-ko NP3), adding definiteness marking on the direct object (so that the configuration is NP1-ne NP2-ko NP3-ko) should result in a significantly longer RT at the verb compared to the fronted IO case.

Although a significant interaction was not found between word order and case marking (F1(1,51)= 2.33, p = 0.1336; F2(1,23)= 1.76 p = 0.1990), the result for the innermost verb discussed in the preceding paragraph (a significantly longer RT at the verb only in the canonical case when case marking is manipulated) is correctly predicted by Lewis’ model. This is independent corroboration for the conclusion in the previous chapter that although the slowdown at the definite-marked NP itself can only be attributed to discourse constraints, any slowdown observed at V2 must

126
be due to similarity-based interference (which subsumes positional similarity-based interference). In this connection, recall that in conditions C and D of Experiment 3, where the definiteness effect was separated entirely from similarity-based interference, no slowdown at V2 was found when the direct object was definite marked (see Section 3.3.2 on page 88).

There are two sets of results relating to the various distance hypotheses of the three models. One concerns sentences with bare direct object NPs (see Section 4.2.4 on page 121). The result here was that in the comparison of sentences with fronted IOs versus canonical order, NPs in positions one and two have significantly longer RTs in the fronted IO case. These longer RTs cannot be accounted for by any of the models, but there is a straightforward explanation.

Regarding position one, the fronted, dative-marked indirect object NP is a possible dative subject. Since dative subjects occur with only experiencer verbs (Verma & Mohanan, 1990), the slowdown at position one could simply be a reflection of the lower frequency of occurrence of dative subjects. The slowdown at position two, however, probably reflects a mild garden path: since a dative NP has been seen first and is in the initial (subject) position, it is assumed to be the subject of the sentence. Then the ergative case-marked NP is seen, and this implies that the NP in position one cannot be a subject, it must be an object. This would explain the slowdown at position two. This explanation is consistent with Gibson’s and Lewis’ models, but not with Hawkins’. Moreover, a longer RT is observed at the innermost verb in the fronted IO sentences. This is consistent with the distance hypotheses of Gibson and Hawkins, but is inconsistent with Lewis’, which predicts that there should be no difference in RT at the verb (since there is no change in the total amount of interference).

In other words, Gibson’s and Hawkins’ distance hypotheses do appear to make several correct predictions. I defer discussion (page 154) of an alternative explanation that does not involve distance but similarity-based interference.
The second result relating to the distance hypothesis concerns sentences with definite-marked direct object NPs. Here, the result was that NPs in position two and three have significantly longer RTs in the fronted IO case, although in position three the ANOVA results show only a near-significant effect. The slowdown at position two can be accounted for in terms of reanalysis, as discussed above. The near-significant slowdown at position three is possibly due to a spillover from the slowdown at position two. All this is consistent with the discussion on page 126, which discusses other evidence that similarity-based interference occurs during retrieval.

Another important result relating to sentences with definite-marked direct object NPs is that, in the fronted IO case (compared to canonical order), no significant difference was found in the RT at V2, but the RT at V1 was longer. It could be that the slowdown at V1 is a spillover from increased processing difficulty at V2; this would be consistent with Hawkins’ and Gibson’s distance hypotheses, but inconsistent with Lewis’ positional similarity hypothesis.

In other words, there appears to be some evidence for positional similarity; but there is also some evidence that not only favors Hawkins and Gibson’s theories but is inconsistent with Lewis’ positional similarity. It cannot be that Lewis’ positional similarity hypothesis is both right and wrong. Experiment 7 resolves this apparent contradiction: I argue there that similarity-based interference is not limited to syntactic similarity. If similarity is construed broadly, the apparent support for the distance hypotheses of Gibson and Hawkins receives an alternative explanation (although this does not mean that the distance hypotheses should be rejected. I defer a detailed discussion of this issue to the next chapter (Section 5.4, page 151)).

The next experiment investigates the distance effect by comparing canonical order sentences with sentence with fronted direct objects.
4.3  Experiment 6: Direct-object fronting

4.3.1  Method

Subjects

Sixty\(^3\) undergraduate and graduate students at Jawaharlal Nehru University, New Delhi participated in the experiment. The subjects were from a variety of majors (Economics, Foreign Languages, Hindi, International Studies, etc.). Each subject was paid Rupees 100 (equivalent approximately to two US dollars in September 2001) for participating. None of the subjects had participated in the previous experiments.

Materials and Procedure

Four lists were constructed with a counterbalanced design as in earlier experiments, with 44 fillers arranged in pseudorandomized order between 32 target sentences (i.e., six items per condition). The four conditions are illustrated below.

\[(71) \hspace{2em} \text{a. A. Canonical, no case marking on direct object NP} \]

\[
\begin{align*}
\text{riinaa-ne} & \mid \text{siitaa-ko} \mid \text{kitaab} \mid \text{khariid-neko} \mid \text{kahaa} \\
\text{Rina-erg} & \mid \text{Sita-dat} \mid \text{book} \mid \text{buy-inf} \mid \text{told} \\
\text{‘Rina told Sita to buy a book.’}
\end{align*}
\]

\[(71) \hspace{2em} \text{b. B. DO Fronted, no case marking on direct object NP} \]

\[
\begin{align*}
\text{kitaab} & \mid \text{riinaa-ne} \mid \text{siitaa-ko} \mid \text{khariid-neko} \mid \text{kahii} \\
\text{book} & \mid \text{Rina-erg} \mid \text{Sita-dat} \mid \text{buy-inf} \mid \text{told} \\
\text{‘It was the book that Rina told Sita to buy.’}
\end{align*}
\]

\(^3\)The data from eight subjects was excluded from the data analysis since they had fewer than 70\% correct responses to the yes/no comprehension questions.
c. C. Canonical, case marking on direct object NP

riinaa-ne | siitaa-ko | kitaab-ko | khariid-neko | kaha
Rina-erg Sita-dat book-acc buy-inf told
‘Rina told Sita to buy the book.’

d. D. DO Fronted, case marking on direct object NP

kitaab-ko | riinaa-ne | siitaa-ko | khariid-neko | kaha
book-acc Rina-erg Sita-dat buy-inf told
‘It was the book that Rina told Sita to buy.’

As in earlier experiments, all but the direct object NP were proper names; the direct object NP was always an inanimate noun. The fillers consisted of various syntactic structures, such as relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses (all the stimuli are in Appendix A). This was a self-paced moving window reading task (Just et al., 1982); the procedure was as described on pages 71 to 71 for Experiment 2.
4.3.2 Results: The effect of definiteness marking (canonical order)

The main regions of interest were the third NP and the innermost verb; the third NP was either unmarked or marked for definiteness (see (72)).

(72) a. A. Canonical, no case marking on direct object NP

riinaa-ne | siitaa-ko | kitaab | khariid-neko | kahaa
Rina-erg  Sita-dat  book   buy-inf   told
   ‘Rina told Sita to buy a book.’

b. C. Canonical, case marking on direct object NP

riinaa-ne | siitaa-ko | kitaab-ko | khariid-neko | kahaa
Rina-erg  Sita-dat  book-acc buy-inf told
   ‘Rina told Sita to buy the book.’

For each of the two regions, two sets of contrast analyses were carried out (one using raw reading times and another using residuals). Each set involved two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 4.14 to 4.18, and are summarized graphically in Figures 4.8 and 4.9. The contrast analysis for the final, direct-object NP showed that the RT was significantly longer when the NP was case marked in the case of raw mean RTs; this was only true for the by-subjects analysis. The linear mixed-effects model analysis also showed a main effect of case marking on the final NP, and there was a main effect of word length. The contrast analysis for the innermost verb showed a significant increase in RT when the final NP was case marked, but only in the by-subjects analysis. In addition, Figures 4.8 and 4.9 suggest a longer RT in Condition C at the first NP and the final verb, so contrast analyses were performed for these regions as well. There was a significantly longer RT at NP1 and at V1; in the latter case, this was only true for the by-subjects analysis.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>4.36</td>
<td>0.0422</td>
<td>21.31</td>
<td>0.0002</td>
</tr>
</tbody>
</table>

Table 4.14: Expt. 6 (Conditions A vs. C); contrast analysis for NP1

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>20.38</td>
<td>&lt; 0.0001</td>
<td>1.00</td>
<td>0.3291</td>
</tr>
<tr>
<td>Residuals</td>
<td>2.72</td>
<td>0.1059</td>
<td>1.22</td>
<td>0.2823</td>
</tr>
</tbody>
</table>

Table 4.15: Expt. 6 (Conditions A vs. C); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,50)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on NP3</td>
<td>16.74328</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Word length of NP3</td>
<td>4.36637</td>
<td>0.0371</td>
</tr>
</tbody>
</table>

Table 4.16: Expt. 6 (Conditions A vs. C); ANOVA based on linear mixed-effects model for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>4.52</td>
<td>0.0386</td>
<td>2.34</td>
<td>0.1418</td>
</tr>
</tbody>
</table>

Table 4.17: Expt. 6 (Conditions A vs. C); contrast analysis for V2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>4.77</td>
<td>0.0339</td>
<td>0.20</td>
<td>0.6603</td>
</tr>
</tbody>
</table>

Table 4.18: Expt. 6 (Conditions A vs. C); contrast analysis for V1

132
Figure 4.8: Expt. 6, canonical order, case marking absent or present on final NP (Conditions A vs. C); raw RTs, with 95% confidence intervals

Figure 4.9: Expt. 6, canonical order, case marking absent or present on final NP (Conditions A vs. C); residual RTs, with 95% confidence intervals
4.3.3 Results: The effect of definiteness marking (direct object fronted)

The main regions of interest were the direct-object NP in first position and the
innermost verb; the NP was either unmarked or marked for definiteness (see (73)).

(73) a. B. DO Fronted, no case marking on direct object NP

kitaab | riinaa-ne | siitaa-ko | khariid-neko | kahii
book Rina-erg Sita-dat buy-inf told
‘It was the book that Rina told Sita to buy.’

b. D. DO Fronted, case marking on direct object NP

kitaab-ko | riinaa-ne | siitaa-ko | khariid-neko | kahaa
book-acc Rina-erg Sita-dat buy-inf told
‘It was the book that Rina told Sita to buy.’

For each of the two regions, two sets of contrast analyses were carried out (one
using raw reading times and another using residuals). Each set involved two separate
repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time
(RT).

The statistical results are as shown in Tables 4.19 to 4.21, and are summarized
graphically in Figures 4.10 and 4.11. The contrast analysis for the direct-object NP
(here in position one) showed that RT was significantly longer when the NP was case
marked for raw mean RTs; the difference approaches significance in the by-subjects
analysis of the residuals mean RTs. However, the linear mixed-effects model analysis
showed no effect of case marking on the direct-object NP; no main effect of word length
was found. The contrast analysis for the innermost verb showed a near-significant
increase in RT when the fronted direct object NP was definite-marked, but only in
the by-subjects analysis.
<table>
<thead>
<tr>
<th>Fixed Effect</th>
<th>F1(1,50)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case marking on first NP position</td>
<td>0.5130</td>
<td>0.4741</td>
</tr>
<tr>
<td>Word length of first NP position</td>
<td>1.0428</td>
<td>0.3844</td>
</tr>
</tbody>
</table>

Table 4.20: Expt. 6 (Conditions B vs. D); ANOVA based on linear mixed-effects model

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>3.41</td>
<td>0.0709</td>
<td>2.34</td>
<td>0.1418</td>
</tr>
</tbody>
</table>

Table 4.21: Expt. 6 (Conditions B vs. D); contrast analysis for V2
Figure 4.10: Expt. 6, DO fronted, case marking absent or present on final NP (Conditions B vs. D); raw RTs, with 95% confidence intervals

Figure 4.11: Expt. 6, DO fronted, case marking absent or present on final NP (Conditions B vs. D); residual RTs, with 95% confidence intervals
4.3.4 Results: The effect of direct-object fronting (bare direct object)

The main region of interest was the innermost verb (see (74)).

(74)  

a. **Canonical, no case marking on direct object NP**

riinaa-ne | siitaa-ko | kitaab | khariid-neko | kahaa  
Rina-erg | Sita-dat | book | buy-inf | told  
‘Rina told Sita to buy a book.’

b. **DO Fronted, no case marking on direct object NP**

kitaab | riinaa-ne | siitaa-ko | khariid-neko | kahii  
book | Rina-erg | Sita-dat | buy-inf | told  
‘It was the book that Rina told Sita to buy.’

A contrast analysis were carried out, with two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Table 4.22, and are summarized graphically in Figure 4.12. The contrast analysis for the innermost verb showed a significantly longer RT in the non-canonical order sentence.
Table 4.22: Expt. 6 (Conditions A vs. B); contrast analysis for V2

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>12.18</td>
<td>0.0010</td>
<td>10.93</td>
<td>0.0035</td>
</tr>
</tbody>
</table>

Figure 4.12: Expt. 6, canonical versus fronted DO, bare final NP (Conditions A vs. B); raw RTs with 95% confidence intervals
4.3.5 Results: The effect of direct-object fronting (definite-marked direct object)

The main region of interest was the innermost verb (see (75)).

(75)  a. **C. Canonical, case marking on direct object NP**

riinaa-ne | siitaa-ko | kitaab-ko | khariid-neko | kaha
Rina-erg  Sita-dat  book-acc  buy-inf    told
‘Rina told Sita to buy the book.’

b. **D. DO Fronted, case marking on direct object NP**

kitaab-ko | riinaa-ne | siitaa-ko | khariid-neko | kaha
book-acc  Rina-erg  Sita-dat  buy-inf    told
‘It was the book that Rina told Sita to buy.’

The statistical results are as shown in Table 4.23, and are summarized graphically in Figure 4.13. A contrast analysis were carried out, with two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT). The contrast analysis for the innermost verb showed a significantly longer RT in the non-canonical order sentence.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>8.71</td>
<td>0.0049</td>
<td>7.65</td>
<td>0.0119</td>
</tr>
</tbody>
</table>

Table 4.23: Expt. 6 (Conditions C vs. D); contrast analysis for V2

Figure 4.13: Expt. 6, canonical versus fronted DO, case-marked final NP (Conditions C vs. D); raw RTs in single embeddings, with 95% confidence intervals
4.3.6 Discussion

The definiteness effect was replicated: in canonical order sentences the definite-marked direct object NP had a longer RT than the bare direct object NP. As I have argued in detail earlier (see Section 4.2.6, page 126) the longer RT observed at the innermost verb is plausibly due to similarity-based interference. The longer RT observed at the final verb is possibly a spillover effect. Regarding the longer RT at the first NP, the stimuli presentation was partly counterbalanced, so the items being compared in position one are identical and no difference in RT was expected. This significant slowdown at position one can only be ascribed to order effects, since the only difference between the two conditions for position one is the order of occurrence of the target sentence in the list of sentences presented to the subject.

Interestingly, in sentences with fronted direct objects, there was no difference in RT for bare versus definite-marked NPs in position one. The definiteness effect has been replicated in the previous experiments already, but in each case the direct object NP was always in third or fourth position. However, this effect is no longer seen when, as is the case in Experiment 6, this NP is in position one. The critical difference between these two situations is the number of discourse referents seen before the definite-marked NP is seen. In the previous experiments, two or three NPs were seen before the definite NP, but in the second case, the definite NP was the first one seen. This suggests that the number of discourse referents already seen is an important factor in determining processing complexity.

In other words, this result is consistent with the discourse-based component of DLT’s complexity metric and provides independent justification for counting discourse referents in computing processing complexity. Note, however, that merely

\[ \text{\footnotesize (4.3.6 Discussion)} \]

\[ \footnotesize \text{Interestingly, in sentences with fronted direct objects, there was no difference in RT for bare versus definite-marked NPs in position one.} \]

\[ \footnotesize (4.3.6 Discussion) \]

\[ \footnotesize The definiteness effect has been replicated in the previous experiments already, but in each case the direct object NP was always in third or fourth position. However, this effect is no longer seen when, as is the case in Experiment 6, this NP is in position one. The critical difference between these two situations is the number of discourse referents seen before the definite-marked NP is seen. In the previous experiments, two or three NPs were seen before the definite NP, but in the second case, the definite NP was the first one seen. This suggests that the number of discourse referents already seen is an important factor in determining processing complexity.} \]

\[ \footnotesize (4.3.6 Discussion) \]

\[ \footnotesize In other words, this result is consistent with the discourse-based component of DLT’s complexity metric and provides independent justification for counting discourse referents in computing processing complexity. Note, however, that merely} \]
counting discourse referents does not suffice, since I have argued using corpus evidence (Chapter 1) that the contribution of a discourse referent’s new or old status crucially depends on whether it occurs in subject or direct object position.

The distance hypotheses of Hawkins and Gibson are also supported by the fact that in canonical order versus fronted DO sentences, RT is longer at the innermost verb in the fronted DO case. For such sentences with bare versus definite direct object NPs, \textit{contra} Lewis’ positional similarity hypothesis, RT at the innermost verb is longer.

There is other evidence as well that could be construed as either inconsistent or consistent with Lewis’ interference theory, depending on the assumption of the model. On the one hand, if superficially identical case marking is assumed to cause interference, Lewis’ model predicts that the canonical order NP1-ne NP2-ko NP3-ko will result in one unit of proactive interference, along with one unit of positional similarity-based interference. By contrast, the sequence NP1-ne NP2-ko NP3 involves no interference of any kind. By contrast, the sequence NP3-ko NP1-ne NP2-ko involves one unit of proactive interference (from the fronted, case-marked NP3), but no (or insignificant amounts of) positional similarity-based interference, whereas the sequence NP3 NP1-ne NP2-ko involves no interference of any kind. It follows that there should be an interaction between case marking and word order. But no such interaction was found ($F_{1(1,52)} = 0.01, p = 0.9406; F_{2(1,23)} = 0.00, p = 0.9927$). On the other hand, it could be the case that superficial identity of case marking is irrelevant in Lewis’ model; in that case, the absence of an interaction is correctly predicted by his model. However, if we assume that superficial identity of case markers does not cause interference at all, it may then be difficult to account for the results in Experiment 3 (see page 103), where some evidence was found for non-adjacent but superficially similarly case-marked NPs (possibly) resulting in longer RTs at verbs.
As mentioned earlier, the experiments discussed so far lead to apparently contradictory conclusions about Lewis’ model: on the one hand, there is support for RIT; on the other, RIT cannot account for the apparent distance effects of the present experiment. I resolve this issue with Experiment 7 in the next chapter.

Finally, DLT’s predicts that fronting direct objects would result in greater processing difficulty at the innermost verb compared to fronted indirect objects. Although the difference was in the predicted direction, a t-test for reading times in Condition B of Experiment 5 and 6 gave a null result: \( t = -0.498, p = .621 \). It is of course possible that the prediction is correct and the between-subjects comparison simply did not have sufficient power, but with the present design nothing can be concluded regarding DLT’s prediction.

I present next the final experiment, which investigates the distance hypothesis (in its various instantiations) from a different perspective than reordering arguments.
CHAPTER 5

INCREASING DISTANCE CAN FACILITATE PROCESSING:
EXPERIMENT 7

When getting my nose in a book
Cured most things short of school,
It was worth ruining my eyes
To know I could still keep cool,
And deal out the old right hook
To dirty dogs twice my size.

Later, with inch-thick specs,
Evil was just my lark:
Me and my cloak and fangs
Had ripping times in the dark.

From A Study of Reading Habits, Philip Larkin, 1960.

This chapter investigates the distance hypothesis in sentences with canonical order. Early Immediate Constituents and Retrieval Interference Theory (RIT-01) predict an increase in processing difficulty at the innermost verb if an adverb intervenes between the final NP and verb (see pages 50, 63). For EIC, this is due to an increase in the number of words that must be seen before the head of the verb phrase, the innermost verb, is recognized. RIT-01 assumes that if an adverb is present, at the innermost verb the items in the focus of attention are the verb itself and the adverb. As a consequence, interference will have a strong effect on the retrieval of an NP, resulting in greater processing difficulty at the verb. The DLT (see page 53) assumes that inserting an adverb should either have no effect on processing complexity (since
no discourse referent is introduced), or should result in increased processing difficulty (if other factors, other than discourse status, affect distance (Gibson, 2000, 107)). These predictions suggest an obvious experiment, which I describe next.

5.1 Method

5.1.1 Subjects

Sixty\(^1\) undergraduate and graduate students at Jawaharlal Nehru University, New Delhi participated in the experiment. The subjects were from a variety of majors (Economics, Foreign Languages, Hindi, International Studies, etc.). Each subject was paid Rupees 100 (equivalent approximately to two US dollars in September 2001) for participating. None of the subjects had participated in any of the previous experiments.

5.1.2 Materials and Procedure

Four lists were prepared in a counterbalanced, Latin Square design, and 44 fillers were inserted between 32 target sentences in pseudorandomized order. The sentence types are as shown in (76) and all the stimuli appear in Appendix A.

(76) a. A. **Single embedding, no intervening adverb:**

<table>
<thead>
<tr>
<th>Sita-erg</th>
<th>Hari-dat</th>
<th>book-acc</th>
<th>buy-inf</th>
<th>told</th>
</tr>
</thead>
</table>
| ‘Sita told Hari to buy the book.’

\(^1\)The data from eight subjects were excluded since they had fewer than 70% correct responses to the yes/no comprehension questions.
b. B. Single embedding, adverb intervenes:

Sita-ne Hari-ko [kitaab-ko jitne-jaldi-ho-sake kharid-neko] kahaa
Sita-erg Hari-dat book-acc as-soon-as-possible buy-inf told
‘Sita told Hari to buy the book as soon as possible.’

c. C. Double embedding, no intervening adverb:

Sita-ne Hari-ko Ravi-ko [kitaab-ko kharid-neko] bol-neko kahaa
Sita-erg Hari-dat Ravi-dat book-acc buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book.’

d. D. Double embedding, adverb intervenes:

Sita-ne Hari-ko Ravi-ko [kitaab-ko jitne-jaldi-ho-sake
Sita-erg Hari-dat Ravi-dat book-acc as-soon-as-possible
kharid-neko] bol-neko kahaa
buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book as soon as possible.’

In each of the test sentences, all but the final NPs were proper names; the final NP was always an inanimate common noun, such as ‘book’ or ‘letter’. The fillers consisted of various syntactic structures, such as relative clauses, medial gapping constructions, simple declaratives, and sentences with that-clauses. This was a self-paced moving window reading task (Just et al., 1982); the procedure was as described on pages 71 to 71 for Experiment 2.
5.2 Results: Adverb insertion in single embeddings

The main region of interest was the innermost verb (see (77)).

(77) a. **Single embedding, no intervening adverb:**

Sītāa-ne ḫaṛi-ko [kītab-ko ḫariid-neko] kāhā

Sītā-erg ḫaṛi-dat book-acc buy-inf told

‘Sītā told ḫaṛi to buy the book.’

b. **Single embedding, adverb intervenes:**

Sītāa-ne ḫaṛi-ko [kītab-ko jītnē-jalḍī-ho-sake khariid-neko] kāhā

Sītā-erg ḫaṛi-dat book-acc as-soon-as-possible buy-inf told

‘Sītā told ḫaṛi to buy the book as soon as possible.’

A contrast analysis was carried out, with two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 5.1 and 5.2, and are summarized graphically in Figure 5.1. The contrast analysis for the innermost verb showed no significant difference in RT for the two conditions. Since Figure 5.1 suggests that the third NP position may have significantly different RTs in the two conditions, a contrast analysis for this position was done as well. The results show a significantly shorter RT in the sentence with the adverb present, but this was only true for the by-subjects analysis.
<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>5.91</td>
<td>0.0189</td>
<td>3.48</td>
<td>0.0768</td>
</tr>
</tbody>
</table>

Table 5.1: Expt. 7 (Conditions A vs. B); contrast analysis for NP3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>2.77</td>
<td>0.1026</td>
<td>2.99</td>
<td>0.0989</td>
</tr>
</tbody>
</table>

Table 5.2: Expt. 7 (Conditions A vs. B); contrast analysis for V2

Figure 5.1: Expt. 7 (Conditions A vs. B); raw RTs, with 95% confidence intervals
5.3 Results: Adverb insertion in double embeddings

The main region of interest was the innermost verb (see (78)).

(78) a. C. Double embedding, no intervening adverb:

Sita-ne Hari-ko Ravi-ko [kitaab-ko khariid-neko] bol-neko kahaa
Sita-erg Hari-dat Ravi-dat book-acc buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book.’

b. D. Double embedding, adverb intervenes:

Sita-ne Hari-ko Ravi-ko [kitaab-ko jitne-jaldi-ho-sake]
Sita-erg Hari-dat Ravi-dat book-acc as-soon-as-possible
khariid-neko bol-neko kahaa
buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book as soon as possible.’

A contrast analysis was carried out, with two separate repeated measures ANOVAs, for subject (F1) and item (F2) means of reading time (RT).

The statistical results are as shown in Tables 5.3 and 5.4, and are summarized graphically in Figure 5.2. The contrast analysis for the innermost verb showed a significantly shorter RT in the sentence with the adverb present, but only in the by-subjects analysis. Since Figure 5.2 suggests that the fourth NP position may have significantly different RTs in the two conditions, a contrast analysis for this position was done as well. However, no significant difference was found at the fourth NP.

\[\text{\footnotesize{\textsuperscript{2}One subject, subject 31, was found to have an abnormally high reading time (2434 msec) for the innermost verb in the adverb-absent case. This subject was removed from the dataset and the ANOVA was done again. The results in this case were: F1(1,51) = 4.3827, p = 0.04150.}}\]
Table 5.3: Expt. 7 (Conditions C vs. D); contrast analysis for NP4

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>2.32</td>
<td>0.1339</td>
<td>1.71</td>
<td>0.2053</td>
</tr>
</tbody>
</table>

Table 5.4: Expt. 7 (Conditions C vs. D); contrast analysis for V3

<table>
<thead>
<tr>
<th>Reading time</th>
<th>F1(1,51)</th>
<th>p-value</th>
<th>F2(1,31)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raw</td>
<td>5.08</td>
<td>0.0288</td>
<td>2.21</td>
<td>0.1526</td>
</tr>
</tbody>
</table>

Figure 5.2: Expt. 7 (Conditions C vs. D); raw RTs, with 95% confidence intervals
5.4 Discussion

Single embeddings failed to show a significant difference at the innermost verb when the adverb was present. This is a null result, with power less than .80, and is therefore inconclusive (Aron & Aron, 1999). But if this were a conclusive result (i.e., if power had been greater than .80) it would be consistent with one instantiation of the DLT (Gibson, 2000, 107). However, in double embeddings, there was a significantly shorter RT at the innermost verb, and this is consistent only with Lewis’ RIT-02, not with any of the other models under consideration.\(^3\)

There are two possible alternative explanations for the speedup at the innermost verb in double embeddings. The first one is that in a self-paced reading task, subjects read words faster and faster as they progress through the sentence (Ferreira & Henderson, 1993).\(^4\) Although this “position” effect is a result of an inadvertent mis-citation of previous research (see Chapter 6 for a detailed discussion), the correct version of the position effect, due originally to (Aaronson & Scarborough, 1976), is still a possible explanation for some of the results in the single embedding case.

Aaronson and Scarborough (1976) argued that for “natural” sentences,\(^5\) subjects tend to read longer sentences faster overall (and not that each successive position became faster and faster, as Ferreira and Henderson (1993) suggest). The Aaronson et al. hypothesis would explain the significantly shorter RT at the final NP in the single embedding sentence with the adverb (see Figure 5.1, page 148). Of course, this is not entirely true for the single embeddings because of the longer RT at the adverb, but it is still a plausible explanation for the overall faster reading time in the single embeddings case. Note that (Aaronson & Scarborough, 1976) restrict their hypothesis

\(^3\)This result is also consistent with my model, which is presented in the final chapter.
\(^4\)I am grateful to Edson Miyamoto for pointing out this potential explanation.
\(^5\)They do not define what “natural” means, but I assume this has the intuitively obvious meaning, whatever that might be.
to “natural” sentences, so one would expect no such effect in the comprehensible but
less than “natural” double embeddings. This is the correct prediction (Figure 5.2).

The other potential explanation is that adverbs may generally be read faster
than NPs, and therefore any spillover from an adverb to a subsequent verb would
be less than a spillover from an NP to a subsequent verb. However, the assumption
that adverbs are generally read faster than NPs is empirically incorrect. Consider the
reading times for the final NP and the adverb in the single embedding’s Condition
B (Figure 5.1 on page 148). A contrast analysis of the reading times at these
two positions shows that there is a statistically significant slowdown at the adverb
following the final NP (F1(1,51)= 9.4357, p = 0.0035; F2(1.23)= 8.4188, p =
0.008824). Moreover, in the double embeddings, although the reading time for the
adverb appears to be faster than that of the final NP (see Figure 5.2, page 150),
a contrast analysis reveals no significant effect (F1(1,51) = 0.5054, p = 0.48057;
F2(1,23)= 0.9303, p = 0.346295). Finally, Speer points out (personal communication)
that independent evidence from self-paced reading and eyetracking studies suggests
that adjuncts are in general read slower than arguments, not faster (Clifton, Speer,

Thus, the spillover argument cannot account for the double embedding facts.
One might argue that adverbs are longer than NPs (at least in this experiment), and
so factoring out word length might change this result. However, as discussed in the
next chapter, word length does not appear to have much effect on reading time in
these particular experiments. Even if, counterfactually, word length did affect reading
time, the spillover argument is that a speedup at the adverb is passed on to the verb.
But since there is no speedup at the adverb (rather, there is a slowdown in the single
embedding case), the word length issue is orthogonal to the present question.
Turning next to similarity-based interference, recall the previous discussion in the context of Experiments 5 and 6 (pages 128, 143), where some evidence was found for positional similarity-based interference, but other evidence supported EIC and DLT’s distance hypotheses and was inconsistent with positional similarity. Experiments 5 and 6 showed that increasing distance between dependents and heads by reordering arguments results in a longer RT at the innermost verb. By contrast, the results of the present experiment show that increasing distance by inserting an adverb results in a shorter RT at the verb; this result is inconsistent with Hawkins’ and Gibson’s distance hypotheses.

However, there is an alternative explanation for the results of Experiments 5 and 6 that resolves these contradictory conclusions about the three models. Recall from Section 1.2.2 (page 15) that the relative similarity of subjects and non-subjects can be characterized according to the Relational, Animacy, and Definiteness scales, repeated below.

(79)  
  a. **Relational Scale**: Subject > Object

  b. **Animacy Scale**: Human > Animate > Inanimate

  c. **Definiteness Scale**:

  Pronoun > Name > Definite > Indefinite Specific > NonSpecific

Comrie (1979) has hypothesized that objects become more similar to subjects if they are not in their canonical configuration with respect to these scales (the canonical configuration for objects is at the right edge of the scales). It is possible that is another factor that determines the canonical configuration of subjects and non-subjects: position. Just as a non-canonical configuration of an object can be defined according to the scales in (79) and can render an object similar to a subject, relocating an object to a canonical subject position makes it similar to a subject. This implies that fronting an indirect object (Experiment 5) or a direct object (Experiment 6) makes it more similar to a subject. Looking only at the results of Experiments 5 and 6, one cannot
choose between Gibson’s and Hawkins’ distance hypotheses and this similarity-based explanation. But looking at the results of Experiment 7 in conjunction with these earlier experiments, it is clear that EIC, DLT’s, and RIT-01’s distance hypotheses make the incorrect prediction for the adverb-insertion results.

Thus, the most plausible explanation for the results of Experiments 5 and 6 is a similarity-based explanation: the fronted element is rendered more similar to a subject, since it is in the canonical position for subjects. This has interesting implications for Lewis’ interference theory. Lewis (1999), following Caplan and Waters (1999), has assumed that syntactic working memory is an independent entity in working memory, and that the constraints in his model(s) apply only, or primarily, to syntactically similar items. The results here suggest that other kinds of similarity are equally relevant in generating interference. As mentioned in Chapter 2 (page 55, footnote 9), Gordon et al. (in press) also present independent evidence that is inconsistent with the view that processing syntactic structure utilizes specialized resources in working memory. Their experiments involved a memory-load task: subjects had to remember a list of three proper names or the role descriptions, and then were shown a subject-extracted or object-extract cleft sentence which contained two NPs; both these NPs were either proper names of descriptions. Gordon et al. have shown previously (2001) that object-extracted clefts are harder to process than subject-extracted clefts (object-extracted clefts lead to higher error rates in comprehension and longer reading times than subject-extracted clefts) Memory interference was manipulated by having the load words and critical words be either matched (both proper names or both descriptions) or unmatched.

In Gordon et al.’s experiments, subjects were shown the list of words (LOAD WORDS) and asked to remember them by reading the items out aloud twice. Then, they read a single sentence presented one word at a time. After that, subjects were required to answer a yes/no comprehension question. Then, subjects were asked to recall the items aloud. For example, the set of proper name load words Joel, Greg,
Andy or description load words poet, cartoonist, voter could be followed by (a) a subject-cleft with proper names like It was Tony that liked Joey before the argument began, or an object-cleft with proper names like It was Tony that Joey liked before the argument began, or (b) a subject-cleft with descriptions like It was the dancer that liked the fireman before the argument began, or an object-cleft with descriptions like It was the dancer that the fireman liked before the argument began.

The key result was that there was a significant interaction in comprehension accuracy between cleft type and match between load words and NPs in the sentence. The Caplan-Waters-Lewis specialization view would predict no such interaction between load words and syntactic complexity of the sentence. The significant interaction found by Gordon et al. is also consistent with other results that suggest that semantic/pragmatic information plays an early and important role in human sentence parsing (see, e.g., (Altmann & Steedman, 1988), (Steedman, 2000)). All this provides independent evidence against syntactic specialization of working memory, and is consistent with my argument that other factors can also cause similarity-based interference.

Finally, regarding the issue (introduced in Chapter 1, pages 10-14) of whether Hindi infinitivals are nominal or verbal, if they are nominal, there is no explanation for why reading time is significantly faster at the innermost verb when the adverb precedes it. On the other hand, if infinitivals are verbs, then the explanation for the speedup is straightforward: the adverb generates a stronger expectation for a verb compared to the case where no adverb is present, and therefore the verb is recognized faster.

This concludes the experimental portion of this dissertation. I present next a detailed discussion regarding the effect (or, rather, the absence thereof) of word length and position on reading time.
CHAPTER 6

WORD LENGTH, POSITION, AND READING TIME: MYTHS
AND MISUNDERSTANDINGS

We see the spring breaking across rough stone
And pause to regard the sky;
But we are pledged to work alone,
To serve, bow, nor ask if or why.

From ‘We see the spring breaking across rough stone’, Philip Larkin, 1939

The planned contrasts in Experiments 2, 3, 5, and 6 involve bare versus definite-marked NPs, and this means that the critical regions will differ by three character lengths (a white space and the two characters – the definiteness marker ko – on the definite marked NP). One question that arises is: is a word-length correction necessary? The assumption in the psycholinguistic literature is that word length does significantly affect reading time (see, e.g.,(Trueswell et al., 1994)).

Since the word length assumption is based on research involving English, it is necessary to establish whether word length effects appear in Hindi as well. Perhaps surprisingly, it turns out that although word length has an effect on reading time in Hindi, it not as large as in English. One important consequence is that word length correction should in general be carried out, but in the particular case of the experiments presented here, the extra three-character lengths of the definite-marked NP are unlikely to contribute much to reading time.
Position has also been claimed to affect reading time significantly: for example, (Ferreira & Henderson, 1993) speculate that reading time becomes shorter as the reader advances further in a sentence. If true, this has important implications for the result in Experiment 7, where a faster reading time was seen at the innermost verb in position six, compared to the innermost verb in position five. However, I show in this chapter that (a) the claim about position effects resulted from a misinterpretation of the actual claims; and (b) consistent with the existing literature on position effects in word-by-word self-paced reading, there is no evidence that position has a significant effect on reading time. Accordingly, the results in Experiment 7 cannot be attributed to position.

The purpose of this chapter is to provide the details behind these conclusions. The structure of this chapter is as follows. Section 6.1 presents relevant details about Hindi orthography. Section 6.2 presents evidence that word length does affect reading time, but to a smaller extent than in English. Section 6.3 shows that, contra the conjecture in (Ferreira & Henderson, 1993), there is no evidence in the Hindi experiment data that position affects reading time.

6.1 Hindi orthography

This brief discussion of the history and details of Hindi orthography is based on (Singh, 1991), (Vaid & Gupta, to appear), (Pandey, 2000), and (ISCI, 1991); the reader interested in the full details of Devanagari\(^1\) should consult these references.

The Brahmi script emerged around the third century in Ashokan inscriptions, and gave rise to northern and southern variants (Gupta and Grantha respectively). Devanagari, belongs to the northern group, and has been in use for languages like Sanskrit, Marathi, Gujarati, Nepali, and Hindi. The orthographic unit of Devanagari

\(^{1}\)The etymology probably is: Deeva, ‘heavenly’, and nagari, ‘script of the city’; but see (Singh, 1991) for alternative explanations.
is the akśara; this corresponds to the modern mora. A partial inventory is shown in Figure 6.1. There is no upper-case and lower-case distinction in Devanagari, and it is written from left to right. In Hindi, spaces separate adjacent words and certain bound morphemes like postpositions/case markers.

An important fact about Devanagari is that certain CCV clusters, such as the sj in masjid, मस्जिद, ‘mosque’, are represented by a conjunct consonant, स्ज, which is assembled from the two akśara स and ज. Furthermore, CV clusters are not always pronounced in the linear order in which they occur. For example, in kitaaba, किताब, ‘book’, the short vowel i (see Figure 6.1) is written before the consonant k that it follows phonologically. Both situations can hold, as in masjid, मस्जिद. Here, the s, स, and j, ज, combine together to form sj, स्ज, and the short vowel i is written to the left of this conjunct consonant.

These quirks of Hindi orthography – the conjunct consonants and the occasionally non-linear nature of consonant-vowel assembly – prompted Vaid and Gupta (to appear) to ask whether Hindi readers delimit akśara along phonemic boundaries or syllable boundaries. If the former, then greater processing cost should be incurred for words in which characters are conjunct and/or not linearly aligned. If readers engage in syllabic segmentation, then there should be no processing cost for such character segments. Through a series of experiments involving naming latency, they found that there is a significant processing cost associated with conjunct and/or non-linearly aligned consonants (this was independent of the frequency of the various kinds of words; see (Vaid & Gupta, to appear) for details); this suggests that at least adult Hindi speakers engage in phonemic segmentation.
<table>
<thead>
<tr>
<th>Vowels</th>
<th>Occlusives</th>
<th>Sibilants</th>
<th>Aspirate</th>
</tr>
</thead>
<tbody>
<tr>
<td>a &amp; i</td>
<td>ka kha</td>
<td>ya ra</td>
<td>ha h</td>
</tr>
<tr>
<td>e &amp; i</td>
<td>ca cha</td>
<td>la va</td>
<td></td>
</tr>
<tr>
<td>o &amp; u</td>
<td>ta tha</td>
<td>ya ra</td>
<td></td>
</tr>
<tr>
<td></td>
<td>nga gha</td>
<td>ya ra</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sa ssa</td>
<td>ya ra</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sonorants</th>
<th>Sibilants</th>
<th>Aspirate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ya ra</td>
<td>sa ssa</td>
<td>ha h</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Supplementary Consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>qa kha</td>
</tr>
</tbody>
</table>

Figure 6.1: A partial inventory of Devanagari characters
6.2 Does word length affect reading time in Hindi?

6.2.1 Computing word length for Hindi

In 1991 The Bureau of Indian Standards formed a standard known as the Indian Script Code for Information Interchange (ISCII) for the use in all computer and communication media, which allows the usage of 7- or 8-bit characters. In an 8-bit environment, the lower 128 characters are the same as defined in the ISO 646 IRV (International Reference Version) 7-bit coded character set for information interchange, also known as the ASCII character set. The top 128 characters cater to all the Indian Scripts based on the Brahmi script.

ISCII code was used for the display of characters on the computer screen in the experiments presented in this dissertation. For example, the sequence हङ्ग स्व, haasya kavi, “comic poet”, is encoded as a sequence of nine bytes: 0xd8 0xda 0xd7 0xe8 0xd6 0x20 0xb3 0xd4 0xdb. Each byte in isolation would also correspond to a character. Since the ISCII encoding is based on the phonology of a syllable, and Hindi readers engage in phonemic segmentation, it is justifiable to use a byte count to represent phonemic length. This counting was automated by means of a shell script (along with tools like awk and sed).

6.2.2 Reading time and word length – results from Experiment 2

A compelling argument that word length does affect reading time is presented in (Trueswell et al., 1994, 310-315), which demonstrates that “reading times for a scoring region are generally longer when the length of the region is longer.” Using reading time data from an eyetracking experiment, first pass reading times were plotted against word length for all subjects and Pearson R was found to be 0.360, F(1,2521)=375.84, p < 0.001. From this Trueswell et al. correctly concluded that future eye-movement research using reading time measurements should adjust for
word length. Subsequently, this became a common assumption in the literature, and word length is now standardly corrected, not just for eye-tracking studies, but even for self-paced reading time studies such as those described in this dissertation.

Keeping in mind that position has been argued to affect reading time (Ferreira & Henderson, 1993) (on which more in the next section), position was held constant when reading time was plotted against word length using the critical items and fillers (this is described in more detail below). For comparison, plots were also constructed of reading time against word length using all words seen by subjects in all positions.

The plots for Experiment 2 are presented next. First, the results of ANOVA results of linear models (with word length as the independent variable) and the corresponding plot are shown. Then, the ANOVA result and plot are shown for all words in all positions. Pearson correlation coefficients and best-fit lines (the best-fit lines were computed using the R (Ihaka & Gentleman, 1996) library eda (exploratory data analysis), which follows the criteria for robust line-fitting discussed in (Tukey, 1977)). The results were identical for the other experiments, so I restrict the discussion to Experiment 2.

For convenience, the critical sentence types from Experiment 2 are shown below.

(80)  
  a. **Condition A:**

  Siitaa-ne Hari-ko [kitaab khariid-neko] kahaa
  Sita-erg Hari-dat book buy-inf told
  ‘Sita told Hari to buy a book.’

  b. **Condition B:**

  Siitaa-ne Hari-ko [Ravi-ko [kitaab khariid-neko] bol-neko] kahaa
  Sita-erg Hari-dat Ravi-dat book buy-inf tell-inf told
  ‘Sita told Hari to tell Ravi to buy a book.’
c. **Condition C:**

Siitaa-ne Hari-ko [kitaab-ko khariid-neko] kahaa
Sita-erg Hari-dat book-acc buy-inf told
‘Sita told Hari to buy the book.’

d. **Condition D:**

Sita-erg Hari-dat Ravi-dat book-acc buy-inf tell-inf told
‘Sita told Hari to tell Ravi to buy the book.’

Table 6.1 shows the F and p values for the plot of reading time for the fillers and critical words in position three (see Conditions A and C in (80)) against word length. The corresponding figure is Figure 6.2, where the correlation coefficient, 0.2141, is shown. Table 6.2 shows F and p values for fillers and critical words in position four (see Conditions B and D in (80)), and the corresponding plot is Figure 6.3; the correlation coefficient was 0.2337. Table 6.1 shows F and p values for all words in all positions, and the corresponding plot is shown in Figure 6.4; the correlation coefficient was 0.1893.
Table 6.1: Expt. 2, ANOVA of word length (wl), using stimulus words in Conditions A and C and filler words in position three, as a predictor of reading time

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wl</td>
<td>1</td>
<td>25700578.00</td>
<td>25700578.00</td>
<td>84.47</td>
<td>0.0000</td>
</tr>
<tr>
<td>Residuals</td>
<td>1758</td>
<td>534898447.82</td>
<td>304265.33</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.2: Expt. 2; plot of RT against word length (using words from Conditions A, C, and from fillers in position three)
Table 6.2: Expt. 2, ANOVA of word length (wl), using stimulus words in Conditions B and D and filler words in position four, as a predictor of reading time

<table>
<thead>
<tr>
<th></th>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>wl</td>
<td>1</td>
<td>37211088.56</td>
<td>37211088.56</td>
<td>101.54</td>
<td>0.0000</td>
</tr>
<tr>
<td>Residuals</td>
<td>1758</td>
<td>644253297.44</td>
<td>366469.45</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.3: Expt. 2; plot of RT against word length (using words from Conditions B, D, and from fillers in position four)
Table 6.3: Expt. 2, ANOVA of word length (wl), using all words in all positions, as a predictor of reading time

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Residuals</td>
<td>14342</td>
<td>23605318.31</td>
<td>6349707557.87</td>
<td>533.06</td>
</tr>
</tbody>
</table>

Figure 6.4: Expt. 2; plot of RT against word length (all words, ignoring position)
The correlation coefficients of raw reading time against word length for Experiments 2, 3, 5, and 6 are shown below in Table 6.4.

The first column in Table 6.4 gives the experiment number; the second column shows correlation coefficients for the various relevant positions; and the third column shows correlation coefficients from plotting reading time against word length for all words seen by all subjects. Specifically, in Experiment 2, position three (single embeddings) and position four (double embeddings) are relevant and the two correlation values in the second column refer to plots limited to position three and four respectively. Experiment 3 was concerned only with position 4, so there is only one entry in the second column. Experiment 5 involved position three, but there were two distinct conditions, A, C, versus B, D, so two sets of correlation coefficients were computed. In Experiment 6 position three and position one are relevant, and so correlation coefficients in the second column refer to those positions.

<table>
<thead>
<tr>
<th>Expt.</th>
<th>Position held constant</th>
<th>All positions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.2141, 0.2334</td>
<td>0.1893</td>
</tr>
<tr>
<td>3</td>
<td>0.2117</td>
<td>0.2193</td>
</tr>
<tr>
<td>5</td>
<td>0.1366, 0.1423</td>
<td>0.1939</td>
</tr>
<tr>
<td>6</td>
<td>0.1632, 0.2244</td>
<td>0.1977</td>
</tr>
</tbody>
</table>

Table 6.4: Raw reading time correlation coefficients (a) when position is held constant; and (b) when words in all positions are considered

The correlation values in the table show that there is not much difference between keeping position constant and ignoring position when determining the effect of word length on reading time. This suggests that position may not be an important factor in determining word length (this is discussed in the next section). Although there is certainly a correlation between word length and reading time, it is a relatively small effect, certainly not large enough to affect reading time significantly when the difference between the key stimuli words (bare versus definite-marked NPs) consists of three extra characters (one white space and the two characters for the case marker $ko$).
6.2.3 Correcting for word length

The preceding section showed that word length *per se* does have an effect on reading time in the experiments reported here; however, the correlation is not large enough to affect a three-character difference between critical conditions. Even so, in order to obtain interpretable results for planned comparisons of words with unequal length, one could argue that it is necessary to factor out any effect of word length, however small, when carrying out the analysis for the planned comparison. One standard way to do this is to use residuals to carry out an ANOVA: Residual reading time are calculated for each region by subtracting from raw reading times each subject’s predicted reading time for regions with the same numbers of characters; this in turn is calculated from a linear regression equation across all of a subject’s sentences in the experiment (Ferreira & Clifton, 1986; Trueswell et al., 1994).

However, carrying out an ANOVA of residuals results in a model in which the parameter estimates used may not be least-squares estimates of that model (Maxwell et al., 1985). This has the consequence that the test statistic is not appropriately distributed. One practical consequence is that results based on an ANOVA of residuals are misleading: in the experiments presented in Chapters 3 and 4, often the critical comparison between definite-marked and bare NPs shows no significant effect of case marking, even though the ANOVA based on raw reading times shows a large effect of case marking. Given that word length is not highly correlated with reading time, it is rather surprising that three extra characters can cause the effect of case marking on direct objects to disappear.

An alternative is to use ANCOVA (analysis of covariance), which is a model that relates a continuous response (raw reading time in the experiments) to both a classification factor (case marked versus not case marked), and to a continuous covariate (word length). However, it is necessary to distinguish random versus fixed effects in such a model; subjects are random effects but word length and raw reading
times are fixed effects. One way to explicitly incorporate random and fixed effects in a model is to use the (linear) mixed-effects (LME) model (Pinheiro & Bates, 2000). Using LME gives results that are consistent with the known facts about the effect of word length on reading time. In most of the planned comparisons, there was no main effect of word length. In Experiment 2, although a main effect of word length was found (F=4.45, p= 0.04), this was due to a single outlier. Once that outlier was removed, the new values were: F=2.51, p=0.11.

The LME model used was the following. In (81), \( y_{ij} \) is the \( j \)th observation on the \( i \)th group of data and \( x_{ij} \) is the corresponding value of the covariate.

\[
(81) \quad y_{ij} = \beta_i + b_i + \beta_x x_{ij} + \epsilon_{ij}
\]

where \( i = 1, \ldots, M, j = 1, \ldots, n_i, b_i \sim \mathcal{N}(0, \sigma_b^2), \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2) \).

6.3 Does position affect reading time?

It is frequently claimed that the reading time of each successive word becomes shorter as participants progress through a sentence. To quote Ferreira and Henderson:

“Gernsbacher (1990) has summarized a variety of evidence indicating that processing speed increases as comprehenders proceed through a sentence or text: The initial sentence of a passage is read more slowly than subsequent sentences; the initial words of a sentence are read more slowly than later words; phoneme monitoring times are shorter the later the position of the phoneme in the sentence; and event-related potentials indicating difficulty of processing are less pronounced for later words than for earlier words. This tendency is even evident in the processing of nonverbal materials: The initial picture of a series of pictures comprising a story is examined longer than later pictures. Thus, it seems quite likely
that reading speed increases throughout a sentence, at least until the sentence-final word is encountered.” (Ferreira & Henderson, 1993, 261)

Three results ((Aaronson & Scarborough, 1976), (Aaronson & Ferres, 1983), (Chang, 1980)) are cited by Gernsbacher as the source of this claim. To quote Gernsbacher (1990, 7):

“In some experiments, researchers measure how long it takes comprehenders to read each word of a sentence. In these experiments, each word appears in the center of a computer monitor, and subjects press a button to signal when they have finished reading each word. After each word disappears, another one appears. In this way, researchers can measure how long subjects need to read each word.

“A consistent finding in these word-by-word experiments is that initial words take longer to read than later-occurring words (Aaronson & Scarborough, 1976), (Aaronson & Ferres, 1983), (Chang, 1980). In fact, the same words are read more slowly when they occur at the beginning of their sentences or phrases than when they occur later.”

Gernsbacher qualifies this with a footnote (1990, 47):

“This effect is not manifested when subjects are required to memorize (as opposed to comprehend) the stimulus sentences. Neither is the effect manifested when subjects are required to perform a second task immediately after they finish reading each sentence (e.g., answer a question or press a key to signal an anomaly). In preparation for this second task, subjects often delay their reading of the last words of the sentences.”
In fact, the original findings regarding reading time and position appear to be rather different from Gernsbacher’s summary above. The original study was by Aaronson and Scarborough and its conclusions are summarized in the quote below ((Aaronson & Scarborough, 1976, 60); emphasis mine):

“subjects might decrease their RT over serial positions within a sentence. RTs for all sentences would start out about the same, but longer sentences would have more words with fast RTs to decrease the average RT over the sentence. *This is not the case*, as suggested by the data in Figures … Rather, RTs throughout the sentence were faster for longer sentences. This is consistent with … findings that longer “natural” sentences are generally more redundant than short ones throughout the entire sentence.”

The actual claim is thus that among “natural” sentences, in each region longer sentences will have shorter reading times compared to shorter sentences.

The second reference, (Chang, 1980, 64) found “a linear increase in reading time over serial positions.” Chang speculates that the divergent result may be due to that fact that (a) his experiment included probe recognition followed by a comprehension question, whereas the (Aaronson & Scarborough, 1976) study required only a yes/no comprehension question; and (b) his stimuli were in general simpler than those used in Aaronson et al.’s study.

Finally, the third reference cited by Gernsbacher, (Aaronson & Ferres, 1983), mentions position effects on page 678, where they write (emphasis mine): “Reading for memory requires a faithful coding of the “surface” information: Subjects attempt to code every word completely, *they do not speed up with more redundant text* …” This is perhaps the source of the comment in Gernsbacher’s footnote, where she mentions that position effects are not manifested when subjects are required to memorize the stimulus sentences rather than reading for comprehension. Gernsbacher may have
been drawing a conclusion about position effects based on the practice effect described on page 689 of (Aaronson & Ferres, 1983): “The mean RTs decrease significantly with increased skill in button pressing and familiarity with the display of characters . . .”. However, even from this, nothing can be concluded about position affecting reading time, since a global improvement in button-pressing speed over time does not imply anything about changes in button-pressing speed when going from position $n$ to $n+1$.

In sum, Aaronson et al. originally concluded that among “natural” sentences, longer ones will have shorter reading times compared to shorter sentences in each region. They did not claim that reading time becomes faster and faster as the reader progresses through a sentence. Aaronson et al.’s original conclusion does not hold for the single versus double center embeddings in, for example, Experiment 2, but perhaps this is because single embeddings are “natural”, but double embeddings less so. This is clear from Figures 6.5 and 6.6.

Furthermore, as shown in Figures 6.7 and 6.8, although the effect found by Aaronson et al. for “natural” sentences could be argued to be partly borne out in the single embeddings case, the opposite effect is seen in double embeddings: regions in the longer sentences take longer to read. The most important implication of the above discussion is that the Aaronson et al. explanation cannot account for the speedup at position six in double embeddings in the adverb-present case (Figure 5.2), compared to position five in the adverb-absent case. Even if, counterfactually, the Aaronson et al. effect had been seen in double embeddings, it could not have accounted for the speedup at position six, since Aaronson et al. make no claim about non-identical positions in sentences of differing lengths; their claim is only about identical positions.
Figure 6.5: Expt. 2; raw RTs by position in single versus double embeddings, non-case-marked final NP

Figure 6.6: Expt. 2; raw RTs by position in single versus double embeddings, case-marked final NP
Figure 6.7: Expt. 7; single embeddings with adverb present versus absent

Figure 6.8: Expt. 7; double embeddings with adverb present versus absent
In spite of all this, it is still possible in principle that the position effects invoked in (Ferreira & Henderson, 1993) do exist. In order to verify this, I chose three relatively long but easy-to-comprehend filler sentences (see (82)), from different locations (first, thirty eighth, and fiftieth, from a total of seventy six sentences) in the list of sentences seen by each subject in Experiment 6 (this experiment was chosen at random). Three distinct positions were chosen, keeping in mind that practice effects might result in faster and/or more efficient button-pressing during the course of the experiment, and relatively “natural” sentences were chosen in order to reduce the likelihood that increasing processing difficulty masks any position effect. For each sentence, reading times were available from fifty-two subjects. I ran a multiple linear regression, with indicator variables for each position (Winer, Brown, & Michels, 1991, 389, 954-963), (Nolan & Speed, 2000, 203-204).

(82) a. Ravi-ko | lagtaa | hai | ki | vo | haasya kavi | kabhii-naa-kabhii |
Ravi-dat seems is that that comic poet one-day
mashuur | hogaa
famous will-be
‘Ravi feels that that comic poet will become famous one day.’

b. Kaantaa-ne | vo akhbaar | tumhaare-liye | bhejii thii | jo | usko
Kanta-erg that newspaper you-for sent had which him
| kabiir-ne | dii thii
Kabir-erg given had
‘Kanta had sent that newspaper for you which Kabir had given to her.’

c. mujhe | aaj | baink-se | paise | nikaal-ne the | lekin | samay |
me-dat today bank-from money take-out-inf were but time
naa milaa
not got
‘I had to take out money from the bank today, but I didn’t get the time.’

On the next three pages, I show Tukey box-and-whiskers plots (Tukey, 1977); each plot is accompanied with the result of the multiple regression. The variables i1 to i8 are indicator variables.
Figure 6.9: Tukey box-and-whiskers plot for sentence (82a))

|     | Estimate | Std. Error | t value | Pr(>|t|) |
|-----|----------|------------|---------|----------|
| (Intercept) | 699.7692 | 83.5578    | 8.37    | 0.0000   |
| i1  | 183.3846 | 118.1686   | 1.55    | 0.1214   |
| i2  | -174.4423| 118.1686   | -1.48   | 0.1406   |
| i3  | -36.6346 | 118.1686   | -0.31   | 0.7567   |
| i4  | 16.6731  | 118.1686   | 0.14    | 0.8879   |
| i5  | 375.4038 | 118.1686   | 3.18    | 0.0016   |
| i6  | 156.5000 | 118.1686   | 1.32    | 0.1860   |
| i7  | -73.0000 | 118.1686   | -0.62   | 0.5370   |
| i8  | 135.1731 | 118.1686   | 1.14    | 0.2533   |

Table 6.5: t- and p-values for each successive position in the sentence (82a)); Multiple R-squared: 0.1357, Adjusted R-squared (penalizing for higher p): 0.1207.
Figure 6.10: Tukey box-and-whiskers plot for sentence (82b)

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| 806.1923 | 75.3061    | 10.71   | 0.0000   |
| 197.2308 | 106.4989   | 1.85    | 0.0648   |
| -183.3077| 106.4989   | -1.72   | 0.0860   |
| 113.5577 | 106.4989   | 1.07    | 0.2869   |
| -247.5769| 106.4989   | -2.32   | 0.0206   |
| -98.4423 | 106.4989   | -0.92   | 0.3559   |
| 106.2308 | 106.4989   | 1.00    | 0.3191   |
| 225.5962 | 106.4989   | 2.12    | 0.0348   |

Table 6.6: t- and p-values for each successive position in the sentence (82b); Multiple R-squared: 0.0582, Adjusted R-squared (penalizing for higher p): 0.04204.
Figure 6.11: Tukey box-and-whiskers plot for sentence (82c))

|       | Estimate | Std. Error | t value | Pr(>|t|) |
|-------|----------|------------|---------|----------|
| (Intercept) | 527.0000 | 37.0006    | 14.22   | 0.0000   |
| i1     | 30.7308  | 52.4116    | 0.59    | 0.5580   |
| i2     | 72.5962  | 52.4116    | 1.39    | 0.1668   |
| i3     | -97.3462 | 52.4116    | -1.86   | 0.0640   |
| i4     | 45.8269  | 52.4116    | 0.87    | 0.3824   |
| i5     | 50.0577  | 52.4116    | 0.96    | 0.3401   |
| i6     | -92.9615 | 52.4116    | -1.77   | 0.0769   |
| i7     | -18.9423 | 52.4116    | -0.36   | 0.7180   |

Table 6.7: t- and p-values for each successive position in the sentence (82c)); Multiple R-squared: 0.02471, Adjusted R-squared (penalizing for higher p): 0.007976.
The above results show that there is no evidence for the kind of position effects that (Ferreira & Henderson, 1993) posit. Furthermore, in order to visualize the data from a slightly different perspective, for four experiments (Expts. 2, 3, 5, and 6) I plotted reading time against position, using all words in all positions, and aggregating across subjects. If increasing position results in decreasing reading time, one should see a large negative correlation. The plots below (Figures 6.12 to 6.15) show that there is almost no correlation. This provides further evidence against the claim. The plots are shown in the following pages.2

This concludes the discussion about word length and position effects. The next and final chapter summarizes the experiment results, discusses the wider implications of the present work for psycholinguistic research, and presents a new theory of sentence processing that is based on abductive inference.

---

2Van Dyke informs me that in her ongoing research (Van Dyke & Lewis, 2002), (Van Dyke, 2002), a position-related speedup was found in self-paced reading experiments involving English sentences. Although the details of these experiments are not currently available, there are at least two critical differences in the kind of sentences she was investigating compared to the experiments discussed here. First, the English sentences were much longer than the ones my experiments use. Second, the important reading time comparisons were not at position n versus n + 1 as in my Experiment 7, but at positions 7 and 15, and 14 and 18. This suggests that the Ferreira and Henderson’s hypothesis applies only to positions relatively distant from each other and not to adjacent ones; I leave this question for future research.
Figure 6.12: Expt. 2; plot of RT against position

Figure 6.13: Expt. 3; plot of RT against position
Figure 6.14: Expt. 5; plot of RT against position

Figure 6.15: Expt. 6; plot of RT against position
CHAPTER 7

CONCLUDING REMARKS, AND A NEW SENTENCE PROCESSING MODEL

The trees are coming into leaf
Like something almost being said;
The recent buds relax and spread,
Their greenness is a kind of grief.
Is it that they are born again
And we grow old? No, they die too.
Their yearly trick of looking new
Is written down in rings of grain.
Yet still the unresting castles thresh
In fullgrown thickness every May.
Last year is dead, they seem to say,
Begin afresh, afresh, afresh.

_The Trees_, Philip Larkin, 1967

The structure of this chapter is as follows. Section 7.1 presents the overall picture that emerges from the experiment results. This leads to Section 7.2, which describes a new sentence processing model based on abductive inference.

7.1 Multiple constraints in language processing: A summing up

I have shown that EIC cannot account for the effect of definiteness marking and the effect of adverb insertion. Although some evidence, from Experiments 5 and 6, does appear to be consistent with EIC, there is an alternative explanation for these results
in terms of similarity-based interference. The strong position, that pragmatic and semantic effects are epiphenomenal and derivable from EIC, cannot be maintained. Kurz (2000) has independently reached the same conclusion, from a German text-corpus study.

There is some evidence consistent with DLT’s discourse-based metric of integration distance. However, as for Hawkins’ EIC, these results have an alternative explanation which is more plausible given the results of the adverb-insertion experiment: this experiment constitutes evidence against Gibson’s distance-based metric as the sole or main determiner of integration difficulty. That said, there is some evidence supporting the use of the number of discourse referents introduced so far as an index of processing difficulty. Gibson’s assumptions (based on the English experiment described in (Warren, 2001)), that indefinites are harder to process than definites, does not apply to Hindi, where the opposite effect is seen. The asymmetry between English and Hindi is shown to fall out of the observation that, for reasons relating to cross-linguistic defaults in grammatical relations, subjects tend to be definites whereas direct objects tend to be indefinites. This means the DLT’s discourse complexity assumption must be re-formulated: rather than stating that definites are easier to process than indefinites, the assumption should be that subjects tend to be definites and direct objects tend to be indefinites. Specifically, a distinction such as new vs. old, or the number of presuppositions etc. does not suffice.

Some evidence was found for the similarity-based interference hypothesis of RIT-01, but there was not much evidence that adjacency (positional similarity) is an important factor. RIT-02 makes the correct prediction for the adverb-insertion experiment, but RIT-01 does not. Recall that, unlike RIT-01, RIT-02 assumes that retrieval requests are feature bundles, not just structural positions. The only way that RIT can account for the results for Experiments 5 and 6 is by assuming that apart from structural (syntactic) similarity, other kinds of similarity can also cause interference. Since this is consistent with the RIT-02 assumption that feature bundles are involved
in retrieval, the 2002 version of RIT is able to account for all the facts. The superficial similarity effect mentioned above can also be explained under the feature-bundle view of RIT-02, since phonological features can then also cause interference during retrieval. I summarize the empirical results and my conclusions in Table 7.1.

<table>
<thead>
<tr>
<th>Only syntax matters</th>
<th>EIC</th>
<th>DLT</th>
<th>RIT-01</th>
<th>RIT-02'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Increasing distance matters (NP reordering)</td>
<td>✔</td>
<td>✗</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Increasing distance matters (adverb insertion)</td>
<td>✔</td>
<td>✔</td>
<td>✓</td>
<td>✗</td>
</tr>
<tr>
<td>Similarity matters</td>
<td>✗</td>
<td>✗</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Discourse status matters</td>
<td>✗</td>
<td>✓</td>
<td>?</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 7.1: Summary of conclusions. Each cell indicates the prediction of each theory regarding a particular issue; a tick mark (✔) stands for an affirmative answer to questions like “Does only syntax matter?”, while a cross (✗) stands for an answer in the negative. A boxed tick mark or cross indicates that the Hindi results suggest that the answer is either not exactly correct, or is wrong. An up-arrow (↑) indicates that a slowdown in reading time is predicted (at the innermost verb), while a down-arrow (↓) indicates that a speedup is predicted. A question mark indicates that the theory is underspecified regarding that particular issue.

The table has some oversimplifications. First, Gibson’s earlier version of the DLT, syntactic prediction locality theory, does not entirely rule out similarity effects (see Gibson, 1998, 4), footnote 4). However, in more recent work (2000, 100), Gibson argues that theories involving similarity-based metrics of processing difficulty do not make the correct predictions. Accordingly, I am assuming here that in DLT, similarity-based interference plays no significant role. Second, DLT actually predicts no effect due to adverb insertion, but admits the possibility of a slowdown (Gibson, 2000, 107); it does not simply predict a slowdown as shown in the table. Third, regarding the column for RIT-02, I conflate the following assumptions into the model; these assumptions may not necessarily be shared by Lewis, so I refer to Lewis’s model as RIT-02′ (“RIT-02 prime”) in the table: (i) similarity-based interference subsumes
any kind of similarity; (ii) discourse information plays an independent and early role in
determining processing complexity; (iii) objects can become more similar to subjects
if they are in the prototypical subject position, and the resulting similarity between
fronted NPs and subjects can result in similarity-based interference; and (iv) there is
no syntactic specialization in working memory.

The above conclusions regarding these theories lead to a meta-theoretical point
about accounts of human sentence processing. Distance, computed by Hawkins’
and Gibson’s metrics, may indeed play a role in sentence processing,\(^1\) but so does
similarity-based interference. This is not an impossible result. Theories are often
presented as if they are mutually exclusive in the sense that one explanation is
assumed to preclude another. Obviously, they may not be mutually exclusive, and the
evidence presented in this dissertation suggests that each of these theories have some
value. It is of course necessary to consider, and if possible eliminate, alternative
hypothesis about a given phenomenon,\(^2\) but a more plausible assumption about
human sentence parsing would be to consider the possibility that a set of diverse
but principled constraints apply. Human cognition is complex, and the vast amount
of evidence from cognitive psychology of different factors affecting human perception
cannot be ignored when determining the constraints on parsing.

If the parsing problem is seen from this perspective, the question is no longer,
Which theory is right? The problem now becomes a matter of addressing two
questions; (a) Given that there is some independent evidence for several distinct
complexity metrics stated in terms of a given cognitive constraint, what are the
relative roles of these different explanations? (b) Given that human cognition utilizes
similar cognitive strategies for perception in different modalities, can the parsing

\(^1\)It is not clear whether distance is separable from or distinct from decay; it is possible that the
putative distance effects are actually decay effects (Gibson, 2000, 103).

\(^2\)This is a methodology that is conspicuous by its absence in most psycholinguistic work; instead
of taking a model-comparison approach, one particular model is often pursued without regard to
the elimination of alternative explanations.
problem be stated in the most general terms possible, at the same time taking into account these broadly applicable cognitive strategies? So far, this dissertation has addressed the first question; the next section turns to the second one. In the following section, I present a model of parsing cast in terms of abductive inference. This section is based on my computer science Master’s thesis (Vasishth, 2002a).

Before turning to the model, I should explain my approach to the human parsing question. The model I present currently accounts only for processing facts in four languages: Dutch, German, Hindi, and Japanese. Moreover, the discussion is restricted to center-embedding constructions. The reason for restricting the discussion in this way is that I am interested in determining a unified account of a phenomenon which occurs in these languages. This phenomenon is the center-embedding construction in which sentences have verb-final matrix and embedded clauses. A central goal of the present model is to find a cross-linguistic parsing strategy and complexity metric that would apply in such a context.

An alternative starting point in the sentence parsing literature has been that “the” human parsing mechanism uses precisely one strategy to parse all constructions in all languages. This is certainly an interesting assumption (e.g., (Mazuka & Lust, 1987)), and it may even turn out to be the correct one. However, the problem with starting with an all-encompassing universalist theory is that its ontology tends to get fixed too early in the process of “discovery”, and as a result it ultimately may end up explaining nothing. For example, assumptions about a Universal Grammar or a Universal Parsing Mechanism are intended to be powerful theoretical explanatory mechanisms, but when faced with variation in language, such assumptions retreat to giving pseudo-explanations which have nothing to do with the allegedly universally applicable theory that was the starting point. See footnote 4 below for a concrete illustration of this point.

In the discussion to follow, I do not address the processing of “garden-path” sentences and ambiguity in language. This is not as strange a limitation as it may
seem. Beginning with the now-classic works like (Bever, 1970) and (Frazier, 1979), the psycholinguistics literature has paid a great amount of attention to the problem of parsing English “garden-path” sentences like the infamous *The horse raced past the barn fell*, and ambiguous prepositional phrase attachment sentences like *The girl saw the boy with the telescope*. Research problems such as garden-pathing often gain prominence as a result of the material one begins with, and it is merely a historical accident that English, rather than some other language, came under focus first.\(^3\) for a recent example.

Consider, for example, that a relatively well-developed theory of parsing like Construal (Frazier & Clifton, 1996), (Frazier & Clifton, 1997) addresses ambiguity etc. (and that too primarily\(^4\) in English) but not constructions like center embeddings. This is not surprising, given the intractable nature of the ambiguity problem in human sentence parsing. But it is also no more surprising that the present model does not address garden-pathing etc., and limits itself to center embeddings in these four languages.

This does not mean that the model described below cannot account for ambiguity or garden-pathing, or cannot be extended to other languages like English where verbs

---

\(^{3}\) Perhaps for sociological reasons having nothing to do with science, research on English often appears to have greater importance than other languages, and often defines the research direction in other languages, even to the extent that one apparently needs to refer to English in order to understand how other languages work. An amusing thought experiment is to consider how seriously the psycholinguistic community would take a theory that presented an account for processing difficulty in two sentence-types of Pitjantjatjara, where one of the two predictions were wrong. Yet, such theories, dealing with English, are routinely proposed in the literature and are taken quite seriously; see pages 38-40

\(^{4}\) There is a discussion about some cross-linguistic support for Construal in (Frazier & Clifton, 1997), but for the unexplained cross-linguistic fact that some languages have high relative-clause attachment (e.g., German, Dutch, Spanish, French), and others low (e.g., English, Arabic, Romanian, Swedish, Brazilian-Portuguese), the explanation provided is: “Languages differ from each other in many respects that have been argued to be relevant for relative clause interpretation preferences, …independent and inescapable differences among the grammars of various languages may ultimately explain the distinct preferences perceivers exhibit across languages for which NP is chosen among the NPs in the current thematic processing domain” (p. 290). This is not intended as a criticism, but as an illustration of the difficulties of trying to bring all languages under one umbrella – the “explanation” degenerates to saying that languages differ because, well, languages differ.
do not normally occur in the final position in the sentence. Within the same general framework, it is possible to specify a different complexity metric that would account for the different kinds of processing difficulties that occur in ambiguous sentences. However, I leave this specification for future work and limit myself here to presenting an account of center-embedding sentences in these four languages.

7.2 A model of sentence processing based on abductive inference

Humans routinely engage in automatic or deliberative reasoning in order to process information about the world around them. In the course of almost any information processing task, one or more hypotheses are developed that “best explain” the facts, and these are either verified or discarded as further evidence emerges. This general cognitive mechanism is evident (as both a conscious and deliberative process as well as an automatic and fast process) in diverse activities like diagnosing illness, solving crimes, developing scientific theories, visual processing, and comprehending language. Peirce was the first to discuss this third kind of reasoning process (distinct from deduction and induction); he called it ABDUCTION (see (Josephson & Josephson, 1996) and the references cited there for a detailed discussion). Abduction is the following pattern of reasoning, and often is characterized as an “inference to the best explanation” (Harman, 1965).

\[ D \text{ is a collection of information available to us} \]

Hypothesis \( h \) explains \( D \)

No other hypothesis explains \( D \) as well as \( h \)

Therefore, \( h \) is probably correct.

Abductive inference has been proposed as an explanatory mechanism for language perception and processing; to give a few examples: text comprehension (Smith, 2000), text generation (Klabunde & Jansche, 1998), semantic processing
(Strigin, 1998), (Strigin, 2001), pragmatic interpretation (Hobbs, Stickel, Appelt, & Martin, 1993), and (chart parser-based) grammar-rule acquisition/learning (Cussens & Pulman, 2000). I propose that sentence comprehension difficulty results from a series of automatic abductive inferences. I also show that, when human parsing is viewed from this perspective, phrase-by-phrase processing difficulty in center embeddings can be accounted for with as much or greater accuracy than is possible with other models.

I will call the proposed model the abductive inference model, or AIM. The main intuition behind AIM is that one or more hypotheses regarding the structure of the sentence are generated as each word in a sentence is processed; the hypothesis-generation process is constrained by a principle called MINIMAL CONSISTENCY (an application of Ockham's razor), which says that no words of a certain syntactic category will be predicted if there is no evidence for them. These hypotheses are assumed to be stored in human working memory and overload working memory capacity under certain well-defined conditions (discussed later).

7.2.1 The underlying competence grammar $G$

I assume a grammar $G$ for a particular natural language $L$. $G$ defines what types of functional categories (or predicate-valuation frames) we can encounter in $L$, and how these functions can combine with their arguments. Throughout, I assume that $G$ is a categorial grammar along the lines of (Kruijff, 2001). In a lexicalist approach like categorial grammar, the lexicon determines how words can be put together. Structural rules, like those in Multi-Modal Logical Grammar (MMLG) (Moortgat, 1997), only vary the order in which words occur grammatically.\footnote{Moortgat calls the system Multimodal Categorial Grammar; here, I follow the terminology used in Kruijff (2001).} The grammar system assumed here is not necessarily limited to the dependency grammar based framework of Kruijff;
it could in principle be a system along the lines of an appropriate formalization
of Chomskyan minimalism (Vermaat, 1999), (Retoré & Stabler, to appear), Tree-
adjointing grammar (Abeillé & Rambow, 2000), Combinatorial Categorial Grammar
(Steedman, 2000), or Head-driven Phrase Structural Grammar (Pollard & Sag, 1994).
Any choice of grammar system will carry its own implications.

7.2.2 The set $\mathcal{H}$ of hypothesis formers

A set of hypothesis formers $\mathcal{H}$ is created on the basis of a compilation of
the lexicon.\footnote{Compilation here does not mean the creation of a lexicon; rather, given a lexicon, a set of
procedures are applied to it in order to compile out information present in the lexicon.} This compilation is based on a procedure proposed for linear logic
in (Hepple, 1998), and extended in (Kruijff, 1999) to cover a larger range of the
multiplicative resource logics used in MMLG.\footnote{Compilation was originally proposed for the purposes of efficient chart parsing with Lambek-
style grammars, in order to overcome problems with earlier approaches, e.g., (Hepple, 1992) and
(König, 1990).} The result of the procedure is a set of
first-order functions to represent categories (i.e., there are no higher-order formulas).\footnote{See (Console, Portinale, & Dupré, 1996) for a similar approach for rendering diagnosis more
efficient.}

$\mathcal{H}$ essentially provides us with schemas that encode information about how words
(of particular categories) can be combined. For example, intransitive verbs give us
the schema $f(NP1)$, transitive verbs $f(NP1, NP2)$, and so on. These hypotheses
can be considered to be tree structures such as those shown in Figure 7.1, but for
typographical convenience I will use the functional notation.

I assume that $\mathcal{H}$ is finite and closed.\footnote{This is not to be confused with the fact that the set of sentences in $\mathcal{L}$ is infinite.} These are all reasonable assumptions since
$\mathcal{H}$ cannot have any redundant hypothesis formers (having been created from the
grammar rules), and the schemas extracted from the grammar will be finite (if this
were not so, the set of grammar rules would be infinite).
The set of hypothesis formers $\mathcal{H}$ is assumed to be partially ordered by a SIMPLICITY CRITERION: simpler schemas appear before the more complex ones.\footnote{Instead of partial ordering, one could consider activation levels of various hypotheses as determining relative “simplicity”; in that case, one can refer to most active, as opposed to least active, hypotheses. This would also allow other factors to be included in determining the relative activation strength of a given hypothesis, factors such as degree of commitment to a hypothesis, the frequency of a particular schema as a function of corpus data. I leave the specification of the exact implementation of the simplicity criterion to future research.} For example, monoclausal schemas are simpler than biclausal ones, the intransitive verb schema is simpler than the ditransitive verb schema. This assumption is based on experimental evidence from Yamashita (1997), which showed that (Japanese) subjects prefer to complete sentences with verbs that are simpler (i.e., verbs that result in monoclausal structures) rather than more complex ones. I take this result to suggest that simpler structures are accessed before more complex ones, and model this assumption by the partial ordering (the ordering is partial because it is possible that there is no way to specify relative simplicity between a given pair of hypotheses).

The precise ordering criteria may vary from language to language and even within a language, and it is also possible that a more sophisticated set of decision criteria are needed (such as the frequency of certain syntactic structures) in order to determine the ordering. As an example of different results in two unrelated languages, consider Japanese and Hindi. For Japanese, Uehara has found (M. Nakayama, personal communication) that (contra Yamashita’s findings above) Japanese subjects prefer

Figure 7.1: Tree representation of schemas
two clauses over one in incomplete sentences beginning with two nominative-case marked NPs; in Japanese, this sequence could be continued either with a stative verb or a bi-clausal structure (another relevant criterion for determining ordering could be based on the claim by Miyamoto (2002) that Japanese nominative case markers force a new clause boundary to be constructed). The simplicity criterion (see above) wrongly predicts that the stative verb (monoclusal structure) is ordered before the biclusal one. However, research on Hindi gives a different result. First, in a pilot study that preceded Experiment 1, subjects were asked to complete sentences beginning with three NPs. Subjects overwhelmingly preferred to use ditransitive verbs rather than transitive one (which would entail a biclusal structure). Second, in a between-subjects comparison of Experiments 2 and 3, the t-test gave the following results: a longer reading time was found at the innermost transitive verb in sentences beginning with four NPs (data from Experiment 2, double embeddings), the reading time was 1029 msec (SE = 74.1) compared to the reading time for ditransitive verbs preceded by four NPs (Condition A of Experiment 3); the reading time in the latter case was 876.4 msec (SE = 65.66). Although the numerical difference in reading time did not reach significance, t = -1.672, p = 0.101, the direction of the difference is suggestive.

7.2.3 Some definitions

Next, I define some terms that I use in the proposed algorithm.

**Abducible structure(s):**

An ABDUCIBLE STRUCTURE is a schema/hypothesis that can be abduced based on (a) the principle of minimal consistency, and (b) the information available so far.
New information results in previous hypotheses being replaced. Abduced functions $f_i$ are part of the abducible structures that are taken from $\mathcal{H}$, and thus predict the presence of a word with a particular syntactic category. For example, in Japanese, if only a nominative NP (this is represented as $NP[nom]$) has appeared so far, $f_i(NP[nom])$ is an abducible structure that says: an intransitive verb $f_i$ with the nominative NP will give a sentence.\footnote{\label{fn11}The subscript on $f$ is merely a notational device used in the derivations to distinguish one abduced function from another.}

Although a nominative case marked NP is in principle consistent with an infinite set of possible continuations, the model allows for the selection of only those hypotheses from the hypothesis formers $\mathcal{H}$ that are \textit{minimally consistent} with the nominative NP. Minimal consistency is defined as follows:

\textbf{Minimal consistency:}

There are two cases: (i) the current word\footnote{\label{fn12}For simplicity, I treat a noun phrase (NP) as consisting of a single word.} is an NP, and (ii) the current word is a verb.

(i) If the current word is an NP: A list\footnote{\label{fn13}This is actually a set, ordered by the simplicity criterion (see page 190).} of hypotheses $H \subset \mathcal{H}$, is minimally consistent with the current list of words iff each hypothesis $h \in H$ can instantiate the NPs and verbs seen so far (in the current list of words) as arguments and functions respectively, without positing any new, unseen arguments.

(ii) If the current word is a verb: A list of hypotheses $H \subset \mathcal{H}$, is minimally consistent with a verb iff each hypothesis $h \in H$ satisfies one of the following two conditions:

Subcase 1: The hypothesis $h$ is able to take all the preceding words as arguments and/or functions and can MATCH (matching is defined below) the current verb with a function in $h$.\footnote{The subscript on $f$ is merely a notational device used in the derivations to distinguish one abduced function from another.}
Subcase 2: If the matching process fails, the current verb becomes part of a hypothesis $h$ which saturates\textsuperscript{14} all the arguments and verbs seen previously. If the current verb requires any new, unseen argument(s) and/or is necessarily an argument of another as-yet-unseen function $f_i$, the unseen argument(s) $n$ and/or the function $f_i$ are posited. Any unseen arguments that the function $f_i$ would then require must also be present in each hypothesis $h$.

An example illustrating the first clause above of minimal consistency is as follows. Suppose that, during the course of processing a Japanese sentence, the reader has seen only one nominative NP so far. In that case, a hypothesis satisfying minimal consistency is $f_i(NP[nom])$, and one violating minimal consistency is: $f_i(NP[nom], n)$, where $n$ is a hypothesized, new, unseen NP. By contrast, if, after the first nominative NP is seen, a second nominative NP is seen, the minimally consistent hypotheses are now $f_i(NP_1[nom], NP_2[nom])$, where $f_i$ is a stative verb, and $f_{i+1}(NP_1[nom], f_{i+2}(NP_2[nom]))$, i.e., a center-embedded structure. This is because Japanese allows only these two structures with two nominative case-marked NPs.

Subcase 1 of the second clause, in which a verb is seen after some NPs, requires a process of matching the verb to a hypothesized function; matching is defined as follows.

**Matching:** A verb $V$ matches with a function $f_i$ iff $V$ has a valency that is identical with that of $f_i$, and all the arguments in the subcategorization frame of the verb MATCH those of $f_i$.

\textsuperscript{14}Saturation is to be understood in the standard linguistic sense of a verb instantiating all of its arguments.
An NP can match with a posited NP argument iff its case marking, person, number, gender, and other information, is subsumed by (informally, “is consistent with”; see (Shieber, 1986)) that of the posited argument.

I illustrate this with the Japanese example above. If, after seeing two nominative case marked NPs, an intransitive embedded verb is seen, this verb is matched successively with the functions $f_i$ to $f_{i+2}$ in the abduced hypotheses, $f_i(NP_1[nom], NP_2[nom])$, and $f_{i+1}(NP_1[nom], f_{i+2}(NP_2[nom]))$. The match succeeds with $f_{i+2}$.

Subcase 2 of the second clause involves a situation where matching has failed; it can be exemplified as follows. Suppose that the Hindi sentence in (83) is being processed. The first item seen is a verb V1, kahaa, ‘said[past].’ At this point, the hypothesis will be V1($n1, n2, f_i(n2)$), because V1 necessarily has a subject $n1$ and an object $n2$; in addition, since V1 is an obligatory subject control verb, an embedded clause headed by a function (verb) $f_i$ must be posited, and its subject is $n2$.

(83) kahaa Siita-ne Hari-ko so-neko
     said  Sita-erg Hari-dat sleep-inf
     ‘Sita told Hari to (go to) sleep’.

The minimal consistency constraint is motivated by the production study by Yamashita (1997), and the pilot production study mentioned earlier.

With these definitions in place, we turn next to the algorithm, and then the complexity metric (which depends on the algorithm).

7.2.4 The algorithm

The top level of the processing algorithm is a Scan-Lookup-Process loop that terminates when there is no more remaining input. Scan() is a function that returns the next word $w$ to be processed, Lookup($w$) is a function that returns the lexical entry $L_w$ of word $w$, and Process($S$) is a function that carries out the abductive or
deductive step using $S$, a list containing the lexical representations of the input so far. The results of **Process** affect the contents of a list data structure, $M$ (emory); $M$ models the role of working memory.

**Algorithm 1** Main loop of the parser

| S ← [] (Initialized to empty list) |
| M ← [] (Initialized to empty list) |
| while remaining input nonnull do |
| w ← Scan() |
| $L_w$ ← Lookup($w$) |
| $S$ ← append($S, L_w$) |
| Process($S$) |
| end while |

**Algorithm 2** The main function **Process($S$)**

| Process($S$) |
| if lastitem($S$) == nominal then |
| if $\exists$ abduced nominals in $M$ then |
| match(lastitem($S$), $M$) |
| else |
| $M$ ← Abduce($S, \mathcal{H}$) |
| end if |
| else if lastitem($S$) == verbal then |
| if $\exists$ abduced function in $M$ then |
| match(lastitem($S$), $M$) |
| else |
| $M$ ← FindHyp($S, \mathcal{H}$) |
| end if |
| end if |

**Abduce** returns all hypotheses $h \in \mathcal{H}$ such that $S$ and $h$ are minimally consistent as defined earlier. **FindHyp** searches the set $\mathcal{H}$ of hypothesis formers for a hypothesis $h$ that is a schema for the latest item in $S$ (a verb).

To repeat the earlier example from Japanese: two nominative case marked NPs starting a sentence could be followed either by a stative predicate (84a), or a nested dependency construction with a single level of embedding (84b).
(84) a. $f_2(NP_1[nom], NP_2[nom])$

b. $f_3(NP_1[nom], f_4(NP_2[nom]))$

These are two hypotheses selected from $\mathcal{H}$. No other hypotheses are selected because these are the only two that are minimally consistent, given the information so far. These hypotheses are based on the grammatical possibilities in Japanese, and since a single clause sentence has a simpler structure than a sentence with an embedded clause, the hypotheses are ordered as shown above. Next, the appearance of an accusative case marked NP will result in these hypotheses being discarded and the new hypothesis shown in (85) being selected:

(85) $f_5(NP_1[nom], f_6(NP_2[nom], NP_3[acc]))$

Since the number of hypotheses decreases from two to one, the model predicts faster processing at the accusative NP. This prediction is borne out, as discussed later. We turn next to the complexity metric.

7.2.5 The complexity metric

The complexity metric has two components: **ABDUCTION COST**, the cost associated with the abductive process, and **MISMATCH COST**, the cost associated with a mismatch between an encountered verb and abduced functions.

**Abduction cost**: This reflects the increasing processing load as sentence fragments appear incrementally. The abduction cost is the sum of the number of NPs seen so far, the number of functions $f_i$ that are posited (in all the the current abduced hypotheses), and the total number of current distinct hypotheses. These three sub-components are intended to reflect the load in working memory of: (a) storing an increasing number of NPs; (b) positing functions; and (c) storing hypotheses.

**Mismatch cost**: I assume that the (queued) hypotheses initially are unanalyzed units. By the term “unanalyzed units” I simply mean that when a hypothesis like
\( f_i(NP_1, f_j(NP_2)) \) is present in working memory and a verb is encountered, any attempt to match the verb with any of the functions \( f \) present in the hypothesis must be a left to right depth first search; the verb cannot directly match the right function. During this search process, every time a verb fails to match with a hypothesized function, there is a mismatch cost of one. Another way of saying this is that matching must be preceded by an unpackaging of the hypothesis in order to access the possible candidates for a match in the hypothesis. This unpackaging has a cost associated with it, which I quantify in terms of mismatch cost. The use of mismatch cost is merely a simplifying assumption; the computation of the unpackaging cost could equally be linked to some other plausible measure of complexity.

The numerical value associated with each sub-component in the metric is assumed to be 1, and the components are assumed to be additive. This is merely a convenience, and nothing crucial hinges on this assumption. In a fully implemented version of this model, the unit costs associated with each component would need to be correlated with precise reading time predictions.

The complexity metric depends crucially on the control structure of the parser: at each stage when the parser incrementally builds/revises the list of possible hypotheses, the complexity metric is used to compute the processing cost at that point. This is significant because it is well known that left-corner parsing, as opposed to pure top-down parsing, or head corner parsing, etc., is exactly the right strategy for differentiating the processing difficulty associated with center embeddings, left-branching, and right-branching structures (Resnik, 1992).

Before ending this discussion of the complexity metric, I would like to briefly justify the assumption that the hypotheses in working memory are assumed to be unanalyzed units. This assumption is based on existing research in cognitive psychology that suggests that humans have difficulty disengaging components of items (such as words and visual patterns) in memory (Johnson, 1977), (Johnson, 1981), (Johnson et al., 1986), (Johnson, 1991). Of particular relevance is the pattern-unit
model, whose core assumption (Johnson, 1991, 83) is that precognitive representations of small visual patterns are always processed by first attempting to assign a single pattern-level cognitive representation of the array, and that step always succeeds if participants know the rule system or have a prelearned code that can be used for the assignment. The holistic encoding or representation is functionally motivated: the internal components of a pattern are functionally unnecessary for storing the pattern. As Johnson (1991, 84) puts it: “If participants would need to identify a component of an array, they always would be delayed in doing so by the initial attempts to encode the pattern as a whole.”

In sum, the pattern-unit model has considerable support from visual and language processing experiments (see the references cited above), and is the basis for assuming an initially holistic representation of hypotheses during the stage of encoding in working memory.

One possible objection to this assumption is that no search may be occurring in working memory at all; instead, memory could simply be content addressable. I show next that the evidence for this view (McElree, 2000) does not affect my assumption above regarding the unpackaging of encoded representations of hypotheses. To clarify the issue, I need to first present McElree’s empirical results, and then discuss the difference between the task involved in (a) “unpacking” a hypothesis and (b) the “search” process (or lack thereof) in McElree’s experiment design. McElree begins by defining two terms: AVAILABILITY, which is the probability that a memory representation is retained, and ACCESSIBILITY, the time it takes to retrieve a representation (assuming that it is available). He measured both these processing components using a series of speed-accuracy tradeoff (SAT) tasks, and found that availability, but not accessibility, is affected by interpolated material in serial recall.

The details of his experiments makes the above statement clearer. He used filler-gap constructions like the ones shown below.

(86)  a. This was the book, that the editor admired $t_i$.  

198
b. This was the book$_i$ that the editor$_j$ who the receptionist married $t_j$
admired $t_i$.

c. This was the book$_i$ that the editor$_j$ who the receptionist who quit married
$t_j$ admired $t_i$.

In each of the sentences in (86), the filler book is coindexed with its trace $t_i$, and
increasing amounts of material intervene between the filler and the gap.\footnote{A note to readers who may be unfamiliar with the notation used here: in syntactic theory, it is
commonly assumed that “displaced” words, like book in the sentence above, are coindexed (this is
represented by identical subscripts) to their “original” position, in this case the position following
the last verb, admired. This is a controversial and theoretically loaded assumption, but I use this
convention here in a theory-neutral manner.} Assuming
that the filler must be retrieved at the gap, McElree measured speed and accuracy of
retrieval in the following manner. Each trial began with a 500 msec fixation point,
followed by a word-by-word presentation of each sentence for 250 msec. Then, a
1000 Hz tone was sounded for 50 msec at one of six response lags: 50, 300, 500,
800, 1200, or 3000 msec after the onset of the final word in the string. At the tone,
subjects responded with a “yes/no” grammaticality judgement response by pressing
a designated key. The assumption here was that more interpolated material would
cause the availability of an item to degrade. He found that the more the interpolated
material between the filler and the gap, the lower the response accuracy. However,
the time to resolve the gap, when it could be resolved, remained constant. In other
words, the availability of an item decreased with increasing interpolation, but the
accessibility remained unchanged.

If a search process were being used in resolving the gap, the accessibility (the time
taken to retrieve a representation) should have decreased with increasing interpolation
since increasingly long lists would have had to be searched. McElree’s results are
taken to imply that memory representations are content addressable during real-
time sentence processing. This could be argued to suggest that the present model’s
assumptions about a serial search of hypotheses in working memory are incorrect.
However, there is a crucial difference between the retrieval process in McElree’s stimuli versus the search process within the current active hypotheses of AIM. In McElree’s sentences, the gap is entailed once the verb is encountered. In these sentences the connection between each filler and its gap is a unique and necessary one. For example, in (86b), the moment the verb *admired* is encountered, the gap $t_i$ is filled: 
*the editor$_j$ who the receptionist married* $t_j$ . . . . Any putative search would begin after the verb was seen. In AIM, the search occurs when an item currently being processed (typically a verb) could match one of a variety of hypothesized nodes (functions) in the active hypotheses – in general there is no unique or necessary connection between the item just seen and the uninstantiated node with which matching is to occur. The holistic encoding of a given hypothesis or hypotheses is the barrier to content addressability. The serial “search” procedure (assumed when multiple hypotheses are being examined for a possible match) could equally be restated in terms of the relative accessibility of each hypothesis (which could in turn be a function of the frequency of occurrence of each verb schema in the language).

In sum, the assumptions regarding “search” in AIM are not necessarily falsified by the McElree results, since McElree has not shown that *all* working memory representations are content addressable, but rather that fillers can fill their gaps in a content-addressable manner, without working memory engaging in a full search of the items currently in memory.

### 7.2.6 Computational properties of the abductive parser

The proposed parser is a variant of the shift-reduce parser, LR(0), the LR parser with no lookahead (Aho & Ullman, 1993), (Sikkel, 1997). That is, the parser has a hybrid bottom-up/top-down discipline for building hypotheses. Although the process of hypothesis generation may appear to be indistinguishable from prediction (the top-down step(s) in the LR parser), prediction is a consequence of the generation
of explanatory hypotheses. It is due to this distinction that I refer to the model as an *abductive* inference based model. Unlike existing LR parsing algorithms, however, the present parsing model builds predictions incrementally as subsequences successively become available. These predictions are not *all* possible completions of the subsequence (which would be an infinite set), but only the ones satisfying minimal consistency. For an input string of length $n$, and given a set of hypothesis formers $\mathcal{H}$, where $|\mathcal{H}| = k$, parsing will take at most $n \times k$ time, since each new item in the string will require searching the set $\mathcal{H}$ for hypotheses.\(^{16}\) Since $k$ is constant for a given language, the worst-case time complexity of the algorithm is $O(n)$.\(^{17}\) Worst-case space complexity is as for left-corner parsers, but in the present model space complexity does not correspond directly to processing difficulty. Instead, the memory cost computation algorithm determines moment-by-moment difficulty.\(^{18}\)

The present parsing model also raises a question heatedly debated in the literature: does the human parser require nondeterministic parsing? Although existing research (Marcus, 1980), (Shieber, 1983), (Briscoe, 1987), (Briscoe, 2000) suggests that a nondeterministic parser is unnecessary for natural language, deterministic parsing models face problems when we evaluate such parsers with respect to experimental data. For example, Marcus’ parser employs a lookahead for making parsing decisions, and Shieber’s parser employs a related notion called “preterminal delay”.\(^{19}\) This particular brand of bounded nondeterminism

\(^{16}\)It could take less time if the set $\mathcal{H}$ of hypothesis formers is not assumed to be a flat data structure but is structured as a hierarchy of hypotheses, so that adopting one hypothesis rules out entire subclasses of hypotheses as possible candidates. This detail does not affect the model presented here.

\(^{17}\)In the actual Prolog implementation of this model (Vasishth, 2002a), time complexity is $O(n \times d)$, where $d$ is the number of embeddings in the sentence being processed; this is because for each level of embedding, the entire hypothesis space is searched. This can also be optimized and does not affect the details of the model.

\(^{18}\)Note that worst-case time complexity results are not directly useful; much more useful is average-case time complexity. However, as far as I know, there are no mathematically determined average-case results for left-corner parsers, only empirically determined ones (Joshi, 1996, 407).

\(^{19}\)Shieber’s points out that his deterministic shift-reduce parser is actually “a simulation of a nondeterministic machine” (Shieber, 1983, 701).
actually goes against the fact that humans process sentences strictly incrementally (Tyler & Marslen-Wilson, 1977). Briscoe (1987) believes that other cues, such as prosody, always disambiguate sentences, but this cannot account for the Hindi center-embedding processing facts, particularly the results of the adverb-insertion experiment and the definiteness effect, where prosody would not facilitate processing.\(^{20}\) Furthermore, it is not clear how the nondeterminism associated with dislocated constituents in English can be disambiguated with prosody. These are sentences such as Mary, I like to tell jokes to e. Here, the gap for the dislocated element Mary cannot be identified until the very end of the string is reached (Johnson-Laird, 1983). It seems difficult to provide a prosodic contour that would help identify the gap in the sentence above.

The present model differs from these others approaches to bounded nondeterminism in the following way. More than one hypothesis can be generated by the parser, depending on the grammar of a particular language. The model is thus technically a nondeterministic pushdown automaton, and it simulates nondeterminism in a manner quite similar to “preterminal delay” in (Shieber, 1983). However, the difference is that the present model allows incremental interpretation at every stage in the processing; this is possible because the hypotheses that are generated contain all the information needed to build partial semantic interpretations. Thus, the parser ends up being equivalent to a deterministic pushdown automaton.

I present next the empirical coverage of the model, using data from Japanese, Dutch, German, and Hindi.

\(^{20}\)Note that this is merely a conjecture based on my intuitions about Hindi.
7.2.7 Empirical coverage of AIM

Japanese

In the following discussion, the verbs in nested sentences are numbered in reverse order of occurrence, i.e., the matrix verb, which appears last, is V1. The numbers do not reflect the verbs’ valencies; this reverse numbering convention is merely in order to highlight the contrasting situation in Dutch (where the matrix verb is the first one seen in center embeddings).

Gibson (1998) reports that (87a) is less acceptable than (87b) (no information is provided in the paper as to how the results were obtained; presumably, these are intuitive acceptability judgements).

(87) a. obasan-ga bebiisitaa-ga ani-ga imooto-o izimeta-to
    aunt-nom babysitter-nom brother-nom sister-acc teased-comp.
    itta-to omotteiru
    said-comp. thinks
    ‘The aunt thinks that the babysitter said that the elder brother teased
     the younger sister.’

b. bebiisitaa-ga ani-ga imooto-o izimeta-to itta-to
    babysitt.-nom brother-nom sister-acc teased-comp. said-comp.
    obasan-ga omotteiru
    aunt-nom thinks
    ‘The aunt thinks that the babysitter said that the elder brother teased
     the younger sister.’

First, consider the application of the algorithm for (87a). For this first derivation, I incrementally build up the parse for each step. Subsequent derivations will be shown in their final form with the entire derivation history.

Step 1:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>f_i(NP1)</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
</tbody>
</table>

203
Here, given only the first NP (oobasan-ga), a sentence with an intransitive verb (IV), denoted by \( f_1 \), is abduced. This contributes 3 to our cost so far (abduction cost, composed of the number of NPs seen so far (1), plus the number of functions abduced (1), plus the number of hypotheses abduced (1); mismatch cost is currently 0).

Step 2:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>( 1+1+1=3 )</td>
<td>0</td>
</tr>
<tr>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>( 2+3+2=7 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Given the second NP (bebisitaa-ga), and given that both the NPs seen so far are nominative case marked, the abducible structures are: a stative predicate taking two nominative arguments (\( f_2(NP1,N2) \)), and a center embedded construction (\( f_3(N1,f_4(N2)) \)). The abduction cost here is 7: 2 NPs, 3 functions, and 2 hypotheses.

Step 3:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>( 1+1+1=3 )</td>
<td>0</td>
</tr>
<tr>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>( 2+3+2=7 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP3-ga</td>
<td>( f_5(NP1,f_6(NP2,NP3)) )</td>
<td>( 3+5+2=10 )</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_7(NP1,f_8(NP2,f_9(NP3)) )</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We now have three nominative NPs, and so we either have an embedded stative predicate, as in \( f_5(NP1,f_6(NP2,NP3)) \), or a center embedding, as in \( f_7(NP1,f_8(NP2,f_9(NP3)) \). The abduction cost is now 10.
Step 4:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP3-ga</td>
<td>( f_5(NP1,f_6(NP2,NP3)) )</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_7(NP1,f_8(NP2,f_9(NP3)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP4-o</td>
<td>( f_{10}(NP1,f_{11}(NP2,f_{12}(NP3,NP4))) )</td>
<td>4+3+1=8</td>
<td>0</td>
</tr>
</tbody>
</table>

The function \( f_{10}(NP1,f_{11}(NP2,f_{12}(NP3,NP4))) \) is abduced because the fourth NP is marked with accusative case, and so there must be at least one embedding with a transitive embedded verb. The abduction cost is now 8; i.e., the model predicts that processing will take less time at this fourth NP, compared to the third NP.

Step 5:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP3-ga</td>
<td>( f_5(NP1,f_6(NP2,NP3)) )</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>( f_7(NP1,f_8(NP2,f_9(NP3)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP4-o</td>
<td>( f_{10}(NP1,f_{11}(NP2,f_{12}(NP3,NP4))) )</td>
<td>4+3+1=8</td>
<td>0</td>
</tr>
<tr>
<td>V3</td>
<td>( f_{10}(NP1,f_{11}(NP2,V3(N3,NP4))) )</td>
<td>4+2+1=7</td>
<td>2</td>
</tr>
</tbody>
</table>

Here, the current word, V3 is *izimeta-to*, ‘teased-complementizer’, and a deduction is performed in the following manner:

(i) V3 tries to match \( f_{10} \) in

\[
 f_{10}(NP1,f_{11}(NP2,f_{12}(NP3,NP4))) \Rightarrow \text{this leads to failure.}
\]

This matching attempt fails because the outermost function \( f_{10} \) has a valency frame that does not match that of the actual verb.
(ii) V3 tries to match $f_{s1}$ in

$$f_{i0}(NP1,f_{s1}(NP2,f_{x2}(NP3,NP4))) \Rightarrow \text{this also leads to failure.}$$

Here, again, the failure occurs due to the valency frame of the verb not matching that of the next function.

(iii) V3 tries to match $f_{s2}$ in

$$f_{i0}(NP1,f_{s1}(NP2,f_{x2}(NP3,NP4))) \Rightarrow f_{i0}(NP1,f_{x1}(NP2,V3(N3,NP4)))$$

This succeeds because the valency frame of the verb matches that of the next function. The cost now is the sum of the abduction cost (7) plus the number of failed matches (2): 9. The number of abduced functions is now 2, since one of the abduced functions has already been resolved by its matching with V3.

Step 6:

<table>
<thead>
<tr>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP1-ga</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>NP2-ga</td>
<td>$f_2(NP1,NP2)$</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$f_3(NP1,f_4(NP2))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP3-ga</td>
<td>$f_5(NP1,f_6(NP2,NP3))$</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>$f_7(NP1,f_8(NP2,f_9(NP3))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP4-o</td>
<td>$f_{i0}(NP1,f_{x1}(NP2,f_{x2}(NP3,NP4)))$</td>
<td>4+3+1=8</td>
<td>0</td>
</tr>
<tr>
<td>V3</td>
<td>$f_{i0}(NP1,f_{x1}(NP2,V3(N3,NP4)))$</td>
<td>4+2+1=7</td>
<td>2</td>
</tr>
<tr>
<td>V2</td>
<td>$f_{i0}(NP1,V2(NP2,V3(NP3,NP4)))$</td>
<td>4+1+1=6</td>
<td>1</td>
</tr>
</tbody>
</table>

The matching process goes as follows:

(i) V2 tries to match $f_{i0}$ in

$$f_{i0}(NP1,f_{x1}(NP2,V3(N3,NP4))) \Rightarrow \text{failure.}$$

206
(ii) V2 tries to match $f_{11}$ in

$$f_{10}(NP1, f_{11}(NP2, V3(N3, NP4))) \Rightarrow f_{10}(NP1, V2(NP2, V3(NP3, NP4)))$$

V2 fails to match $f_{10}$, but successfully matches $f_{11}$. The cost is now 7 (the abduction cost, 6, plus the mismatch cost, 1).

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ga</td>
<td>$f_2(NP1, NP2)$</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3-ga</td>
<td>$f_3(NP1, f_4(NP2))$</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>NP4-o</td>
<td>$f_{10}(NP1, f_{11}(NP2, f_{15}(NP3, NP4)))$</td>
<td>4+3+1=8</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>V3</td>
<td>$f_{10}(NP1, f_{11}(NP2, V3(N3, NP4)))$</td>
<td>4+2+1=7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>V2</td>
<td>$f_{10}(NP1, V2(NP2, V3(NP3, NP4)))$</td>
<td>4+1+1=6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>V1</td>
<td>$V1(NP1, V2(NP2, V3(NP3, NP4)))$</td>
<td>4+0+0=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.2: Complete derivation history for (87a); total cost = 48

Step 7 marks the end of the derivation, and is shown in Figure 7.2. The deduction in this case is immediate:

V1 tries to match $f_{10}$ in

$$f_{10}(NP1, V2(NP2, V3(N3, NP4))) \Rightarrow V1(NP1, V2(NP2, V3(NP3, NP4)))$$

Here, V1 matches the outermost abduced function $f_{10}$ immediately, and the parse is completed. The cost at this stage is 4. The overall complexity of the sentence relative to other sentences is the total processing cost in the sentence. So, in this case, the maximal cost is 48. The local complexity at each point is the cost (Abduction Cost + Mismatch Cost) at that point, and constitutes a prediction about relative reading time difficulty.

By contrast, (87b)’s processing, shown in Figure 7.3, yields a lower total cost of 39. Note that in Step 5 in Figure 7.3, the appearance of an embedded verb results in an abduced hypothesis involving a matrix verb and a nominal argument. This is
because V2 has the complementizer -to, which requires it to be an embedded verb. That is, the second clause in the definition of minimal consistency applies in this case.

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NP3-o</td>
<td>( f_5(NP1,f_6(NP2,NP3)) )</td>
<td>3+2+1=6</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>V3</td>
<td>( f_5(NP1,V3(NP2,NP3)) )</td>
<td>3+1+1=5</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>V2</td>
<td>( f_7(x,V2(NP1,V3(NP2,NP3))) )</td>
<td>4+1+1=6</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>NP4-ga</td>
<td>( f_7(NP4,V2(NP1,V3(NP2,NP3))) )</td>
<td>4+1+1=6</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>V1</td>
<td>( V1(NP4,V2(NP1,V3(NP2,NP3)) )</td>
<td>4+0+1=5</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.3: Complete derivation history for (87b); total cost = 39

Nakatani et al. (2000) conducted several acceptability rating questionnaire experiments with Japanese; their results may be summarized as follows:21

Nakatani et al. found that double embeddings are less acceptable than left branching structures. The examples below illustrate the relevant structures.

(88) a. [obasan-wa [bebisitaag-ga [imooyo-ga naita-to] itta-to] aunt-top babysitter-nom sister-nom cried-comp. said-comp. omotteiru] thinks ‘The aunt thinks that the babysitter said that the younger sister cried.’

b. [imooyo-ga naita-to] bebisitaag-ga itta-to obasan-wa sister-nom cried-comp. babysitter-nom said-comp. aunt-top omotteiru] thinks ‘The aunt thinks that the babysitter said that the elder brother teased the younger sister.’

The model makes the correct prediction about this set of examples, as the derivations in Figures 7.4 and 7.5 show.
<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-wa</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ga</td>
<td>$f_2(NP1,NP2)$</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_3(NP1,f_4(NP2))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NP3-ga</td>
<td>$f_5(NP1,f_6(NP2,NP3))$</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_7(NP1,f_8(NP2,f_9(NP3)))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V3-to</td>
<td>$f_7(NP1,f_8(NP2,V3(NP3)))$</td>
<td>3+2+1=6</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>V2-to</td>
<td>$f_7(NP1,V2(NP2,V3(NP3)))$</td>
<td>3+1+1=5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>V1</td>
<td>$V1(NP1,V2(NP2,V3(NP3)))$</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.4: Complete derivation for (88a); total cost = 38

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>V3-to</td>
<td>$f_2(V3(NP1),x)$</td>
<td>2+1+1=4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP2-ga</td>
<td>$f_2(V3(NP1),NP2)$</td>
<td>2+1+1=4</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>V2-to</td>
<td>$f_3(V2(V3(NP1),NP2),y)$</td>
<td>3+1+1=5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>NP3-ga</td>
<td>$f_3(V2(V3(NP1),NP2),NP3)$</td>
<td>3+1+1=5</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>V1</td>
<td>$V1(V2(V3(NP1),NP2),NP3)$</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.5: Complete derivation for (88b); total cost = 25

209
Moreover, Nakatani et al. found that in double embeddings intransitive V3’s are more acceptable than transitive V3’s. Examples of these structures are shown below.

(89) a. haha-ga titi-ga fukigen-na akatyan-ga naita-to itta-to mother-nom father-nom fussy baby-nom cried-comp. said-comp. omotteiru
   thinks
   ‘My mother thinks that my father said that the fussy baby cried.’

   b. obasan-ga syoojiki-na bebisitaa-ga ani-ga imooto-o
      aunt-nom honest babysitter-nom brother-nom sister-acc
      izimeta-to itta-to omotteiru
      teased-comp. said-comp. thinks
      ‘My aunt thinks that the honest babysitter said that my brother teased my sister.’

The model makes the correct prediction since (89)a has cost 40 and (89)b has cost 48; see Figures 7.6 and 7.7.

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>( f_1(NP1) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ga</td>
<td>( f_2(NP1,NP2) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( f_3(NP1,f_4(NP2)) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NP3-ga</td>
<td>( f_5(NP1,f_6(NP2,NP3)) )</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( f_7(NP1,f_8(NP2,f_9(NP3))) )</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V3-to</td>
<td>( f_7(NP1,f_8(NP2,V3(NP3))) )</td>
<td>3+2+1=6</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>V2-to</td>
<td>( f_7(NP1,V2(NP2,V3(NP3))) )</td>
<td>3+1+1=5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>V1</td>
<td>( V1(NP1,V2(NP2,V3(NP3))) )</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.6: Derivation history for (89a); total cost = 38

Yamashita (1997) investigated the effect of word order and case marking on the processing of Japanese. One of her experiments is a moving window task involving three conditions:

\(^{21}\)Note: the English glosses are sometimes different from (Nakatani et al., 2000).
<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ga</td>
<td>$f_2(NP1,NP2)$</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_3(NP1,f_4(NP2))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>NP3-ga</td>
<td>$f_5(NP1,f_6(NP2,NP3))$</td>
<td>3+5+2=10</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_7(NP1,f_8(NP2,f_9(NP3))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NP4-o</td>
<td>$f_{10}(NP1,f_{11}(NP2,f_{12}(NP3,NP4)))$</td>
<td>4+3+1=8</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>V3-to</td>
<td>$f_{10}(NP1,f_{11}(NP2,V3(N3,NP4)))$</td>
<td>4+2+1=7</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>V2-to</td>
<td>$f_{10}(NP1,V2(NP2,V3(NP3,NP4)))$</td>
<td>4+1+1=6</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>V1</td>
<td>$V1(NP1,V2(NP2,V3(NP3,NP4)))$</td>
<td>4+0+0=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.7: Complete derivation history for (89b); total cost = 48

A. Canonical order, with 4NPs and 2 verbs:

[NP1-nom NP2-dat [NP3-nom NP4-acc V2] V1]

B. Same structure as in Condition A, but scrambled NP3 and NP4:

[NP1-nom NP2-dat [NP4-acc NP3-nom V2] V1]

C. Same structure as in Condition A, but scrambled NP1, NP2, NP3 and NP4:

[NP2-dat NP1-nom [NP4-acc NP3-nom V2] V1]

The results for Condition A are interesting in the context of the present model;\textsuperscript{22} consider the example below.

\textsuperscript{22}I do not discuss the effect of word order variation here since this introduces issues of pragmatics that the model currently does not take into account. The model can, however, be extended to incorporate constraints from pragmatics into the abductive process.

211
(90) [denwa-de hansamu-na gakusei-ga sensei-ni [tunetai koibito-ga phone-on handsome student-nom teacher-dat cold girlfriend-nom nagai tegami-o yabutta-to] itta] long letter-acc tore-comp. said ‘On the phone, a handsome student told the teacher that the cold-hearted girlfriend had torn up the letter.’

Yamashita found that reading times rose steadily in such examples until the NP preceding the accusative marked NP, and then fell at the accusative NP. The present model predicts this pattern, as shown below.

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ga</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ni</td>
<td>$f_2(NP1, NP2)$</td>
<td>2+1+1=4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3-ga</td>
<td>$f_3(NP1, NP2, f_4(NP3))$</td>
<td>3+4+2=9</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$f_5(NP1, f_6(NP2, NP3))$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>NP4-o</td>
<td>$f_7(NP1, NP2, f_8(NP3, NP4))$</td>
<td>4+2+1=7</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>V2</td>
<td>$f_7(NP1, NP2, V2(NP3, NP4))$</td>
<td>4+1+1=6</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>V1</td>
<td>$V1(NP1, NP2, V2(NP3, NP4))$</td>
<td>4+0+0=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.8: Derivation for (90)

Before step 4, the reading time is predicted to rise steadily. At step 4, a fall in reading time is predicted since the number of hypotheses falls from two to one, and the number of functions is now one.

Dutch and German

Turning next to Dutch, Kaan and Vasić (2000) conducted several self-paced reading studies and found the following. First, double embeddings are more difficult for native speakers to process than single embeddings; examples of each type are shown below:
(91) a. De leider heeft Paul Sonya het kompas helpen leren gebruiken tijdens
de bergtocht
during
de hike
‘The leader helped Paul teach Sonya to use the compass during the hike.’

b. Met aanwijzingen van de leider heeft Paul Sonya het kompas helpen
with directions of the leader has Paul Sonya the compass teach
gebruiken tijdens de bergtocht
use during the hike
‘With the leader’s directions Paul taught Sonya to use the compass during
the hike.’

Double embeddings have a cost of 50, as shown in Figure 7.9, and the moment-by-
moment reading time predictions match the observed effects.

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1</td>
<td>( f_4(NP1) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2</td>
<td>( f_2(NP1,NP2), f_3(NP1,f_4(NP2)) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3</td>
<td>( f_5(NP1,NP2,NP3) ) ( f_6(NP1,f_7(NP2,NP3)) )</td>
<td>3+6+6=18</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( f_8(NP1,f_9(NP2,f_{10}(NP3)) ) ( f_{11}(NP1,f_{12}(NP2,NP3,NP4)) )</td>
<td>4+5+2=11</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>NP4</td>
<td>( f_{14}(NP1,f_{12}(NP2,NP3,NP4)) ) ( f_{13}(NP1,f_{14}(NP2,f_{15}(NP3,NP4))) )</td>
<td>4+2+1=7</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>V1</td>
<td>( V1(NP1,f_{14}(NP2,f_{15}(NP3,NP4))) )</td>
<td>4+1+1=6</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>V2</td>
<td>( V1(NP1,V2(NP2,f_{15}(NP3,NP4))) )</td>
<td>4+0+0=1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>V3</td>
<td>( V1(NP1,V2(NP2,V3(NP3,NP4))) )</td>
<td>4+0+0=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.9: Derivation for (91a); total cost = 50

Single embeddings have a lower cost of 28, as Figure 7.10 shows.

Kaan and Vasić also found that RTs increased with each incoming NP, and fell
at the innermost verb, which is what this model predicts. In the present model, the
NP reading times are predicted to rise due to the increase in the number of abducted
functions, and a fall in reading time is predicted at the first verb due to the elimination
of some hypotheses.

Bach et al. (1986) showed that Dutch crossed dependencies are easier to process for native Dutch speakers than German nested dependencies are for native German speakers.

Examples of crossed Dutch and nested German dependencies are shown below:

(92)  a. Jan Piet Marie zag laten zwemmen
       Jan Piet Marie saw make swim
       ‘Jan saw Piet make Marie swim.’

       b. ... dass Hans Peter Marie schwimmen lassen sah
           ... that Hans Peter Marie swim make saw
           ‘... that Hans saw Peter make Marie swim.’

The Dutch SCEs are called crossed dependencies because the verbs and the subjects they link with form crossing chains (NP1 NP2 NP3 V1 V2 V3), and the German SCEs are called nested dependencies since the pattern is NP1 NP2 NP3 V3 V2 V1.

The model correctly predicts that Dutch center embeddings will be more acceptable than German ones: as shown in Figures 7.11 and 7.12, in Dutch, there will be a mismatch cost of one at the first verb seen; but in the German example, there will be a mismatch cost of five at the first verb seen.

As shown in Figures 7.11 and 7.12, in Dutch, there will be a mismatch cost of one at the first verb seen; in the German example, however, there will be a mismatch
cost of five at the first verb seen. The model thus predicts that reading times at the innermost verb in German will be higher than in Dutch. Unfortunately, there exist no reading time studies on German center embeddings at present that would help us (dis)confirm this prediction.

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1</td>
<td>( f_4(\text{NP1}) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2</td>
<td>( f_2(\text{NP1},\text{NP2}), f_3(\text{NP1},f_4(\text{NP2})) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3</td>
<td>( f_5(\text{NP1},\text{NP2},\text{NP3}) )</td>
<td>3+6+3=12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V1</td>
<td>( f_7(\text{NP2},\text{NP3}) )</td>
<td>3+3+2=8</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>V2</td>
<td>( f_9(\text{NP2},f_{10}(\text{NP3})) )</td>
<td>3+1+1=5</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>V3</td>
<td>( f_10(\text{NP2},\text{NP3}) )</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.11: Derivation for Dutch crossed embedding (92a); total cost = 41

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1</td>
<td>( f_4(\text{NP1}) )</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2</td>
<td>( f_2(\text{NP1},\text{NP2}), f_3(\text{NP1},f_4(\text{NP2})) )</td>
<td>2+3+2=7</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3</td>
<td>( f_5(\text{NP1},\text{NP2},\text{NP3}) )</td>
<td>3+6+3=12</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>V3</td>
<td>( f_8(\text{NP1},\text{NP2},f_{10}(\text{NP3})) )</td>
<td>3+2+1=6</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>V2</td>
<td>( f_9(\text{NP2},f_{10}(\text{NP3})) )</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>V1</td>
<td>( f_10(\text{NP2},\text{NP3}) )</td>
<td>3+0+1=4</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.12: Derivation for German nested embedding (92b); total cost = 41

Hindi

First, the model is consistent with the results relating to the definiteness effect: processing cost due to semantic/pragmatic factors can exist independently of the actual process of abducing hypotheses, as in sentence pairs like (93).

215
(93)  a. Siitaa-ne Hari-ko [kitaab khariid-neko] kahaa
    Sita-erg Hari-dat book buy-inf told
    ‘Sita told Hari to buy a/the book.’

    b. Siitaa-ne Hari-ko [kitaab-ko khariid-neko] kahaa
    Sita-erg Hari-dat book-acc buy-inf told
    ‘Sita told Hari to buy the book.’

The model predicts that in the case of both (93a,b) there will be only one hypothesis by the time the third NP is processed. However, the effects of discourse context are consistent with the model since no assumptions are made about pragmatic information. Thus, definite object NPs could be harder to process than indefinites due to the independent effect of pragmatics on processing (recall Givón’s hypothesis).

<table>
<thead>
<tr>
<th>Step</th>
<th>Input</th>
<th>Abduction/deduction</th>
<th>Abduction Cost</th>
<th>Mismatch Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>NP1-ne</td>
<td>$f_1(NP1)$</td>
<td>1+1+1=3</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>NP2-ko</td>
<td>$f_2(NP1,NP2)$</td>
<td>2+1+1=4</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>NP3(-ko)</td>
<td>$f_3(NP1,f_4(NP2,NP3))$</td>
<td>3+2+1=6</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>V2</td>
<td>$f_2(NP1,V2(NP2,NP3))$</td>
<td>3+1+1=5</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>V1</td>
<td>$V1(NP1,V2(NP2,NP3))$</td>
<td>3+0+0=3</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 7.13: Derivation for Hindi examples (93a,b)

Second, the adverb-insertion results are also consistent with AIM: the abducted hypotheses are predictions regarding the next expected word, and since an adverb predicts a verb (the adverb normally attaches to a verb projection (Mahajan, 1990, 87-88)), the degree of confidence in an abducted hypothesis increases if an adverb is present after the final NP, compared to the case where the final NP is followed immediately by a verb.

Finally, the model predicts that a ditransitive verb following four NPs will be read faster than a transitive verb. This is because the ditransitive verb hypothesis is ordered before the transitive verb hypothesis, and should therefore result in a
lower mismatch cost. Although a between-subjects comparison of means for pairs of sentence structures like (94) yielded \( p = .101 \), there was a speedup in reading time at the ditransitive verb in (94a) compared to the transitive verb in (94b).

(94) a. Siitaa-ne Hari-ko Ravi-ko kitaab de-neko kahaa
   Sita-erg Hari-dat Ravi-dat book  give-inf told
   ‘Sita told Hari to give a book to Ravi.’

   b. Siitaa-ne Hari-ko Ravi-ko kitaab pañh-neko bol-neko kahaa
   Sita-erg Hari-dat Ravi-dat book read-inf tell-inf told
   ‘Sita told Hari to tell Ravi to read a book.’

I discuss next the connection between AIM and its complexity metric.

7.2.8 The connection between the abduction-based model and the complexity metric; and integrating AIM, DLT, and RIT

One question that arises is: Why choose abduction as a way to characterize the parsing process? One answer is that abduction is a general cognitive mechanism that is instantiated in a particular way in human sentence parsing. One may argue that predictive parsing could be defined differently, rather than being a part of the abductive process. But abduction is the more general process, and it suggests a complexity metric that can correctly characterize parsing difficulty, as discussed earlier.

Another question is: Are the parser and the complexity metric inseparable? In other words, could the complexity metric be applied to theories without importing any of the ideas in AIM into them, or could other complexity metrics replace the one I propose? The proposed metric depends crucially on the number and nature of the hypotheses abducted, and on the search strategy. This means that there is no
straightforward way to separate the complexity metric from the abductive process. Of course, any other complexity metric could replace the current one, but choosing an alternative depends on whether such a metric could make the correct predictions regarding the behavioral data available. That is an empirical question.

Finally, there is the question of whether similarity-based interference-based, distance-based, and abductive-inference models can be integrated together. Distance, if construed as decay (Gibson, 2000, 103), can certainly be incorporated in an implementation of AIM which explicitly includes activation levels of items that decay over time in working memory. The question here would be defining the extent to which decay effects are more (or less) important than (for example) the facilitating effects of predicting the most likely next word. Integrating interference effects with AIM may be less straightforward: the crucial assumption in RIT is that a verb (for example) sets retrieval cues for an NP, which experiences interference due to similarity with other NPs. AIM, however, assumes that an attempt is made to match the verb to functions (verb-slots) in a current set of hypotheses about sentence structures — specifically, there is no attempt at the verb to set retrieval cues for NPs that would satisfy the verb’s subcategorization frame. However, one possibility is that the Matching stage in AIM could include a retrieval-cue setting component; the resulting interference would (in effect) quantify the extent to which the verb’s subcategorization frame matches the relevant hypothesis. More difficult matches would be those where NPs are similar and mutually more confusable, and such difficult matches would translate to a longer reading time at the verb. I leave these further refinements of AIM to the future.
APPENDIX A: STIMULI FOR THE EXPERIMENTS

The design for all the experiments presented here had the following general pattern. I illustrate the design with experiment 1, which has a $2 \times 2$ design, i.e., there are two factors with two levels each. Let us call these four conditions A, B, C, and D. I first prepared 16 variants or tokens of condition A. These sentences were used to create 16 sentences for each of the remaining conditions B, C, and D. For example, if condition A varied from condition B in terms of whether the sentence was a single embedding (condition A) or a double embedding (condition B), then the 16 scrambled versions were created for condition B from the 16 sentence in condition A. Similarly for conditions C and D, which also varied from condition A along one dimension (i.e., all the conditions constitute minimal pairs when considered pair-wise).

Then, four lists or groups were prepared, with each group containing a total of 16 stimuli. Each of these stimuli were interleaved pseudorandomly with 32 fillers, yielding a total of 48 sentences. I ensured that there was at least one filler (usually two to four) between two stimulus sentences. The resulting presentation order of the stimuli was as shown below. The design is only partially counterbalanced because, for example, a condition A sentence is never the immediately following stimulus (ignoring fillers) after a condition B sentence.
<table>
<thead>
<tr>
<th>Sentence No.</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>2</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>5</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>6</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>8</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>9</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>10</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>11</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>12</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>13</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
</tr>
<tr>
<td>14</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>A</td>
</tr>
<tr>
<td>15</td>
<td>C</td>
<td>D</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>16</td>
<td>D</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
</tbody>
</table>

Apart from minimizing order effects (Cowart, 1997), another advantage of using a counterbalanced design is that each lexical item appears in all the different conditions of the experiment. As a result, any inherent effect on acceptability of a given item (e.g., its frequency) is distributed equally over all the conditions.

A.1 Experiment 1

I present conditions A (single embeddings, no case marking on final NP) and B (double embeddings, no case marking on final NP); the stimuli for the other two conditions can be inferred from the ones presented below since conditions C and D differ from A
and B in having case marking on the final NP. For each of the 16 stimulus sentences, I first present the condition A version and then the condition B version.

(1) a. सीता ने हृरि को किताब खरीदने को कहा।
b. राणा ने सीता को हृरि को किताब खरीदने को कहने की सलाह दी।

(2) a. रितु ने उमेश को गाड़ी चलाने को कहा।
b. प्रकाश ने रितु को उमेश को गाड़ी चलाने को कहने की नसीहत दी।

(3) a. रमेश ने इला को सामान खोलने का ह्रास दिया।
b. सीमा ने रमेश को इला को सामान खोलने का ह्रास देने को बोला।

(4) a. अलक ने समीर को कलम छिपाने को कहा।
b. राहुल ने अलक को समीर को कलम छिपाने को कहने को उकसाया।

(5) a. प्रवल ने अनिल को चाबी उठाने का आदेश दिया।
b. शिलपा ने प्रवल ने अनिल को चाबी उठाने का आदेश देने की राय दी।

(6) a. कपिल ने संदीप को कंबल ओढ़ने की नसीहत दी।
b. जयंत ने कपिल को संदीप को कंबल ओढ़ने की नसीहत देने का सुझाव दिया।

(7) a. मालती ने इंद्र को दूध उबालने का आदेश दिया।
b. कमलेश ने मालती को इंद्र को दूध उबालने का आदेश देने की सलाह दी।

(8) a. पूजा ने सुदीप को पतंग उड़ाने का आदेश दिया।
b. लता ने पूजा को सुदीप को पतंग उड़ाने का आदेश देने को कहा।

(9) a. उमिला ने सुनीता को रोटी खाने को बोला।
b. पूजा ने उमिला को सुनीता को रोटी खाने को बोलने की इजाजत दी।

(10) a. लता ने दिलीप को दुकान बेचने की सलाह दी।
b. अजय ने लता को दिलीप को दुकान बेचने की सलाह देने को बोला ।

(11) a. नरेंद्र ने राव को पत्र लौटने का सुझाव दिया ।
   b. प्रमोद ने नरेंद्र को राव को पत्र लौटने का सुझाव देने का हुक्म दिया ।

(12) a. ज्योति ने राजू को गीत सुनाने की राय दी ।
   b. गिरीश ने ज्योति को राजू को गीत सुनाने की राय देने का हुक्म दिया ।

(13) a. सीमा ने राव को घर बेचने को उकसाया ।
   b. रीता ने सीमा को राव को घर बेचने को उकसाने का सुझाव दिया ।

(14) a. बिल्लू ने अरुण को झूठ पीने को उकसाया ।
   b. गीता ने बिल्लू को अरुण को झूठ पीने को उकसाने को कहा ।

(15) a. कमला ने प्रीति को कागज जलाने की सलाह दी ।
   b. अशोक ने कमला को प्रीति को कागज जलाने की सलाह देने का आदेश दिया ।

(16) a. कार्तिक ने कुसुम को पानी उबालने की इजाजत दी ।
   b. अभय ने कार्तिक को कुसुम को पानी उबालने की इजाजत देने को कहा ।

A.2 Experiment 2

I present conditions A (single embeddings, no case marking on final NP) and B (double
embeddings, no case marking on final NP); the stimuli for the other two conditions
can be inferred from the ones presented below since conditions C and D differ from A
and B in having case marking on the final NP. For each of the 16 stimulus sentences,
I first present the condition A version and then the condition B version.

(1) a. सीता ने हरि को किताब सरायेदेने को कहा ।
   b. राव के सीता को हरि को किताब सरायेदेने को कहने की सलाह दी ।
(2) a. सीमा ने रवि को घर बेचने को उकसाया ।
    b. रीता ने सीमा को रवि को घर बेचने को उकसाने का सुझाव दिया ।

(3) a. कांता ने सुनीता को रोटी खाने को बोला ।
    b. पुनीता ने कांता ने सुनीता को रोटी खाने को बोलने की इजाजत दी ।

(4) a. प्रवं ने अनिल को चाबी उठाने का आदेश दिया ।
    b. जिलपा ने प्रबल को अनिल को चाबी उठाने का आदेश देने की राय दी ।

(5) a. रितु ने उमेश को कार चलाने को कहा ।
    b. प्रकाश ने रितु को उमेश को कार चलाने को कहने की नसीहत दी ।

(6) a. अभय ने अरुणा को जहर पीने को उकसाया ।
    b. गीता ने अभय को अरुणा को जहर पीने को उकसाने को कहा ।

(7) a. लता ने दिलीप को दुकान बेचने की सलाह दी ।
    b. अजय ने लता को दिलीप को दुकान बेचने की सलाह देने को बोला ।

(8) a. कपिल ने संदीप को कंबल आदने की नसीहत दी ।
    b. जयंत ने कपिल को संदीप को कंबल आदने की नसीहत देने का सुझाव दिया ।

(9) a. रमेश ने इला को सामान खोलने का हुक्म दिया ।
    b. सीमा ने रमेश को इला को सामान खोलने का हुक्म देने को बोला ।

(10) a. कमला ने अनुज को कागज जलाने की सलाह दी ।
    b. अशोक ने कमला को अनुज को कागज जलाने की सलाह देने का आदेश दिया ।

(11) a. नरेश ने रवि को पत्र खोजने का सुझाव दिया ।
    b. प्रमोद ने नरेश को रवि को पत्र खोजने का सुझाव देने का हुक्म दिया ।

223
(12) a. मालती ने इंद्र को चाय बनाने का आदेश दिया।
   b. कमलेश ने मालती को इंद्र को चाय बनाने का आदेश देने की सलाह दी।

(13) a. अलक ने सुधीर को कलम छिपाने को कहा।
   b. राहुल ने अलक को सुधीर को कलम छिपाने को कहने को उकसाया।

(14) a. जयंत ने कुमुम को पानी उबालने की इजाजत दी।
   b. अभय ने जयंत को कुमुम को पानी उबालने की इजाजत देने को कहा।

(15) a. ज्योति ने राज को गीत सुनाने की राय दी।
   b. गिरीश ने ज्योति को राज को गीत सुनाने की राय देने का इतना दिया।

(16) a. तनुजा ने सुधीप को पतंग उड़ाने का आदेश दिया।
   b. लता ने तनुजा को सुधीप को पतंग उड़ाने का आदेश देने को कहा।

A.3 Experiment 3

This experiment had six conditions, which I call A, B, C, D, E, G. The sentences for conditions A, D, E, G are presented below; the other conditions can be inferred from these.

(1) a. चमन ने दिव्या को जनक को कविता सुनाने को उकसाया।
   b. चमन ने दिव्या से जनक से कविता को सुनाने को कहा।
   c. चमन ने दिव्या से जनक से कविता को सुनाने को कहा।
   d. चमन ने दिव्या को जनक से कविता को सुनाने को कहा।

(2) a. शिलपा ने प्रवल को अनिल को चाबी देने को कहा।
   b. शिलपा ने प्रवल से अनिल से चाबी को लेने को कहा।
   c. शिलपा ने प्रवल से अनिल को चाबी को देने को कहा।
   d. शिलपा ने प्रवल को अनिल से चाबी को लेने को कहा।

224
(3) a. रीटा ने सीमा को रवि को घर वापस करवाने को कहा।

b. रीटा ने सीमा से रवि से घर को बिकवाने को कहा।

c. रीटा ने सीमा से रवि को घर को वापस करवाने को कहा।

d. रीटा ने सीमा को रवि से घर को बिकवाने को कहा।

(4) a. गीता ने अभय को अरुणा को पेड़ दिखाने को कहा।

b. गीता ने अभय से अरुणा से पेड़ उठवाने को कहा।

c. गीता ने अभय से अरुणा को पेड़ को दिखाने को कहा।

d. गीता ने अभय को अरुणा से पेड़ को कटवाने को कहा।

(5) a. जयंत ने कपिल को संदीप को कंबल देने का निर्देश किया।

b. जयंत ने कपिल से संदीप से कंबल को लेने का निर्देश किया।

c. जयंत ने कपिल से संदीप को कंबल को दिखाने का निर्देश किया।

d. जयंत ने कपिल को संदीप से कंबल को लेने का निर्देश किया।

(6) a. अशोक ने कमला को अनुज को मसाला चिस्तवाने का आदेश दिया।

b. अशोक ने कमला से अनुज से मसाले को विच्छेदन किया।

c. अशोक ने कमला से अनुज को मसाले को चिस्तवाने को कहा।

d. अशोक ने कमला को अनुज से मसाले को चिस्तवाने का आदेश दिया।

(7) a. कमलेश ने मालती को इंदर को चाय पिलवाने को कहा।

b. कमलेश ने मालती से इंदर से चाय को बिच्छेदन किया।

c. कमलेश ने मालती से इंदर को चाय को पिलवाने को कहा।

d. कमलेश ने मालती को इंदर से चाय को बिच्छेदन को कहा।

(8) a. अभय ने जयंत को कुमुद को पानी देने को कहा।

b. अभय ने जयंत से कुमुद से पानी को लेने को कहा।

c. अभय ने जयंत से कुमुद को पानी को देने को कहा।
d. अभय ने जयंत को कुमुद से पानी को लेने को कहा।

(9) a. लता ने तनुजा को सुदीप को पतंग देने को कहा।
   b. लता ने तनुजा से सुदीप से पतंग को उड़वाने को कहा।
   c. लता ने तनुजा से सुदीप को पतंग को देने को कहा।
   d. लता ने तनुजा को सुदीप से पतंग को उड़वाने को कहा।

(10) a. रूचि ने सीता को रवि को किताब पढ़वाने का निवेदन किया।
    b. रूचि ने सीता से रवि से किताब को पढ़वाने का निवेदन किया।
    c. रूचि ने सीता से रवि को किताब को पढ़वाने का निवेदन किया।
    d. रूचि ने सीता को रवि से किताब को पढ़वाने का निवेदन किया।

(11) a. पुनिता ने कांता को सुनीता को रोटी चखवाने को कहा।
    b. पुनिता ने कांता से सुनीता से रोटी को चखवाने को कहा।
    c. पुनिता ने कांता से सुनीता को रोटी को चखवाने को कहा।
    d. पुनिता ने कांता को सुनीता से रोटी को चखवाने को कहा।

(12) a. इंद्रानी ने हेमंत को तपन को कमरा दिखाने का मुझाव दिया।
    b. इंद्रानी ने हेमंत से तपन से कमरे को साफ करवाने को कहा।
    c. इंद्रानी ने हेमंत से तपन को कमरे को दिखाने को कहा।
    d. इंद्रानी ने हेमंत को तपन से कमरे को साफ करवाने का मुझाव दिया।

(13) a. संघ्या ने उतम को चमन को चान लौटाने को कहा।
    b. संघ्या ने उतम से चमन से पान को बनवाने को कहा।
    c. संघ्या ने उतम से चमन को पान को लौटाने को कहा।
    d. संघ्या ने उतम को चमन से पान को बनवाने को कहा।

(14) a. प्रकाश ने रितु को उमेश को कार साफ करवाने को कहा।
    b. प्रकाश ने रितु से उमेश से कार को साफ करवाने को कहा।
c. प्रकाश ने रितु से उमेश को कार को दिखाने को कहा ।
d. प्रकाश ने रितु को उमेश से कार को साफ करवाने को कहा ।

(15)  a. प्रमोद ने नरेश को रवि को पत्र पढ़वाने की बिनती की ।
b. प्रमोद ने नरेश से रवि से पत्र को पढ़वाने की बिनती की ।
c. प्रमोद ने नरेश से रवि को पत्र को पढ़वाने की बिनती की ।
d. प्रमोद ने नरेश को रवि से पत्र को पढ़वाने की बिनती की ।

(16)  a. अजय ने लता को दिलीप को दुकान दिखाने का निवेदन किया ।
b. अजय ने लता से दिलीप से दुकान को खुलवाने का निवेदन किया ।
c. अजय ने लता से दिलीप को दुकान को दिखाने का निवेदन किया ।
d. अजय ने लता को दिलीप से दुकान को खुलवाने का निवेदन किया ।

(17)  a. सीमा ने रमेश को इला को चौराहा पार करवाने को कहा ।
b. सीमा ने रमेश से इला से चौराहे को पार करवाने को कहा ।
c. सीमा ने रमेश से इला को चौराहे को पार करवाने को कहा ।
d. सीमा ने रमेश को इला से चौराहे को पार करवाने को कहा ।

(18)  a. राहुल ने अलक को सुधीर को सितार दिखाने का आदेश दिया ।
b. राहुल ने अलक से सुधीर से सितार को ठीक करवाने को कहा ।
c. राहुल ने अलक से सुधीर को सितार को दिखाने को कहा ।
d. राहुल ने अलक को सुधीर से सितार को ठीक करवाने का आदेश दिया ।

(19)  a. गिरीश ने ज्योति को राज को गीत गावाने का हुक्म दिया ।
b. गिरीश ने ज्योति से राज से गीत को गावाने का निवेदन किया ।
c. गिरीश ने ज्योति से राज को गीत को गावाने का निवेदन किया ।
d. गिरीश ने ज्योति को राज से गीत को गावाने का निवेदन किया ।

(20)  a. नीति ने मनीष को केतन को सवाल समझाने को कहा ।

227
b. नीति ने मनीष से केंद्र से सवाल को सुलझाने को कहा।

c. नीति ने मनीष से केंद्र से सवाल को समझाने को कहा।

d. नीति ने मनीष को केंद्र से सवाल को सुलझाने को कहा।

(21) a. जयदेव ने किरण को ललित को पैसे देने कहा।

b. जयदेव ने किरण से ललित से पैसे को लेने को बोला।

c. जयदेव ने किरण से ललित को पैसे को लौटाने को कहा।

d. जयदेव ने किरण को ललित से पैसे को लेने को कहा।

(22) a. विभा ने चेतन को गौरी को साइकिल देने को कहा।

b. विभा ने चेतन से गौरी से साइकिल को तुड़वाने को कहा।

c. विभा ने चेतन से गौरी को साइकिल को देने को कहा।

d. विभा ने चेतन को गौरी से साइकिल को तुड़वाने को कहा।

(23) a. लीना ने दीपक को माला को चावल चिप्कियाने को बोला।

b. लीना ने दीपक से माला से चावल को चिप्कियाने को कहा।

c. लीना ने दीपक से माला को चावल को चिप्कियाने को कहा।

d. लीना ने दीपक को माला से चावल को चिप्कियाने को बोला।

(24) a. हंसराज ने रमा को गौरव को फोटो देने को कहा।

b. हंसराज ने रमा से गौरव से फोटो को लेने को कहा।

c. हंसराज ने रमा से गौरव को फोटो को देने को कहा।

d. हंसराज ने रमा को गौरव से फोटो को लेने को कहा।

(25) a. बीना ने मित्थल को नेहा को खिलौना देने का आदेश दिया।

b. बीना ने मित्थल से नेहा से खिलौने को लेने को कहा।

c. बीना ने मित्थल से नेहा को खिलौने को देने को कहा।

d. बीना ने मित्थल को नेहा से खिलौने को लेने का आदेश दिया।
(26) a. जयंत ने राज को हूमा को मेज दिखाने का निर्देशन किया।
b. जयंत ने राज से हूमा से मेज को बनवाने का निर्देशन किया।
c. जयंत ने राज से हूमा को मेज का दिशाने का निर्देशन किया।
d. जयंत ने राज को हूमा से मेज को बनवाने का निर्देशन किया।

(27) a. तारा ने पदमा को सुधीर को पंसल वापस करने का निर्देशन किया।
b. तारा ने पदमा से सुधीर से पंसल को तेज करवाने का निर्देशन किया।
c. तारा ने पदमा से सुधीर को पंसल को वापस करने का निर्देशन किया।
d. तारा ने पदमा को सुधीर से पंसल को तेज करवाने का निर्देशन किया।

(28) a. सारिता ने परि को महीप को इश्तेहार पढ़वाने को कहा।
b. सारिता ने परि से महीप से इश्तेहार को पढ़वाने को कहा।
c. सारिता ने परि से महीप को इश्तेहार को पढ़वाने को कहा।
d. सारिता ने परि को महीप से इश्तेहार को पढ़वाने को कहा।

(29) a. सारिका ने गगन को तनुजा को दाल देने को कहा।
b. सारिका ने गगन से तनुजा से दाल को लेने को कहा।
c. सारिका ने गगन से तनुजा को दाल को देने को कहा।
d. सारिका ने गगन को तनुजा से दाल लेने को कहा।

(30) a. अनीश ने जया को भरत को कंगन पढ़वाने की राय दी।
b. अनीश ने जया से भरत से कंगन को बनवाने को कहा।
c. अनीश ने जया से भरत को कंगन को बनवाने को कहा।
d. अनीश ने जया को भरत से कंगन को पढ़वाने को कहा।

(31) a. जनक ने ज्योत्सना को तपन को कमीज पहनाने को बोला।
b. जनक ने ज्योत्सना से तपन से कमीज को साफ करवाने को बोला।
c. जनक ने ज्योत्सना से तपन को कमीज को पहनवाने को बोला।
d. नक ने ज्योत्सना को तपन से कमीज को साफ करवाने को बोला ।

(32) a. यामिनी ने जीतेंद्र को कमल को उपयोग करवाने को कहा ।
b. यामिनी ने जीतेंद्र से कमल से उपयोग करवाने को कहा ।
c. यामिनी ने जीतेंद्र से कमल को उपयोग करवाने को कहा ।
d. यामिनी ने जीतेंद्र को कमल से उपयोग करवाने को कहा ।

(33) a. ओमप्रकाश ने कली को एकादाक को कहानी सुनाने को कहा ।
b. ओमप्रकाश ने कली से एकादाक से कहानी को सुनाने को कहा ।
c. ओमप्रकाश ने कली से एकादाक को कहानी को सुनाने को कहा ।
d. ओमप्रकाश ने कली को एकादाक से कहानी को सुनाने को कहा ।

(34) a. हृदेनंत ने पायल को शिव को अखारा दिखाने को कहा ।
b. हृदेनंत ने पायल से शिव से अखारा को जलवाने को कहा ।
c. हृदेनंत ने पायल से शिव को अखारा को दिखाने को कहा ।
d. हृदेनंत ने पायल को शिव से अखारा को जलवाने को कहा ।

(35) a. मालिका ने यश को अदित्य को कारोबार बेचने का निदेशन किया ।
b. मालिका ने यश से अदित्य से कारोबार को बिकवाने का निदेशन किया ।
c. मालिका ने यश से अदित्य को कारोबार को बेचने का निदेशन किया ।
d. मालिका ने यश को अदित्य से कारोबार को बिकवाने का निदेशन किया ।

(36) a. लोकेश ने तेजल को रमिता को कोट देने का इच्छुक दिया ।
b. लोकेश ने तेजल से रमिता से कोट को लेने को कहा ।
c. लोकेश ने तेजल से रमिता को कोट को देने को कहा ।
d. लोकेश ने तेजल को रमिता से कोट को लेने का इच्छुक दिया ।
A.4 Experiment 4

In this experiment there were 16 distinct stimuli sentences; based on these, the sentences for the four conditions were created. These 16 sentences are shown below.

1. सीता ने हूरि को किताब खरीदने को कहा।
2. रितु ने उमेश को गाड़ी चलाने को कहा।
3. रमेश ने इला को सामान खोलने का दुकान दिया।
4. अलक ने समीर को कलम छिपाने को कहा।
5. प्रबल ने अनिल को चारी उठाने का आदेश दिया।
6. कपिल ने संदीप को कंबल औढ़ने की नसीहत दी।
7. मालती ने इंद्र को दूध उबालने का आदेश दिया।
8. पूजा ने सुदीप को पतंग उड़ाने का आदेश दिया।
9. उमेंदा ने सुनीता को रोटी खाने को बोला।
10. लता ने दिलीप को दुकान बेचने की सलाह दी।
11. नरेंद्र ने रवि को पत्र खोलने का सुझाव दिया।
12. ज्योती ने राजु की गीत सुनाने की राय दी।
13. सीमा ने रवि को घर बेचने को उकसाया।
14. बिल्लू ने अरवण को जहर पीने को उकसाया।
15. कमला ने ग्रीति को कागज जलाने की सलाह दी।
16. कार्तिक ने कुसुम को पानी उबालने की इजाजत दी।
A.5 Experiments 5 and 6

The stimuli for these two sentences were identical, the only difference was that in Experiment 5 one of the conditions involved fronting the indirect object, whereas in Experiment 6 the corresponding condition involved fronting the direct object. There were six items per condition; since there were four conditions, this means that there were 24 unique sentences. These are shown below; these sentences are the unscrambled, non-case-marked versions.

1. सीता ने हृद को किताब खरीदने को कहा।
2. रितु ने उमेश को कार चलाने को कहा।
3. रमेश ने इला को सामान खोलने का हुस्स दिया।
4. अलक ने समीर को कलम छिपाने को कहा।
5. ज्योति ने यशपाल को कोट उतारने की नसीहत दी।
6. तपन ने पायल को चादर सुखाने को कहा।
7. सीमा ने रवि को घर बेचने को उकसाया।
8. अभय ने अरुण को जहर पीने को उकसाया।
9. कमला ने अनुज को कागज जलाने की सलाह दी।
10. जयंत ने कुमुद को पानी उबालने की इजाजत दी।
11. दिव्या ने उत्तम को बस रोकने को बोला।
12. जीतेंद्र ने रमिता को रस्सी तोड़ने को उकसाया।
13. गीता ने सुनीता को रोटी खाने को बोला।
14. लता ने दिलीप को दुकान देखने की सलाह दी।
(15) नरेंद्र ने रवि को पत्र खोजने का सुझाव दिया।
(16) ज्योति ने राज को गीत सुनाने की राय दी।
(17) ज्योतिष्णा ने ओमप्रकाश टायर को बदलने का हुक्म दिया।
(18) तारा ने पदमा को गिलास जोड़ने की इजाजत दी।
(19) प्रब्ल ने अनिल को चाबी उठाने का आदेश दिया।
(20) कपिल ने संदीप को कब्ल ओढ़ने की नसीहत दी।
(21) मालती ने इंद्र को चाव बनाने का आदेश दिया।
(22) तनुजा ने सुंदरेश को पतंग उड़ाने का आदेश दिया।
(23) लोकेश ने तेजल को दाल पकाने का सुझाव दिया।
(24) निती ने मनीष को सेव छोड़ने को बोला।

A.6 Experiment 7

There were six items per condition; since there were four conditions, this means that there were 24 unique sentences. These are shown below; these sentences are the single and double embedding versions.

(1) a. सीता ने दुर्ग को किताब को जितनी जल्दी ही सके खरीदने को कहने की सलाह दी।
   b. कपिल ने सीता को दुर्ग को किताब को जितनी जल्दी ही सके खरीदने को कहने की सलाह दी।

(2) a. सीमा ने रवि को घर को छो: वज्र से पहले बेबने को उकसाया।
   b. रवि ने सीमा को रवि को घर को छो: वज्र से पहले बेबने को उकसाने को कहा।
(3) a. कांता ने सुनीता को रोटी को जितनी जल्दी हो सके खाने को बोला।
b. पुष्पा ने कांता को सुनीता को रोटी को जितनी जल्दी हो सके खाने को बोलने को कहा।

(4) a. प्रवंल ने अनिल को चाकी को बहुत ध्यान से उठाने का आदेश दिया।
b. शिल्पा ने प्रवंल को अनिल को चाकी को बहुत ध्यान से उठाने का आदेश देने को कहा।

(5) a. अभय ने अरुणा को ज़हर को सबके सामने पीने को उकसाया।
b. रमन्ता ने अभय को अरुणा को ज़हर को सबके सामने पीने को उकसाने को बोला।

(6) a. रितू ने उमेश को कार को बिना जल्दी किये चलाने को कहा।
b. प्रकाश ने रितू को उमेश को कार को बिना जल्दी किये चलाने को कहने की नसीहत दी।

(7) a. लता ने दिलीप को दुकान को बिना जल्दी किये देखने की सलाह दी।
b. अजय ने लता को दिलीप को दुकान को बिना जल्दी किये देखने की सलाह देने को बोला।

(8) a. कपिल ने संदीप को कंबल को ठीक तरह से आधार की नसीहत दी।
b. जयांत ने कपिल को संदीप को कंबल को ठीक तरह से आधार की नसीहत देने का सुझाव दिया।

(9) a. रमेश ने इला को सामान को मेज पर रखकर खोलने का हुक्म दिया।
b. सीमा ने रमेश को इला को सामान को मेज पर रखकर खोलने का हुक्म देने को बोला।

(10) a. नरेंद्र ने रवि के पत्र को हर कमरे में खोजने का सुझाव दिया।
b. प्रभादेव ने नरेंद्र को रवि के पत्र को हर कमरे में खोजने का सुझाव देने का हुक्म दिया।
(11) a. कमला ने अनुज को कागज को धाती पर रखकर जलाने की सलाह दी ।
b. प्रमोद ने कमला को अनुज को कागज को धाती पर रखकर जलाने की सलाह देने को कहा ।

(12) a. मालती ने इंद्र को चाय को रसोई घर में बनाने का आदेश दिया ।
b. कमलेश ने मालती को इंद्र को चाय को रसोई घर में बनाने का आदेश देने की सलाह दी ।

(13) a. अलक ने सुधीर को कलम को ठीक तरह से छिपाने को कहने को उकसाया ।
b. राहुल ने अलक को सुधीर को कलम को ठीक तरह से छिपाने को कहने को उकसाया ।

(14) a. जयंत ने कुसुम को पानी को अच्छी तरह से उवालने को कहा ।
b. राहुल ने जयंत को कुसुम को पानी को अच्छी तरह से उवालने को कहने को बोला ।

(15) a. तनुजा ने सुदीप को पतंग को खुले मैदान में उड़ाने का आदेश दिया ।
b. लता ने तनुजा ने सुदीप को पतंग को खुले मैदान में उड़ाने का आदेश देने को कहा ।

(16) a. ज्योति ने राज को गीत को दिल लगाकर सुनाने की राय दी ।
b. गिरीश ने ज्योति को राज को गीत को दिल लगाकर सुनाने की राय देने का हुक्म दिया ।

(17) a. ज्योति ने यज्ञपाल को कोट को जितनी जल्दी हो सके उतारने की नसीहत दी ।
b. हरि ने ज्योति को यज्ञपाल को कोट को जितनी जल्दी हो सके उतारने की नसीहत देने को कहा ।

(18) a. लोकेश ने तेजस को दाल को छुः बजे से पहले पकाने को कहा ।
b. हरपाल ने लोकेश को तेजल को दाल को छः बजे से पहले पकाने को कहने का मुहूर्त दिया।

(19) a. ज्योत्सना ने ओमप्रकाश को टायर को ठीक तरह से बदलने का हुक्कम दिया।
b. सीमा ने ज्योत्सना को ओमप्रकाश को टायर को ठीक तरह से बदलने का हुक्कम देने को बोला।

(20) a. दिव्या ने उतम को बस को खुले मैदान में रोकने को बोला।
b. गीता ने दिव्या को उतम को बस को खुले मैदान में रोकने को बोलने को कहा।

(21) a. नीति ने मनीष को सेब को थाली पर रखकर छीलने को बोला।
b. चमन ने नीति को मनीष को सेब को थाली पर रखकर छीलने को बोलने को कहा।

(22) a. जीतेंद्र ने रमिता को रस्सी को पेड़ पर लटकाकर तोड़ने को उकसाया।
b. हृदि ने जीतेंद्र को रमिता को रस्सी को पेड़ पर लटकाकर तोड़ने को उकसाने को कहा।

(23) a. तारा ने पदमा को गिलास को मेज पर रखकर जोड़ने की इजाजत दी।
b. कपिल ने तारा को पदमा को गिलास को मेज पर रखकर जोड़ने की इजाजत देने को कहा।

(24) a. तपन ने पायल को चादर को अच्छी तरह से सुखाने को कहा।
b. काजल ने तपन को पायल को चादर को अच्छी तरह से सुखाने को कहने को बोला।
Index

Aaronson, D., 151
abducible structure, 191
abduction, 187
abduction cost, 196
Abney, S., 152
accessibility, 198
ACT-R, 55
affixes, 13
Layer I, 13
Layer II, 13
Layer III, 13
Aho, A. V., 200
Aissen, J., 15
Altmann, G., 49, 80, 155
Anderson, J., 55
Aron, A., 151
availability, 198
Axiom of Constructibility, 45
Babylonyshev, M., 53
Bach, E., 4, 47, 214
background, 20, 23
Baddeley, A., 7, 56
Bever, T., 39, 186
Bickel, B., 9, 11
Bossong, G., 15, 19
Brants, T., 38
Briscoe, E., 201, 202
Brown, J., 7
Butt, M., 10
Carlson, G., 32
case-marked object, 16
center embeddings, see self-center embeddings
Christiansen, M. H., 6
Cohen, J. D., 71
Comrie, B., 15, 16, 19, 29, 153
Console, L., 189
Constellations, 14
Constituent Recognition Domain, 44
Construal, 186
content addressable, 198
tcontrol constructions, 9
Cowan, N., 7, 58
Croft, W., 16
crossed dependencies, 48
Cussens, J., 188
Dayal, V., 31, 32
decay, 7, 218
definite-marked object, 16
definiteness effect, 104, 108, 111, 126, 141
Dependency Locality Theory, 51
Dickey, M., 4
Differential object marking, 15
discourse complexity assumption, 53
discourse complexity assumption (DLT), 52
discourse referents, 52
distance hypothesis, 106
DLT, see Dependency Locality Theory
DOM, see Differential object marking
Early Immediate Constituents, 44
encoding, 6
Estes, W. K., 54
Ferreira, F., 73, 151, 161, 167
focus of attention, 58, 60, 62
foreground, 20, 23
Frazier, L., 186

Gambhir, V., 15, 112
garden-path sentences, 3, 39, 185, 186
Gibson, E., 3, 39, 51–53, 55, 145, 151, 183, 184, 203, 218
Givón’s hypothesis, 20, 68, 81, 104, 216
Gordon, P. C., 6, 39, 40, 54
Grammatical Hierarchy Prediction, 46
Hale, J., 38
Harman, G., 187
Hawkins, J., 3, 41, 43–47, 49, 50
Hepple, M., 189
Hindi, 2
Hobbs, J. R., 188
Hopper, P. J., 19–21, 23, 30
Humphreys, M. S., 54, 59
Hyönä, J., 108
hypothesis formers, 189

Ihaka, R., 161
immunity from interference, 58, 60, 62
Individuation, 19
inference to the best explanation, 187
infinitivials as nominals, 10–14, 105, 155
integration, 6
interference, 7

Janda, R., 14
Johnson, N., 74, 197
Johnson-Laird, P. N., 202
Joseph, B. D., 10
Josephson, J., 187
Joshi, A., 7, 201
Jurafsky, D., 38, 81
Just, M. A., 71

Kaan, E., 4, 212
Kaiser, E., 49
Keppel, G., 7
Kidwai, A., 15
Klabunde, R., 187
König, E., 189
Kratzer, A., 32
Kruijff, G. J. M., 188, 189
Kuno, S., 56
Kurz, D., 49, 182
Layer I, see affixes
Layer II, see affixes
Layer III, see affixes
Lewis, R. L., 3, 7, 8, 53, 57, 58, 69, 108
Linguistic Data Consortium, 22

Mahajan, A., 15, 216
Marcus, M., 201
marked object, 16
Masica, C., 10, 13, 15, 30
matching, 193
Maxwell, S. E., 73, 167
Mazuka, R., 185
McElree, B., 198
medial gapping constructions, 66
Melton, A. W., 55
Miller, G., 6
minimal consistency, 188, 191, 192
minimal consistency, 192
mismatch cost, 196
Miyamoto, E. T., 108, 191
MMLG, see Multi-Modal Logical Grammar
Mohanen, T., 11, 17, 31, 34
Moortgat, M., 188
Mother Node Construction, 45
Müller, G. E., 7
Multi-Modal Logical Grammar, 188

Nakatani, K., 53, 208
Nakayama, M., 3
Narayanan, S., 38
nested dependencies, 48
Norman, D. A., 7

238
Peterson, L. R., 7  
Petersohn, M. J., 7  
PI, see proactive interference  
Pinzelter, A., 7  
Pinheiro, J. C., 73  
Pitmanjtotara, 186  
Pollard, C., 189  
Porterfield, L., 31  
position effect, 151  
positional similarity hypothesis, 126  
Prasad, R., 36  
Prince, A., 16  
proactive interference, 7, 54  
probabilistic parsing, 38  
Rambow, O., 7  
Resnik, P., 197  
resource bounded (working memory), 7  
retrieval, 6  
retrieval cues, 60  
retrieval requests, 60  
retroactive interference, 7, 54  
RI, see proactive interference  
Richardson, J. T. E., 7  
Ross, J., 66  
SCE, see self-center embeddings  
self-center embeddings, 4  
Dutch-German contrast, 47  
Shieber, S. M., 194, 201, 202  
short-term memory, see working memory  
Sikkel, K., 200  
similarity  
phonological, 56  
positional, 57, 64, 111, 126  
superficial, 56, 64, 183  
syntactic, 55, 57, 182  
similarity hypothesis, 82  
simplicity criterion, 190  
Singh, M., 30  
Smith, E., 187  
speed-accuracy tradeoff, 198  
Stabler, E., 8  
Steedman, M., 49, 80, 155, 189  
stochastic context-free grammars, 39  
storage, 6  
Strigum, A., 188  
Transitivity, 19  
Trueswell, J., 73, 160, 167  
Tukey, J., 161  
Tyler, L. K., 202  
Uehara, K., 4, 58, 59  
Underwood, B. J., 7  
Urdu, 2  

Vaid, J., 3, 73  
Van Dyke, J., 178  
Vasishth, S., 15, 60, 185  
Verma, M. K., 31, 127  
Vermaat, W., 189  
Warren, T., 52, 53, 80, 182  
Waugh, N. C., 7  
Wescoat, M., 17  
working memory, 4  
working memory, syntactic, 55, 154  
Yamashita, H., 190, 194, 210  
Yngve, V. H., 6


239
REFERENCES


& D. Dowty (Eds.), *Proceedings of the Second Conference on Semantics and 
Linguistic Theory* (pp. 39–58).

Dayal, V. (1999). Bare NPs, reference to kinds, and incorporation. In T. Matthews 
& D. Strollovitc (Eds.), *Proceedings of the Ninth Conference on Semantics and 
Linguistic Theory* (pp. ??–??).

study.* (MS)

Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens’ handbook of experimental 

*Journal of Memory and Language, 25,* 348–368.


Unpublished doctoral dissertation, University of Massachusetts, Amherst.


Gambhir, V. (1981). *Syntactic restrictions and discourse functions of word order in 
standard Hindi.* Unpublished doctoral dissertation, University of Pennsylvania, 
Philadelphia.

UK: Lawrence Erlbaum Associates.

New Jersey: Lawrence Erlbaum.

*Cognition, 68,* 1–76.

complexity. In A. Marantz, Y. Miyashita, & W. O’Neil (Eds.), *Image, language, 
brain: Papers from the first mind articulation project symposium.* Cambridge, 
MA: MIT Press.


Jurafsky, D. (2002). *Probabilistic models of human linguistic processing.* (Talk presented at The Ohio State University, May 17)


246


Miyamoto, E. T. (2002). *Case markers as clause boundary inducers in Japanese.* (MS, under review; cited with permission)

Miyamoto, E. T., & Takahashi, S. (2001). Sources of difficulty in processing japanese. In M. Nakayama (Ed.), *Sentence processing in East Asian Languages* (pp. ??– ??). Stanford, CA: CSLI.


Van Dyke, J., & Lewis, R. L. (2002). *Distinguishing separate effects of structure and decay on attachment and repair of temporarily ambiguous sentences.* University of Michigan. (MS)


