How listeners resolve reference: Effects of pitch accent, edge tones, and lexical contrast.

Dissertation

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

By
Anouschka Foltz, M.A.

Graduate Program in Linguistics
The Ohio State University
2010

Dissertation Committee:
Professor Shari R. Speer, Advisor
Professor Cynthia Clopper
Professor Judith Tonhauser
Speakers use prosodic structure to group words together into prosodic phrases and to highlight individual words by making them perceptually more salient. This dissertation examined the impact of prosody on the process of identifying the referents of definite noun phrases (e.g. *the purple bottle*) across utterances. Participants’ eye movements were monitored as they used a mouse to select objects on a computer screen in response to auditory instructions (e.g. *Click on the purple bottle*). The research explored two primary aspects of intonation, pitch accents (tonal targets associated with metrically stressed syllables) and edge tones (pitch movements associated with the edge of a phrase). While previous studies focused on the role of local pitch accents (i.e. in the target instruction), especially L+H* or “contrastive” accents, the current work considers the effect of prosodic context - non-local edge tones and pitch accents (i.e. in preceding instructions) - on referent identification. This dissertation poses two novel questions about the function of prosody in referent identification: How do local and non-local pitch accents act in combination to affect referent identification? How do pitch accents interact with phrasal edge tone patterns to affect referent identification? Across utterances, pitch accents were placed in locations that appropriately or inappropriately marked a lexical difference (e.g. a difference only in color adjective such as *yellow bottle ⇒ purple bottle*). Expanding on previous findings, appropriate and inappropriate pitch accent locations were expected to speed up and slow down referent identification, respectively. Edge tone patterns were expected to modulate the effects of pitch accents by connecting or separating successive utterances. Results indicate that both local and non-local pitch accents can influence the process of referent identification in spoken discourse processing. In addition, this effect is conditioned by the structure
of prosodic phrasing (edge tone type) in the discourse. Implications of the results for models of discourse processing will be discussed.
Acknowledgements

Writing this dissertation would not have been possible without the help and support of countless people. My advisor, Shari Speer, supported and encouraged me throughout my graduate studies and during the writing of this dissertation. My dissertation committee members, Cynthia Clopper and Judith Tonhauser, helped me look at my data from different angles. Shari Speer, Cynthia Clopper, Judith Tonhauser, and Kiwako Ito carefully read and commented on numerous drafts of my dissertation proposal and dissertation. Laurie Maynell and Liz McCullough produced the stimuli for the experiments. Jon Dehdari, DJ Hovermale, and Andy Plummer provided \LaTeX\ help and troubleshooting. Jon Dehdari wrote the style file that I used as a template for this dissertation. Mary Beckman and Cong Liu provided R help and wrote a sizeable chunk of the R code used for graphing and statistics in this dissertation. Ping Bai wrote the Perl script that converted my raw data into the format needed for graphing and statistical analyses. Dan Mirman patiently answered every single one of my questions about growth curve analysis. Sarah Bibyk helped tag soundfiles for phonetic measurements and ToBI annotated all stimuli. Shari Speer ToBI annotated a subset of the stimuli. Liesel Kautz, Tanya Senne, Katie Carmichael, Nancy Foltz, Becky Hoak, Shawn Foltz, and Jordan Main provided countless hours of babysitting. I would like to heartily and sincerely thank all of the people just mentioned (and everybody I may have forgotten to mention) for their contributions to the work that went into finishing this dissertation. Without them, there would be no dissertation. Any remaining errors, inconsistencies, and flaws are my own.
Vita

2003 ..........................  M.A. Anglistik, Universität Mannheim
2010 ..........................  M.A. Linguistics, The Ohio State University

Fields of Study

Major Field: Linguistics
# Table of Contents

Abstract .................................................................................................................. ii  
Acknowledgements ............................................................................................... iv  
Vita ........................................................................................................................... v  
List of Figures .......................................................................................................... ix  
List of Tables ........................................................................................................... xix  

1 Introduction ......................................................................................................... 1  
1.1 Definite Noun Phrases ..................................................................................... 5  
1.2 Intonational Structure and Meaning ............................................................... 7  
1.2.1 Phrases ........................................................................................................ 8  
1.2.2 Pitch Accents ............................................................................................... 9  
1.2.3 Edge Tones .................................................................................................. 13  
1.3 Focus Structure ............................................................................................... 14  
1.4 Effects of Pitch Accent in Psycholinguistic Studies .......................................... 17  
1.5 Effects of Edge Tones in Psycholinguistic Studies ............................................ 24  
1.6 Thesis Overview .............................................................................................. 26  

2 Methods ............................................................................................................... 27  
2.1 The Methodological Framework .................................................................... 28  
2.1.1 The Visual World Paradigm ...................................................................... 28  
2.1.2 Eyetracking ................................................................................................. 29  
2.1.3 Eye Movements .......................................................................................... 32  
2.2 Participants ...................................................................................................... 34  
2.3 Visual Materials ............................................................................................... 34  
2.3.1 Objects and Colors .................................................................................... 34  
2.3.2 Object Characteristics .............................................................................. 36  
2.3.3 Object Location .......................................................................................... 36  
2.3.4 Choice of Competitor Objects ................................................................... 37  
2.3.5 Display Complexity ................................................................................... 38  
2.3.6 Display Preview .......................................................................................... 38  
2.4 Spoken Materials ............................................................................................ 39  
2.4.1 Color Adjectives and Nouns ...................................................................... 39  
2.4.2 Tunes .......................................................................................................... 41
8 General Discussion and Conclusions .......................................................... 235
  8.1 Pitch Accents ................................................................. 240
  8.2 Edge Tones ................................................................. 244
  8.3 Lexical Information .......................................................... 247
  8.4 Implications for Theories of Focus Structure ..................................... 249
  8.5 General Conclusions ......................................................... 250

References ............................................................................................... 254
List of Figures

2.1 Sample visual display used in the two large EXPERIMENTS. ................. 35
2.2 Object type and color combinations used in the large EXPERIMENTS. ........ 35
2.3 ToBI annotation of *Click on the purple pencil* ↘ (HHNP LL, where ↘ indicates a L-L% pattern), produced by Speaker 1. ........................... 47
2.4 ToBI annotation of *Click on the PURPLE pencil* ↘ (LHA LL, where ↘ indicates a L-L% pattern), produced by Speaker 1. ........................... 48
2.5 ToBI annotation of *Click on the purple PENCIL* ↘ (LHN LL, where ↘ indicates a L-L% pattern), produced by Speaker 1. ......................... 48
2.6 ToBI annotation of *Click on the purple pencil* ↘ (HHNP LL, where ↘ indicates a L-L% pattern), produced by Speaker 2. ......................... 49
2.7 ToBI annotation of *Click on the purple pencil* ↗ (HHNP LH, where ↗ indicates a L-H% pattern), produced by Speaker 2. ......................... 49
2.8 ToBI annotation of *Click on the PURPLE pencil* ↘ (LHA LL, where ↘ indicates a L-L% pattern), produced by Speaker 2. ......................... 50
2.9 ToBI annotation of *Click on the PURPLE pencil* ↗ (LHA LH, where ↗ indicates a L-H% pattern), produced by Speaker 2. ......................... 50
2.10 ToBI annotation of *Click on the purple PENCIL* ↘ (LHN LL, where ↘ indicates a L-L% pattern), produced by Speaker 2. ......................... 51
2.11 ToBI annotation of *Click on the purple PENCIL* ↗ (LHN LH, where ↗ indicates a L-H% pattern), produced by Speaker 2. ......................... 51
2.12 Sample visual display (see Figure 2.1 above) with AOIs and AOI numbers overlaid. ................................................................. 71
2.13 Sample overview graph from EXPERIMENT A. Proportion of fixations to the target object in the display in response to hearing *Click on the PURPLE bottle, Click on the purple bottle* or *Click on the purple BOTTLE* either in a repetition or *color contrast* sequence. .................................................. 73

2.14 Sample graph of a pairwise comparison from EXPERIMENT A. Comparison of the repetition *HHNP LL* and the repetition *LHN LL* conditions. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 75

2.15 Schematic relationship of the order of orthogonal polynomial terms and relevant captured curve shapes. The top panels show schematic target fixation curves, and the bottom panels show schematic competitor fixation curves. ... 80

2.16 Grand mean of looks to the target object cell in the six critical conditions presented in Figure 2.13 in Section 2.9. The arrows indicate the beginning and end of the analysis window for the GCA. .................................................. 83

2.17 Data from EXPERIMENT A. Proportion, log odds, and empirical logit of fixations to the target object (top panels) or color competitor object (bottom panels) of data from Experiment 1 (see Chapter 3 for an explanation of the conditions names). .................................................. 84

3.1 Experiment 1, data from EXPERIMENT A: Proportion of looks to the target object cell in the six critical conditions presented in Tables 3.1 and 3.3. .... 97

3.2 Experiment 1, data from EXPERIMENT A: Grand mean of looks to the target object cell in the six critical conditions presented in Tables 3.1 and 3.3. The arrows indicate the beginning and end of the analysis window for the GCA. 98

3.3 Experiment 1, data from EXPERIMENT A: Comparison of the *felicitous* and *neutral* conditions presented in Table 3.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............ 101
3.4 Experiment 1, data from EXPERIMENT A: Comparison of the felicitous and infelicitous conditions presented in Table 3.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............... 101
3.5 Experiment 1, data from EXPERIMENT A: Comparison of the neutral and infelicitous conditions presented in Table 3.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............... 102
3.6 Experiment 1, data from EXPERIMENT B: Proportion of looks to the target object cell in the six critical conditions presented in Tables 3.2 and 3.4. .... 105
3.7 Experiment 1, data from EXPERIMENT B: Grand mean of looks to the target object cell in the six critical conditions presented in Tables 3.2 and 3.4. The arrows indicate the beginning and end of the analysis window for the GCA. 106
3.8 Experiment 1, data from EXPERIMENT B: Comparison of the felicitous and neutral conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............... 108
3.9 Experiment 1, data from EXPERIMENT B: Comparison of the felicitous and infelicitous conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............... 109
3.10 Experiment 1, data from EXPERIMENT B: Comparison of the neutral and infelicitous conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ............... 109
3.11 Experiment 1, data from EXPERIMENT A: Proportion of looks to the color-competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3. ................................................................. 112
3.12 Experiment 1, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3. The downward arrows indicate the beginning and end of the analysis window for the GCA. The upward arrow represents the middle of the mountain of the curve. ......................... 113

3.13 Experiment 1, data from EXPERIMENT A: Comparison of the garden-path and neutral conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. ................. 116

3.14 Experiment 1, data from EXPERIMENT A: Comparison of the garden-path and infelicitous conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. ............... 117

3.15 Experiment 1, data from EXPERIMENT A: Comparison of the neutral and infelicitous conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. ............... 117

3.16 Experiment 1, data from EXPERIMENT B: Proportion of looks to the color competitor object cell in the six critical conditions presented in Tables 3.2 and 3.4. ................................................. 120

3.17 Experiment 1, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3. The downward arrows indicate the beginning and end of the analysis window for the GCA. The upward arrow represents the middle of the mountain of the curve. ......................... 121

3.18 Experiment 1, data from EXPERIMENT B: Comparison of the garden-path and neutral conditions presented in Table 3.4. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. ............... 124
3.19 Experiment 1, data from EXPERIMENT B: Comparison of the garden-path and infelicitous conditions presented in Table 3.4. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. 125

3.20 Experiment 1, data from EXPERIMENT B: Comparison of the neutral and infelicitous conditions presented in Table 3.4. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits. 125

4.1 Experiment 2, data from EXPERIMENT A: Proportion of looks to the target object cell in the four critical conditions presented in Tables 4.1 and 4.3. 138

4.2 Experiment 2, data from EXPERIMENT A: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 4.1 and 4.3. The arrows indicate the beginning and end of the analysis window for the GCA analyses. 139

4.3 Experiment 2, Color-contrast sequence from EXPERIMENT A: Comparison of the felicitous and anticipatory conditions presented in Table 4.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. 142

4.4 Experiment 2, No-contrast sequence from EXPERIMENT A: Comparison of the garden-path and anticipated garden-path conditions presented in Table 4.3. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. 144

4.5 Experiment 2, data from EXPERIMENT B: Proportion of looks to the target object cell in the four critical conditions presented in Tables 4.2 and 4.4. 145

4.6 Experiment 2, data from EXPERIMENT A: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 4.2 and 4.4. The arrows indicate the beginning and end of the analysis window for the GCA analyses. 146
4.7 Experiment 2, Color-contrast sequence from EXPERIMENT B: Comparison of the felicitous and the anticipatory conditions presented in Table 4.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 147

4.8 Experiment 2, Non-contrastive sequence from EXPERIMENT B: Comparison of the garden-path and the anticipated garden-path conditions presented in Table 4.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 150

4.9 Experiment 2, data from EXPERIMENT A: Proportion of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3. ................................................................. 151

4.10 Experiment 2, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3. The arrows indicate the beginning and end of the analysis window for the GCA analyses. ................................................................. 152

4.11 Experiment 2, No-contrast sequence from EXPERIMENT A: Comparison of the garden-path and the anticipated garden-path conditions presented in Table 4.1. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits. ................................................................. 155

4.12 Experiment 2, Color-contrast sequence from EXPERIMENT A: Comparison of the felicitous and the anticipatory conditions presented in Table 4.3. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits. ................................................................. 156

4.13 Experiment 2, data from EXPERIMENT B: Proportion of looks to the color competitor object cell in the four critical conditions presented in Tables 4.2 and 4.4. ................................................................. 158
4.14 Experiment 2, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3. The arrows indicate the beginning and end of the analysis window for the GCA analyses. ................................................................. 159

4.15 Experiment 2, No-contrast sequence from EXPERIMENT B: Comparison of the garden-path and the anticipated garden-path conditions presented in Table 4.3. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits. ................................................................. 161

4.16 Experiment 2, Color-contrast sequence from EXPERIMENT B: Comparison of the felicitous and the anticipatory conditions presented in Table 4.1. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits. ................................................................. 163

5.1 Experiment 3a, data from EXPERIMENTS A and B: Proportion of looks to the target object cell in the four critical conditions presented in Tables 5.1 and 5.2. 175

5.2 Experiment 3a, data from EXPERIMENTS A and B: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 5.1 and 5.2. The arrows indicate the beginning and end of the analysis window for the GCA analyses. ................................................................. 176

5.3 Experiment 3a, Early L+H* sequence from EXPERIMENTS A and B: Comparison of the early L+H* / finality and early L+H* / continuation conditions presented in Table 5.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 179

5.4 Experiment 3a, Late L+H* sequence from EXPERIMENTS A and B: Comparison of the late L+H* / finality and late L+H* / continuation conditions presented in Table 5.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 181
5.5 Experiment 3a, *Continuation* sequence from EXPERIMENT B: Comparison of the *early L+H*/* continuation* and *late L+H*/* continuation* conditions presented in Tables 5.1 and 5.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. .......................................................... 182

5.6 Experiment 3a, *Finality* sequence from EXPERIMENT A: Comparison of the *early L+H*/* finality* and *late L+H*/* finality* conditions presented in Tables 5.1 and 5.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. .......................................................... 183

5.7 Instruction 1, data from EXPERIMENTS A and B: Proportion of looks to the target object cell in the four conditions described above. ........................................... 189

5.8 Instruction 1, data from EXPERIMENTS A and B: Comparison of the curves derived from all six conditions of EXPERIMENTS A and B where Instruction 1 is pronounced with either a L+H* on the adjective or the noun (*Six Cond*) with the best fitting curve derived from only one of those conditions (*Best Fit*). 191

5.9 Instruction 1, data from EXPERIMENTS A and B: Grand mean of looks to the target object cell in the four conditions described above. The arrows indicate the beginning and end of the analysis window for the GCA analyses. .... 192

5.10 Instruction 1, Data from EXPERIMENTS A and B: Comparison of the conditions with a L+H* accent on the adjective (*LHA*) across EXPERIMENTS. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. .......................................................... 193

5.11 Instruction 1, Data from EXPERIMENTS A and B: Comparison of the conditions with a L+H* accent on the noun (*LHN*) across EXPERIMENTS. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. .......................................................... 195
<table>
<thead>
<tr>
<th>Section</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Experiment 3b, data from EXPERIMENT A: Proportion of looks to the target</td>
<td>205</td>
</tr>
<tr>
<td></td>
<td>object cell in the four critical conditions presented in Tables 6.1 and 6.2.</td>
<td></td>
</tr>
<tr>
<td>6.2</td>
<td>Experiment 3b, data from EXPERIMENTS A and B: Grand mean of looks to the</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>target object cell in the four critical conditions presented in Tables 6.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and 6.2. The arrows indicate the beginning and end of the analysis window</td>
<td></td>
</tr>
<tr>
<td></td>
<td>for the GCA analyses.</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Experiment 3b, Target 3 sequence from EXPERIMENT A: Comparison of the</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>early L+H*, target 3 and late L+H*, target 3 conditions presented in Tables</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.1 and 6.2. Left: Empirical logit of looks to the target object cell and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>error bars. Right: Model fits.</td>
<td></td>
</tr>
<tr>
<td>6.4</td>
<td>Experiment 3b, Target 2 sequence from EXPERIMENT A: Comparison of the</td>
<td>209</td>
</tr>
<tr>
<td></td>
<td>early L+H*, target 2 and late L+H*, target 2 conditions presented in Tables</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.1 and 6.2. Left: Empirical logit of looks to the target object cell and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>error bars. Right: Model fits.</td>
<td></td>
</tr>
<tr>
<td>6.5</td>
<td>Experiment 3b, Early cue sequences from EXPERIMENT A: Comparison of the</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>early L+H*, target 3 and early L+H*, target 2 conditions presented in Table</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6.1. Left: Empirical logit of looks to the target object cell and error bars.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right: Model fits.</td>
<td></td>
</tr>
<tr>
<td>6.6</td>
<td>Experiment 3b, Late cue sequences from EXPERIMENT A: Comparison of the</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>late L+H*, target 3 and late L+H*, target 2 conditions presented in Table 6.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Left: Empirical logit of looks to the target object cell and error bars.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Right: Model fits.</td>
<td></td>
</tr>
<tr>
<td>7.1</td>
<td>Experiment 4, data from EXPERIMENTS A and B: Proportion of looks to the</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>target object cell in the four critical conditions presented in Tables 7.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>and 7.2.</td>
<td></td>
</tr>
</tbody>
</table>
7.2 Experiment 4, data from EXPERIMENTS A and B: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 7.1 and 7.2. The arrows indicate the beginning and end of the analysis window for the GCA. ................................................................. 223

7.3 Experiment 4, Finality sequence from EXPERIMENT A: Comparison of the *early L+H* / *finality* and *no early L+H* / *finality* conditions presented in Tables 7.1 and 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 225

7.4 Experiment 4, Continuation sequence from EXPERIMENT B: Comparison of the *early L+H* / *continuation* and *no early L+H* / *continuation* conditions presented in Tables 7.1 and 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 226

7.5 Experiment 4, Early L+H* sequence from EXPERIMENTS A and B: Comparison of the *early L+H* / *finality* and *early L+H* / *continuation* conditions presented in Table 7.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 228

7.6 Experiment 4, No early L+H* sequence from EXPERIMENTS A and B: Comparison of the *no early L+H* / *finality* and *no early L+H* / *continuation* conditions presented in Table 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits. ................................................................. 228
List of Tables

2.1 Colors and objects used in the sentence materials. .................................................. 40
2.2 Adjective and noun combinations used in the sentence materials for all experiments. Greyed text indicates combinations that occurred in filler trials only. ................................................................. 41
2.3 Summary of ToBI labels for Speaker 1 (EXPERIMENT A): Number of utterances for each tune, where the pitch accents and edge tones on both adjective and noun were on target (On Target), where the pitch accent on the adjective deviated from the target (Adj off Target), where the pitch accent on the noun deviated from the target (Noun off Target), and where the edge tones ending the noun deviated from the target (Edge off Target). Labels in parentheses provide the identity of deviating pitch accents or edge tones. .......................... 46
2.4 Summary of ToBI labels for Speaker 2 (EXPERIMENT B): Number of utterances for each tune, where the pitch accents and edge tones on both adjective and noun were on target (On Target), where the pitch accent on the adjective deviated from the target (Adj off Target), where the pitch accent on the noun deviated from the target (Noun off Target), and where the edge tones ending the noun deviated from the target (Edge off Target). Labels in parentheses provide the identity of deviating pitch accents or edge tones. .......................... 46
2.5 Average segment durations of target words for the HHNP LL, LHA LL, and LHN LL tunes produced by Speaker 1 ................................................................. 54
2.6 Average segment durations of target words for the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2 .......................... 55
2.7 Average pitch of target pitch accents for the HHNP LL, LHA LL and LHN LL tunes produced by Speaker 1. ................................. 57
2.8 Average pitch of target pitch accents for the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2. .................. 58
2.9 Pairwise comparisons from the "adjective" two-way ANOVA for Speaker 2. 59
2.10 Pairwise comparisons from the "noun" two-way ANOVA for Speaker 2. ... 60
2.11 Average pitch of target pitch edge tones (last pitch value in the utterance) for the HHNP LL, LHA LL, and LHN LL tunes produced by Speaker 1 and the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2. .................................................. 61
2.12 Pairwise comparisons from the "adjective" two-way ANOVA for Speaker 2. 62
2.13 Lexical Contrast Types used in EXPERIMENTS A and B, with example lexical strings. **Bold type** indicates color contrasts and **SMALL CAPS** indicates object type contrasts. ................................................................. 65
2.14 Filler trial patterns for EXPERIMENTS A and B. ............................. 66
2.15 Pitch Accent Sequence for ColorContrast/ObjectContrast and NoContrast/ObjectContrast Lexical Contrast Types used in EXPERIMENTS A and B. The pitch accent condition HHNP is indicated by *italics*, the L+H* accent of the LHA and LHN conditions is indicated by CAPS. ........................................ 67
2.16 Pitch Accent Sequence for ObjectContrast/ColorContrast and Repetition/ColorContrast Lexical Contrast Types used in EXPERIMENTS A and B. The pitch accent condition HHNP is indicated by *italics*, the L+H* accent of the LHA and LHN conditions is indicated by CAPS. ........................................ 67
2.17 Data shown in Figure 2.13 in Section 2.9: Target fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction. ...................... 88
2.18 Data shown in Figure 2.13 in Section 2.9: Target fixation analysis comparing the repetition HHNP LL condition (model) with the repetition LHA LL and repetition LHN LL conditions.

3.1 Experiment 1, Color-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. The gray instruction is not relevant for this experiment. All instructions end in a L-L% pattern (indicated by a ↘).

3.2 Experiment 1, Color-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).

3.3 Experiment 1, Non-contrastive sequence from EXPERIMENT A: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. The gray instruction is not relevant for this experiment. All instructions end in a L-L% pattern (indicated by a ↘).

3.4 Experiment 1, Non-contrastive sequence from EXPERIMENT B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).
<table>
<thead>
<tr>
<th>Experiment</th>
<th>Data Source</th>
<th>Analysis Details</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Experiment A (see Figure 3.1)</td>
<td>Target fixation analysis: Main effects (<em>Lexical Contrast</em> and <em>Prosody</em>) and interaction.</td>
<td>100</td>
</tr>
<tr>
<td>3.6</td>
<td>Color-contrast sequence from Experiment A (see Figure 3.11)</td>
<td>Target fixation analysis comparing the <em>felicitous</em> condition (model) with the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>103</td>
</tr>
<tr>
<td>3.7</td>
<td>Color-contrast sequence from Experiment A (see Figure 3.11)</td>
<td>Target fixation analysis comparing the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>104</td>
</tr>
<tr>
<td>3.8</td>
<td>Experiment B (see Figure 3.6)</td>
<td>Target fixation analysis: Main effects (<em>Lexical Contrast</em> and <em>Pitch Accent</em>) and interaction.</td>
<td>107</td>
</tr>
<tr>
<td>3.9</td>
<td>Color-contrast sequence from Experiment B (see Figure 3.16)</td>
<td>Target fixation analysis comparing the <em>felicitous</em> condition (model) with the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>110</td>
</tr>
<tr>
<td>3.10</td>
<td>Color-contrast sequence from Experiment B (see Figure 3.16)</td>
<td>Target fixation analysis comparing the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>111</td>
</tr>
<tr>
<td>3.11</td>
<td>Experiment A (see Figure 3.11)</td>
<td>Competitor fixation analysis: Main effects (<em>Lexical Contrast</em> and <em>Pitch Accent</em>) and interaction.</td>
<td>115</td>
</tr>
<tr>
<td>3.12</td>
<td>No-contrast sequence from Experiment A (see Figure 3.11)</td>
<td>Competitor fixation analysis comparing the <em>garden-path</em> condition (model) with the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>119</td>
</tr>
<tr>
<td>3.13</td>
<td>No-contrast sequence from Experiment A (see Figure 3.11)</td>
<td>Competitor fixation analysis comparing the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>120</td>
</tr>
<tr>
<td>3.14</td>
<td>Experiment B (see Figure 3.1)</td>
<td>Competitor fixation analysis: Main effects (<em>Lexical Contrast</em> and <em>Pitch Accent</em>) and interaction.</td>
<td>123</td>
</tr>
<tr>
<td>3.15</td>
<td>No-contrast sequence from Experiment B (see Figure 3.16)</td>
<td>Competitor fixation analysis comparing the <em>garden-path</em> condition (model) with the <em>neutral</em> and <em>infelicitous</em> conditions.</td>
<td>127</td>
</tr>
</tbody>
</table>
3.16 Experiment 1, No-contrast sequence from EXPERIMENT B (see Figure 3.16): Competitor fixation analysis comparing the neutral and the infelicitous conditions.

4.1 Experiment 2, Color-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘).

4.2 Experiment 2, Color-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).

4.3 Experiment 2, No-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘).

4.4 Experiment 2, No-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).
4.5

Experiment 2, data from E XPERIMENT A (see Figure 4.1): Target fixation
analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction. . . 141

4.6

Experiment 2, Color-contrast sequence from E XPERIMENT A (see Figure 4.1):
Target fixation analysis comparing the felicitous and the anticipatory conditions.143

4.7

Experiment 2, No-contrast sequence from E XPERIMENT A (see Figure 4.1):
Target fixation analysis comparing the garden-path and the anticipated gardenpath conditions. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 143

4.8

Experiment 2, data from E XPERIMENT B (see Figure 4.5): Target fixation
analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction. . . 148

4.9

Experiment 2, Color-contrast sequence from E XPERIMENT B (see Figure 4.7):
Target fixation analysis comparing the felicitous and the anticipatory conditions.149

4.10 Experiment 2, No-contrast sequence from E XPERIMENT B (see Figure 4.8):
Target fixation analysis comparing the garden-path and the anticipated gardenpath conditions. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 149
4.11 Experiment 2, data from E XPERIMENT A (see Figure 4.1): Color-competitor
fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 154
4.12 Experiment 2, No-contrast sequence from E XPERIMENT A (see Figure 4.11):
Competitor fixation analysis comparing the garden-path and the anticipated
garden-path conditions. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 155
4.13 Experiment 2, Color-contrast sequence from E XPERIMENT A (see Figure 4.11):
Competitor fixation analysis comparing the felicitous and the anticipatory
conditions. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 157
4.14 Experiment 2, data from E XPERIMENT B (see Figure 4.5): Color-competitor
fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 160

xxiv


4.15 Experiment 2, No-contrast sequence from EXPERIMENT B (see Figure 4.15):

Competitor fixation analysis comparing the \textit{garden-path} and the \textit{anticipated garden-path} conditions. .................................................. 161

4.16 Experiment 2, Color-contrast sequence from EXPERIMENT B (see Figure 4.15):

Competitor fixation analysis comparing the \textit{felicitous} and the \textit{anticipatory} conditions. ................................................................. 163

5.1 Experiment 3a, Early L+H* sequence from EXPERIMENTS A and B: CAPS indicate L+H* accents; \textit{Italics} on an adjective indicate H* accents; \textit{Italics} on a noun indicate !H* accents. The instruction in \textbf{bold face} is the target instruction. Instructions ending in a L-L\% pattern are shown with a \textbullet\text{↘}, whereas instructions ending in a L-H\% pattern are shown with a \textbullet\text{↗}) ............... 172

5.2 Experiment 3a, Late L+H* sequence from EXPERIMENTS A and B: CAPS indicate L+H* accents; \textit{Italics} on an adjective indicate H* accents; \textit{Italics} on a noun indicate !H* accents. The instruction in \textbf{bold face} is the target instruction. Instructions ending in a L-L\% pattern are shown with a \textbullet\text{↘}, whereas instructions ending in a L-H\% pattern are shown with a \textbullet\text{↗}) ............... 172

5.3 Experiment 3a, data from EXPERIMENTS A and B (see Figure 5.1): Target fixation analysis: Main effects (\textit{L+H* Position} and \textit{Edge Tone}) and interaction. 178

5.4 Experiment 3a, \textit{Early L+H*} sequence from EXPERIMENTS A and B (see Figure 5.3): Target fixation analysis comparing the \textit{early L+H*/finality} and \textit{early L+H*/continuation} conditions. ........................................ 180

5.5 Experiment 3a, \textit{Late L+H*} sequence from EXPERIMENTS A and B (see Figure 5.4): Target fixation analysis comparing the \textit{late L+H*/finality} and \textit{late L+H*/continuation} conditions. ........................................ 181
5.6 Experiment 3a, *Continuation* sequence from EXPERIMENT B (see Figure 5.5): Target fixation analysis comparing the *early L+H*/continuation and *late L+H*/continuation conditions......................................................... 184

5.7 Experiment 3a, *Finality* sequence from EXPERIMENT A (see Figure 5.6): Target fixation analysis comparing the *early L+H*/finality and *late L+H*/finality conditions................................................................. 184

5.8 Instruction 1, Data from EXPERIMENTS A and B (see Figure 5.10): Target fixation analysis comparing the conditions with a L+H* accent on the adjective (*LHA*) from EXPERIMENTS A and B................................................................. 194

5.9 Instruction 1, Data from EXPERIMENTS A and B (see Figure 5.10): Target fixation analysis comparing the conditions with a L+H* accent on the noun (*LHN*) from EXPERIMENTS A and B................................................................. 194

6.1 Experiment 3b, Early L+H* sequence from EXPERIMENT A: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘)................................................................. 203

6.2 Experiment 3b, Late L+H* sequence from EXPERIMENT A: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘)................................................................. 203

6.3 Experiment 3b, data from EXPERIMENT A (see Figure 6.1): Target fixation analysis: Main effects (*L+H* Position and *Target Instruction*) and interaction. 207
6.4 Experiment 3b, Target 3 sequence from EXPERIMENT A (see Figure 6.3): Target fixation analysis comparing the early L+H*, target 3 and late L+H*, target 3 conditions. ................................................................. 210
6.5 Experiment 3b, Target 2 sequence from EXPERIMENT A (see Figure 6.4): Target fixation analysis comparing the early L+H*, target 2 and late L+H*, target 2 conditions. ................................................................. 210
6.6 Experiment 3b, Early L+H* sequences from EXPERIMENT A (see Figure 6.5): Target fixation analysis comparing the early L+H*, target 3 and early cye, target 2 conditions. ................................................................. 212
6.7 Experiment 3b, Late L+H* sequences from EXPERIMENT A (see Figure 6.6): Target fixation analysis comparing the late L+H*, target 3 and late L+H*, target 2 conditions. ................................................................. 213

7.1 Experiment 4, Early L+H* sequence from EXPERIMENTS A and B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. Instructions ending in a L-L% pattern are shown with a \(\searrow\), whereas instructions ending in a L-H% pattern are shown with a \(\nearrow\) ........................................ 220
7.2 Experiment 4, No early L+H* sequence from EXPERIMENTS A and B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. Instructions ending in a L-L% pattern are shown with a \(\searrow\), whereas instructions ending in a L-H% pattern are shown with a \(\nearrow\) ........................................ 221
7.3 Experiment 4, data from EXPERIMENTS A and B (see Figure 7.1): Target fixation analysis: Main effects (Edge Tone and Accent Location) and interaction. ................................................................. 224
7.4 Experiment 4, Finality sequence from EXPERIMENT A (see Figure 7.3): Target fixation analysis comparing the early $L+H^*$ / finality and no early $L+H^*$ / finality conditions. .......................................................... 227

7.5 Experiment 4, Continuation sequence from EXPERIMENT B (see Figure 7.4): Target fixation analysis comparing the early $L+H^*$ / continuation and no early $L+H^*$ / continuation conditions. .......................................................... 227

7.6 Experiment 4, Early $L+H^*$ sequence from EXPERIMENTS A and B (see Figure 7.5): Target fixation analysis comparing the early $L+H^*$ / finality and early $L+H^*$ / continuation conditions. .......................................................... 229

7.7 Experiment 4, No early $L+H^*$ sequence from EXPERIMENTS A and B (see Figure 7.6): Target fixation analysis comparing the no early $L+H^*$ / finality and no early $L+H^*$ / continuation conditions. .......................................................... 230
Prosody is the rhythm and melody in spoken language (see Beckman & Venditti, in press, for a more formal definition of prosody). It comprises aspects of timing, pitch, rate, stress, loudness etc. in spoken language. There are measurable acoustic correlates of prosody in the sound signal. These include F0, duration, intensity, and spectral changes. Speakers and listeners use aspects of prosody, including rhythm, tone, and emphasis, to structure their conversations. Components of this structure include pitch accents (local prominences), which allow speakers to foreground individual lexical items and constituents, providing information about the items’ status in discourse structure. Similarly, edge tones (perceived junctures that contribute to the perception of phrasing) let speakers group or separate lexical items. Evidence has been accumulating to show that listeners use prosodic information immediately during language processing and that prosodic processing operates in parallel with processing at other levels of linguistic analysis. Prosodic structure has also been shown to trigger anticipatory processing at other levels of linguistic analysis, and to interfere with processing when prosodic and other linguistic structures conflict in the spoken language stream (see Arnold, 2008; Carreiras & Clifton, 1999; Dahan et al., 2002; Ito & Speer, 2008; Kjelgaard & Speer, 1999; Schafer, 1997; Venditti et al., 2001; Watson et al., 2006; Weber et al., 2006a,b, among others).

At the sentence level, prosodic phrase boundaries play a major role in disambiguating the syntactic structure of sentences that may be ambiguous (e.g. Lehiste, 1973; Schafner et al., 2000) – either temporarily (where the ambiguity is resolved by the end of the sentence, as in Whenever the lady cleans the room..., where continuing the sentence with is empty vs. it's
empty disambiguates the syntactic role of the NP the room) or globally (where the ambiguity is not resolved, as in Someone shot the servant of the actress who was on the balcony, where either the servant or the actress can be on the balcony).

At the discourse level, pitch accents affect the ease with which referents can be identified. For example, eyemovement monitoring data showed that when listeners in a Christmas tree decoration task heard Hang the red drum. Now hang the BLUE drum (where CAPS indicate a prominent L+H* pitch accent, characterized by a steep rise in pitch), they looked at the blue drum earlier than if instructions were pronounced with a neutral H* !H* pattern (as in Hang the red drum. Now hang the blue drum, cf. Ito & Speer, 2008). In particular, looks to the blue drum were planned while hearing BLUE, i.e. before lexical information from drum was heard. This suggests that listeners used prosodic information immediately and even predictively. In this case, the L+H* accent on BLUE evoked a contrast set of different-colored drums, and listeners expected a repetition of the previous noun drum. Pitch accents can also slow referent identification. For example, when listeners in the same experiment heard the infelicitous Hang the blue angel. Now hang the BLUE drum, looks were delayed as compared to a felicitous case with a repeated noun (Hang the red drum. Now hang the BLUE drum).

The effects in the example above depended on the type of one pitch accent – that on the adjective in the target utterance (the second of two utterances). However, pitch accents and edge tones are pervasive in spoken language. Little research has been conducted to investigate whether or not the interpretation of a local pitch accent (i.e. a pitch accent in the target instruction) is influenced by other relevant pitch accents and/or edge tones (but see Weber et al., 2006a). Compare, for example, the instruction sequence Click on the yellow bottle. Click on the PURPLE bottle to the sequence Click on the YELLOW bottle. Click on the PURPLE bottle. In the first case, the previous findings would lead us to predict that listeners would look to the purple bottle even before hearing bottle (cf. Ito & Speer, 2008, above). But what about the second case? Might the pitch accent on YELLOW "announce"
that the noun *bottle* will be repeated, leading to even earlier looks to the purple bottle?

Also consider the three-part instruction sequence *Click on the yellow scissors. Click on the YELLOW bottle. Click on the PURPLE bottle*. Here, the L+H* accent on the adjective YELLOW in the second utterance creates an infelicitous comparison with the preceding NP *yellow scissors*. However, the L+H* accents on YELLOW and PURPLE felicitously mark the difference in color term between the last two instructions (*yellow bottle ⇒ purple bottle*, where ⇒ is used as a shorthand for noun phrases mentioned in successive instructions).

In cases like these, can information from edge tones influence comprehension, grouping the utterances to determine which accents will be interpreted together? More specifically, could a rise in pitch ending the second instruction (a L-H% continuation rise, which should connect the last two instructions (cf. Pierrehumbert & Hirschberg, 1990)) appropriately group the lexical items differing in color-adjective? If so, listeners should look to the purple bottle earlier than when they hear the second instruction ending in a fall in pitch (L-L% sentence-final intonation, which indicates separation between the last two instructions).

This dissertation presents novel evidence that listeners consider successive pitch accent information as well as edge tone information during the resolution of reference. This means that listeners consider different components of prosody together in processing, suggesting that the language processing system operates on a complex prosodic representation. In particular, pitch accents and edge tones are not interpreted singly or in isolation, but instead their interpretation relies on their position in prosodic structure and its relation to the broader discourse context.

Compared to lexical and syntactic processing, prosodic processing has been neglected in the psycholinguistic literature. In addition, prosody has often been given a minor role in models on language processing, and the importance of prosody for language processing has only been emphasized recently (e.g. Frazier et al., 2006). This dissertation contributes to a growing body of literature documenting the contribution of prosody to discourse processing (cf. Dahan et al., 2002; Kjelgaard & Speer, 1999; Ito & Speer, 2008; Schafer, 1997;
Schafer et al., 2005; Slowiaczek, 1980; Speer et al., 1996; Weber et al., 2006a, see Speer & Blodgett, 2007 for an overview).

The experiments conducted for the dissertation used the visual world paradigm (VWP) (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Participants in such experiments listen to spoken instructions and respond to target phrases by selecting real-world objects or object images on a computer screen. In the current experiments, participants responded to instructions like *Click on the purple bottle* by clicking on the appropriate colored drawing on a computer screen. The results support the hypothesis that listeners build a global prosodic representation during discourse processing such that the interpretation of local prosodic cues is influenced by neighboring prosodic and lexical information. Listeners’ responses showed effects of relevant neighboring pitch accents or edge tones as they interpreted novel and repeated lexical items. That is, lexical information, pitch accents, and edge tones together affected the cognitive ease with which listeners could use definite noun phrases to select a referred-to item in a discourse context. For example, listeners’ use of the instruction to locate the purple bottle in a display of objects was a function of the lexical and prosodic information they heard in successive instructions to click on objects of different colors (e.g. a succession of instructions to click on yellow scissors, a yellow bottle, and a purple bottle).

In order to study the effect of intonation on the resolution of noun phrase referents in spoken language processing, theoretical elements from a broad range of areas, including semantics (treatment of definite noun phrases, see Section 1.1), phonetics and phonology (prosodic structure, see Section 1.2), and psycholinguistics (prosodic processing, see Sections 1.4 and 1.5), are needed.
1.1 Definite Noun Phrases

In the experiments presented below, listeners heard instructions to click on objects, each described by a definite noun phrase that specified its color and type (e.g. the purple bottle). Definite noun phrases like the purple bottle are used in discourse when there is a unique familiar discourse referent in the context that the noun phrase in question refers to (cf. Heim, 1982; Roberts, 2003). Roberts (2003) emphasizes that a definite noun phrase can refer to a previously unmentioned discourse referent as long as the referent is weakly familiar in the context. In particular, a previously unmentioned referent is weakly familiar if interlocutors can see the referred-to entity (see Roberts, 2003, for other circumstances in which a previously unmentioned referent is weakly familiar). Thus, when listeners in the experiments below hear Click on the purple bottle without previous mention of the purple bottle, the use of a definite noun phrase is felicitous because listeners can see a purple bottle on the computer screen and the discourse referents corresponding to the object images are weakly familiar.

In semantic theories of linguistic structure, the reference of a definite noun phrase only depends on there being a uniquely identifiable, familiar entity, regardless of the number of other available entities. This differs from psycholinguistic models of language processing, where the process of reference resolution is affected by the number of available competitors, including the number of referent competitors, especially cohort competitors (words with the same onset as the target). For example, when hearing Click on the purple..., not just the purple bottle, but other purple objects on the screen, are temporarily considered. Eye movement patterns are further affected by the visually available entities. In particular, eye movement patterns may differ if the location of a referred-to object is likely to be already known than if the location has to be determined through visual search. This difference between semantic theories and psycholinguistic models is related to goal differences. The semantic theories aim to specify the meaning(s) of a string of words and their constituent
structures. Psycholinguistic models aim to uncover how listeners interpret these linguistic objects in the real world (and the experimental context that represents it).

The definite noun phrases used in the instructions for the current experiments contained a head noun naming the type of the object and an adjectival modifier specifying the color that the object exhibits. Color adjectives used in this way are gradable. This means that they are interpreted on a scale. For example, a bottle can be very purple, somewhat purple, extremely purple, more purple than another bottle, not purple enough etc.

Gradable color adjectives have two subreadings (cf. Kennedy & McNally, 2010), a color-quantity reading and a color-quality reading. The quantity reading conveys how much of an object is the color in question, for example, how much of the bottle is purple. Proportional modifiers like half and completely can be used with this reading, as in The bottle is half purple or The bottle is completely purple. All items used in the visual displays of the experiments reported in this dissertation are completely the color denoted by the adjective (see Section 2.3 for the object types and colors used in the experiments).

The quality reading expresses closeness or distance to a prototype. For example, it states how much the hue, saturation, brightness etc. of the purple of a particular bottle approximates the color that people have in mind when they think of a bottle being purple. Modifiers that license this reading are perfectly, somewhat etc., as in The bottle is perfectly purple or The bottle is somewhat purple. Notice that these prototypes vary with the entity they describe. For example, the red of a prototypical red bing cherry is different from the red of prototypical red hair. In addition, there are color prototypes not associated with any object, as in The color red. An attempt was made in this dissertation to evoke general color prototypes not connected to any object types so that the same shade of a color could be used across object types. This was done by using only unconventional colors for the objects. For example, a bottle is not usually purple and an apple is not usually pink.
1.2 Intonational Structure and Meaning

In spoken discourse, definite noun phrases are situated in utterances that are structured by intonation. This dissertation assumes an autosegmental-metrical (AM) framework (cf. Beckman & Pierrehumbert, 1986; Ladd, 2008; Pierrehumbert, 1980; Beckman & Venditti, in press), where the intonation of an English utterance is made up of pitch accents (tonal targets associated with the metrically stressed syllables of words) and edge tones (pitch events associated with the edge of phrases). The approach is called autosegmental because, like vowels and consonants, tones are treated as autonomous segments; it is called metrical because hierarchically organized prosodic constituents form the meter (or rhythm) of an utterance. The framework allows decomposing the tune of an utterance into smaller linguistic units and assumes tonal targets (pitch accents and edge tones) as the primitives of intonation systems.

Consistent with previous work in AM frameworks, this dissertation employs a ToBI (Tones and Break Indices) labelling system. ToBI was originally created for Standard American English (Beckman & Ayers, 1997; Silverman et al., 1992) and has since been adapted to describe the intonational patterns of other varieties of English (e.g. Glasgow English (GlaToBI), Mayo et al., 2000, see also Jun, 2005) as well as various other languages (e.g. German, Greek, Dutch, Serbo-Croatian, Japanese, Korean, Mandarin, Cantonese, Italian etc., see Jun, 2005). ToBI annotates prosodic events as a linear sequence on parallel tiers, including a word, tone, and break index tier (see Figures 2.3 through 2.11). The word tier provides the beginnings and ends of individual words. All tonal events (pitch accents and edge tones) are marked on the tone tier. The tonal events most relevant for this dissertation will be described in subsections 1.2.1 through 1.2.3. The break index tier uses the numbers from 0 to 4 to indicate perceived break strength between consecutive words. Most relevant are the break indices 1 (indicating a phrase-medial word boundary), 3 (for an intermediate phrase boundary), and 4 for an intonational phrase boundary).
1.2.1 Phrases

Phrases structure discourse by highlighting individual lexical items and by separating or grouping together adjacent words. The building blocks of phrases are tonal targets, the tones proposed by phonological analysis. There are two types of tonal targets in English: pitch accents and edge tones. Both are stated in time (alignment) and fundamental frequency (scaling). Pitch accents are associated with stressed syllables, whereas edge tones are associated with the edge of phrases. There are two types of edge tones in English: Phrase accents control the pitch movements from the last pitch accent of the phrase to the end of the phrase (or to a boundary tone). Boundary tones are pitch movements associated with the pitch pattern at the edge of an intonational phrase.

Phrases are organized hierarchically, as illustrated in Example (1.1). An intermediate phrase (ip) ends in a phrase accent (L- or H-) and must contain at least one pitch accent (H*, !H* etc.). The last pitch accent in an intermediate phrase is the nuclear accent (in Example (1.1), the H* on click and the !H* on candle), which is assumed to be especially prominent (cf. Ayers, 1996). In English, the nuclear accent is often on the last content word of the intermediate phrase. Words that do not receive a pitch accent are unaccented. An unaccented word that occurs after the nuclear pitch accent and out of context would be expected to be accented is called deaccented. An intonational phrase (IPh or IP) contains one or more intermediate phrases and ends in a boundary tone (L% etc.). The right edge of an intonational phrase coincides with that of the last intermediate phrase it contains, such that an intonational phrase always ends in both a phrase accent and a boundary tone. Intonational phrase boundaries are perceived as stronger than intermediate phrase boundaries.

(1.1) \[ \text{Click on the } |_{ip} \text{ blue candle } |_{ip} \cdot \]
\[H^* \quad H^- \quad H^* \quad !H^* \quad L- \quad L% \]

The term tune refers to the pitch patterns of intonational phrases. For example, \(H^* L-L\%\) represents a declarative tune, whereas \(L^* H-H\%\) corresponds to a yes-no-question tune.
Pierrehumbert & Hirschberg (1990) proposed a compositional approach to tune meaning, where tunes that shared certain features also shared certain aspects of meaning. For example, the tunes H* L-H%, L+H* L-H% and L*+H L-H% share the low phrase accent followed by the high boundary tone and also share the sense that the speaker will continue with the utterance. A compositional approach to tune meaning captures these similarities. Pierrehumbert & Hirschberg (1990) proposed that tune meaning is built from the meanings of pitch accents, phrase accents, and boundary tones, all of which have scope over different domains of interpretation. In particular, pitch accents provide information about the discourse status of the words with which they are associated; phrase accents provide information about the degree of relatedness of adjacent intermediate phrases; boundary tones provide information about whether an intonational phrase should be interpreted with respect to the following phrase. The tunes used in the current experiments manipulate the position and type of pitch accents (and thus the discourse status of different lexical items) and the type of boundary tone (and thus whether or not an intonational phrase should be interpreted with respect to the following phrase).

1.2.2 Pitch Accents

As mentioned above, pitch accents are aligned with metrically stressed syllables. Stressed syllables carrying a pitch accent tend to be more prominent than stressed syllables not carrying a pitch accent; the latter, in turn, are more prominent than unstressed syllables. Several factors contribute to this prominence. Prominent syllables are generally long in duration (especially the vowel part), high in intensity, have considerable pitch movement, and are articulated clearly (especially the vowels’ formant structures) (cf. Ayers, 1996; Cooper et al., 1985; de Jong, 1995; Fear et al., 1995; Fry, 1955, 1958; Lehiste, 1970). Thus, stressed syllables that also carry a pitch accent are usually long in duration, high in intensity, hyperarticulated (especially the vowel) and contain a local pitch excursion. Stressed, but unaccented syllables are usually slightly shorter in duration, lower in intensity, phonetically
reduced (especially the vowel) and contain no local pitch excursion. Unstressed syllables are usually much shorter than stressed syllables and are characterized by vowel reduction. Notice that even though stress and pitch accents have similar effects on the phonetic realization of segments, stress is related to the prominence of different syllables within a word whereas pitch accents are related to the prominence of the word within a phrase. Stress is assigned by lexical-phonological rules, whereas pitch accents are assigned by syntactic and discourse rules.

Pitch accents, as well as edge tones (see Section 1.2.3), can be low (L) and high (H) in fundamental frequency (F0). High pitch and low pitch are local phenomena, not absolute notions. Therefore, whether a pitch accent is labelled as high or low does not depend on the absolute level of pitch, but on factors such as the speaker’s pitch range, the utterance’s pitch range, the neighboring pitch accents and edge tones etc. English has two simple pitch accents (H* and L*) and three bitonal pitch accents (L+H*, L*+H and H+!H*) (cf. Pierrehumbert, 1980). The * indicates which tone is associated with the vowel of the stressed syllable of the word receiving the pitch accent. For example, the L+H* accent has a leading low-tone and a high tone associated with the vowel of the stressed syllable. L*+H, on the other hand, has a low tone associated with the vowel of the stressed syllable and a trailing high-tone. Thus, these tones differ only in their alignment (cf. Beckman & Pierrehumbert, 1986, but see Ladd, 1983). The high tone in English can be downstepped (!H*). Downstep compresses the pitch range, so a downstepped high tone is perceptually lower in pitch than the preceding high tone. A bitonal pitch accent earlier in the same ip, for example, triggers downstep (cf. Beckman & Pierrehumbert, 1986, but see Ladd, 1983). The relevant pitch accents for this dissertation are the H*, !H* and L+H* accents.

**H* and !H* Accents**

H* accents are local F0 maxima. Pitch in a H* accent rises slightly onto the stressed syllable and peaks approximately in the middle of the word receiving the accent (see Figures 2.3,
A H* accent is often used to convey new information, such that the hearer should add the item receiving the H* accent and the proposition conveyed by the phrase containing the H* accent(s) to his or her mutual belief space (Pierrehumbert & Hirschberg, 1990). A H* accent is therefore often found on the first mention of an item in a discourse. In contrast, unaccented or deaccented items usually refer to entities that are already given in the discourse or are easily accessible. Hirschberg & Pierrehumbert (1986) propose that deaccenting can be used to ‘intonationally pronominalize’ repetitions of lexical items.

Like H* accents, !H* accents are local F0 maxima (again, see Figures 2.3, 2.6, and 2.7). However, they are perceptually lower in pitch that a preceding pitch maximum. Numerous functions of downstepped accents have been proposed in the literature. Pierrehumbert and Hirschberg speculated that downstepped accents mark discourse topic structure, are used for inferable discourse entities (Pierrehumbert & Hirschberg, 1990), and provide an alternative marker for given information (Hirschberg & Pierrehumbert, 1986). Ladd (2008) suggested that downstepped accents are used for given information when deaccentuation would change the focus structure of an utterance in an undesirable way. In spontaneous speech, downstepped contours occur most frequently in NPs that are hearer-inferable (as compared to hearer-given or hearer-new NPs, cf. Hirschberg et al., 2005).

In the experiments below, target instructions contained a H* !H* pattern on the NP , with the color adjective carrying the H* accent and the noun carrying the !H* accent (e.g. Click on the purple bottle). This pitch accent pattern presents a typical intonation contour for a declarative sentence, and the instruction to click on the purple bottle is rather neutral.

L+H* Accents

L+H* accents are characterized by a low pitch target before the accented syllable and a steep rise onto the accented syllable (see Figures 2.4, 2.5, 2.8, 2.9, 2.10, and 2.11). The peak of the L+H* accent is often realized late in the accented syllable or even slightly after the accented syllable. (L*+H accents differ from L+H* accents in that the pitch falls onto
the accented syllable in L\(^*\)+H accents, but rises or peaks in the accented syllable in L+H* accents.) Speakers use L+H accents to convey salience on some scale, where the accented item is linked to other items salient in the hearer’s mutual beliefs. The L+H* accent, in particular, conveys that the accented entity should be mutually believed as opposed to some alternative related entity (Pierrehumbert & Hirschberg, 1990).

The L+H* accent has been of particular interest to psycholinguists. While a H* accent is assumed to convey relatively inaccessible information, the L+H* accent is said to evoke a contrast set of similar items. A production study by Ito & Speer (2006) found that L+H* accents occurred more commonly in contexts where lexical items were contrasted with one another compared to when they were not. Furthermore, prenominal adjectival modifiers were produced with a L+H* accent when contrasted 45% of the time compared to nouns, which were produced with a L+H* accent when contrasted only 20% of the time. Ito & Speer (2008) also found that a L+H* accent, but not a H* accent, on a prenominal adjectival modifier evoked a contrast set (of the same item with similar attributes) and caused listeners to assume that the noun from a previous instruction was repeated, as in red drum ⇒ BLUE drum.

In the experiments below, target instructions may be pronounced with a L+H*-accented adjective and a deaccented noun (e.g. Click on the PURPLE bottle), or with an unaccented adjective and a L+H*-accented noun (e.g. Click on the purple BOTTLE). A L+H* accent on a lexical items highlights a difference with similar lexical items. For example, if the L+H* accent is on the adjective PURPLE, a difference in color is highlighted and a contrast-set of bottles of different colors is evoked. If the L+H* accent in on the noun BOTTLE, a difference in object type is highlighted and a contrast-set of different purple objects is evoked.

**H* vs. L+H* **

The phonological distinction between H* and L+H* has been debated (cf. Ladd & Morton, 1997; Ladd & Schepman, 2003). Some researchers do not consider H* and L+H* to be
categorically distinct, instead positing a continuum. Ladd & Morton (1997), for example, used an emphasis-rating task and stimuli with systematically modified F0 peak ranges and found that emphasis ratings increased gradually as F0 peak range increased, suggesting continuous perception. However, a forced-choice task between “everyday occurrence” and “unusual experience” showed a sharp increase in “unusual experience” choices in the mid F0 range, suggesting a categorical boundary between H* and L+H* accents. Despite this, a same/different discrimination task did not show the pattern typical for categorical perception. Ladd & Morton (1997) suggested that pitch range changes were categorically interpreted, but not categorically perceived.

Watson et al. (2006) investigated whether H* and L+H* are associated with different meanings and found that the domain of interpretation for H* and L+H* overlap. Their eyetracking experiment revealed that L+H* was restricted to a ”contrastive” interpretation whereas a H* was interpreted either as referring to a ”contrastive” or ”discourse-new” referent. Thus, L+H* was interpreted more restrictively than H*.

This dissertation assumes, but does not argue, that there is a categorical distinction between H* and L+H* accents. The H* and L+H* accents in the stimuli produced for this dissertation are clearly acoustically distinct. If the distinction between the two accents is continuous, the productions used here would be near either end of the continuum and should not be ambiguous.

1.2.3 Edge Tones

As mentioned above, there are two types of edge tones in English: Phrase accents and boundary tones. This dissertation focuses on intonational phrases, which end in both a phrase accent and a boundary tone and are characterized by phrase-final lengthening. The phrase accent of the intonational phrase controls the pitch from the last pitch accent to the boundary tone, and the boundary tone controls the pitch at the very edge of the phrase. Phrase accents are either high or low and can be downstepped (H-, L- or !H-) and boundary
tones are either high or low (H% or L%). This gives us the following possible tone patterns:
L-L%, L-H%, H-L%, H-H%, !H-L%, !H-H%. The patterns of interest here are L-L% and L-H%. The L-L% pattern is characterized by a drop in pitch at the end of the intonational phrase. The pitch falls for the low phrase accent and then often falls even further for the low boundary tone. The L-H% pattern is characterized by a low-high pattern at the end of the intonational phrase. The pitch first falls for the low phrase accent and then rises again for the high boundary tone (see Figures 2.3 through 2.11 for examples of L-L% and L-H% patterns).

Notice that these patterns differ only in boundary tone, which plays an important role in discourse segmentation. A high boundary tone (H%) relates an utterance to the following utterance, whereas a low boundary tone (L%) separates an utterance from the one following it (cf. Pierrehumbert & Hirschberg, 1990, and above). The L-L% pattern is often called sentence-final intonation and indicates that the speaker is done speaking or has finished a discourse segment. The L-H% pattern is often called continuation rise and indicates that the speaker is not done speaking and that the current utterance should be interpreted with respect to the one following it.

Instructions in the experiments below either ended in L-L% sentence-final intonation or a L-H% continuation rise. Instructions ending in a L-H% pattern should be connected more closely to the following instruction than those ending in a L-L% pattern.

1.3 Focus Structure

The focus of an utterance is the part of the utterance that answers the immediate question under discussion (QUD, see Roberts, 1998). Focus can be and is often prosodically signaled in English. The focus of an utterance can be narrow or broad. In a narrow focus utterance, only one item in the utterance answers the question under discussion, as in Click on the PURPLE bottle in response to the overt question What bottle should I click on? or as
a correction of the statement Click on the yellow bottle. In both cases, PURPLE is the focus of the utterance because it answers the QUD. Thus, the narrow focus is signaled by the use of a pitch accent on the narrowly focused word. When more than one item provides the answer to the QUD, focus is considered to be broad, as in Click on the purple bottle in response to the question What should I click on?

Semantic accounts of focus often concern the mapping between pitch accent location and focus structure. Selkirk (1995) used F-marking and F-projection rules to link pitch accent location to focus structure. She proposed that pitch-accented words are F-marked. If the accented word is the head of the phrase or an internal argument of the head of the phrase, the whole phrase can be F-marked. The focus of an utterance is an F-marked constituent that is not dominated by another F-marked constituent. Target instructions in the current experiments are pronounced either with a H* !H* pattern on the NP (Click on the purple bottle), or a L+H* accent either on the adjective or the noun (Click on the PURPLE bottle or Click on the purple BOTTLE). Based on Selkirk’s (1995) approach, an instruction like Click on the purple bottle is compatible with the purple bottle or Click on the purple bottle as the focus of the utterance. Click on the PURPLE bottle is only compatible with a narrow focus (on PURPLE) interpretation. It presupposes that sentence content other than (PURPLE) is given in the discourse. Click on the purple BOTTLE is also compatible with a narrow focus (in this case, on BOTTLE) interpretation. This interpretation presupposes that content other than (BOTTLE) has a given status. Focus may also project in this case, such that Click on the purple BOTTLE is compatible with the purple bottle or Click on the purple bottle as the focus of the utterance.

Schwarzschild (1999) also made use of an F-marking feature. He presented a constraint-based approach (Prince & Smolensky, 1993) to pitch accent location and focus structure. The main constraints were GIVENness and AVOIDF. GIVENness says that constituents that are not F-marked (i.e. that are not pitch-accented) are given (see Schwarzschild, 1999, for a definition of given). AVOIDF says to not F-mark (i.e. to use pitch accents sparingly).
According to Schwarzschild (1999), neither the adjective nor the noun of Click on the purple bottle are given (since both are accented). In Click on the PURPLE bottle and Click on the purple BOTTLE, on the other hand, the unaccented/deaccented word is given (bottle and purple, respectively).

Féry & Samek-Lodovici (2006) provided another constraint-based approach. They proposed discourse-structure and prosodic constraints, with discourse-structure constraints ranked higher than prosodic constraints. The discourse-structure constraints are STRESS-FOCUS (SF) and DESTRESS-GIVEN (DG). SF assigns the highest prosodic prominence to a focused phrase, and is split up into two related constraints, SF\textsubscript{contrast}, which requires contrastive focus to receive the highest prosodic prominence, and SF\textsubscript{new}, which requires new information focus to receive the highest prosodic prominence. SF\textsubscript{contrast} is ranked higher than SF\textsubscript{new}, such that contrastive information receives higher prominence than new information. DG states that given phrases are prosodically nonprominent. The lower-ranked prosodic constraints (STRESSXP, HP, and HI) place pitch accents in the rightmost position in a phonological and intonational phrase (where phrase structure is determined from syntactic structure, e.g. Selkirk, 1984).

Féry & Samek-Lodovici’s (2006) approach assumes that an instruction like Click on the purple bottle consists of two phonological phrases (Click on and the purple bottle). In an out-of-the-blue instruction Click and bottle carry prominence, with bottle being more prominent than Click. The same accent pattern emerges for instructions where a difference in object type is highlighted, as in Click on the purple BOTTLE. Click and BOTTLE are prominent, and BOTTLE is more prominent than Click. Instructions highlighting a difference in color, e.g. Click on the PURPLE bottle, receive prominence on Click and PURPLE, with PURPLE being more prominent than Click. All three approaches make reasonable assumptions about the relationship between pitch accent location and focus structure. However, the approaches neglect pitch accent type.
1.4 Effects of Pitch Accent in Psycholinguistic Studies

Psycholinguistic studies have investigated the effects of pitch accents on discourse processing. Unlike semantic approaches to focus structure, which are concerned with accent location, psycholinguistic studies have focused on pitch accent type. Sedivy et al. (1999), Weber et al. (2006a), and Ito & Speer (2008) studied how the pitch accent pattern found on definite noun phrases modified by an adjective, as in the purple bottle, affected the ease with which listeners could resolve reference. These studies revealed that prosody can be processed in parallel with segmental information, that prosody is processed immediately and can be used predictively, that discourse properties are updated incrementally, and that pitch accent type affects the ease with which listeners can resolve reference. Evidence from these studies has implications for the structure of discourse processing systems.

Sedivy et al. (1999) used eyetracking to study adjective interpretation. In their Experiment 1B participants followed instructions to touch real objects in a visual display. Each target display contained a target object (e.g. a yellow comb), an object sharing the object type with the target object (e.g. a pink comb), and object sharing the color, shape, or material with the target object (e.g. a yellow bowl), and an unrelated distractor object (e.g. a metal knife). During each trial, the experimenter read aloud two instructions from a script. Instructions told participants to touch objects in the display, e.g. Touch the pink comb. Now touch the yellow comb. The referent of the second instruction either differed only in attribute (color, shape, or material) from that of the first instruction (contrastive referent, e.g. pink comb ⇒ yellow comb) or differed in both attribute and object type from that of the first instruction (non-contrastive referent, e.g. pink comb ⇒ yellow bowl). In addition, there were two target prosodies for the second instruction. The experimenter either pronounced the adjective with a “contrastive” L+H* accent (YELLOW comb/bowl) or the noun with a H* accent

\(^1\)See Dahan et al. (2002) and Arnold (2008) for studies using cohort competitors (lexical items with identical onsets, e.g. candle vs. candy) to study the effect of noun phrase accentuation of reference resolution.
Sedivy et al. (1999) measured eye movement latencies from noun onset and found a main effect of referent. Participants looked at the target object earlier in the contrastive referent conditions compared to the non-contrastive referent conditions. There was neither an effect of prosody nor a referent / prosody interaction. Sedivy et al. (1999) also listed the percentage of trials that included a look to the competitor. Listeners were numerically most likely to look at the competitor object (e.g. the yellow comb) if the target instruction named the non-contrastive referent and contained a L+H* accent on the adjective (Touch the pink comb. Now touch the YELLOW bowl). Again, however, there was only a main effect of referent, but neither a main effect of prosody nor a referent / prosody interaction.

Sedivy et al. (1999) argue that the lack of a prosodic effect may be a ceiling effect due to a strong effect of referent, even for adjectives without a L+H* accent. The effect of referent may be strong for two reasons. Not all filler items included an adjectival modifier, which may have heightened participants’ sensitivity to adjective presence, and all experimental first instructions included an adjectival modifier, which may have drawn more attention to the contrastive referent in the second instruction. In addition, the prosody of the target instructions may not have been produced consistently enough to reveal an effect of prosody.

Weber et al. (2006a) used eyetracking to investigate how accentuation across utterances influenced referent resolution. This German-language study was conducted in response to Experiment 1B in Sedivy et al. (1999) and used the same task and visual display complexity as Sedivy et al. (1999). In two experiments listeners heard pre-recorded auditory instructions to click on objects on a computer screen, for example, Klicke die lila Schere an. Klicke jetzt die rote Schere an (“Click on the purple scissors. Now click on the red scissors”). Similar to Sedivy et al. (1999), the display contained two objects sharing their object type (e.g. purple scissors and red scissors), an object sharing the color with one of these objects (e.g. a red vase), and an unrelated object (e.g. a clock).
Referents in both experiments were either contrastive (purple scissors ⇒ red scissors) or non-contrastive (purple scissors ⇒ red vase). In Experiment 1, the first instruction was pronounced with a neutral H* accent on the noun (Klicke die lila Schere an). The target instruction was pronounced either with a L+H* accent on the adjective (Klicke jetzt die ROTE Schere/Vase an) or on the noun (Klicke jetzt die rote SCHERE/VASE an).

In contrast to Sedivy et al. (1999), Weber et al. (2006a) found both a main effect of referent and of prosody as well as a referent / prosody interaction. The main effect of referent suggested a general bias for contrastive referents over non-contrastive ones. The main effect of prosody and the interaction revealed that, crucially, looks to the contrastive referent increased significantly faster when the adjective was produced with a L+H* accent than when the noun was produced with a L+H* accent. The data suggested that the presence of an adjectival modifier in the second instruction biased participants to look at the contrastive referent and that the presence of a L+H* pitch accent on the adjective compared to on the noun increased this bias. In particular, the presence of an adjectival modifier may have biased participants to look at the contrastive referent because only the contrastive referent needed the modifier for unique identification. That is, the red vase can be uniquely identified if referred to as the vase (since it is the only vase in the display). But referring to the red scissors as the the scissors is vague because there are two scissors in the display that can only be differentiated by color.

Weber et al. (2006a) suggested that visual display preview and the choice of statistical analysis caused their results to be different from those of Sedivy et al. (1999). Participants in Sedivy et al.’s (1999) study had considerably more display preview before instruction onset than participants in Weber et al.’s (2006a) study. As a result, 80% of Sedivy et al.’s (1999) first fixations after noun onset were to the target object (as compared to 54% for Weber

\footnote{Weber et al.’s (2006a) filler instruction included instances where the same object was clicked on in both the first and second instructions. Thus, referring to the red scissors as scissors is vague even if one pair of scissors has already been mentioned and clicked on.}
et al., 2006a). Participants in Sedivy et al.’s (1999) study were most likely not engaged in visual search, but were mapping lexical input onto objects whose location was known.

Sedivy et al. (1999) analyzed first fixation latencies after noun onset, whereas Weber et al. (2006a) analyzed fixation proportion across 300 ms time windows after adjective onset. Weber et al. (2006a) found anticipatory eye movements in that participants started fixating referent objects prior to noun onset, suggesting that analyzing first fixation latencies after noun onset (rather than after adjective onset) could have obscured possible effects of prosody in Sedivy et al.’s (1999) study. In fact, Weber et al. (2006a) calculated first fixation latencies after noun onset for their data and found an effect of referent, but not of prosody.

Weber et al.’s (2006a) Experiment 2 differed from Experiment 1 only in the prosodic contour used for the first instruction of experimental trials. In this experiment, the first instruction was produced with a L+H* accent on the adjective, as in Klicke die LILA Schere an (“Click on the PURPLE scissors”). The results again revealed an interaction of referent and prosody. Unlike Experiment 1, there was no across-the-board bias for contrastive referents. Rather, the bias towards interpreting adjectives contrastively was neutralized when the target adjective was unaccented, as in Klicke die rote SCHERE/VASE an (“Click on the red SCISSORS/VASE”).

Weber et al. (2006a) proposed that the L+H* accent on the adjective in the first instruction highlighted the contrast set (purple scissors vs. red scissors) causing listeners to expect another accented adjective if the referent was the other member of the contrast set. When the second adjective was not accented, the contrastive object was no longer preferred. But when the second adjective was accented, participants started fixating the contrastive referent as soon as they encountered the adjective, as they did in Experiment 1.

---

3This result has to be taken with caution since the analyses included only trials where first fixations were to the target object. Weber et al.’s (2006a) results are thus based on 54% of trials compared to Sedivy et al.’s (1999) 80% of trials.
Weber et al.’s (2006a) experiments suggested that prosodic information is rapidly exploited for interpretation of referential expressions and that listeners considered the prosody of both instructions when determining a referent for the target instruction. This supports the assumption that semantic processing is incremental, that contrast sets are computed immediately upon encountering a L+H* accent, and that expectations set up in preceding utterances can affect the interpretation of prosodic information.

Ito & Speer (2008) used eyetracking to investigate anticipatory effects of accentuation, especially of L+H* accents. Participants in the experiments were engaged in a Christmas tree decoration task. They followed pre-recorded auditory instructions, such as First hang the green drum. Now hang the blue drum. The study used a rather complex visual display (containing over 40 objects), which engaged participants in visual search for objects. The objects were organized such that objects sharing the object type (e.g. all drums) occurred in the same cell in the display.

Target sequences in Experiment 1 presented either an adjective contrast (green drum ⇒ blue drum) or a noun contrast (blue onion ⇒ blue drum). Target instructions were either produced felicitously with a L+H* accent on the contrast item (green drum ⇒ BLUE drum or blue onion ⇒ blue DRUM) or infelicitously with a L+H* accent on the repeated item (green drum ⇒ blue DRUM or blue onion ⇒ BLUE drum). If the L+H* accent occurred on the adjective, the noun was deaccented, but if it occurred on the noun, the adjective received a H* accent. These prosodies were based on production data from naïve speakers engaged in the same Christmas tree decoration task (Ito & Speer, 2006).

The results showed both a main effect of contrast and accentuation. Overall, earlier eye movements to the target object occurred in the adjective contrast conditions compared to the noun contrast conditions and in the felicitous conditions compared to the infelicitous conditions.

Looks to the target object cell started rising before noun onset in both adjective contrast conditions, suggesting anticipatory eye movements to the target object. Listeners returned
to the same object cell in the display immediately after hearing a non-repeated adjective in the target instruction (as in *Hang the green drum. Now hang the BLUE/blue...*). In addition, a felicitously placed L+H* accent (*BLUE drum*) yielded earlier looks to the target object than an infelicitously placed L+H* accent (*blue DRUM*), suggesting either that a felicitously placed L+H* accent contributed to the expectation of a repeated noun or that an infelicitously placed L+H* accent decreased the expectation of a repeated noun.

Looks to the target object cell started rising after noun onset in the *noun contrast* conditions, suggesting no anticipatory eye movements to the target object. Rather, looks occurred in response to lexical and prosodic information from the noun. In addition, a felicitously placed L+H* accent (*blue DRUM*) showed an advantage compared to an infelicitously placed L+H* accent (*BLUE drum*) only 600 ms after noun onset. Ito & Speer (2008) suggested that the structure of the search task environment, where objects were sorted by type, contributed to the difference in effects between the *adjective contrast* and *noun contrast* conditions.

Ito & Speer’s (2008) Experiment 2 compared the *adjective contrast* conditions from Experiment 1 (*red drum ⇒ blue drum*) with *no contrast* conditions (*red angel ⇒ blue drum*). Unlike Experiment 1, target instructions were produced either with a L+H* accent on the adjective (*BLUE drum*) or a H* !H* sequence on the noun phrase (*blue drum*). These pronunciations, respectively, yielded a felicitous (*green drum ⇒ BLUE drum*) and neutral (*green drum ⇒ blue drum*) *adjective contrast* condition, and an infelicitous (*red angel ⇒ BLUE drum*) and felicitous (*red angel ⇒ blue drum*) *no contrast* condition.

As hinted at above, Experiment 1 did not determine whether the earlier looks in the felicitous *adjective contrast* condition compared to the infelicitous *adjective contrast* condition were a result of a processing advantage for a felicitously placed L+H* accent or of a processing delay for an infelicitously placed L+H* accent. Comparison of the *adjective contrast* conditions of Experiment 2 suggested that the advantage for the felicitous *adjective contrast* condition compared to the infelicitous *adjective contrast* condition in Experiment 1 did result
from a processing advantage for a felicitously placed L+H* accent rather than a processing delay for an infelicitously placed L+H* accent. Participants looked at the target object earlier in the felicitous adjective contrast compared to the neutral adjective contrast condition. The neutral condition can be considered a baseline for the visual search task since it involves neither a repetition nor a L+H* accent. Earlier looks to the target in the felicitous condition compared to the neutral condition then confirmed the anticipatory nature of the felicitously placed L+H* accent. Comparing the neutral adjective contrast condition from Experiment 2 with the infelicitous adjective contrast condition from Experiment 1 could determine if, in addition, an infelicitously placed L+H* accent would cause a processing delay.

Comparison of the no contrast conditions of Experiment 2 confirmed the anticipatory and facilitative effect of a felicitous L+H* on the adjective. Participants were garden-paused to look at the angel cell before noun onset in the infelicitous no contrast condition (red angel ⇒ BLUE drum), but not in the felicitous no contrast condition (red angel ⇒ blue drum). That is, listeners showed increased looks to the blue angel if they heard Hang the red angel. Now hang the BLUE..., but not if they heard Hang the red angel. Now hang the blue.... This suggests that a L+H* accent on a non-repeated adjective evoked a contrast set of objects of the same type and created the expectation of a repeated noun. When the color adjective carried a L+H* accent, participants planned and executed saccades to the just mentioned object type immediately upon hearing the adjective (even when it was not the target object). No such effect was found for adjectives with a neutral, less prominent H* accent.

The experiments suggested that prosodic information was used very early during realtime processing. In particular, anticipatory looks were based on intonational cues and occurred before confirming lexical information had come in. This supports the view that segmental and suprasegmental information is processed immediately and in parallel with other levels of language processing. Furthermore, pitch accents can initiate predictive lexical access.
1.5 Effects of Edge Tones in Psycholinguistic Studies

Psycholinguistic research on edge tones has largely focused on their effect in syntactic rather than discourse processing (but see Silverman, 1987), with recent work demonstrating that prosodic phrasal boundaries are interpreted with respect to the relative strength of preceding relevant boundaries in the sentence domain (Carlson et al., 2001; Clifton et al., 2002, but see Watson & Lee, 2009).

Carlson et al. (2001) and Clifton et al. (2002) investigated how prosodic boundary strengths and locations affected ambiguity resolution. They investigated multiple syntactic structures containing a global ambiguity. Listeners heard these sentences pronounced with varying boundary strengths at two relevant sentence locations, and were asked to provide a sentence interpretation. For example, the sentence *I met the daughter of the colonel who was on the balcony* is globally ambiguous. Either the daughter (high attachment) or the colonel (low attachment) could be on the balcony. The sentence was pronounced with a L- phrase accent on *colonel* and either no edge tone, a L- phrase accent, or a L-H% pattern on *daughter*. Listeners were asked whether the daughter or the colonel was on the balcony. The results showed that listeners’ interpretation of the relative clause (*who was on the balcony*) depended on the relative strengths of the two boundaries, not on the absolute strength of the local boundary at *colonel*. Listeners provided more high attachment interpretations if *daughter* was pronounced without edge tones than if it was pronounced with a L- phrase accent and if *daughter* was pronounced with a L- phrase accent than if it was pronounced with a L-H% pattern. If only the local boundary at *colonel* affected interpretations, the number of high attachment choices would be the same across conditions. Instead, the strength of the earlier boundary at *daughter* modulated interpretation choices. This suggests that ambiguous structures were interpreted based on the relative strengths of relevant boundaries (cf. Schafer et al., 2000) rather than the strength of one critical boundary.
Prosodic boundaries not only differ in strength, but also type. All instructions in the eyetracking experiments mentioned in Section 1.4 necessarily ended in an intonational phrase boundary, but could have differed in tonal type. Instead all instructions in Ito & Speer (2008) and Weber et al. (2006a) ended in L-L% sentence-final intonation4. Except Ito & Speer (2008, who presented 32 trials with each visual display), the studies introduced in Section 1.4 all presented two successive utterances with each visual display to study the effects of accentuation on reference resolution in discourse. However, most natural discourse involves many more utterances, such that there is usually an utterance preceding and following a given utterance. In such a situation edge tone pattern type may contribute to reference resolution.

A given L+H* accent may highlight a difference with a preceding utterance (as in Ito & Speer’s (2008) First hang the green drum. Now hang the BLUE drum) or a following utterance (as suggested for Weber et al.’s (2006a) Click on the PURPLE scissors. Now click on the red SCISSORS). If there were to be both a preceding and following utterance, the choice of edge tone pattern might contribute to listeners’ expectations about which utterance (the preceding or following) should contain the relevant difference. For example, a L-H% continuation rise concluding the instruction containing the L+H* accent may increase connection between this utterance and the one following it, leading participants to expect the relevant difference in the following utterance. But a L-L% sentence-final contour in the instruction containing the L+H* accent may increase separation between the utterance and the one following it, leading participants to look for the relevant difference in a preceding utterance. In this way, edge tone patterns may affect how easily listeners can resolve reference. As of the writing of this dissertation, no study has investigated whether and how

4Instructions in Dahan et al. (2002) and Arnold (2008) ended in either L-L% or H-H% and could contain L-L%, L-H%, H-L% and H-H% patterns instruction-medially. However, edge tone patterns in both studies were not systematically manipulated or compared, but represented random variation in how speakers pronounced the stimuli.
edge tone types contribute to the ease with which listeners can resolve reference.

1.6 Thesis Overview

This chapter presented relevant background information pertaining to the experiments conducted for this study. The semantics of definite noun phrases and the phonetics and phonology of prosody help understand the linguistic structure of the spoken materials created for this study. The semantics of discourse structure and focus help specify the content of the discourse representations that listeners in the experiments should build as they comprehend the language heard during the experimental task.

Chapter 2 presents the experimental design and statistical background common to all experimental chapters. Chapters 3 through 7 present data from 5 Experiments. Experiment 1 (Chapter 3) tests whether facilitation and garden-path effects found in previous studies (in particular Ito & Speer, 2008; Weber et al., 2006a) can be obtained using the current experimental design. Experiment 2 (Chapter 4) investigates whether listeners consider pitch accent information from successive utterances when resolving reference (cf. Weber et al., 2006a). Experiments 3a and 3b (Chapters 5 and 6) explore whether edge tone patterns structure lexical and pitch accent information of utterances and influence the ease with which listeners resolve reference. Finally, Experiment 4 (Chapter 7) investigates whether listeners consider preceding pitch accent and edge tone information during reference resolution. Chapter 8 presents a summary and discussion of the results obtained in the experiments.
Chapter 2

Methods

This dissertation presents data from five experiments (Experiments 1-4 described in Chapters 3 through 7). These experiments were run as components of two large EXPERIMENTS (EXPERIMENT A and EXPERIMENT B). Throughout this dissertation, the five experiments will be referred to using regular font. The two large EXPERIMENTS will be referred to using SMALL CAPS.

When studying the effects of intonation on language comprehension, it is important that participants are not aware of the goals of the experiment and are processing the intonation as they would in typical spoken language processing. Participants should not notice patterns in the tunes of experimental sequences, as they might develop hypotheses about the experiment or response strategies that may mask the effects the experiment is trying to measure. Combining five experiments into two omnibus EXPERIMENTS for data collection had the following advantages: Target stimuli from some experiments could function as filler items for one or more of the other experiments. The function of these filler items was to ensure full counterbalancing of the lexical contrast, pitch accent and edge tone conditions. The counterbalancing ensured that participants could not guess the nature of the experimental hypothesis from recurring patterns in the spoken instructions due to their frequency. In addition, distractor patterns were built into the the filler items to draw participants’ attention away from the manipulated variables. The counterbalancing will be described in detail in Sections 2.5.1 through 2.6. Collecting eyetracking data is also time-consuming and can be stressful for participants. Responsible research practices indicate that the minimum amount of experiment time be used that will allow sufficient power.
to fully address the research questions. By combining four experiments into two larger experiments, data collection was sped up significantly and a minimum number of participants were recruited.

Participants in all experiments listened to pre-recorded spoken instructions to click on pictured objects presented on a computer screen in the "visual world" paradigm. Instructions were carefully constructed to contain natural-sounding productions with specified prosodic structures. Stimuli were ToBI-annotated and phonetically measured to verify the presence of critical intonational characteristics. The independent variables manipulated in the instructions were pitch accent placement, edge tone type, and lexical contrast type (for example a color contrast, such as yellow bottle followed by purple bottle). The dependent measure was the relative proportion over time of eye movements to visual targets and competitors.

This chapter provides information about the experimental paradigm and methods used for this dissertation. Section 2.1 introduces the visual world paradigm, eyetracking, eye movements, and eyetrackers. Section 2.2 describes the participants in the study. Sections 2.3 through 2.7 describe the two large experiments, including visual materials, spoken materials, independent variables, list creation, and experimental procedure. Sections 2.8 through 2.10 describe the data collection, the graphic display of results, and the statistical analyses used.

2.1 The Methodological Framework

2.1.1 The Visual World Paradigm

This dissertation used a version of the visual world paradigm (e.g. Cooper, 1974; Tanenhaus et al., 2000) to examine spoken language processing in a context where objects mentioned are visually available to the listener. In the visual world paradigm, participants are presented with spoken language material and a visual scene, while their eye movements
are monitored. The timing and location of listeners’ looks to objects in the scene provide a reliable measure of language processing as it occurs, time-locked to information in the spoken signal. In the experiments described here, participants were presented with an array of images depicting simple objects on a computer screen. Eye movements were monitored while participants followed spoken instructions to locate particular objects and click on them.

The visual world paradigm was introduced by Cooper (1974), who monitored participants’ eye movements while they listened to short stories and looked at displays that included pictures of nine objects or animals that were either semantically related to words in the stories (experimental group) or irrelevant to the stories (control group). Cooper proposed that when presented with spoken language and a semantically related visual display, participants would spontaneously look at elements in the display most closely associated with the meaning of concurrently heard language. The results of the study showed that participants in the experimental group looked at pictures semantically related to the spoken input significantly more often than participants in the control group looked at (irrelevant) pictures in the same cells. The results confirmed the proposed relationship between auditory words and looks to pictured objects semantically related to those words.

2.1.2 Eyetracking

Eyetracking is the procedure of monitoring a person’s eye movements, usually during a particular task. It is used in a wide variety of disciplines and applications, including medical research, vehicle simulators, web usability, marketing and cognitive science. The methodology is based on the assumption that eye movements are probabilistically related to attentional mechanisms and that attention and changes in attention are closely time-locked to cognitive processes. Eye movement patterns can thus provide insight into such cognitive processes as decision making, mental imagery, memory, as well as language comprehension and production (cf. Findlay, 2004; Kowler, 1999; Liversedge & Findlay, 2001;
Richardson et al., 2007; Richardson & Spivey, 2004a,b; Tanenhaus & Trueswell, 2005, but see Anderson et al., 2004).

Tracking the movements of the eye to study linguistic processing started in the 1970s. In 1978 Rayner (1978) published an influential article detailing eye movement behavior during reading. In 1980 Just & Carpenter (1980) found that fixation duration during reading is influenced by the linguistic characteristics of the text: fixations were longer for longer and rarer words than for shorter and more frequent words. Two years later Frazier & Rayner (1982) first used eye movements to study syntactic processing. By the mid-1990s free-viewing head-mounted eyetrackers were first used to study language processing (Tanenhaus et al., 1995). Since then, research has shown that eye movements can be employed to chart the time course of language comprehension and production (see Henderson & Ferreira, 2004). This research has revealed that spoken word recognition, syntactic processing, spoken language production, natural unscripted conversation etc. are fast and incremental processes that are strongly influenced by the visual context (cf. Brown-Schmidt et al., 2005; Griffin & Bock, 2000; Richardson & Spivey, 2004b; Spivey et al., 2002).

Eyetracking was chosen as a methodology here because it has a number of advantages over traditional psycholinguistic paradigms. Most importantly, eyetracking provides a semi-continuous, implicit, and sensitive measure. The measure is semi-continuous because data is collected over time (in this case, every 17 ms). Since intonational patterns unfold over time, a continuous measure is most appropriate for studying when and how prosody affects spoken language processing. In particular, eyetracking offers a real-time measure of mental activity and provides information about the time course of eye movements in response to unfolding prosodic patterns (cf. Richardson et al., 2007).

Eyetracking is an implicit measure because it involves no requests for metalinguistic judgments. In particular, eye movements can be measured without interrupting processing. This allows for the use of tasks that are more natural and interactive than others typically used in psycholinguistic experiments and gives eyetracking increased ecological
validity. Richardson et al. (2007) further suggest that eye movement responses occur regardless of participants’ intent or experimenters’ instructions. Eyetracking may thus do a better job than more traditional psycholinguistic paradigms of preventing participants from developing explicit strategies that interfere with the effects experiments are trying to uncover.

The sensitivity of the eyetracking measure stems from the nature of eye movements. Eye responses are rapid, quickly corrected, metabolically cheap, and have a very low threshold for being triggered. In particular, eye movements are faster than other motor movements, such as hand responses (people usually fixate on an object before reaching for it, cf. Ballard et al., 1997). As a result, eyetracking may pick up effects that are not detected by traditional psycholinguistic paradigms.

Eyetracking also has disadvantages. Most notably, the language studied has to be picturable and the number of pictured objects and possible actions is limited (closed-set problem, e.g. Trueswell & Tanenhaus, 2005). The picturability problem affects experiment design and choice of stimuli rather than the validity of the paradigm. The closed-set problem, on the other hand, affects the validity of the paradigm. The limitation on pictured objects and possible actions may lead to task-specific strategies and result in language processing behavior that does not generalize to more natural linguistic situations. Limited displays and action choices in visual world experiments may allow bypassing more detailed linguistic processing. This concern has been addressed in the literature (cf. Allopenna et al., 1998; Tanenhaus et al., 2000). For example, studies of word recognition have shown that eyetracking is sensitive to effects deriving from the entire lexicon and does not lead to strategies bypassing lexical items not found in the visual display (cf. Dahan et al., 2001; Magnuson, 2001; Magnuson et al., 2007). This suggests that the entire lexicon determines which lexical candidates are activated, and that the visual display merely serves as a response selection set. Similar results have been found for syntactic processing (e.g. Snedeker & Trueswell, 2004). Head-mounted eyetrackers or table-mounted remote optical
eyetrackers are typically used in visual world studies. An ASL E5000 head-mounted eye-tracker with eye-head integration was used in the current experiments. It measures eye rotation and head position in order to determine where a person is looking. Eye rotation is measured by reflecting infrared light from the eye. Rotation is calculated using the relative positions of the center of the pupil and the corneal reflection (the first Purkinje image). This relation changes when the eye rotates, but not when the head moves while the eye is stable. Thus point-of-regard can be calculated from the reflection as long as the location of the head is known. The ASL E5000 measures head position through a head-mounted magnet, which allows participants to move freely during an experiment.

2.1.3 Eye Movements

Eye movements can provide information about what the mind is doing while it processes language. In order to interpret eye movement data, some basic information about the nature and timing of eye movements as well as eye movement control are needed. The eyes move as a natural and rapid response to even low-threshold signals. Eye movements consist of a succession of saccades and fixations, which allow the human eye to examine small sections of the visual world in rapid sequence (cf. Treue, 2001). (Other kinds of eye movements exist, but they are of little interest for investigators of cognitive processes, cf. Richardson et al., 2007.)

Fixations are periods during which the eye is at rest, and information about where on a computer screen a participant is looking can be measured. During a fixation, the eye monitors about 200° of visual field but receives highly acute visual information only from a small depression in the retina called the fovea (cf. Levi et al., 1985), which corresponds to about 2° of visual field. This means that while participants fixate a given object on the computer screen, they receive some information about objects in the vicinity of the one being fixated. In a search task, this information can help plan the location of the next fixation. It also has implications for the layout of visual displays since more information
is received from objects close to the one currently fixated (cf. Section 2.3). Fixations last about 200-300 ms, depending on the task and the visual display (cf. Richardson & Spivey, 2004a; Viviani, 1990).

Saccades are rapid, ballistic scanning movements that occur between fixations. The eye launches a saccade about three to four times every second (cf. Richardson & Spivey, 2004a). Saccades last from about 20 to 60 ms (Tanenhaus & Trueswell, 2005), during which the fovea reaches speeds of about 500° a second (cf. Matin, 1974; Tanenhaus & Trueswell, 2005; Thiele et al., 2002). Immediately before, during, and immediately after a saccade, the eye’s sensitivity drops, and it is unable to receive visual information (saccadic suppression, cf. Volkmann, 1962; Volkmann et al., 1968). Spatial and temporal distortions which occur around the time of saccades create the illusion of vision as stable and continuous (cf. Bischoff & Kramer, 1968; Cai et al., 1997; Matin & Pearce, 1965; Morrone et al., 2005; Ross et al., 1997). No eyetracking data can be collected during a saccade since the eye is in motion. However, since eye movements are ballistic and cannot be redirected or interrupted once launched, the decision to move the eye to a certain location has been made prior to the onset of the saccade. Therefore when eyemovements are the dependent measure in spoken language experiments, it is assumed that saccades are planned and executed on the basis of speech available before the saccade begins. For data collection purposes, saccades are thus considered as part of the immediately following fixation (cf. Section 2.8).

Programming (that is, planning and launching) a saccade during a visual search task takes about 200 to 300 ms (Matin et al., 1993; Tanenhaus & Trueswell, 2005). This means that eye movement responses in the current experiments occur at least 200 ms after the auditory information that triggered the eye movements. This delay needs to be taken into account when interpreting eye movement responses. For example, effects from a L+H* accent are measurable at the very earliest 200 ms after the beginning of the vowel carrying the L+H* accent. Similarly, a visible rise in eye movements to a certain screen location must be in response to auditory input that occurred at least 200 ms before the rise.
2.2 Participants

Forty-eight native speakers of American English from a large midwestern university participated in each of two large experiments – for a total of 96 participants. They received partial credit toward a Linguistics course requirement for their participation. All participants had normal or corrected-to-normal vision and normal hearing by self-report. Only data from participants who were successfully calibrated at the beginning and at a second calibration midway through a large experiment were included in the study.

2.3 Visual Materials

The display used in visual-world eyetracking studies provides a visual context or referential domain for what participants hear or say. During each trial of the two large experiments, participants saw a visual display containing a central fixation cross and twelve simple, monochrome object drawings. The object drawings were arranged such that they formed a square around the cross (see Figure 2.1).

2.3.1 Objects and Colors

Eight different colors (blue, brown, green, orange, pink, purple, red and yellow) and eight objects (apple, arrow, bottle, candle, glasses, pencil, scissors, whistle) were used in the experiments. Across each experiment, each object appeared in three different colors, and each color was used for three different objects (see Figure 2.2). Each color/object combination appeared equally often in each cell of the display. The pictures were drawn by the author and colored using the GNU Image Manipulation Program (GIMP) 2.4 (The GIMP Development Team, 2008).
Figure 2.1: Sample visual display used in the two large EXPERIMENTS.

Figure 2.2: Object type and color combinations used in the large EXPERIMENTS.
2.3.2 Object Characteristics

Object characteristics, such as color, size, and complexity can affect eye movements patterns. Treisman & Gelade (1980) found that tasks with target objects that seemed to pop out of a display have a flat search function (where response time is independent of the number of objects displayed). Williams (1966) investigated the effect of different object attributes on the search function using a display cluttered with geometric shapes of different color, shape and size. He found that target color prespecification produced faster searches with fixations restricted to competitor shapes of the target color. However, prespecifying target shape or size did not lead to faster searches or eye fixations restricted to competitors. This suggests that differences in color are detected more easily than differences in size or shape. Buswell (1935) further found that viewers will fixate regions with color variation and pay little attention to solid regions of color.

In the present experiments, objects were of roughly equal size, simply drawn, and monochrome to avoid having some objects stick out due to size, complexity, or – especially – choice of color. Colors were bright and chosen such that they were visually clearly distinct. An attempt was made to also have easily distinguishable object shapes.

2.3.3 Object Location

The location of objects relative to one another can also affect the search function. During any given fixation, people receive highly acute visual input from only 2° of the visual field. However, less detailed visual information is available from the periphery. Thus, the eye can pick up visual information from nearby objects that are not currently fixated. To ensure that such peripheral preview influences visual search as little as possible, neighboring objects in the visual displays were approximately equidistant and all objects had one of three similar distances to the central fixation point (a cross in the middle of the screen, see Figure 2.1). Object locations in the experiments don’t need to be exactly equidistant
from one another and from a center cross in order to be equivalently available to the search process because attention is repeatedly directed to the same spatial locations over and over (Henderson & Ferreira, 2004).

### 2.3.4 Choice of Competitor Objects

The choice of objects displayed in a scene – in particular the presence or absence of competitor objects – can also affect eye movement patterns during stimulus presentation. In a classic study, Tanenhaus et al. (1995) showed that syntactic ambiguity resolution is affected by the choice of visual context (i.e. the visual display) at the earliest stages of processing. Participants were presented with a display of four objects and listened to instructions to move one of the objects. Instructions of interest were temporarily ambiguous, containing an ambiguous prepositional phrase, as in *Put the apple on the towel in the box*. When the prepositional phrase *on the towel* is encountered, it can either be interpreted as the intended location of the apple (the syntactically preferred interpretation) or as a modifier of the apple (as in *the apple that’s on the towel*). Only once the second prepositional phrase *in the box* is heard can *on the towel* unambiguously be assigned a modifier interpretation. If the visual display contained only one apple (that was sitting on a towel), an empty towel, and two distractor objects, participants frequently looked at the empty towel (55% of trials) when they heard *on the towel*. But if the visual display contained two apples (one on a towel and one on a napkin), an empty towel, and one distractor object, participants looked at the empty towel reliably less often (under 20% of trials). Thus, the visual context had an immediate effect on eye movement patterns. That is, only when *on the towel* provided information about which apple should be moved was it immediately interpreted as a modifier of *apple* despite it being the syntactically dispreferred interpretation.

In the present study, the visual display always showed one or two objects of the same color and one or two objects of the same type as the target object to ensure that looks to objects were not differentially affected by the presence or absence of competitor objects.
For example, when the yellow bottle was mentioned, at least one other yellow object and at least one other bottle were in the display (see Figure 2.1).

### 2.3.5 Display Complexity

The complexity of the visual display affects visual search patterns. Tasks generally have a search function which increases linearly as the number of objects in the display increases, provided that all objects are equally visually salient (Treisman & Gelade, 1980). In most search tasks, several fixations occur before a target object is found (e.g. Binello et al., 1995; Findlay, 1995; Hooge & Erkelens, 1999; Motter & Belky, 1998; Zelinsky et al., 1997).

A number of visual world studies introduced in Chapter 1 used visual displays with only four objects. Ito & Speer (2008) have argued that such minimal visual displays do not involve search at all. Rather definite noun phrases recovered from the speech signal are mapped onto an object in the display with little need for search and few looks to unrelated items. They argued that a more complex display provides a more realistic test of spoken language processing phenomena, where objects are more likely to be identified as information from the spoken signal unfolds. Displays in the present study contained twelve objects to ensure sufficient visual search.

### 2.3.6 Display Preview

The amount of scene preview can also affect participants’ eye movement behavior during trials. Weber et al. (2006a) used a display with four objects and little preview. When participants were instructed to click on an object (e.g. *Click on the red scissors* in a display with pink scissors, red scissors, a red vase, and a metal knife), only 54% of first fixations after noun onset were to the target object. Using the same task with a longer preview (about 20s), Sedivy et al. (1999) report that 80% of first fixations after noun onset were to the target object. This suggests that the long preview may have eliminated visual search and led participants to fixate the only object where modification with the adjective *red* provided
useful identificatory information. With a more complex visual display, longer preview can still elicit visual search. Ito & Speer (2008), for example, used a complex visual display and only changed the display four times throughout an entire experimental session. In addition to a moderately complex visual display, this study used only 100ms of preview to ensure sufficient visual search.

2.4 Spoken Materials

During each trial of the experiments, participants listened to a set of three instructions, each specifying an object to click on the computer screen. All instructions were of the form Click on the ADJECTIVE NOUN. The adjective always described the color of the object, and the noun always determined the type of object. Thus, all instructions were of the form Click on the COLOR OBJECT. For example, participants may have heard the sequence of instructions in (2.1). Three aspects of an instruction were manipulated: The color and object type mentioned, the type and location of pitch accents, and the type of phrasal edge tones with which the instruction was produced.

(2.1) Click on the yellow bottle.
     Click on the purple bottle.
     Click on the purple pencil.

2.4.1 Color Adjectives and Nouns

Eight color adjectives and eight nouns were used in the instructions. They are given in Table 2.1. The color adjectives were chosen based on a number of constraints. Most importantly, an informal test ensured that all color names used in the study were those that native English speakers most commonly used to refer to the colors in question. The chosen color adjectives also have either one or two syllables and as many voiced sounds as possible. The limitation on the number of syllables ensured that all recorded utterances were
of comparable duration. The avoidance of voiceless sounds (whenever possible) ensured that a pitch track was available for as much of each recorded utterance as possible. Furthermore, color adjectives have as many different word onsets a possible. The avoidance of such cohort competitors (again, whenever possible) ensured that there was little or no initial confusion as to which word was being uttered. Due to visual constraints (see Section 2.3) and the most common color names, it was not possible to choose color adjectives with eight different word onsets. However, the words that share their onset (blue/brown and pink/purple) began with a voiced or voiceless stop consonant, which were relatively short and transitioned quickly into a subsequent continuant (liquid or vowel).

The object types were also chosen based on various constraints. Most importantly, the nouns had to be picturable and the pictures had to be clearly distinct and easily identifiable (see Sections 2.1.3 and 2.3). Again, an informal test ensured that the nouns chosen to refer to the objects were the ones that most speakers used when asked what the object was. All chosen nouns have two syllables. This ensured that all utterances were of similar duration. But more importantly, this ensured that nouns were long enough to carry up to four tonal targets (for example, a LHLH sequence consisting of a L+H* pitch accent, a L- phrase accent and a H% boundary tone). As with the color adjectives, voiceless sounds and shared word onsets were avoided when possible.

Each color adjective was combined with three nouns to describe the object pictures in Figure 2.2. Table 2.2 shows the adjective/noun combinations. The combinations written in gray only occurred in filler trials. Notice that the two nouns that in many dialects of English start with the vowel /æ/ (apple and arrow Wilson (1988)) were combined with different color adjectives to ensure that the sequence Click on the COLOR /æ/… was unambiguous.
Table 2.2: Adjective and noun combinations used in the sentence materials for all experiments. Greyed text indicates combinations that occurred in filler trials only.

2.4.2 Tunes

Participants heard the instructions Click on the COLOR OBJECT with the target tunes shown in (2.2) and (2.3). Instructions were either produced with a H* on the adjective and !H* on the noun (abbreviated as HHNP for H* !H* on the NP), a L+H* accent on the adjective followed by a deaccented noun (abbreviated as LHA for L+H* on the adjective), or an unaccented adjective followed by a L+H* accent on the noun (abbreviated as LHN for L+H* on the noun). The HHNP tune is typical of a broad focus or ‘out of the blue’ utterance, where all content words receive a H* or !H* pitch accent. The LHA and LHN tunes are often used when utterances convey contrastive focus. Here the color or object that receives the L+H* accent is contrasted with some other color or object, respectively. In addition, instructions either ended in L-L% sentence-final intonation (abbreviated as LL), indicating finality (Example (2.2)), or a L-H% continuation rise (abbreviated as LH), indicating that the sentence should be interpreted with respect to the following one (Example (2.3)).

(2.2) Words: Click on the purple bottle.
HHNP LL: H* H* !H* L-L%
LHA LL: H* L+H* L-L%
LHN LL: H* L-H* L-L%

41
(2.3) Words: Click on the purple bottle.
HHNP LH: H* H* !H* L-H%
LHA LH: H* L+H* L-H%
LHN LH: H* L+H* L-H%

2.4.3 Production

Two female phoneticians familiar with the ToBI annotation system produced the stimuli for this study during separate recording sessions. Speaker 1 produced the stimuli for EXPERIMENT A, and Speaker 2 produced the stimuli for EXPERIMENT B. Both speakers were recorded in a soundbooth. Stimuli with the same target tune were recorded together. For example, both speakers produced all utterances with the HHNP LL tune before moving on to all utterances with the LHA LL tune, etc. Both speakers produced several instances of all utterances during the first recording session. Select utterances were rerecorded in later sessions to ensure that the pitch range of utterances matched and that speakers produced the correct target intonation contour as determined by the author. Recordings were digitized at 22 kHz using Audacity 1.2.6 (The Audacity Team, 2006) on a Dell desktop personal computer.

2.4.4 ToBI Labeling

The author selected the instructions used as stimuli in the study by listening to all instructions and visually examining their pitch track. The prosodic contours of the target stimuli were then described using the ToBI system, a method for annotating the intonation of Standard American English and a number of other varieties of English (Beckman & Ayers, 1997; Silverman et al., 1992, see Section 1.2). The target contours, given in Section 2.4.2, contained the pitch accents H*, !H* and L+H* and the edge tone patterns L-L% and L-H% (see Sections 1.2.2 and 1.2.3). The author ToBI annotated all instructions used in the EXPERIMENTS prior to data collection. A steady rise to a local peak was labeled as a H* accent. If the local peak was perceptually lower than a previous local peak, it was labeled
as a !H* accent. A steep rise preceded by a low plateau was labeled as a L+H* accent. A fall to the bottom the speaker’s pitch range starting immediately after the last pitch accent of the phrase was labeled as a L-L% pattern, whereas a fall-rise pattern after the last pitch accent was labeled as a L-H% pattern. All initial ToBI annotations corresponded to the relevant target contours.

As all other aspects of linguistic structure, prosody can be ambiguous. The signal can be ambiguous in two primary ways: the presence or absence of a pitch accent or edge tone in a certain location and the type of pitch accent or edge tone if there clearly is one. Due to the possibility of ambiguity and to provide a more thorough analysis of the stimuli used in the experiments, all instructions were annotated by an independent native-speaker coder after data collection was completed. Cases of disagreement between the author and the independent labeler were given to a second independent native-speaker labeler (see below).

Studies on ToBI inter-labeler agreement showed high levels of agreement on the presence or absence of a pitch accent or edge tone and moderate to high levels of agreement on pitch accent or edge tone type (cf. Syrdal & McGory, 2000; Dilley et al., 2006, see Shattuck-Hufnagel et al., 2010 and Beckman, 1996 for situations in which pitch accent and edge tone presence or type may be uncertain or confusable). The level of agreement often depended on how strictly agreement criteria were defined. For example, agreement of pitch accent type in Dilley et al. (2006) was high (80%) because agreement was defined rather broadly. Two labelers agreed if they both indicated that a syllable had a high pitch accent (H*, L+H*, !H*, L+!H*, H+!H*), a low pitch accent (L*, L*+H, H+L*), or no pitch accent.

The following criteria, which pertain to ToBI labels on the adjectives (e.g. purple) and nouns (e.g. bottle) of instructions, are essential for this study and were used to judge agreement/disagreement of the author’s and first independent labeler’s annotations. Target contours differ in whether the adjective or noun or both are accented. Labelers should therefore agree about the presence or absence of a pitch accent on the adjective or noun.
Target contours also differ in pitch accent type. The adjective and/or noun carry either a H*/!H* accent or a L+H* accent. For clearly pitch accented lexical items, labelers should therefore agree on whether or not a high tone is preceded by a low leading tone. Agreement for high tones is therefore defined as two labelers agreeing that (a) the starred tone is high and (b) that there is (L+H* or L+!H*) or there is no (H*, !H*, H+!H*) low leading tone. Finally, target contours differed in edge tone pattern. The noun either ended in a L-L% or L-H% pattern. Agreement for edge tone patterns was therefore defined as two labelers agreeing that (a) the phrase accent is low (L-) and (b) that the boundary tone is low (L%) or that it is high (H%, !H%).

Syrdal & McGory (2000) observed that labelers most commonly confused the pitch accent types H* and L+H* and the edge tone types L-L% and L-H%. The distinction between H* and L+H* accents may turn on the steepness of the slope leading up to the high target, which is cannonically less steep for H* than L+H* accents; ambiguity occurs when the slope of an individual instance falls between that of the prototypical H* and L+H*. The distinction between L-L% and L-H% patterns may be compromised by production phenomena that may obscure pitch movement at the very end of the phrase, for example, final glottalization. The agreement criteria provided here are thus rather strict (for example, compared to those in Dilley et al., 2006).

Selection of utterances for the author’s initial ToBI annotations was by necessity not blind to condition. The utterances were randomized for the ToBI annotation, such that both labelers (the author and the first independent labeler) were unaware of the target contour for any given utterance and such that labelers did not annotate all utterances with the same target contour in a row. In addition, the independent labeler was unaware of the set of target contours used in the EXPERIMENTS. Utterances for which there was disagreement were labeled independently again by both labelers to ensure that no error occurred in the original labeling. In most cases both labelers confirmed their previous labeling. At the end of this process there were 11 cases of disagreement for the stimuli of EXPERIMENT A.
and 4 cases of disagreement for the stimuli of EXPERIMENT B. These utterances were given to the second independent labeler, who was unaware of the target contour for any given utterance. The second independent labeler’s annotations were compared with those of the other two labelers. For each lexical item, agreement was defined as two labelers agreeing, according to the criteria above. This process left two cases of disagreement, i.e. where all three labelers disagreed (see below).

Tables 2.3 and 2.4 show a summary of how well the final ToBI annotations for EXPERIMENTS A and B, respectively, matched the target contours. The same criteria as above were used, such that, for example, a starred high tone without a low leading tone (H*, !H*, H+!H*) was considered a match for either a target H* on the adjective or a target !H* on the noun. The tables show that both speakers produced the target contours consistently. Table 2.3 shows that most variation was found in how Speaker 1 produced the HHNP LL tune. One utterance did not contain a pitch accent on the adjective and one utterance contained a L* accent on the noun. In two cases, the labelers could not agree on the pitch accent of the noun (indicated by X? – in both cases, one labeler each proposed a !H*, L* and no pitch accent on the noun). Speaker 1 produced the LHA LL and LHN LL tunes very consistently. Only one LHN LL utterance was pronounced with a H* accent on the adjective. Table 2.4 shows that Speaker 2 produced all patterns very consistently. One HHNP LH utterance did not contain a pitch accent on the noun, and one LHN LL and LHN LH utterance each contained a H* accent on the adjective.

The final ToBI annotations crucially revealed no disagreements or deviations from the target contour involving H*/!H* vs. L+H* pitch accents and L-L% vs. L-H% edge tones. In addition, 90% of utterances from Speaker 1 and 97% of utterances from Speaker 2 were produced with the target contour. There is therefore no indication that the small deviations from the target contours should affect the experimental results\(^1\). Note that both Speaker 1

\(^1\)Dahan et al. (2002), for example, found an effect of pitch accent using stimuli with moderate variation. The target noun was either accented or unaccented. All target nouns in the unaccented condition were produced
and Speaker 2 always produced a H* accent on Click.

Figures 2.3 through 2.11 show sample pitch tracks and ToBI annotations for the prosodic patterns produced by the two speakers. The figures exemplify the acoustic differences among the six prosodic patterns (HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH). For easy comparison, productions with the same lexical items are shown for each speaker (Click on the purple pencil for Speaker 1 and Click on the purple bottle for Speaker 2). The choice of lexical item for each speaker was determined by the canonicity with a H+H* accent, but only 14 of 24 target nouns in the accented condition were produced with a L+H* accent (the remaining target nouns were produced with a H* accent).

<table>
<thead>
<tr>
<th>Tune</th>
<th>On Target</th>
<th>Adj off Target</th>
<th>Noun off Target</th>
<th>Edge off Target</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL</td>
<td>12</td>
<td>1 (Ø)</td>
<td>3 (L*/X*)</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHA LL</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHN LL</td>
<td>15</td>
<td>1 (H*)</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2.3: Summary of ToBI labels for Speaker 1 (EXPERIMENT A): Number of utterances for each tune, where the pitch accents and edge tones on both adjective and noun were on target (On Target), where the pitch accent on the adjective deviated from the target (Adj off Target), where the pitch accent on the noun deviated from the target (Noun off Target), and where the edge tones ending the noun deviated from the target (Edge off Target). Labels in parentheses provide the identity of deviating pitch accents or edge tones.

<table>
<thead>
<tr>
<th>Tune</th>
<th>On Target</th>
<th>Adj off Target</th>
<th>Noun off Target</th>
<th>Edge off Target</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>HHNP LH</td>
<td>15</td>
<td>0</td>
<td>1 (Ø)</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHA LL</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHA LH</td>
<td>16</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHN LL</td>
<td>15</td>
<td>1 (H*)</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>LHN LH</td>
<td>15</td>
<td>1 (H*)</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 2.4: Summary of ToBI labels for Speaker 2 (EXPERIMENT B): Number of utterances for each tune, where the pitch accents and edge tones on both adjective and noun were on target (On Target), where the pitch accent on the adjective deviated from the target (Adj off Target), where the pitch accent on the noun deviated from the target (Noun off Target), and where the edge tones ending the noun deviated from the target (Edge off Target). Labels in parentheses provide the identity of deviating pitch accents or edge tones.
Figure 2.3: ToBI annotation of *Click on the* purple pencil \( \downarrow \) (HHNP LL, where \( \downarrow \) indicates a L-L\% pattern), produced by Speaker 1.

Each figure shows time on the x-axis and pitch (in Hz) on the y-axis. The waveform of the utterance is shown above the pitch track. A word and ToBI annotation are shown below the pitch track. The pitch track exemplifies the shape of the target prosodic patterns described above. The waveform shows amplitude (i.e. relative loudness) throughout the utterance.

Figures 2.3, 2.6 and 2.7 show example utterances of the HHNP tunes. The H* accents on *click* and on the adjective are rather level in pitch, followed by a !H* accent on the noun that is visibly lower in pitch than the previous two accents. The remaining figures show utterances with the LHA (2.4, 2.8 and 2.9) and LHN (2.5, 2.10 and 2.11) tunes. In all examples, the peak of the L+H* accent (either on the adjective or the noun) is clearly higher in pitch than the H* accent on *click*. Notice that Speaker 1’s pitch range is larger than Speaker 2’s. Notice also that both speakers fall to the bottom of their pitch range immediately after the last pitch accent such that the pitch rises again for the H% boundary tone, but does not fall any further for a L% boundary tone.
Figure 2.4: ToBI annotation of Click on the PURPLE pencil ↘ (LHA LL, where ↘ indicates a L-L% pattern), produced by Speaker 1.

Figure 2.5: ToBI annotation of Click on the purple PENCIL ↘ (LHN LL, where ↘ indicates a L-L% pattern), produced by Speaker 1.
Figure 2.6: ToBI annotation of *Click on the purple pencil ↘* (HHNP LL, where ↘ indicates a L-L% pattern), produced by Speaker 2.

Figure 2.7: ToBI annotation of *Click on the purple pencil ↗* (HHNP LH, where ↗ indicates a L-H% pattern), produced by Speaker 2.
Figure 2.8: ToBI annotation of *Click on the PURPLE pencil ↘* (LHA LL, where ↘ indicates a L-L% pattern), produced by Speaker 2.

Figure 2.9: ToBI annotation of *Click on the PURPLE pencil ↗* (LHA LH, where ↗ indicates a L-H% pattern), produced by Speaker 2.
Figure 2.10: ToBI annotation of *Click on the purple PENCIL* ↘ (LHN LL, where ↘ indicates a L-L% pattern), produced by Speaker 2.

Figure 2.11: ToBI annotation of *Click on the purple PENCIL* ↗ (LHN LH, where ↗ indicates a L-H% pattern), produced by Speaker 2.
2.4.5 Phonetic Measurements

Duration and $F_0$ measurements for all target utterances were obtained to support the ToBI annotations. In general, we should expect the number of tonal targets to be positively correlated with duration, so that words produced with two tonal targets, e.g. a L+H* accent, generally have longer durations than words produced with one tonal target, such as H* or !H*. In addition, unaccented or deaccented words tend to be shorter than accented words. Alternations of low and high targets, as in L+H* or L-H%, tend to increase duration more than successive low targets, as in L-L%, or successive high targets. Therefore, words produced with a L-H% pattern should have longer durations than those with a L-L% pattern. The $F_0$ peak of a H* accent is generally lower than that of a L+H*. In addition, successive H* accents should show a slight lowering in $F_0$, whereas a !H* accent should be notably lower than the preceding high target. Finally, the pitch at the end of the utterance should be higher for utterances ending in a L-H% pattern compared to a L-L% pattern (Beckman & Ayers, 1997).

Duration Measurements

Duration measurements for use in phonetic analyses were obtained for three parts of each utterance: Click on the, the adjective, and the noun. The measurement used here to support the ToBI annotations is average segment duration (obtained for each part by dividing total duration by number of segments, cf. Nakatani et al., 1981). Average segment duration was chosen over word duration because adjectives and nouns varied in number of segments. The word purple, for example, tends to be longer in duration than the word red. Average segment duration was chosen over duration of the stressed vowel alone in order to capture any duration effects of pitch accents or edge tones they may have occurred on unstressed vowels or other segmental material, such as that between the final stress.
and utterance end, a region associated with edge tone information. Using average segment duration minimizes effects of word length while still capturing duration differences across the whole word. Duration values were extracted automatically in Praat (Boersma & Weenink, 2009). Numbers of segments (phonemes) were determined through the MRC Psycholinguistic Database (Wilson, 1988). Utterances which differed from the target tune in the presence or absence of a pitch accent (four and two utterances from Speakers 1 and 2, respectively) were excluded.

Tables 2.5 and 2.6 show average segment durations of *Click on the*, the adjective, and the noun for utterances of each tune. Table 2.5 shows duration measurements for Speaker 1. The table shows that, as expected, average segment duration of a word carrying a L+H* accent was longer in duration than of a word carrying a H* or !H* accent or an unaccented word. Four one-way ANOVAs were performed with tune as the independent variable and average segment duration of *Click on the*, the adjectives, the nouns, and the complete utterances, respectively, as the dependent variable. The ”adjective” one-way ANOVA showed a reliable effect of tune on average segment duration for adjectives ($F(2, 41) = 7.92, p < 0.01$). Pairwise comparisons showed that average segment durations for adjectives with a L+H* were reliably longer than those for adjectives with either a H* accent ($t(27) = 3.24, p < 0.01$) or no accent ($t(29) = -3.55, p < 0.01$). However, the “noun” one-way ANOVA showed no effect of tune on average segment duration for nouns. Average segment duration for nouns with a L+H* accent was thus only numerically longer than for nouns with a H* or no accent. The absence of differences in noun segment duration across tunes may be related to the presence of edge tones. Adjectives carried either no, one or two tonal targets – all of which were pitch accents. Here, two tonal targets doubled the number of targets compared to one tonal target. Nouns carried either two, three or four tonal targets - two of which were always edge tones. In comparison, four tonal targets compared to three tonal targets only increased the number of targets by one third. In addition, edge tones may contribute more to a longer duration than pitch accents. It is thus possible that the two
edge tones present in all tunes equalized noun segment duration such that the durational differences among noun segments were only numeric. The “Click on the” and “complete utterance” one-way ANOVAs showed no reliable effect of tune on average segment duration of Click on the and of the complete utterance².

Table 2.6 shows average segment duration measurements for Speaker 2. The table shows that the average segment duration for words with a L+H* accent was greater than for words with a H* or !H* accent or for unaccented words when durations are collapsed over edge tones (i.e. when averaging over both HHNP, both LHA, or both LHN patterns). The table contains one unexpected result: Average segment duration for the H* accent on the adjective of the HHNP LL tune is numerically longer than that for the L+H* accent on the adjective of the LHA LL tune. In addition, the table shows that, as expected, average segment duration for nouns with a L-H% edge tone pattern was longer than that for nouns with a L-L% pattern.

²All four ANOVAs were repeated including the utterances that differed from the target tune in the presence or absence of a pitch accent. Including these utterances did not affect the results of the “adjective”, “noun”, and “complete utterance” ANOVAs. The ”Click on the” ANOVA did show a reliable effect of tune on Click on the duration (F(2, 45) = 3.3, p < 0.05). In particular, durations for the LHN LL and HHNP LL tunes differed reliably (t(26) = -2.34, p < 0.05).
<table>
<thead>
<tr>
<th>Tune</th>
<th>Target Tones and Average Segment Duration</th>
<th>Click on the</th>
<th>Adjective</th>
<th>Noun</th>
<th>Complete Utterance</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL</td>
<td>H* (on click)</td>
<td>H*</td>
<td>!H* L-L%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>53.35 ms</td>
<td>99.16 ms</td>
<td>148.17 ms</td>
<td>(SD = 3.41)</td>
<td>(SD = 19.59)</td>
</tr>
<tr>
<td>HHNP LH</td>
<td>H* (on click)</td>
<td>H*</td>
<td>!H* L-H%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>54.06 ms</td>
<td>90.2 ms</td>
<td>157.08 ms</td>
<td>(SD = 3.41)</td>
<td>(SD = 13.03)</td>
</tr>
<tr>
<td>LHA LL</td>
<td>H* (on click)</td>
<td>L+H*</td>
<td>L-L%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>54.71 ms</td>
<td>98.71 ms</td>
<td>141.44 ms</td>
<td>(SD = 1.76)</td>
<td>(SD = 8.66)</td>
</tr>
<tr>
<td>LHA LH</td>
<td>H* (on click)</td>
<td>L+H*</td>
<td>L-H%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>54.05 ms</td>
<td>102.29 ms</td>
<td>147.06 ms</td>
<td>(SD = 2.68)</td>
<td>(SD = 11.93)</td>
</tr>
<tr>
<td>LHN LL</td>
<td>H* (on click)</td>
<td>—</td>
<td>L+H* L-L%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>54.24 ms</td>
<td>90.61 ms</td>
<td>154 ms</td>
<td>(SD = 4.03)</td>
<td>(SD = 15.19)</td>
</tr>
<tr>
<td>LHN LH</td>
<td>H* (on click)</td>
<td>—</td>
<td>L+H* L-H%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>52.34 ms</td>
<td>84.58 ms</td>
<td>172.34 ms</td>
<td>(SD = 2.63)</td>
<td>(SD = 10.69)</td>
</tr>
</tbody>
</table>

Table 2.6: Average segment durations of target words for the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2
Two-way ANOVAs were performed with pitch accent pattern (HHNP, LHA, and LHN) and edge tone pattern (LL and LH) as independent variables and average segment duration of Click on it, the adjectives, the nouns, and the complete utterances, respectively, as the dependent variable. Unexpectedly, all ANOVAs showed neither a reliable main effect of pitch accent pattern or edge tone pattern nor a pitch accent pattern / edge tone pattern interaction. Two effects were marginally reliable. The "adjective" two-way ANOVA revealed a marginal effect of pitch accent pattern on average segment duration of the adjective ($F(1, 90) = 3.3, p = 0.07$) and the "noun" two-way ANOVA revealed a marginal effect of edge tone pattern on average segment duration of the noun ($F(1, 90) = 3.89, p = 0.05$). This suggests marginally longer average adjective segment durations if the adjective carries a L+H* accent compared to a H* accent or no accent as well as marginally longer average noun segment durations if the noun ends in a L-H% pattern compared to a L-L% pattern. Thus, Speaker 2 showed numeric differences in average segment duration that were in the same direction as differences that were significant for Speaker 1.

Overall, the duration measurements for both speakers were consistent with the ToBI annotations. In all but one case, words produced with a L+H* accent were longer than words produced with a H* accent or no accent and words produced with a L-H% pattern were longer than those produced with a L-L% pattern) – either numerically or reliably.

**Pitch Measurements**

Next, $F_0$ values were measured at all sentence locations that had a pitch accent in the target tune. Measurements for high targets with a clear peak were taken at the peak. If the pitch was flat or falling (e.g. for many !H* preceding a L- phrase accent), measurements were

---

$^3$Again, all four ANOVAs were repeated including the utterances that differed from the target tune in the presence or absence of a pitch accent. The "adjective", "noun" and "complete utterance" ANOVAs that included these utterances revealed the same results, including the marginally reliable effects. The "Click on the" ANOVA showed a marginally reliable pitch accent / edge tone pattern interaction ($F(1, 90) = 3.35, p = 0.07$)
Table 2.7: Average pitch of target pitch accents for the HHNP LL, LHA LL and LHN LL tunes produced by Speaker 1.

<table>
<thead>
<tr>
<th>Tune</th>
<th>Target Accent and Average $F_0$</th>
<th>Adjective</th>
<th>Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL</td>
<td>$H^*$</td>
<td>$H^*$</td>
<td>$!H^*$</td>
</tr>
<tr>
<td></td>
<td>342.72 Hz (SD = 12.92)</td>
<td>316.72 Hz (SD = 16.1)</td>
<td>240.31 Hz (SD = 14.78)</td>
</tr>
<tr>
<td>LHA LL</td>
<td>$H^*$</td>
<td>$L+H^*$</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>316.01 Hz (SD = 11.6)</td>
<td>427.68 Hz (SD = 21.47)</td>
<td>—</td>
</tr>
<tr>
<td>LHN LL</td>
<td>$H^*$</td>
<td>—</td>
<td>$L+H^*$</td>
</tr>
<tr>
<td></td>
<td>316.55 Hz (SD = 10.1)</td>
<td>—</td>
<td>422.48 Hz (SD = 38.38)</td>
</tr>
</tbody>
</table>

| Table 2.7 and 2.8 show the $F_0$ values for the pitch accents of the HHNP, LHA, and LHN tunes. Table 2.7 shows the measurements for Speaker 1. As expected, the table reveals that the peak of $L+H^*$ accents is much higher in pitch than the peak of $H^*$ or $!H^*$ accents.

A one-way ANOVA showed a reliable effect of tune on the $F_0$ value of the pitch accent on click ($F(2, 45) = 25.86, p < 0.0001$). Pairwise comparisons (Tukey contrasts) revealed that the $F_0$ on click in the HHNP LL tune is reliably higher than $F_0$ on click in both the LHA LL tune ($t(30) = –6.29, p < 0.0001$) and the LHN LL tune ($t(30) = –6.17, p < 0.0001$). It is possible that the $F_0$ values in the LHA LL and LHN LL tunes are lower because the speaker is anticipating the low target of the umcoming $L+H^*$ accents. Next, Welsh two-sample t-tests revealed that, as expected, the peaks of $L+H^*$ accents were reliably higher.
Table 2.8: Average pitch of target pitch accents for the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2.

than those of the H* or !H* accents for both the adjective (comparing the HHNP LL and LHA LL tunes, t(29) = −15.81, p < 0.0001) and the noun (comparing the HHNP LL and the LHN LL tunes, t(29) = −17.15, p < 0.0001; cf. Ito & Speer, 2008, but see Arnold, 2008).

Table 2.8 shows average $F_0$ values for Speaker 2. The table again shows that L+H* accents peak at a much higher pitch level than H* or !H* accents. Notice, in addition, that Speaker 2 has a lower pitch range than Speaker 1, such that both L+H* and H*/!H* accents for Speaker 2 peak lower than those for Speaker 1. Three two-way ANOVAs were performed (one for $F_0$ of click, one for $F_0$ of the adjective, and one for $F_0$ of the noun) with pitch accent pattern and edge tone pattern as the independent variables and $F_0$ as the dependent variable.

The "click" two-way ANOVA showed a main effect of pitch accent pattern ($F(2, 90) =$
### Table 2.9: Pairwise comparisons from the "adjective" two-way ANOVA for Speaker 2.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHA LL and HHNP LL</td>
<td>88.74</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHA LH and HHNP LL</td>
<td>105.4</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHA LL and HHNP LH</td>
<td>84.57</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHA LH and HHNP LH</td>
<td>101.22</td>
<td>0.0001</td>
</tr>
<tr>
<td>HHNP LL and HHNP LH</td>
<td>-4.17</td>
<td>n.s.</td>
</tr>
<tr>
<td>LHA LL and LHA LH</td>
<td>-16.65</td>
<td>0.05</td>
</tr>
</tbody>
</table>

10.97, p < 0.0001) and edge tone pattern (F(1, 90) = 21.23, p < 0.0001) as well as a pitch accent / edge tone pattern interaction (F(2, 90) = 8.93, p < 0.001) on the F0 value of the H* on click. Pairwise comparisons\(^4\) revealed that F0 on click in the LHA LL and LHN LL tunes is reliably lower than that of the HHNP LL tune (t(30) = -13.4, p < 0.001 and t(30) = -16.84, p < 0.0001, respectively). Furthermore, F0 on click in the LHA LL and LHN LL tunes is reliably lower than in the respective LH tunes (t(30) = -15.3, p < 0.0001 comparing LHA LL with LHA LH and t(30) = -11.28, p < 0.01 comparing LHN LL with LHN LH). Again, in some cases the speaker might have been anticipating the low target of the upcoming L+H* accents.

The "adjective" two-way ANOVA showed a reliable effect of pitch accent pattern (F(1, 60) = 534.51, p < 0.0001) and edge tone pattern (F(1, 60) = 6.43.08, p < 0.05) on the F0 value of the pitch accents on the adjective. Pairwise comparisons are shown in Table 2.9. Importantly, H* accents were reliably lower than L+H* accents in all comparisons. In addition, the F0 peaks of L+H* accents in the LHA LH tune were reliably lower than those in the LHA LL tune, but the F0 of H* accents in the HHNP LL and HHNP LH tunes were not reliably different.

The "noun" two-way ANOVA showed a reliable effect of pitch accent pattern (F(1, 59)

\(^4\)Reliable pairwise comparisons involving a comparison across both pitch accent pattern and edge tone pattern (e.g. comparing HHNP LH with LHA LL) are not reported since it is difficult to meaningfully interpret these effects.
Comparison | $t$ | $p <$
---|---|---
LHN LL and HHNP LL | 112.78 | 0.0001
LHN LH and HHNP LL | 133.17 | 0.0001
LHN LL and HHNP LH | 90.09 | 0.0001
LHN LH and HHNP LH | 110.47 | 0.0001
HHNP LL and HHNP LH | -22.69 | 0.001
LHN LL and LHN LH | -20.38 | 0.001

Table 2.10: Pairwise comparisons from the “noun” two-way ANOVA for Speaker 2.

= 1021.15, $p < 0.0001$) and edge tone pattern ($F(1, 59) = 37.7$, $p < 0.0001$) on the $F_0$ value of the pitch accents on the noun. Pairwise comparisons are shown in Table 2.10. The table reveals that, as predicted, !H* accents were reliably lower than L+H* accents. Furthermore, the $F_0$ of the !H* accents in the HHNP LL tune was reliably lower than that of the HHNP LH tune, and the $F_0$ peak of the L+H* accent in the LHN LL tune was reliably lower than that in the LHN LH tune.

The $F_0$ measurements revealed that both speakers reliably produced L+H* accents with pitch peaks that were higher than those for H* or !H* accents. Notice also that $F_0$ values for H* accents on the adjective and !H* accents on the noun are (numerically) lower than those for the H* accents on click, whereas the $F_0$ values for L+H* accents on either the adjective or the noun are (numerically) higher than those for the H* accents on click. Thus, $F_0$ declines throughout the utterance when speakers produce a series of (!)H* accents. L+H* accents on the other hand represent a clear $F_0$ excursion.

Table 2.11 shows the average $F_0$ values for the last recorded pitch value of each tune produced by each speaker. The table reveals the expected results. Speaker 1 produced all utterances with a L-L% pattern, and the average pitch values for the end of the utterance are comparable (within 8 Hz of each other) for all tunes (HHNP LL, LHA LL, and LHN LL). Speaker 2 produced utterances either with a L-L% or L-H% pattern. Average pitch values for the end of the utterance of LL tunes (HHNP LL, LHA LL, and LHN LL) are again comparable (within 6 Hz of each other). Average pitch values for the end of the utterance
Table 2.11: Average pitch of target pitch edge tones (last pitch value in the utterance) for the HHNP LL, LHA LL, and LHN LL tunes produced by Speaker 1 and the HHNP LL, HHNP LH, LHA LL, LHA LH, LHN LL, and LHN LH tunes produced by Speaker 2.

<table>
<thead>
<tr>
<th>Tune</th>
<th>Target Edge Tones and Average $F_0$</th>
<th>Speaker 1</th>
<th>Speaker 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL</td>
<td>L-L% 204.99 Hz (SD = 7.09)</td>
<td>L-L% 146 Hz (SD = 4.8)</td>
<td></td>
</tr>
<tr>
<td>HHNP LH</td>
<td>L-H% —</td>
<td>L-H% 202.14 Hz (SD = 8.43)</td>
<td></td>
</tr>
<tr>
<td>LHA LL</td>
<td>L-L% 202.04 Hz (SD = 9.14)</td>
<td>L-L% 147.4 Hz (SD = 4.44)</td>
<td></td>
</tr>
<tr>
<td>LHA LH</td>
<td>L-H% —</td>
<td>L-H% 239.26 Hz (SD = 11.44)</td>
<td></td>
</tr>
<tr>
<td>LHN LL</td>
<td>L-L% 197.17 Hz 14.01 (SD = 4.21)</td>
<td>L-L% 142.07 Hz (SD = 4.21)</td>
<td></td>
</tr>
<tr>
<td>LHN LH</td>
<td>L-H% —</td>
<td>L-H% 220.1 Hz (SD = 15.56)</td>
<td></td>
</tr>
</tbody>
</table>

of LH tunes (HHNP LH, LHA LH, and LHN LH) vary somewhat more. Importantly, average pitch values for the LL tunes are consistently lower than those for the LH tunes. Notice also that Speaker 1 ended LL utterances at a higher pitch than Speaker 2. This is consistent with the $F_0$ values for pitch accents, which suggested a higher pitch range for Speaker 1 compared to Speaker 2.

A one-way ANOVA with pitch accent pattern as the independent variable and $F_0$ as the dependent variable was performed on the pitch values from Speaker 1. The results revealed no reliable differences between the pitch values of the HHNP LL, LHA LL, and LHN LL tunes. This confirms that Speaker 1 produced all utterances (which all ended in a L-L% pattern) with comparable final pitch.
A two-way ANOVA with pitch accent pattern and edge tone pattern as the independent variables and $F_0$ as the dependent variable was performed on the pitch values from Speaker 2. The results revealed a reliable effect of pitch accent pattern ($F(2, 90) = 33.89, p < 0.0001$), edge tone pattern ($F(1, 90) = 1518.72, p < 0.0001$), and a pitch accent / edge tone pattern interaction ($F(2, 90) = 28.92, p < 0.0001$). Pairwise comparisons are shown in Table 2.12. Importantly, L-L% utterances ended in reliably lower pitch than L-H% utterances. In addition, there were no reliable differences between the three LL tunes. Unexpectedly, LH tunes ended in reliably different pitch heights, with utterance-final pitch being highest for the LHA LH tune, followed by the LHN LH tune, followed by the HHNP LH tune. The HHNP LH and LHN LH tunes possibly ended in lower pitch than the LHA LH tune because of tonal crowding. When three or four alternating low and high tones, respectively, had to be realized on two syllables, the final H% may not have been fully realized. In this case, however, we would expect the LHN LH tune (and not the HHNP LH tune) to have the lowest utterance-final pitch (since the most alternating tones had to be realized on two syllables).

---

Table 2.12: Pairwise comparisons from the “adjective” two-way ANOVA for Speaker 2.

<table>
<thead>
<tr>
<th>Comparison</th>
<th>$t$</th>
<th>$p &lt;$</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP LL and HHNP LH</td>
<td>−16.77</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHA LL and LHA LH</td>
<td>−27.43</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHN LL and LHN LH</td>
<td>−23.3</td>
<td>0.0001</td>
</tr>
<tr>
<td>HHNP LL and LHA LL</td>
<td>0.42</td>
<td>n.s.</td>
</tr>
<tr>
<td>LHA LL and LHN LL</td>
<td>−1.59</td>
<td>n.s.</td>
</tr>
<tr>
<td>LHN LL and HHNP LL</td>
<td>−1.18</td>
<td>n.s.</td>
</tr>
<tr>
<td>HHNP LH and LHA LH</td>
<td>−11.08</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHA LH and LHN LH</td>
<td>5.72</td>
<td>0.0001</td>
</tr>
<tr>
<td>LHN LH and HHNP LH</td>
<td>5.36</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

5Reliable pairwise comparisons involving a comparison across both pitch accent pattern and edge tone pattern (e.g. comparing HHNP LH with LHA LL) are not reported since it is difficult to meaningfully interpret these effects.
Altogether, the analyses for both speakers revealed, as predicted, that L+H* accents peaked much higher than H* accents and that utterances ending in a L-H% pattern ended in reliably higher pitch than utterances ending in a L-L% pattern. This confirms that the L+H* vs. H*/!H* distinction as well as the L-L% vs. L-H% distinction of the ToBI labeling have measurable acoustic correlates.

2.5 General Design and Independent Variables

The independent variables in the experiments are the types of lexical contrast between successive instructions, the pitch accent patterns with which utterances are produced, and the edge tones that end utterances. Throughout the experiments, participants hear instruction triplets. Triplets differ in lexical contrast types and pitch accent patterns to study effects of felicitous and infelicitous contrast, for example, yellow bottle ⇒ PURPLE bottle compared to yellow bottle ⇒ PURPLE pencil or yellow bottle ⇒ purple BOTTLE. Triplets also differ in edge tone type to study effects of grouping, for example, yellow scissors ↘⇒ YELLOW bottle ↗⇒ PURPLE bottle ↘ compared to yellow scissors ↘⇒ YELLOW bottle ↘⇒ PURPLE bottle ↘ (where a ↗ represents a L-H% pattern that connects successive utterances and a ↘ represents a L-L% pattern that separates successive utterances). Finally, lexical contrast, pitch accent and edge tone patterns are counterbalanced to minimize that participants develop processing strategies or guess the manipulations of the experiment.

2.5.1 Lexical Contrast Types

The words lexical contrast are used here to describe whether successively mentioned color-adjectives and nouns are repeated or novel within a trial. Target trials in the experiments contained four kinds in lexical contrast. In a color contrast, successive color-adjectives were novel, whereas successive nouns were repeated (as in yellow bottle, purple bottle). In an object contrast, successive nouns were novel, whereas successive color-adjectives were
repeated (as in purple bottle, purple pencil). No contrast refers to cases where both the color-adjective and noun are novel (e.g. yellow bottle, purple pencil). Finally, repetition refers to cases where both the color-adjective and noun are repeated (as in purple bottle, purple bottle). Table 2.13 shows the lexical contrast types used in EXPERIMENTS A and B. In the Table, Lexical Contrast Type names refer to the lexical contrasts between Instructions 1 and 2 and Instructions 2 and 3, respectively. For example, the name ColorContrast/ObjectContrast in the first row of Table 2.13 refers to a color contrast between Instructions 1 and 2 (yellow bottle, purple bottle) and an object contrast between Instructions 2 and 3 (purple bottle, purple pencil).

Lexical contrast types were counterbalanced to minimize the predictability of the target object for Instruction 2, and to ensure that all objects of interest across the different lexical contrast and pitch accent conditions were equally often the target of Instruction 2. For a quarter of the lexical contrast types (25% of target trials) each, the target object of Instruction 2 (a) differed from the target object of Instruction 1 in both color and object type; (b) was identical to the target object of Instruction 1; (c) differed from the target object of Instruction 1 only in color; (d) differed from the target object of Instruction 1 only in object type. Depending on the display, there were between two and four objects that shared a feature (color or type) with the target object of Instruction 1 and between seven and nine objects that did not. Due to constraints on the overall size of the experiment, this kind of counterbalancing was not possible for Instruction 3. Instead, the target object of Instruction 3 either differed from the target object of Instruction 2 only in color or only in object type. Therefore, participants may have developed a general expectation that colors and/or object types repeat within a trial.

---

6The terms contrast and contrastive are used to refer to a number of different notions in the literature. Throughout this dissertation, an attempt was made to only use the four terms introduced above (i.e. lexical contrast, color contrast, object contrast, no contrast, and repetition) to refer to contrast as it is used here. Quotes or italics were used if the terms contrast and contrastive were used in the description of previous studies (e.g. if they appeared as part of condition names).
<table>
<thead>
<tr>
<th>Lexical Contrast Type</th>
<th>Instruction 1</th>
<th>Instruction 2</th>
<th>Instruction 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>ColorContrast/ObjectContrast</td>
<td>Click on the yellow bottle.</td>
<td>Click on the purple BOTTLE</td>
<td>Click on the purple PENCIL</td>
</tr>
<tr>
<td>NoContrast/ObjectContrast</td>
<td>Click on the yellow scissors.</td>
<td>Click on the purple BOTTLE</td>
<td>Click on the purple PENCIL</td>
</tr>
<tr>
<td>ObjectContrast/ColorContrast</td>
<td>Click on the purple PENCIL</td>
<td>Click on the purple BOTTLE</td>
<td>Click on the yellow bottle</td>
</tr>
<tr>
<td>Repetition/ColorContrast</td>
<td>Click on the purple bottle.</td>
<td>Click on the purple bottle.</td>
<td>Click on the yellow bottle</td>
</tr>
</tbody>
</table>
Table 2.14: Filler trial patterns for Experiments A and B.

<table>
<thead>
<tr>
<th>Pattern Description</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants clicked on an object located in one of the four corners on the computer screen three times in a row.</td>
<td>18</td>
</tr>
<tr>
<td>Participants click on three adjacent objects either beginning or ending with an object located in one of the four corners on the computer screen.</td>
<td>16</td>
</tr>
<tr>
<td>Participants clicked on an object not located in one of the four corners on the computer screen three times in a row.</td>
<td>12</td>
</tr>
<tr>
<td>Participants clicked on two objects, both located in one of the four corners on the computer screen, in an ABA pattern.</td>
<td>9</td>
</tr>
<tr>
<td>Participants clicked on two objects, with no or one object located in one of the four corners on the computer screen, in an ABA pattern.</td>
<td>6</td>
</tr>
<tr>
<td>No visual pattern.</td>
<td>6</td>
</tr>
<tr>
<td>Participants clicked on three objects located in three different corners on the computer screen.</td>
<td>5</td>
</tr>
</tbody>
</table>

an utterance. HHNP stands for H* !H* on the noun phrase, LHA for L+H* on the adjective, and LHN for L+H* on the noun (see Section 2.4.2 above).

The pitch accent conditions were created such that the prosody of Instruction 2 was not predictable from the prosody of Instruction 1. Although in Instruction 1 the H* !H* pattern occurred more often (50% of target trials) than the L+H* patterns (25% of target trials each), each of these patterns occurs equally often in Instruction 2. That is, each prosodic contour in Instruction 1 was equally likely to be followed by a HHNP, LHA, or LHN pattern in Instruction 2. Predictability in the prosody of Instruction 3 could not be avoided altogether. A HHNP pattern in Instruction 2 was always followed by another HHNP pattern. Both LHA and LHN patterns were equally likely to be followed by either a LHA or a LHN pattern. Instruction 3 was the target utterance in Experiments 3a, 3b, and 4. Importantly, Instruction 2 has a LHA or LHN pattern is all cases, such that the pitch accent pattern of Instruction 3 was not entirely predictable (as would be the case with an HHNP sequence).

Filler trials contained varying pitch accent sequences. More filler trial instructions had a HHNP pattern than LHA or LHN patterns since the HHNP pattern was less common.
### Table 2.15: Pitch Accent Sequence for ColorContrast/ObjectContrast and NoContrast/ObjectContrast Lexical Contrast Types used in EXPERIMENTS A and B. The pitch accent condition HHNP is indicated by *italics*, the L+H* accent of the LHA and LHN conditions is indicated by CAPS.

<table>
<thead>
<tr>
<th>Pitch Accent Sequence</th>
<th>Instruction 1</th>
<th>Instruction 2</th>
<th>Instruction 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP      HHNP      HHNP</td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
</tr>
<tr>
<td>HHNP      LHA       LHN</td>
<td><em>adj noun</em></td>
<td><em>ADJ noun</em></td>
<td><em>adj NOUN</em></td>
</tr>
<tr>
<td>HHNP      LHN       LHN</td>
<td><em>adj noun</em></td>
<td><em>adj NOUN</em></td>
<td><em>adj NOUN</em></td>
</tr>
<tr>
<td>LHA       HHNP      HHNP</td>
<td><em>ADJ noun</em></td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
</tr>
<tr>
<td>LHA       LHA       LHN</td>
<td><em>ADJ noun</em></td>
<td><em>ADJ noun</em></td>
<td><em>adj NOUN</em></td>
</tr>
<tr>
<td>LHA       LHN       LHN</td>
<td><em>ADJ noun</em></td>
<td><em>adj NOUN</em></td>
<td><em>adj NOUN</em></td>
</tr>
</tbody>
</table>

### Table 2.16: Pitch Accent Sequence for ObjectContrast/ColorContrast and Repetition/ColorContrast Lexical Contrast Types used in EXPERIMENTS A and B. The pitch accent condition HHNP is indicated by *italics*, the L+H* accent of the LHA and LHN conditions is indicated by CAPS.

<table>
<thead>
<tr>
<th>Pitch Accent Sequence</th>
<th>Instruction 1</th>
<th>Instruction 2</th>
<th>Instruction 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHNP      HHNP      HHNP</td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
</tr>
<tr>
<td>HHNP      LHA       LHA</td>
<td><em>adj noun</em></td>
<td><em>ADJ noun</em></td>
<td><em>ADJ noun</em></td>
</tr>
<tr>
<td>HHNP      LHN       LHA</td>
<td><em>adj noun</em></td>
<td><em>adj NOUN</em></td>
<td><em>ADJ noun</em></td>
</tr>
<tr>
<td>LHN       HHNP      HHNP</td>
<td><em>adj NOUN</em></td>
<td><em>adj noun</em></td>
<td><em>adj noun</em></td>
</tr>
<tr>
<td>LHN       LHA       LHA</td>
<td><em>adj NOUN</em></td>
<td><em>ADJ noun</em></td>
<td><em>ADJ noun</em></td>
</tr>
<tr>
<td>LHN       LHN       LHA</td>
<td><em>adj NOUN</em></td>
<td><em>adj NOUN</em></td>
<td><em>ADJ noun</em></td>
</tr>
</tbody>
</table>

67
than the LHA or LHN patterns in target trial instructions.

2.5.3 Edge Tone Patterns

Edge tone pattern was a between-experiment manipulation. All instructions in Experiment A ended in L-L% sentence-final intonation. In Experiment B, Instructions 1 and 3 ended in L-L% sentence-final intonation, but Instruction 2 ended in a L-H% continuation rise. Filler trial instructions in Experiment A also all ended in a L-L% pattern. In Experiment B, either all instructions in a filler trial ended in a L-L% pattern or Instruction 1 ended in a L-H% pattern while Instructions 2 and 3 ended in a L-L% pattern.

2.6 List Creation

Experimental lists contained 72 filler trial – target trial – target trial sequences for a total of 216 trials (72 filler trials and 144 target trials). Each lexical contrast type occurred every six trials throughout each list in the following order: Filler trial, ColorContrast/ObjectContrast trial, NoContrast/ObjectContrast trial, Filler trial, ObjectContrast/ColorContrast trial, Repetition/ColorContrast trial, with filler trial types used to vary the sequence of visual and spoken patterns across the list. The six pitch accent conditions associated with each lexical contrast type were arranged in a Latin square counterbalancing design, so that any effect of order would contribute equally to all conditions – for a total of six experimental lists. Each lexical contrast type / pitch accent condition combination occurred six times in each list.

2.7 Procedure

Participants were seated in front of a computer screen. Their eye movements were monitored with an Applied Sciences Laboratory (ASL) E6000 head-mounted eyetracker with
eye-head integration. Participants wore a lightweight plastic headband with a magnetic receiver and a 60 Hz camera providing a view of their right eye. The magnetic receiver provided information about the participants’ head location and orientation while the camera provided information about the location of the pupil and corneal reflection. Another 60 Hz camera was mounted to the wall behind them and provided a view of the computer screen. Separate screen and eye displays were available for monitoring by the experimenter throughout the experiment. The visual and auditory stimuli were controlled by a Dell laptop computer running E-Prime, Version 1.0 (Psychology Software Tools, Inc.). Eye movement data collection was synchronized with the beginning and end of each trial.

Due to the large number of trials (216) in the experiment, participants completed two successive experimental sessions (i.e. two blocks of trials with a break in between). They experienced the first 108 trials of the experiment in the first session and the last 108 trials in the second. Each session lasted about 20 minutes. Participants experienced only one of the six lists, and heard only one of the two speakers. Eye position was calibrated before each session and calibration was checked after each session. During each trial in a session, participants clicked on a cross in the middle of the computer screen below the written instruction Please click on the cross to start the next trial. Participants then saw 12 pictures of objects arranged in a grid on the computer screen (see Figure 2.1) and heard a series of three instructions of the form: Click on the COLOR OBJECT-TYPE, e.g., Click on the purple bottle. Using the mouse to click on an object triggered the next instruction in the series. After participants had responded to the three consecutive instructions, one of two things happened: (a) Participants saw the screen with the cross and started the next trial by clicking on the cross; (b) Participants performed a secondary task intended to ensure their attention: They saw a screen with an empty grid showing the locations of objects during the previous trial. Written instructions asked them where a specific object had been, e.g. Where was the yellow bottle? Participants responded by clicking on the object’s previous location. Then, participants saw the screen with the cross and started the next trial by click-
ing on the cross. Participants never performed the secondary task following a target trial. After completing the experimental sessions, participants filled in a brief questionnaire for purposes of course credit and were debriefed by the experimenter.

2.8 Data Collection and Preparation

Several kinds of data were collected. E-Prime provided time stamps for the beginnings and ends of auditory stimuli. The ASL E6000 eyetracker and integrated software calculated gaze location every 17 ms, starting with the beginning of each trial. Fixation onset times were automatically coded from the onset of the saccade since saccades are ballistic and listeners had already planned the eye movement during the time of the saccade (cf. Section 2.1.3 and Altmann & Kamide, 2004). ASL E6000 software was used to create Areas of Interest (AOIs) on the computer screen. For example, the area surrounding the object displayed in the top left corner of the screen was labeled as AOI 1. All gazes falling into this area could then be associated with AOI 1 (see Figure 2.12, where the yellow bottle is in AOI 1).

In addition, for each trial the AOI numbers of the target object (e.g. the yellow bottle in AOI 1 in Figure 2.12), the color competitor objects (the objects sharing the color with the target object; e.g. the yellow scissors in AOI 4 in Figure 2.12), and the object competitor objects (the objects sharing the object type with the target object; the purple bottle in AOI 8 and the red bottle in AOI 10 in Figure 2.12) were coded. Thus, for every trial, gaze location was recorded every 17 ms and linked to an AOI coded for whether it showed an object of interest.

A Perl script converted the coded data into a series of 1s and 0s such that for each trial and each object of interest a 1 was recorded if a participant looked at the cell containing the object of interest at a given time and a 0 was recorded if s/he did not. If, for example, a participant looked at one of the color competitor objects for 200 ms, every 17 ms during
this interval, a 1 was recorded in the color competitor row, but a 0 was recorded in the target object and object competitor rows.

Due to experimenter error, 4 of the 216 trials (1.85%) of Experiment A had to be excluded from the analysis.

2.9 Graphing Eye Movement Data

Data collected in the experiments are presented in a series of graphs: an initial overview graph of results for all conditions, a graph of the grand mean function for the experiment, and a series of paired graphs showing the data for critical conditions and the fit of the statistical models created for them. This section describes each graph type in succession, and provides a justification for the methodological and statistical decisions used to create them. The following section provides a brief overview of the statistics used for all analyses.
in the dissertation, with additional information about the display and interpretation of models and effects.

All graphs of the results show time (in milliseconds) on the x-axis and proportion of looks to the target or competitor object on the y-axis. Looks are presented as fixation proportions or an empirical logit of fixation proportion (see Section 2.10), both collapsed over items as appropriate for conditions displayed. This graphic presentation style is appropriate for participants’ responses to the auditory instructions as they unfold over time.

The initial graph for each set of results displays fixation proportions to the target or competitor object(s) in all conditions, providing an overview of the data. Fixation proportion for each condition was calculated from all items and all participants. That is, at any given point in time fixation proportion was calculated by dividing the number of looks to an object of interest by the total number of trials in that condition. Fixation proportion is used on the y-axis of the overview graph because fixation proportion values are easily interpretable. For example, a value of 0.6 on the y-axis means that 60% of the trials in a certain condition elicited looks to the object of interest at a given point in time. In contrast, the paired data and model-fit graphs present the same results using an empirical logit scale. This scale shows empirical log-transformed odds, so that 0 corresponds to a 0.5 likelihood of fixation. Another way of stating this is that positive logit values indicate more than a 50% likelihood of looks to the object of interest while negative values indicate less than a 50% likelihood.

An example overview graph using data from EXPERIMENT A is shown in Figure 2.13. The graph shows the proportion of fixations to the target object on the screen as participants listened to the second instruction of a sequence like purple bottle ⇒ purple bottle (repetition) or purple pencil ⇒ purple bottle (object contrast), where the first instruction was pronounced with a HHNP LL pattern and the target instruction was pronounced either with a HHNP LL pattern (Click on the purple bottle), a LHA LL pattern (Click on the PURPLE bottle), or a LHN LL pattern (Click on the purple BOTTLE). The legend shows which
Figure 2.13: Sample overview graph from EXPERIMENT A. Proportion of fixations to the target object in the display in response to hearing *Click on the purple bottle*, *Click on the purple bottle* or *Click on the purple BOTTLE* either in a repetition or color contrast sequence.

The fixation proportion data are aligned for each item and each participant at a time point between the two successive critical parts of the instruction (the adjective and the noun). That is, 0 on the x-axis (indicated by the solid vertical line) indicates the end of the adjective (i.e. the beginning of the noun) for each item for each participant. The closer critical words are to the point of alignment, the more reliably can effects be said to have happened at a certain time or in a certain time window. This kind of alignment is especially important when prosodic accentuation is manipulated since accented words are generally
longer than unaccented words. Aligning between the two critical parts of the instruction allows the best possible alignment for both adjectives and nouns: Adjectives will be perfectly aligned at their offset, nouns at their onset. Adjective onsets and noun offsets will be estimated using adjective and noun durations, respectively, averaged across like tokens. This kind of display allows for processing differences in both the adjective and the noun to be displayed in the same graph.

The dashed vertical line occurs 200 ms after the beginning of the noun and indicates a best estimate of the time point when looks in response to hearing the beginning of the noun (bottle) could start occurring, given that it takes approximately 200 ms to plan and execute an eye movement when listeners follow auditory search instructions for objects in a relatively complex display (cf. Allopenna et al., 1998; Dahan et al., 2001; Ito & Speer, 2008; Matin et al., 1993).

The boxes at the top of the graph show the average duration of the adjective and noun in the three prosodic conditions (HHNP LL, LHA LL, and LHN LL). The left edge of the adjective-duration boxes shows the average beginning of the adjective in each prosodic condition; the right edge of the noun-duration boxes shows the average end of the noun in each prosodic condition. Colored object drawings close to a function or set of functions show the object in the visual display that corresponds to the function(s). That is, the boxes show the words participants heard during the second instruction, whereas the drawings show the object to which they looked. In this case, looks to the target object are shown, so the drawings show the object to which the instruction referred.

The example graph in Figure 2.13 shows that looks to the target object rise much earlier and from a much higher baseline in the repetition conditions compared to the object contrast conditions. The difference in baseline is most likely a result of participants still fixating the purple bottle at the beginning of the trial in the repetition sequences (since it was also the target object of the first instruction). The earlier rise of looks to the target object in the repetition compared to the object contrast sequences is likely related to participants always
knowing the location of the purple bottle in the repetition conditions (since they fixated it during the first instruction), but less frequently so in the object contrast conditions (i.e. only if they fixated it during the first instruction).

Following the overview graph is a graph showing the grand mean data from all conditions in an experiment. This graph will be used to determine an appropriate analysis window for the statistical analyses, and will be described in more detail in the following section (Section 2.10). Finally, a series of paired graphs zoom in on individual comparisons. For each pair, the left panel shows data points and error bars and the right panel shows model fits from a growth curve analysis for the two conditions being compared. The y-axis in both panels shows an empirical logit of the fixation proportions since transformed data was used in the statistical analyses (see Section 2.10). An example graph showing an individual comparison from Figure 2.13 is shown in Figure 2.14. The repetition HHNP LL and repetition LHN LL conditions are compared in the graph.
The left panel of the graph shows that the curve for the repetition HHNP LL condition seems to have a steeper rise to the peak of the curve than the curve for the repetition LHN LL condition. In addition, we can see that there is a small time window (ranging from 238 ms to 289 ms) with reliably more looks to the target object in the repetition HHNP LL condition compared to the repetition LHN LL condition. The right panel of the graph shows that the model fits for the two curves also differ mainly in the steepness of the rise of the curves. This panel relates to the statistical analyses and will be explained in more detail in Section 2.10.

2.10 Statistical Analyses

The data from Experiments 1 through 4 will be analyzed using growth curve analysis (see Mirman et al., 2008; Singer & Willett, 2003). Traditionally, repeated measures ANOVAs have been performed on visual world data, with fixation proportions as the dependent measure. A common approach compares main effects and interactions among the independent variables, with separate ANOVAs conducted for participants and stimulus items as random variables. Such an approach has recently been criticized in the literature for several reasons (cf. Baayen et al., 2008; Barr, 2008; Jaeger, 2008; Mirman et al., 2008): Time (which is a continuous variable) is treated as categorical, with mean fixation proportions calculated across arbitrary “windows”. In addition, time is not treated as a factor in the analyses, missing potentially interesting information about the time course of effects. The dependent measure, fixation proportion, is treated as continuous when in fact it is categorical – participants are either looking or not looking at an area of interest.

For the data presented here, the criticism regarding the treatment of time is most relevant. Eyetracking allows obtaining fine-grained time course data. For example, eye movement location data used in this dissertation were collected at a camera rate of 60Hz, approximately every 17 ms. Condition means for traditional repeated measures ANOVAs
are most commonly obtained by calculating mean fixation proportions for each condition either for the whole time window of interest (e.g. Chambers et al., 2004; Magnuson et al., 2003) or for successive smaller time windows (e.g. Ito & Speer, 2008; Weber et al., 2006a, who chose 300 ms time windows). In the first case, time course information is lost completely. In the second case, time course information is greatly reduced. In particular, information about the trajectory of change over time is lost (see Mirman et al., 2008). In addition, the choice of window location and size can affect statistical significance. Without a priori justification for particular window locations and sizes, the analysis remains somewhat arbitrary.

Recently, mixed-effects models, logistic regression models, and growth curve analysis (GCA) have been proposed as alternatives to the traditional repeated measures ANOVAs (Baayen et al., 2008; Barr, 2008; Jaeger, 2008; Mirman et al., 2008), primarily on the grounds that they produce fewer errors and are more powerful. The data presented here will be analyzed using GCA. This analysis was chosen because it provides insight into the trajectory of change over time. In particular, it allows comparison of differences in curve shape that reflect changes in the dependent measure over time. This is especially advantageous when dealing with an independent measure like prosody, which unfolds over time. In other words, changes in the dependent measure can more easily be linked to information from the unfolding prosodic structure. GCA has traditionally been applied to longitudinal data. Even though the time scales in eyetracking data (milliseconds) and typical longitudinal data (months or years) are vastly different, the data structures are similar.

Growth curve modeling is a standard regression approach, that uses hierarchically-related submodels. Importantly, the first submodel (level-1) measures effects of time. As in standard regression models, there is an intercept, a slope, and an error term. The second submodel (level-2) captures variation across individuals and/or experimental conditions on the intercept and/or slope. For example, the level-2 model measuring variation across individuals on the intercept is the population average for the intercept and individuals’
deviation from this population average. Orthogonal power polynomials (linear transformations of time\(^n\) polynomial terms) will be used to expand the model to represent lines that are not straight (another approach would be to use a logistic function). The main advantage of using orthogonal power polynomials as opposed to non-orthogonal polynomial terms is that lower order terms are independent of higher order terms. This independence is important, for example, when testing an effect of condition on the linear term, while also adding a quadratic and cubic term to the model. If the terms were not independent, the higher order terms (quadratic and cubic) would change the estimated effect of the lower order term (linear). This is undesirable if the effect on the linear term is the one of interest. The most appropriate analysis for the current data involves exactly the situation just described. As will become clear below, the highest-order polynomial term included in each model is never the one for which the estimated effect is of interest.

Figure 2.15 relates the order of orthogonal power polynomials included in a model to the shape of the curve captured by that model. Only curve shapes that may be observed in target and competitor fixation data are considered. The top panels of Figure 2.15 show possible target fixation shapes. A model with an intercept and linear term captures lines that increases linearly (panel 1; a horizontal line would be captured by the intercept term only). A model with an intercept, linear, and quadratic term describes curves that increase linearly and then level off or curves that start level and then increase linearly (panels 2 and 3). Finally, a model with an intercept, linear, quadratic, and cubic term describes curves that start level, then rise linearly, then level off again (panel 4). The bottom panels of Figure 2.15 show possible competitor fixation curves. In this case, a model with an intercept, linear, and quadratic term describes curves that increase linearly, then decrease linearly (panel 5). A model with an intercept, linear, quadratic, and cubic term captures curves that start level, then increase linearly, and then decrease linearly (panel 6) or curves that increase linearly, then decrease linearly, and then level off (panel 7). A model with an intercept, linear, quadratic, cubic, and quartic term presents curves that start level, then
rise and fall, and level off again (panel 8).

Target fixation curves usually rise steadily throughout the window of analysis. Depending on how much of the tails of the curves are included in the analysis window, target fixation curves can be level before and after the steady rise. A model describing target fixations should thus be able to capture the curves shown in panels 1 through 4. This means that the cubic term should be the highest order polynomial term included in GCA models of target fixations. (If looks away from the target object are also included in the analysis window, the quartic term should also be included in the GCA.) Competitor fixation curves usually have a symmetric rise-fall pattern throughout the window of analysis. Depending on how much of the tails of the curves are included in the analysis window, competitor fixation curves can be level before and after the rise-fall pattern. A model describing competitor fixations should thus be able to fit the curves shown in panels 5 through 8, and the quartic term should be the highest polynomial term included in GCA models of competitor fixations.

When looking at target fixations, only the intercept and linear terms relate meaningfully to cognitive processes (even though the quadratic and cubic terms are needed for curve fitting). The intercept term captures differences in average curve height and thus corresponds to the area under the curve. The linear term captures differences in slope. A reliable effect on the intercept term with no effect on the linear term corresponds to more earlier looks to the target in some conditions over others. A reliable effect on the linear term with no effect on the intercept term reflects a difference in slope such that curves cross towards the middle of the analysis window (i.e. some conditions have more looks to the target than other conditions early in the time window, but fewer later in the time window). A reliable effect on both the intercept and linear terms indicates a difference in slope and area under the curve. Here, the relevant curves either cross towards the beginning or end of the analysis window. In the first case, looks to the target start rising around the same time across conditions, but rise faster for some conditions than others. In the second case,
Figure 2.15: Schematic relationship of the order of orthogonal polynomial terms and relevant captured curve shapes. The top panels show schematic target fixation curves, and the bottom panels show schematic competitor fixation curves.
looks to the target start rising earlier, but with a shallower slope, in some conditions than others. Overall, then, reliable differences on the intercept term correspond to a processing advantage or delay, and reliable differences on the linear term correspond to a change in (degree of) processing advantage over time. As can be seen in Figure 2.15, the quadratic and cubic terms mainly capture small differences in the asymptotic tails of target fixation curves. Because, for target trials, the primary focus of this dissertation concerns looks to the target object that begin during the noun phrase that describes it, effects that may be reflected in asymptotic tails will not be considered further here.

In the case of competitor fixations, the intercept and quadratic terms relate meaningfully to cognitive processes (but the model should include a linear, cubic, and quartic term). Both terms capture differences in curve height, which corresponds to how strongly a competitor is considered to be the target. The intercept term captures differences in average height of the curve, and the quadratic term captures differences in rise-fall rate. (If the rise-fall pattern is not symmetric, however, then the linear and/or cubic terms would capture differences in rise-fall rate.) A reliable effect on the intercept, but not the quadratic term, suggests that there are more looks to the competitor in some conditions over others throughout the time window. A reliable effect on the quadratic term, but not the intercept term, would suggest fewer looks to the competitor in some conditions over others towards the ends of the analysis window, but more looks in the middle of the analysis window. A reliable effect on the intercept and quadratic terms indicates equal looks to the competitor towards the ends of the analysis window, but more looks to the competitor in some conditions over others in the middle of the analysis window. Here, the linear captures asymmetric differences in rise-fall pattern. The cubic and quartic terms capture small differences in the asymptotic tails of the curve.

GCA does not require data aggregation within the analysis window. Nevertheless, an analysis window must be chosen. Eye movements driven by the auditory input usually appear in a graph as a sudden rise in fixations to the target object. Barr (2008) proposes
to plot the grand mean data for all conditions and start the analysis window at the point where looks to the target object begin rising. The end of the analysis window will be the peak of looks to the target object in the grand mean data. For competitor fixations, the beginning of the analysis window will also be the point where looks to the competitor object begin rising in the grand mean data. The end of the analysis window will be determined based on the peak of looks to the competitor object. The point in time which approximately splits the mountain of the curve in half will be considered the peak. The end of the analysis window will be the same number of data points to the right of the peak than the beginning of the analysis window is to the left of the peak. For example, if the curve starts rising at 0 ms and peaks at 170 ms, then the end of the analysis window will be 340 ms. Choosing the analysis window based on the grand mean avoids biasing towards a given hypothesis (cf. Barr, 2008). This method also avoids including too much of the asymptotic tails in the analyses, which tends to result in a poor GCA model fit (cf. Mirman et al., 2008).

Thus, the appropriate analysis window will be chosen individually for each experiment by plotting and visually inspecting the grand mean data. Figure 2.16 shows the grand mean data for the conditions shown in Figure 2.13 in Section 2.9. Figure 2.16 shows that looks to that target object start rising at -238 ms and start dropping off again at 731 ms. Therefore, the analysis window for GCAs run on this particular data would range from -238 ms to 731 ms.

The dependent variable for the GCA will be an empirical logit of fixation proportion rather than fixation proportion as used in Figure 2.13. Fixation proportion as the dependent variable may underestimate differences at the ends of the range (cf. Baayen et al., 2008; Barr, 2008; Jaeger, 2008, but see Mirman et al., 2008). This mostly affects the asymptotic tails, which are not of much interest for the analysis. Using fixation proportion rather than a more appropriate logistic scale may, however, affect the slope of the curves. In particular, two curves may differ in slope on the probability scale but not on the log odds scale and vice versa (cf. Barr, 2008). Using fixation proportion may thus provide incorrect
Figure 2.16: Grand mean of looks to the target object cell in the six critical conditions presented in Figure 2.13 in Section 2.9. The arrows indicate the beginning and end of the analysis window for the GCA.

results on the linear term, either showing a reliable effect when there is none (Type I error) or not detecting an effect that is actually there (Type II error). However, using GCA with a logistic regression instead of fixation proportions may greatly inflate differences in the asymptotic tails, especially if there are few trials and/or participants. As a compromise, an empirical logit link function (Barr, 2008) is used here. It corrects for underestimating differences at the ends of the range, but does not inflate differences in the asymptotic tails. This can be seen in Figure 2.17, which shows target and competitor data from Experiment 1 (see Chapter 3 below) with fixation proportion, log odds, and empirical logit on the y-axis. Notice that differences at the bottom of the range are greatly inflated in log odds, but not using an empirical logit transformation$^7$.

$^7$The slopes of the curves on the empirical logit scale do, however, resemble those on the probability scale more than those on the log odds scale. This suggests that the empirical logit transformation may only reduce, but not eliminate, potential Type I and Type II errors on the linear term.
Figure 2.17: Data from EXPeriment A. Proportion, log odds, and empirical logit of fixations to the target object (top panels) or color competitor object (bottom panels) of data from Experiment 1 (see Chapter 3 for an explanation of the conditions names).
Each GCA reported in the upcoming chapters starts with a base model with the empirical logit of fixation proportion as the dependent variable and the orthogonal time polynomials, subjects, and subjects-by-time-polynomials as fixed effects. Subjects-by-condition is added as a random effect. Items are not included as a random effect (traditional in psycholinguistic analyses using ANOVA) as this can lead to under-reporting of effects when experimental conditions are matched and counterbalanced, as they are here (Raaijmakers et al., 1999). The deviance statistic (minus two times the log likelihood, -2LL) is calculated for the base model. Then, a parameter is added to the model and the deviance statistic for this model is calculated. The change in deviance ($\Delta D$, i.e. the difference between the -2LL of both models) reveals whether addition of the parameter improved model fit. For example, a fixed effect of condition on the intercept is added to the base model. A reliable change in deviance would suggest that this addition improved model fit and that the conditions differed reliably in average curve height. Then, a fixed effect of condition on the linear term is added to the intercept model etc. The analyses will be performed using the lme function in R (R Development Core Team, 2010). A model fit using lme is typically of the form \( \text{lme}(\text{response} \sim \text{covariate(s)} + (\text{covariates} | \text{group}), \text{data}) \). Example (2.4) a. shows the R code for a sample base model for target fixations. \( \text{elogtarget} \) refers to the empirical logit transformed data, \( \text{ot1} \) through \( \text{ot3} \) refer to orthogonal time polynomials, which model effects of time, and \( \text{subj:ot1} \) through \( \text{subj:ot3} \) provide subjects-by-time-polynomials, which model subjects effects. \( \text{random} = \sim \text{ot1} | \text{subcond} \) models subjects-by-condition as a random effect. \( \text{method} = \text{"ML"} \) means that the model is fit using Maximum Likelihood Estimation. Example (2.4) b. shows the R code for a model where an intercept, linear, quadratic, and cubic term of condition (\( \text{Cond} \)) has been added to the base model.

(2.4) a. \( \text{t.base = lme}(\text{elogtarget} \sim \text{ot1} + \text{ot2} + \text{ot3} + \text{subj} + \text{subj:ot1} + \text{subj:ot2} + \text{subj:ot3}, \text{data=data1, random = \sim ot1 | subcond, method=\"ML\")} \)

b. \( \text{t.3 = lme}(\text{elogtarget} \sim \text{ot1} + \text{ot2} + \text{ot3} + \text{subj} + \text{subj:ot1} + \text{subj:ot2} + \text{subj:ot3} + \)
Experiment 1 has a 2x3 design, and Experiments 2 through 4 have a 2x2 design. The following GCA are performed for each experiment: First, an omnibus GCA tests for significance of each factor and the interaction factor. This is done by first adding one factor to the base model, then the other factor, and finally the interaction term. All terms are added step-by-step from the lowest to highest polynomial term. For example, if the two factors are lexical contrast and pitch accent pattern, *Lexical Contrast* will be added step by step to the base model (starting with the intercept term and ending with the cubic or quartic term), then *Pitch Accent* will be added step by step to the full *Lexical Contrast* model, and finally, the *Lexical Contrast / Pitch Accent* interaction will be added step by step to the full *Lexical Contrast + Pitch Accent* model. Since the design is fully crossed, it does not matter which factor is added first. The results from each omnibus GCA will be presented in a table showing the deviance statistic (-2LL) for each model, the change in deviance (\(\Delta D\)) of each model and the previous one, and a \(p\)-value indicating whether addition of the last term improved model fit.

The results of an omnibus GCA of the target fixation data from Figure 2.13 in Section 2.9 are shown in Table 2.17. For target fixation analyses, the intercept and linear terms relate meaningfully to cognitive processes. Table 2.17 shows a main effect of *Lexical Contrast* (i.e. repetition vs. color contrast) on both the intercept and linear terms. This means that lexical contrast type affects both the area under the curve and the steepness of the curve. The reliable effect on the intercept term shows that there are reliably earlier looks to the target object in the repetition conditions compared to the color contrast conditions. The reliable effect on the linear term shows that once looks to the target object start rising in the color contrast condition, they rise steeper than in the repetition conditions. Table 2.17 also shows a main effect of *Pitch Accent* (i.e. HHNP LL vs. LHA LL vs. LHN LL) on the linear term, but
not the intercept term. This suggests that the pronunciation of the target instruction affects how steeply looks to the target object rise.

Next, pairwise comparisons are obtained by performing GCA on appropriate subsets of the data. For experiments with a 2x2 design, the results of the pairwise comparisons will be presented in a table containing the same information as that for the omnibus GCA (-2LL, \( \Delta D \), and \( p \)). Experiment 1 has a 2x3 design with two levels of Lexical Contrast and three levels of Pitch Accent. To minimize the number of models run when comparing the three levels of Pitch Accent, one level will first be compared to the other two levels in one analysis. Tables showing the results from this kind of GCA comparison will show the model fit for the first level by providing -2LL, \( \Delta D \), and \( p \). Whether this level differs reliably from the other two levels can be seen in the parameter estimate columns, which provide a parameter estimate, a \( t \)-value and an associated \( p \)-value for each polynomial term. The other two levels will be compared directly as in the experiments with a 2x2 design.

Table 2.18 shows the results from a direct comparison of the repetition HHNP LL condition of Figure 2.13 with the repetition LHA LL and repetition LHA LL conditions. The \( p \)-values under LHA LL would reveal reliable effects between the repetition HHNP LL and repetition LHA LL conditions. Those under LHN LL would reveal reliable effects between the repetition HHNP LL and repetition LHN LL conditions. The table shows no reliable effects on either the intercept or linear terms, suggesting that looks to the target object in a repetition sequences are not affected by the pitch accent pattern of the target instruction.

Figure 2.14 in Section 2.9 showed data points and error bars, as well as model fits, comparing the repetition HHNP and repetition LHN LL conditions. Figure 2.14 corresponds to the GCA results under LHN LL in Table 2.18. The figure shows two kinds of model fits. The solid lines show the model fit for the full model (i.e. including the intercept through cubic terms), whereas the dashed lines show the model fits without the quadratic and cubic terms. Figure 2.14 reveals that the curve for the repetition HHNP LL condition is numerically steeper than that for the repetition LHN LL condition. Table 2.18 reveals that
Table 2.17: Data shown in Figure 2.13 in Section 2.9: Target fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
the difference in steepness of the curves in not reliable.

The presentation of graphs and statistical analyses in the upcoming chapters can be summarized as follows. Overview graphs are presented for comparability with previous work that has typically used this kind of display and to give an intuitively accessible presentation of the results. Aligning graphs between the adjective and noun maximizes the accuracy of the critical time portion of the data. Examining the grand mean function of the data determines the appropriate overall window of analysis (without any biasing reference to experimental condition). Omnibus GCAs on empirical logit transformed data test for the reliability of factors and the interaction term. This is followed by relevant pairwise GCA comparisons and corresponding graphs.
<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>LHA LL</td>
</tr>
<tr>
<td></td>
<td>ΔD</td>
<td>Est.</td>
</tr>
<tr>
<td>Base</td>
<td>12079.8</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>12079.26</td>
<td>-0.54 n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>12077.18</td>
<td>-2.09 n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>11998.51</td>
<td>-78.66 0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>11997.97</td>
<td>-0.53 n.s.</td>
</tr>
</tbody>
</table>

Table 2.18: Data shown in Figure 2.13 in Section 2.9: Target fixation analysis comparing the repetition HHNP LL condition (model) with the repetition LHA LL and repetition LHN LL conditions.
Experiment 1 tests whether previously obtained, robust facilitation and garden-path effects such as those found in Weber et al. (2006a) and Ito & Speer (2008) can be obtained when participants are exposed to more trials in much faster succession than in previous studies. It compares target utterances with a L+H* in felicitous and infelicitous sentence positions as well as target utterances with a H* !H* pattern.

Both Weber et al. (2006a) and Ito & Speer (2008) demonstrated that for adjective noun sequences in succession, such as green bottle ... purple bottle, a L+H* accent on the second color adjective restricts the referent of the upcoming noun to that of the previous noun phrase. In Ito & Speer’s (2008) study, for example, when listeners heard utterances like Hang the green drum. Now hang the BLUE... they looked at the blue drum (the color-contrast object, differing only in color from the one mentioned just previously), regardless of the identity of the following noun. This led to a facilitation effect when listeners heard Hang the green drum. Now hang the BLUE drum: Listeners started looking at the blue drum before lexical information from the word drum was heard. But it also led to a prosodic “garden-path” effect when listeners heard Hang the green drum. Now hang the BLUE angel: Listeners often looked at the blue drum and only corrected themselves when enough lexical information from the noun angel had been heard. The absence of a L+H* accent on the second noun phrase did not cause referent restriction to the color-contrast object. When listeners heard Hang the green drum. Now hang the blue drum, looks to the blue drum started increasing later than when listeners heard Hang the green drum. Now hang the BLUE drum. In addition, when listeners heard Hang the green drum. Now hang the blue angel, they were
not garden-pathed to look at the blue drum. Finally, when listeners heard an infelicitous sequence like *Hang the green drum. Now hang the blue DRUM*, where the L+H* is placed on the repeated noun, looks to the blue drum also started increasing later than when the adjective contained the L+H* pitch accent. Weber et al. (2006a) found similar results for sequences with felicitous and infelicitous L+H* sequences. However, they found a general preference for the color-contrast object that led to a small garden-path effect even when the L+H* was on the noun, with looks to the red scissors for sequences such as *Click on the purple scissors. Click on the red VASE.*

Experiment 1 tests whether the above mentioned effects from Weber et al. (2006a) and Ito & Speer (2008) can be obtained with the current experimental design. Effects are expected to be comparable to previous studies for several reasons: Previous studies suggest that prosody is processed immediately and in parallel with lexical input. In addition, eye movements are fast, automatic, and metabolically cheap. The speed and duration of the experimental task should therefore not alter or interfere with participants’ eye movement responses. Therefore, a L+H* accent on a color adjective is expected to evoke a contrast set of different colors and to restrict the second referent to one differing only in color with the previous one. This restriction should lead to earlier eye movements (compared to looks in other conditions) to the color-contrast object. Looks to the color-contrast object are expected to start increasing before noun onset. In a color-contrast sequence, i.e. when successive instructions refer to objects differing only in color, the early looks will be to the target object and thus anticipatory; In a no-contrast sequence, i.e. when successive instructions refer to objects differing in color and object type, the early looks will be to the competitor object differing only in color from the previously mentioned object (garden-path effect).
3.1 Experimental Conditions

Experiment 1 included auditory conditions like those found in Weber et al.’s (2006a) Experiment 1 and Ito & Speer’s (2008) Experiments 1 and 2. Tables 3.1 through 3.4 show the critical comparison conditions for this experiment. In all cases, the context utterance is Instruction 1 and the target utterance is Instruction 2. (Instruction 3 is never relevant for Experiment 1.) Data from both Experiment A and Experiment B are included because the prosody in the two large experiments differed only in the edge tone patterns ending the second instruction, that is, prosody differed after participants heard the relevant L+H* accent or H*!H* pattern of Instruction 2.

Tables 3.1 and 3.3 show the prosodic conditions used in Experiment 1 that were taken from Experiment A, and Tables 3.2 and 3.4 show those from Experiment B. Tables 3.1 and 3.2 present color-contrast sequences, where the target object differed from the previous object only in color (yellow bottle ⇒ purple bottle). It compares trials where the target adjective carried a L+H* accent and a deaccented noun (LHA) to those where the target NP carried a H*!H* pattern (HHNP) and those where the target noun carried a L+H* accent and followed an unaccented adjective (LHN). The condition names for these sequences are felicitous, neutral, and infelicitous, respectively (cf. condition names in Ito & Speer, 2008).

The felicitous condition received this name because the L+H* accent on the adjective of the target instruction is in the appropriate sentence location. The objects referred to in the context and target instructions differed only in color and the L+H* accent on the color adjective of the target utterance felicitously marked this lexical contrast. In the neutral condition, the target instruction carried a H*!H* pattern, which neither felicitously marked the color contrast between the context and target instructions nor provided information conflicting with it. The condition is called neutral because the prosody of the target instruction provided no information concerning the identity of the noun. The unaccented adjective followed by the L+H* accent on the noun of the infelicitous condition provided pitch accent
<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Context</td>
<td>Color Contrast</td>
<td>Object Contrast</td>
</tr>
<tr>
<td>felicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the PURPLE bottle ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
<tr>
<td>neutral</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the purple bottle ↘</td>
<td>Click on the purple pencil ↘</td>
</tr>
<tr>
<td>infelicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the purple BOTTLE ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
</tbody>
</table>

Table 3.1: Experiment 1, Color-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All instructions end in a L-L% pattern (indicated by a ↘).

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Context</td>
<td>Color Contrast</td>
<td>Object Contrast</td>
</tr>
<tr>
<td>felicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the PURPLE bottle ↗</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
<tr>
<td>neutral</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the purple bottle ↗</td>
<td>Click on the purple pencil ↘</td>
</tr>
<tr>
<td>infelicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the purple BOTTLE ↗</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
</tbody>
</table>

Table 3.2: Experiment 1, Color-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).

Information conflicting both with a non-repeated adjective and with a repeated noun. This kind of pitch accent pattern would felicitously mark the object contrast in a sequence like *Click on the yellow bottle. Click on the yellow SCISSORS.* Findings would be consistent with previous effects if looks to the purple bottle occurred earlier in the felicitous compared to the neutral and infelicitous conditions.

Tables 3.3 and 3.4 present no-contrast sequences, where the target shared neither color nor object type with that of the previous instruction (*green pencil ⇒ purple bottle*). The conditions again compared trials where the target adjective carried a L+H* accent that pre-
ceded a deaccented noun (LHA) to those where the NP carried a H* !H* pattern (HHNP) and those where the target noun carried a L+H* accent and followed an unaccented adjective (LHA). These conditions are named garden-path, neutral, and infelicitous, respectively.

The L+H* accent on the target adjective of the garden-path condition carried the expectation of an upcoming repeated noun. However, the target noun was not repeated in the no-contrast sequences in Tables 3.3 and 3.4. Therefore, the L+H* accent did not felicitously indicate the identity of the target noun. Rather it should have lead listeners up the garden-path by suggesting that the noun of the context instruction (scissors) would be repeated.

The H* !H* pattern on the target instruction of the neutral condition was again neutral in the sense that it provided no information about the identity of the target noun. The L+H* on the target noun of the infelicitous condition was again infelicitous because it felicitously marked an object contrast rather than no lexical contrast. Findings would be consistent with previous effects if participants were garden-pathed to look at the color competitor (the purple pencil) in the garden-path condition, but not or less so in the neutral and infelicitous conditions.
Table 3.4: Experiment 1, Non-contrastive sequence from EXPERIMENT B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in bold face is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>garden-path</td>
<td>Click on the green pencil ↘</td>
<td>Click on the PURPLE bottle ↗</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
<tr>
<td>neutral</td>
<td>Click on the green pencil ↘</td>
<td>Click on the purple bottle ↗</td>
<td>Click on the purple pencil ↘</td>
</tr>
<tr>
<td>infelicitous</td>
<td>Click on the green pencil ↘</td>
<td>Click on the purple BOTTLE ↗</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
</tbody>
</table>

3.2 Results

3.2.1 Target Fixations

Data from EXPERIMENT A

Figure 3.1 provides an overview of the data. It shows looks to the target object cell in the six critical conditions of EXPERIMENT A (see Tables 3.1 and 3.3). The figure shows that, as predicted, looks to the target object started rising earliest in the felicitous condition. In addition, looks to the target object in the no-contrast sequences started increasing later than looks to the target object in the color-contrast sequences.

Figure 3.2 shows the grand mean of looks to the target object for the six critical conditions. The figure shows a fixation proportion function that has a single peak, with a shape that is level, begins to rise, and then starts to fall after peaking at 867 ms. The beginning of the curve’s rise was chosen as the beginning of the analysis window. It was set at the -170 ms data point because it is the last data point that is level with the previous one. The peak at 867 ms was chosen as the end of the analysis window. The beginning and end of the analysis window for the GCA are marked by arrows.
Figure 3.1: Experiment 1, data from EXPERIMENT A: Proportion of looks to the target object cell in the six critical conditions presented in Tables 3.1 and 3.3.
Figure 3.2: Experiment 1, data from EXPERIMENT A: Grand mean of looks to the target object cell in the six critical conditions presented in Tables 3.1 and 3.3. The arrows indicate the beginning and end of the analysis window for the GCA.

Based on the curve shown in Figure 3.2, the analysis window for all GCAs in this section will range from -170 ms to 867 ms. The results from an omnibus GCA are shown in Table 3.5. Of interest for fixations to a target object are the intercept term and linear term (see Section 2.10). The table shows a main effect of *Lexical Contrast* on the intercept term, but not the linear term. This means that the area under the curve for the color-contrast sequences is larger than that for the no-contrast sequences, but that the curves in the two sequences do not differ in slope. This pattern of results suggests a constant advantage for the color-contrast conditions over the no-contrast conditions. In other words, throughout the analysis window, there were earlier looks to the target object in the color-contrast conditions compared to the no-contrast conditions. The table also shows a main effect of *Pitch Accent* on the intercept term, but not the linear term. This suggests that, overall, instructions with a L+H* on the adjective (that is, the *felicitous* and *garden-path* conditions combined) elicited earlier looks to the target object than instructions with a H* !H* pattern.
or a L+H* on the noun. Finally, the table shows a Lexical Contrast / Pitch Accent interaction on the intercept term, but not the linear term. This suggests that looks to the target object rose earlier in the felicitous condition than other conditions. In other words, resolving reference was facilitated when there was both a color-contrast and a L+H* accent on the adjective of the target instruction.

Figures 3.3 through 3.5 show direct comparisons of the three color-contrast sequences. (Direct comparisons of the three no-contrast sequences will not be shown here because these sequences were predicted to affect looks to the color-competitor object, not the target object.) The graphs on the left show data points with error bars and the graphs on the right show model fits for the curves predicted by the model. All graphs show time (ms) on the x-axis and an empirical logit of looks on the y-axis (see Sections 2.9 and 2.10). Within the color-contrast sequence, we will consider the neutral condition as a prosodic baseline since both instructions in this condition were pronounced with a H* !H* pattern on the NP, which provided no (helpful or misleading) cues as to the identity of the target object. Figures 3.3 and 3.4 compare the felicitous condition to the neutral and infelicitous conditions, respectively. They show earlier looks to the target object in the felicitous condition, both compared to the neutral and to the infelicitous conditions. Examination of the error bars shows that this advantage is reliable from -102 ms, where the two lines start to split, to 340 ms for the felicitous compared to the neutral condition, and from -68 ms to 238 ms for the felicitous compared to the infelicitous condition. This suggests that the felicitous condition presented a processing advantage compared to the neutral and infelicitous conditions. The advantage occurred before noun onset and must therefore have been based solely on auditory information from the adjective PURPLE. Figure 3.5 compares the neutral and infelicitous conditions. It shows no advantage between the two conditions across the analysis window. This suggests that the infelicitous L+H* on the noun presented no delay in processing compared to the H* !H* pattern of the baseline condition.

Results from a GCA comparing the felicitous condition to the neutral and infelicitous
<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Model fit</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
<td>-2LL</td>
</tr>
<tr>
<td>Base</td>
<td>29269.35</td>
<td>—</td>
<td>—</td>
<td>28474.42</td>
</tr>
<tr>
<td>Intercept</td>
<td>29214.36</td>
<td>-54.99</td>
<td>0.0001</td>
<td>28468</td>
</tr>
<tr>
<td>Linear</td>
<td>29212.78</td>
<td>-1.58</td>
<td>n.s.</td>
<td>28467.03</td>
</tr>
<tr>
<td>Quadratic</td>
<td>28483.77</td>
<td>-729</td>
<td>0.0001</td>
<td>28461.17</td>
</tr>
<tr>
<td>Cubic</td>
<td>28474.42</td>
<td>-9.35</td>
<td>0.01</td>
<td>28406.6</td>
</tr>
</tbody>
</table>

Table 3.5: Experiment 1, data from EXPERIMENT A (see Figure 3.1): Target fixation analysis: Main effects (Lexical Contrast and Prosody) and interaction.
Figure 3.3: Experiment 1, data from EXPERIMENT A: Comparison of the felicitous and neutral conditions presented in Table 3.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Figure 3.4: Experiment 1, data from EXPERIMENT A: Comparison of the felicitous and infelicitous conditions presented in Table 3.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.
conditions are shown in Table 3.6. Again, of interest are the intercept and linear terms. The table shows a reliable effect on the intercept term both comparing the felicitous condition to the neutral condition and to the infelicitous condition. This confirms the advantage of the felicitous over the neutral and infelicious conditions observed in Figures 3.3 and Figures 3.4. The absence of a reliable effect on the linear term shows that there is no difference in slope between the felicitous and the other two conditions. Notice that the GCA also reveals a reliable effect on the cubic and quartic terms. These do not translate into any cognitive processes (see Section 2.10). As can be seen in Figures 3.3 and 3.4, they reflect differences in the asymptotic tails of the curves. Table 3.7 shows results from a GCA comparing the neutral and infelicitous conditions. The analysis reveals no reliable differences between the two conditions. This confirms that the infelicitous condition presented no processing delay compared to the neutral condition.
### Table 3.6: Experiment 1, Color-contrast sequence from EXPERIMENT A (see Figure 3.11): Target fixation analysis comparing the felicitous condition (model) with the neutral and the infelicitous conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Parameter estimates</th>
<th>Parameter estimates</th>
<th>Parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>neutral</td>
<td>infelicitous</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ΔD</td>
<td>Est.</td>
<td>t</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>13632.67</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>13616.45</td>
<td>-16.22 0.001</td>
<td>-0.27026 -3.34 0.01</td>
<td>-0.27463 -3.39 0.01</td>
</tr>
<tr>
<td>Linear</td>
<td>13615.75</td>
<td>-0.7 n.s.</td>
<td>0.34409 0.82 n.s.</td>
<td>0.12773 0.3 n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>13612.13</td>
<td>-3.62 n.s.</td>
<td>0.16803 1.66 n.s.</td>
<td>0.1628 1.61 n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>13551.25</td>
<td>-60.88 0.0001</td>
<td>-0.73209 -7.23 0.0001</td>
<td>-0.60558 -5.98 0.0001</td>
</tr>
</tbody>
</table>

Table 3.6: Experiment 1, Color-contrast sequence from EXPERIMENT A (see Figure 3.11): Target fixation analysis comparing the felicitous condition (model) with the neutral and the infelicitous conditions.
Table 3.7: Experiment 1, Color-contrast sequence from EXPERIMENT A (see Figure 3.11): Target fixation analysis comparing the neutral and infelicitous conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8555.2</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>8555.19</td>
<td>-0.01</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>8554.78</td>
<td>-0.41</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>8554.78</td>
<td>-0.00</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>8553.06</td>
<td>-1.72</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Data from EXPERIMENT B

Figure 3.6 presents an overview of the data from EXPERIMENT B. It shows looks to the target object cell in the six critical conditions of EXPERIMENT B (see Tables 3.2 and 3.4) and shows a pattern very similar to that found in EXPERIMENT A. Looks to the target object started to rise earliest in the felicitous condition, and looks to the target object occurred earlier in the color-contrast compared to no-contrast sequences.

Figure 3.7 shows the grand mean of looks to the target object for the six critical conditions of EXPERIMENT B. The curve first starts rising from -289 ms, peaks from 765 ms to 833 ms, and then starts falling. The analysis window for the following GCA ranges from -289 ms to 833 ms. -289 ms was chosen as the beginning of the analysis window because it represents the first sign of an upward trend in the curve. 833 ms (rather than 765 ms) was chosen as the end of the analysis window because the curve shows a clear downward trend immediately after 833 ms, but not after 765 ms.

The results from an omnibus GCA are shown in Table 3.8. The GCA reveals results remarkably similar to those that were found in EXPERIMENT A. There is a main effect of Lexical Contrast, a main effect of Pitch Accent, and a Lexical Contrast / Pitch Accent interaction on the intercept term, but no reliable effects on the linear term. These results confirm that looks to the target occurred earlier in the color-contrast sequences compared to the no-contrast sequences (main effect of Lexical Contrast on the intercept). In addition, looks to the target started rising earlier if the target utterance was pronounced with a L+H* on the
Figure 3.6: Experiment 1, data from EXPERIMENT B: Proportion of looks to the target object cell in the six critical conditions presented in Tables 3.2 and 3.4.
Figure 3.7: Experiment 1, data from EXPERIMENT B: Grand mean of looks to the target object cell in the six critical conditions presented in Tables 3.2 and 3.4. The arrows indicate the beginning and end of the analysis window for the GCA.

an adjective compared to a H* !H* pattern on the NP or a L+H* accent on the noun (main effect of Pitch Accent and Lexical Contrast / Pitch Accent interaction on the intercept). Finally, none of the conditions differed in the slope of the curve (no reliable effects on the linear term).

Figures 3.8 through 3.10 show direct comparisons of the three color-contrast conditions, analogous to Figures 3.3 through 3.5 for EXPERIMENT A. Figures 3.8 and 3.9 again show an advantage for looks to the target object in the felicitous condition over both the neutral and infelicitous conditions. Compared to the neutral condition, the advantage is reliable from -51 ms to 306 ms. Compared to the infelicitous condition, the advantage is reliable from -170 ms to 374 ms. Again, the advantage occurred before noun onset and was therefore based solely on auditory information from the adjective PURPLE. Figure 3.10 shows no advantage between the neutral and infelicitous conditions. Thus, as in EXPERIMENT A, the felicitous condition presented a processing advantage compared to the neutral and infelicitous conditions.
Table 3.8: Experiment 1, data from EXPERIMENT B (see Figure 3.6): Target fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.

<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Model fit</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
<td>-2LL</td>
</tr>
<tr>
<td>Base</td>
<td>31083.32</td>
<td>—</td>
<td>—</td>
<td>30513.85</td>
</tr>
<tr>
<td>Intercept</td>
<td>31041.76</td>
<td>-41.56</td>
<td>0.0001</td>
<td>30505.54</td>
</tr>
<tr>
<td>Linear</td>
<td>31041.15</td>
<td>-0.61</td>
<td>n.s.</td>
<td>30505.19</td>
</tr>
<tr>
<td>Quadratic</td>
<td>30522.50</td>
<td>-518.65</td>
<td>0.0001</td>
<td>30470.96</td>
</tr>
<tr>
<td>Cubic</td>
<td>30513.85</td>
<td>-8.66</td>
<td>0.01</td>
<td>30451.87</td>
</tr>
</tbody>
</table>
Figure 3.8: Experiment 1, data from EXPERIMENT B: Comparison of the felicitous and neutral conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Again, the infelicitous condition did not slow down resolving reference compared to the neutral condition.

A GCA comparing the felicitous condition to the other two conditions is shown in Table 3.9. Again, of interest are the intercept and linear terms. The table shows a reliable effect on the intercept term comparing the felicitous and neutral as well as the felicitous and infelicitous conditions. This confirms the processing advantage for the felicitous condition compared to both the neutral and infelicitous conditions. The absence of any reliable effect on the linear term shows that there was no difference in slope between the felicitous condition and the other two conditions. Notice that the reliable effects on the quadratic and cubic terms again reflect differences in the asymptotic tails of the curves (see the right panel in Figures 3.8 and 3.9) rather than differences related to cognitive processes.

Table 3.10 shows results from a GCA comparing the neutral and infelicitous conditions and reveals no reliable differences between the two conditions. This confirms that pro-
Figure 3.9: Experiment 1, data from EXPERIMENT B: Comparison of the felicitous and infelicitous conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Figure 3.10: Experiment 1, data from EXPERIMENT B: Comparison of the neutral and infelicitous conditions presented in Table 3.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.
<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>( \Delta D )</th>
<th>( p &lt; )</th>
<th>Parameter estimates</th>
<th>( p &lt; )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>14255.47</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>14238.68</td>
<td>—16.79</td>
<td>0.001</td>
<td>—0.24255</td>
<td>-2.65</td>
</tr>
<tr>
<td>Linear</td>
<td>14238.49</td>
<td>-0.19</td>
<td>n.s.</td>
<td>0.19715</td>
<td>0.4</td>
</tr>
<tr>
<td>Quadratic</td>
<td>14120.33</td>
<td>-118.16</td>
<td>0.0001</td>
<td>0.85136</td>
<td>8.68</td>
</tr>
<tr>
<td>Cubic</td>
<td>14085.36</td>
<td>-34.97</td>
<td>0.0001</td>
<td>-0.51907</td>
<td>-5.29</td>
</tr>
</tbody>
</table>

Table 3.9: Experiment 1, Color-contrast sequence from EXPERIMENT B (see Figure 3.16): Target fixation analysis comparing the felicitous condition (model) with the neutral and infelicitous conditions.
Table 3.10: Experiment 1, Color-contrast sequence from EXPERIMENT B (see Figure 3.16): Target fixation analysis comparing the neutral and infelicitous conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>9206.51</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>9204.49</td>
<td>-2.02</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>9204.34</td>
<td>-0.15</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>9202.72</td>
<td>-1.62</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>9202.49</td>
<td>-0.23</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

nouncing the target utterance with an infelicitous L+H* accent on the noun did not delay reference resolution compared to a neutral H* !H* pattern on the target utterance.

### 3.2.2 Color-Competitor Fixations

**Data from EXPERIMENT A**

Figure 3.11 provides an overview of the data from EXPERIMENT A. It shows looks to the color-competitor object cell in the six critical conditions of EXPERIMENT A (see Tables 3.1 and 3.3). The figure shows that in all conditions looks to the color-competitor object, the purple pencil, started rising before lexical information from bottle or BOTTLE was heard. In addition, the no-contrast sequences elicited larger garden-path effects than the color-contrast sequences. Numerically the garden-path condition elicited the largest garden-path effect.

Figure 3.12 shows the grand mean of looks to the color competitor object for the six critical conditions. The figure shows that looks to the color competitor object start rising at -238 ms. This point was chosen as the beginning of the analysis window because it is the last data point which is level with the one preceding it. 238 ms was chosen as the peak of the curve because it approximately cuts the mountain of the curve in half. To create an analysis window that is symmetrical around the peak, 714 ms was calculated as the end of the window.
Figure 3.11: Experiment 1, data from EXPERIMENT A: Proportion of looks to the color-competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3.
Figure 3.12: Experiment 1, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3. The downward arrows indicate the beginning and end of the analysis window for the GCA. The upward arrow represents the middle of the mountain of the curve.
The results from an omnibus GCA with an analysis window ranging from -238 ms to 714 ms are shown in Table 3.11. Of most interest for fixations to a competitor object are the intercept and quadratic terms (see Section 2.10). The table shows a main effect of *Lexical Contrast* on both the intercept and quadratic terms. This combination of effects suggests a larger up and down excursion in the no-contrast conditions over the color-contrast conditions. In other words, the no-contrast conditions elicited a larger garden-path effect than the color-contrast conditions, regardless of how the target utterance was pronounced. Section 3.2.1 showed that the no-contrast sequences also elicited later looks to the target object than the color-contrast sequences. Together these results suggest that listeners expected the noun of the first instruction to be repeated in the target instruction regardless of the prosodic pattern the target instruction carried.

The table also shows a main effect of *Pitch Accent* on the quadratic term, but not the intercept term. This suggests a difference in the slope of the rise and fall of the curves without a difference in area under the curve. This means that those prosodic conditions that have a larger up-down excursion (i.e. steeper slope) also have slightly lower tails, resulting in a total area under the curve similar to those conditions with less steep slopes. There is also a *Lexical Contrast / Pitch Accent* interaction on the quadratic term, but not the intercept term. This suggests that *Lexical Contrast* and *Pitch Accent* interact to affect the slope of the curves, but not the area under the curve. A reliable effect on the quadratic term, but not the intercept term, is more difficult to interpret than a reliable effect on both terms. In particular, there may not be any meaningful cognitive correlates to this pattern of results.

Figures 3.13 through 3.15 show direct comparisons of the three no-contrast conditions. Only the no-contrast conditions will be compared since a difference in looks to the competitor object was predicted for the no-contrast condition, not the color-contrast conditions. Again, the graphs on the left show data points with error bars and the graphs on the right show model fits for the curves. Figure 3.13 shows that the curve in the garden-path condi-
<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>27267.09</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>27233.99</td>
<td>−33.1</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>27231.41</td>
<td>−2.58</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>26877.51</td>
<td>−353.91</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>26790.67</td>
<td>−86.84</td>
<td>0.0001</td>
</tr>
<tr>
<td>Quartic</td>
<td>26779.12</td>
<td>−11.55</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 3.11: Experiment 1, data from EXPERIMENT A (see Figure 3.11): Competitor fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
Figure 3.13: Experiment 1, data from EXPERIMENT A: Comparison of the garden-path and neutral conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits.

The model fit for the garden-path condition has a higher peak than the model fit for the neutral condition. This suggests that both conditions elicit equally frequent looks to the color-competitor object. Thus, contrary to predictions, the garden-path condition does not elicit a larger garden-path effect than the neutral condition. Figure 3.14 shows that there is a small window from 323 ms to 357 ms for which there are reliably more looks to the color-competitor in the garden-path condition than in the infelicitous condition. In addition, the model fit for the garden-path condition has a higher peak than the model fit for the infelicitous condition. However, across most of the time window, the two lines are not reliably different. Figure 3.15 shows no differences around the peak of the curve between the neutral and infelicitous conditions.

Results from a GCA comparing the garden-path condition to the other two conditions are shown in Table 3.12. Again, of interest are the intercept and quadratic terms. The table
Figure 3.14: Experiment 1, data from EXPERIMENT A: Comparison of the garden-path and infelicitous conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits.

Figure 3.15: Experiment 1, data from EXPERIMENT A: Comparison of the neutral and infelicitous conditions presented in Table 3.3. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits.
shows no reliable effects on the intercept term, revealing no differences in area under the
curve. There is a reliable effect on the quadratic term comparing the garden-path condition
to the neutral condition and to the infelicitous condition. Thus, the garden-path condition has
a steeper slope than the other two conditions, but this did not lead to reliably more looks
to the color-competitor object over a considerable time window. Table 3.13 shows results
from a GCA comparing the neutral and infelicitous conditions. The analysis also reveals no
reliable differences on the intercept term, but a reliable effect on the quadratic term. Thus
the neutral condition has a steeper rise-fall rate, but an area under the curve that is compa-
rable to that of the infelicitous condition. The model fits in Figures 3.13 through 3.15 reveal
that the differences in slope between the conditions seem to be related to how quickly par-
ticipants looked away from the competitor object after lexical information from the noun
had come in. Altogether, the results suggest that all conditions elicited an equally large
garden-path effect. Contrary to prediction, it is not the case that the garden-path condition
elicited a larger garden-path effect than the neutral or infelicitous conditions.

**Data from Experiment B**

Figure 3.16 shows an overview of the data from Experiment B. It shows looks to the color-
competitor object cell in the six critical conditions of Experiment B (see Tables 3.2 and 3.4).
Overall, the figure shows the same trend as Experiment A. Looks to the color-competitor
started rising before noun onset, and the garden-path condition has the numerically largest
up and down excursion.

Figure 3.17 shows the grand mean of looks to the color competitor object for the six
critical conditions. The figure shows that looks to the color competitor object start rising at
-238 ms. This point was chosen because it is the last data point which is level with the one
preceding it. 187 ms was chosen as the peak of the curve because it approximately cuts the
mountain of the curve in half. In addition, it is actually the highest data point. The end of
the analysis window was calculated to be 612 ms.
<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Parameter estimates</th>
<th>Parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>12832.35</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>12831.30</td>
<td>-1.06</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>12828.17</td>
<td>-3.12</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>12800.41</td>
<td>-27.76</td>
<td>0.001</td>
</tr>
<tr>
<td>Cubic</td>
<td>12788</td>
<td>-12.42</td>
<td>0.01</td>
</tr>
<tr>
<td>Quartic</td>
<td>12775.02</td>
<td>-12.98</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 3.12: Experiment 1, No-contrast sequence from EXPERIMENT A (see Figure 3.11): Competitor fixation analysis comparing the garden-path condition (model) with the neutral and the infelicitous conditions.
Table 3.13: Experiment 1, No-contrast sequence from EXPERIMENT A (see Figure 3.11): Competitor fixation analysis comparing the neutral and infelicitous conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>7633.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>7632.85</td>
<td>-0.58</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>7629.94</td>
<td>-2.91</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>7622.58</td>
<td>-7.37</td>
<td>0.01</td>
</tr>
<tr>
<td>Cubic</td>
<td>7613.43</td>
<td>-9.14</td>
<td>0.01</td>
</tr>
<tr>
<td>Quartic</td>
<td>7612.81</td>
<td>-0.62</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Figure 3.16: Experiment 1, data from EXPERIMENT B: Proportion of looks to the color competitor object cell in the six critical conditions presented in Tables 3.2 and 3.4.
Figure 3.17: Experiment 1, data from Experiment A: Grand mean of looks to the color competitor object cell in the six critical conditions presented in Tables 3.1 and 3.3. The downward arrows indicate the beginning and end of the analysis window for the GCA. The upward arrow represents the middle of the mountain of the curve.
Table 3.14 shows results from an omnibus GCA with an analysis window from -238 ms to 612 ms. Again, there is a main effect of *Lexical Contrast* on both the intercept and quadratic terms. This suggests a steeper rise and fall rate and larger area under the curve in the no-contrast sequences compared to the color-contrast sequences. Thus, the no-contrast sequences, as expected, showed a larger garden-path effect than the color-contrast sequences. The table also shows a main effect of *Lexical Contrast* on the linear and cubic terms, suggesting that there are asymmetric differences in the rise-fall pattern of the curves across conditions. Table 3.14 also shows a main effect of *Pitch Accent* on the quadratic term without a reliable effect on the intercept term, suggesting a slope difference without a difference in area under the curve. Notice there is also a main effect of *Pitch Accent* on the linear and cubic terms, again suggesting asymmetric differences in rise-fall pattern between curves. It looks like the rise and fall of curve for the two conditions with a L+H* on the adjective of the target utterance (the *felicitous* and *garden-path* conditions) occurs slightly earlier in time than for the other conditions. Finally, there is a *Lexical Contrast / Pitch Accent* interaction on the quadratic term, but not the intercept term, revealing that *Lexical Contrast* and *Pitch Accent* interact to affect the rise and fall of the slope, but not the area under the curve.

Figures 3.18 through 3.20 show direct comparisons of the three no-contrast conditions for EXPERIMENT B. Figure 3.18 reveals a small window from 204 ms to 306 ms with reliably more looks to the color competitor in the *garden-path* condition than in the *neutral* condition. The model fit also shows a curve with a higher peak for the *garden-path* condition than the *neutral* condition. Figure 3.19 shows reliably more looks to the color-competitor in the *garden-path* compared to the *infelicitous* condition from -85 ms to 340 ms. This clear effect is also reflected in the model fit: The peak of the curve for the *garden-path* condition is much higher than that for the *infelicitous* condition. This means that, as predicted, the *garden-path* condition elicited a larger garden-path effect than the *infelicitous* condition.

Since the curves differ reliably before lexical information from the noun was heard, the
<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>23239.99</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>23211.99</td>
<td>-28</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>23203.61</td>
<td>-8.38</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>23017.88</td>
<td>-185.73</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>22965.93</td>
<td>-51.94</td>
<td>0.0001</td>
</tr>
<tr>
<td>Quartic</td>
<td>22954.53</td>
<td>-11.4</td>
<td>0.001</td>
</tr>
</tbody>
</table>

Table 3.14: Experiment 1, data from Experiment B (see Figure 3.1): Competitor fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
larger garden-path effect in the garden-path condition compared to the infelicitous condition must be a result of having heard a L+H* accent compared to no pitch accent on the adjective of the target instruction. Figure 3.20 shows no differences around the peak of the curves between the neutral and infelicitous conditions. However, there is an early advantage, from -68 ms to 85 ms, for looks to the color-competitor object in the neutral condition compared to the infelicitous condition. Again, this difference occurs before noun onset and in response to having heard a H* accent compared to no pitch accent on the adjective.

Table 3.15 shows results from a GCA comparing the garden-path condition to the other two conditions. The table confirms the patterns seen in Figures 3.18 through 3.20. It shows a reliable effect on the quadratic term comparing the garden-path and neutral conditions, suggesting a higher rise and fall rate in the garden-path condition, but no difference in area under the curve. Overall, the data suggest that the garden-path condition did not elicit a larger garden-path effect than the neutral condition. The curve for the garden-path condition
Figure 3.19: Experiment 1, data from EXPERIMENT B: Comparison of the garden-path and infelicitous conditions presented in Table 3.4. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits.

Figure 3.20: Experiment 1, data from EXPERIMENT B: Comparison of the neutral and infelicitous conditions presented in Table 3.4. Left: Empirical logit of looks to the color competitor object cell and error bars. Right: Model fits.
is reliably higher than that for the neutral condition only for a small time window, and the GCA reveals no difference in area under the curve for the two conditions. Table 3.15 also reveals a reliable effect on both the intercept and quadratic terms for the garden-path condition compared to the infelicitous condition. This pattern of results confirms that there is a larger garden-path effect in the garden-path compared to the infelicitous condition.

Table 3.16 shows results from a GCA comparing the neutral and infelicitous conditions. The analysis reveals reliable differences on both the intercept and quadratic terms, suggesting a larger garden-path effect in the neutral compared to the infelicitous condition. Figure 3.20 shows that the two conditions also differ in that the curve for the neutral condition starts rising earlier than that for the infelicitous condition. This is also reflected in the GCA, which reveals reliable effects on the linear and cubic terms between the two conditions in addition to the effects on the intercept and quadratic terms. The linear and cubic terms reflect asymmetric differences in rise-fall pattern between conditions.

### 3.3 Discussion

The results presented above revealed some general trends. Data from both large experiments confirmed the facilitation effect found in Ito & Speer (2008) and Weber et al. (2006a): Looks to the target object (the purple bottle) in a color-contrast sequence (yellow bottle ⇒ purple bottle) rose reliably earlier if purple was produced with a prominent L+H* accent than if it was produced either with a less prominent H* accent or without a pitch accent. This early rise in looks occurred before lexical information from the noun was available, and was thus purely based on information from the adjective. This confirms that listeners used prosodic information on the adjective to anticipate an upcoming referent\(^1\). In partic-

---

\(^1\)The current experiments can not distinguish among several descriptions of the entity anticipated by the listener, which might be the visual form of the referent, a noun phrase that specifies the referent, the speaker’s intention to produce a particular noun phrase, or the speaker’s intention to refer to a particular object. This is an open research question.

126
<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Parameter estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
</tr>
<tr>
<td>Base</td>
<td>11247.64</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>11238.98</td>
<td>-8.66</td>
</tr>
<tr>
<td>Linear</td>
<td>11232.92</td>
<td>-6.06</td>
</tr>
<tr>
<td>Quadratic</td>
<td>11061.90</td>
<td>-171.02</td>
</tr>
<tr>
<td>Cubic</td>
<td>11045.53</td>
<td>-16.38</td>
</tr>
<tr>
<td>Quartic</td>
<td>11037.37</td>
<td>-8.16</td>
</tr>
</tbody>
</table>

Table 3.15: Experiment 1, No-contrast sequence from EXPERIMENT B (see Figure 3.16): Competitor fixation analysis comparing the garden-path condition (model) with the neutral and infelicitous conditions.
Table 3.16: Experiment 1, No-contrast sequence from EXPERIMENT B (see Figure 3.16): Competitor fixation analysis comparing the neutral and the infelicitous conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>6805.34</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>6801.25</td>
<td>−4.09</td>
<td>0.05</td>
</tr>
<tr>
<td>Linear</td>
<td>6793.34</td>
<td>−7.9</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>6749.99</td>
<td>−43.36</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>6735.17</td>
<td>−14.82</td>
<td>0.0001</td>
</tr>
<tr>
<td>Quartic</td>
<td>6735.03</td>
<td>−0.13</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

ular, a L+H* accent on the adjective evoked a contrast set (in this case of bottles of different colors) such that listeners expected a repeated noun.

The data for looks to the color-competitor object (the purple pencil) in a no-contrast sequence (*yellow scissors ⇒ purple bottle*) are less clear. Numerically, we see the expected pattern: There are more looks to the competitor object if purple is pronounced with a L+H* accent than if it is produced with a H* accent or no pitch accent. However, in EXPERIMENT A, the garden-path condition does not elicit a larger garden-path effect than the neutral or infelicitous conditions. The differences between conditions occur mainly in the rise-fall rate of the curve, not in the area under the curve. EXPERIMENT B, on the other hand, does reveal a more expected pattern of results. In particular, the garden-path condition elicited a larger garden-path effect than the infelicitous condition. In addition, looks to the color-competitor object seemed to start rising later in the infelicitous condition compared to the garden-path or neutral conditions. This suggests that the prosody of the target utterance does affect the size of the garden-path effect as expected.

Notice that the data coming from EXPERIMENT A showed smaller effects overall than the data coming from EXPERIMENT B. That is, relevant curves from EXPERIMENT A tended to diverge less than those from EXPERIMENT B. This may be a result of the faster rate of speech for the auditory stimuli used in EXPERIMENT A compared to EXPERIMENT B (see Section 2.4.5). However, even responses to the faster stimuli yielded reliable differences,
confirming the robust nature of intonational effects on eye movements as a dependent variable.

The stimuli duration difference also seemed to affect when looks to the object of interest started to rise. Data from EXPERIMENT A, for example, showed that looks to the target object in the felicitous condition started rising about 150 ms before noun onset. In EXPERIMENT B, looks to the target in the same condition started rising just a little after 250 ms before noun onset. A similar pattern can be found for looks to the color-competitor object: Looks in the garden-path condition started rising about 50 ms before noun onset in EXPERIMENT A and about 150 ms before noun onset in EXPERIMENT B. The longer duration of the adjective in EXPERIMENT B may have given listeners more time to process lexical and prosodic information from the adjective and plan and execute anticipatory eye movements. Notice also that the earlier rise in looks in EXPERIMENT B compared to EXPERIMENT A occurred across prosodic conditions and was therefore not related to the prosody on the adjective.

In conclusion, even though the results from EXPERIMENTS A and B differ in various details from each other as well as from those of Weber et al. (2006a) and Ito & Speer (2008), they do confirm the major effects found in these studies: Looks to the target object were reliably earlier when the adjective in a color-contrast sequence was produced with a L+H* accent compared to either a H* or no accent. Looks to the color-competitor object revealed either a numerically or reliably higher curve when the adjective in a no-contrast sequence was pronounced with a L+H* accent than when it was pronounced with either a H* or no accent. The finding of patterns highly similar to the previous results suggests that the current study tapped into the same processing mechanisms as earlier studies and confirms that listeners processed prosody immediately and used it to make predictions about upcoming referents.
Chapter 4

Experiment 2

The pattern of effects found in Experiment 1 closely followed the basic results found previously by Ito & Speer (2008) and Weber et al. (2006a). These results show that prosodic information is used as soon as it becomes available to the processing system and that it can be used to make predictions about upcoming sentence material. The experiment confirmed that in a two-instruction sequence like *Click on the yellow bottle. Click on the purple bottle*, the prosodic contour of the second (i.e. target) instruction affects how quickly listeners can identify the speaker’s intended referent. Earlier looks were found if the target instruction was pronounced with a L+H* accent on the adjective (as in *Click on the PURPLE bottle*), compared to a L+H* on the noun (as in *Click on the purple BOTTLE*) or a H* accent on the adjective and a !H* accent on the noun (as in *Click on the purple bottle*). This result suggests that the L+H* on the adjective (PURPLE) evoked a contrast set of bottles of different colors, such that listeners expected a repetition of the previous noun bottle. The result crucially depends on listeners, at a minimum, keeping in memory a representation of the previous noun (that is, the information that the previous utterance referred to a bottle) as well as making immediate use of the prosodic information on PURPLE.

Rather, listeners likely maintain the denotation (i.e. the set of applicable entities) of both adjectives and nouns in the ongoing discourse representation, along with aspects of their spoken form, including prosodic information. What the previous experiments do not establish is whether listeners maintain prosodic information, in addition to color and object information, from preceding utterances. It is possible that prosodic aspects of the representations decay faster than other aspects. In this case, the prosodic information from
Instruction 2 might be processed with access to color and object, but not prosodic information from Instruction 1. Weber et al.’s (2006a) experiments suggest that this is not the case. They found that the prosody of the first instruction affected looks to the target object if the target instruction was pronounced with a L+H* accent on the noun. If the first instruction did not contain a L+H* accent, there were earlier looks to a contrastive target object (Click on the pink scissors. Click on the red SCISSORS) compared to a target object that contrasted in neither adjective or noun (Click on the pink scissors. Click on the red VASE). If, however, the first instruction contained a L+H* accent on the adjective, there was no difference in looks to contrastive targets (Click on the PINK scissors. Click on the red SCISSORS) compared to no-contrast targets (Click on the PINK scissors. Click on the red VASE). Weber et al. (2006a) proposed that the L+H* accent in the first instruction highlighted the contrast set, and listeners expected the adjective in the second instruction to also carry a L+H* accent. When the second adjective was unaccented, listeners no longer preferred the contrastive target. When listeners’ expectations of a L+H* accent on the second adjective were confirmed (Click on the PINK scissors. Click on the RED ...), they fixated the contrastive target as soon as information from the second adjective arrived. This result suggests that listeners also kept a representation of the focus structure (i.e. their interpretation of the prosodic structure) of the preceding discourse in working memory.

According to Weber et al.’s (2006a) interpretation of the data, the first instruction triggered the expectation of a L+H* accent on the second adjective by highlighting the contrast set scissors of different colors. This interpretation matched their findings: Looks to the target object were only affected by the presence or absence of a L+H* accent in the first instruction if the second instruction did not carry the expected L+H* accent. If the second instruction carried the expected L+H* accent, the prosody of the first instruction did not affect looks to the target object. In other words, only if expectations were broken, were looks to the target object affected. If the L+H* accent of the first instruction did highlight a contrast set of different-colored scissors, it is surprising that this triggered the expecta-
tion of an upcoming L+H* accent on the adjective, but not of an upcoming repeated noun. If "highlighting" the contrast set led listeners to also expect a repeated noun, a different pattern of looks would have been expected. In particular, if the adjective of the second instruction carried the expected L+H* accent, looks to a **contrastive** target object would have been expected earlier when the first instruction also highlighted the contrast set than when it did not, especially with only one object in the display which shared the object type with that of the first instruction. It is possible that Weber et al. (2006a) did not find even earlier looks for two successive L+H* accents on the adjective compared to a L+H* accent on the target adjective only in a **contrastive** sequence because an overall bias for the **contrastive** target object created a ceiling effect. A bias towards the **contrastive** target object could have been created by the visual display used in the experiments. Each target display contained four objects: a target object (e.g. *red scissors*) sharing its noun with one object (e.g. *pink scissors*) and its adjective with another (e.g. *a red vase*) as well as an unrelated object (e.g. *a green clock*). Notice that in this case only the scissors needed to be modified by a color adjective in order to be uniquely identified. The red vase and green clock can be uniquely identified when referred to as *the vase* and *the clock*. The expression *the scissors*, on the other hand, is infelicitous in this context because the discourse context contains two pairs of scissors, a pink and a red pair. It is thus possible that the use of color adjectives in the instructions created an overall bias for the **contrastive** target objects, which may have obscured differences arising from the presence or absence of a L+H* in the first instruction.

Experiment 2 includes conditions similar to those from Weber et al.'s (2006a) experiments, but uses a display where all mentioned objects require a description containing both color and object terms in order to be uniquely identified. The experiment investigates whether earlier looks to a **contrastive** target object can be found with this kind of display if both adjectives in a color-contrast sequence carry a L+H* accent than if only the second adjective carries a L+H* accent. Such a result would suggest that the L+H* accent on the target adjective of the first instruction highlighted an upcoming color-contrast which led
listeners to expect not only the adjective of the second instruction to carry a L+H* accent, but also the noun of the second instruction to be identical to that of the first.

Thus, Experiment 2 asks the following question: Does a L+H* accent on the adjective of the first instruction highlight an upcoming color contrast such that listeners expect the noun of the second instruction to be identical to that of the first instruction? If the L+H* accent on the adjective of the first instruction leads to the expectation of a repeated noun, earlier eye movements to the target object are expected in a color-contrast sequence if both instructions carry a L+H* accent on the adjective compared to only the second instruction carrying a L+H* accent in this position. By the same token, a larger garden-path effect for looks to the color-competitor object is expected in a no-contrast sequence if both instructions compared to just the target instruction carry a L+H* accent. In other words, the L+H* accent of both the first and the second instruction are predicted to carry the expectation of a repeated noun. If both instructions carry a L+H* accent this expectation is created earlier and looks to the color-contrast object should occur earlier than if only the second instruction carries a L+H* accent. If the L+H* accent in the first instruction does not lead listeners to expect a repeated noun, trials with L+H* accents on two successive adjectives compared to only a L+H* accent on the adjective of the target instruction should yield no differences in looks to the target or color-competitor objects.

4.1 Experimental Conditions

Tables 4.1 through 4.4 show the critical conditions for Experiment 2. Again, data from both Experiment A and Experiment B are presented. Tables 4.1 and 4.2 show color-contrast sequences, where the target object was of the same type as the previously mentioned object. Table 4.1 shows sequences from Experiment A, and Table 4.2 shows sequences from Experiment B.
The *felicitous* condition in Tables 4.1 and 4.2 also occurred in Experiment 1. Here, the first instruction (i.e. the context instruction) was pronounced with a H* !H* sequence on the NP (HHNP), yielding a neutral, out-of-the-blue, instruction. The second instruction (i.e. the target instruction, shown in bold face) was pronounced with a L+H* accent on the adjective (LHA), evoking a color-contrast with the previous instruction. Previous work suggests that once listeners heard the L+H*-accented color term of the target instruction, they should have expected the upcoming noun to be repeated. Since the noun was, indeed, repeated, the condition is called *felicitous*. Experiment 1 showed earlier looks to the target object in this condition compared to conditions with either a L+H* on the noun (LHN) or a H* !H* pattern on the NP (HHNP) of the target instruction.

The context instruction of the *anticipatory* condition Tables 4.1 and 4.2 was pronounced with a L+H* accent on the adjective (LHA). This was predicted to evoke a color-contrast with the upcoming instruction. In particular, the L+H* in the first instruction should have led listeners to expect a L+H* on the adjective of the target instruction and a repetition of the noun from the context instruction in the target instruction. The target instruction confirms both expectations: The adjective carries a L+H* accent and the noun is repeated. Since this expectation is assumed to be created earlier than in the *felicitous* condition, the condition is named *anticipatory*. Notice, however, that both the *felicitous* and *anticipatory* conditions are expected to involve anticipation of the repeated noun.

Notice also that up to and including the target adjective, there is no difference between the prosodic structures of the *felicitous* and *anticipatory* conditions taken from EXPERIMENTS A and B. The target instruction, however, differs in the prosody of the noun. It carries L-L% sentence-final intonation in the conditions from EXPERIMENT A, but a L-H% continuation rise in the conditions from EXPERIMENT B. This difference occurs after the critical parts of the target instruction, in particular, after the L+H*-accented target adjective is heard, and is therefore not expected to influence looks to the target instruction.

Tables 4.3 and 4.4 show no-contrast sequences, where the target object shares neither
<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Context</td>
<td>Color Contrast</td>
<td>Object Contrast</td>
</tr>
<tr>
<td>felicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the PURPLE bottle ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
<tr>
<td>anticipatory</td>
<td>Click on the YELLOW bottle ↘</td>
<td>Click on the PURPLE bottle ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
</tbody>
</table>

Table 4.1: Experiment 2, Color-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘)

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Context</td>
<td>Color Contrast</td>
<td>Object Contrast</td>
</tr>
<tr>
<td>felicitous</td>
<td>Click on the yellow bottle ↘</td>
<td>Click on the PURPLE bottle ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
<tr>
<td>anticipatory</td>
<td>Click on the YELLOW bottle ↘</td>
<td>Click on the PURPLE bottle ↘</td>
<td>Click on the purple PENCIL ↘</td>
</tr>
</tbody>
</table>

Table 4.2: Experiment 2, Color-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; *italics* on an adjective indicate H* accents; *italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The gray instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗)
color nor type with the previously mentioned object. Table 4.3 shows sequences from Experiment A, and Table 4.4 shows sequences from Experiment B.

The garden-path condition from Tables 4.3 and 4.4 also occurred in Experiment 1. The prosodic structure of the context and target instructions of the felicitous and garden-path conditions is identical: The context instruction is pronounced with a neutral H* !H* sequence on the NP (HHNP), and the target instruction is pronounced with a L+H* accent on the adjective (LHA). As in the felicitous condition, the L+H* on the adjective of the target instruction should evoke a color-contrast with the previous instruction and create the expectation of a repeated noun. However, the noun in the garden-path condition is novel. Upon hearing the L+H* accent on the adjective of the target instruction, listeners should be “garden-pathed” – they are predicted to look at the object sharing its type with that of the previously mentioned object. Only once lexical information from the novel noun is heard, should listeners recover from the garden-path and look to the mentioned noun.

The prosodic structure of the context and target instructions of the anticipatory and anticipated garden-path conditions is also identical: Both the context and target instructions have a L+H*-accented adjective (LHA). Again, the first L+H* accent (i.e. that of the context instruction) should create the expectation of a L+H* accent on the adjective of the target instruction and of a repeated noun. As in the anticipatory condition, the expectation of a L+H*-accented adjective is confirmed in the anticipated garden-path condition. But unlike the anticipatory condition, the target instruction of the anticipated garden-path condition has a novel noun and should garden-path listeners. Thus, both the garden-path and anticipated garden-path conditions are predicted to yield a garden-path effect. The anticipated garden-path condition receives this name because the expectations that led listeners up the garden-path are predicted to occur earlier than in the garden-path condition. In addition, if the expectation of a repeated noun comes from two cues in the anticipated garden-path condition (both L+H* accents), but only one cue (the L+H* accent of the target instruction) in the garden-path condition, then the anticipated garden-path condition should cause
Table 4.3: Experiment 2, No-contrast sequence from EXPERIMENT A: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The *gray* instruction is not relevant for this experiment. All Instructions end in a L-L% pattern (indicated by a ↘).

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>garden-path</td>
<td>Context</td>
<td><strong>Click on the</strong></td>
<td><strong>Click on the</strong></td>
</tr>
<tr>
<td></td>
<td>green pencil ↘</td>
<td><strong>PURPLE bottle ↘</strong></td>
<td>purple PENCIL ↘</td>
</tr>
<tr>
<td>anticipated garden-path</td>
<td>Click on the</td>
<td><strong>Click on the</strong></td>
<td><strong>Click on the</strong></td>
</tr>
<tr>
<td></td>
<td>GREEN pencil ↘</td>
<td><strong>PURPLE bottle ↗</strong></td>
<td>purple PENCIL ↘</td>
</tr>
</tbody>
</table>

Table 4.4: Experiment 2, No-contrast sequence from EXPERIMENT B: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. The *gray* instruction is not relevant for this experiment. Instructions 1 and 3 end in a L-L% pattern (indicated by a ↘), whereas Instruction 2 ends in a L-H% pattern (indicated by a ↗).

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1:</th>
<th>Instruction 2:</th>
<th>Instruction 3:</th>
</tr>
</thead>
<tbody>
<tr>
<td>garden-path</td>
<td>Context</td>
<td><strong>Click on the</strong></td>
<td><strong>Click on the</strong></td>
</tr>
<tr>
<td></td>
<td>green pencil ↘</td>
<td><strong>PURPLE bottle ↗</strong></td>
<td>purple PENCIL ↗</td>
</tr>
<tr>
<td>anticipated garden-path</td>
<td>Click on the</td>
<td><strong>Click on the</strong></td>
<td><strong>Click on the</strong></td>
</tr>
<tr>
<td></td>
<td>GREEN pencil ↘</td>
<td><strong>PURPLE bottle ↗</strong></td>
<td>purple PENCIL ↗</td>
</tr>
</tbody>
</table>

a larger garden-path effect than the garden-path condition. The garden-path and anticipated garden-path conditions from EXPERIMENTS A and B also differ in the edge tone patterns of the target instructions, which end in L-L% sentence-final intonation in EXPERIMENT A and a L-H% continuation rise in EXPERIMENT B. These prosodic differences again occur after the critical parts of the target instruction and are thus not expected to affect the results of the experiment.

Experiment 2 thus has a 2 x 2 design, with the two factors **Lexical Contrast** and **Pitch Accent**. The factor **Lexical Contrast** has the levels color-contrast sequence and no-contrast sequence, and the factor **Pitch Accent** has the levels L+H* accent on both adjectives and L+H* accent on target adjective only.
4.2 Results

4.2.1 Target Fixations

Data from Experiment A

Figure 4.1 provides an overview of the results from Experiment A. It shows looks to the target object cell in the four critical conditions presented in Tables 4.1 and 4.3. The figure shows earlier looks to the target object in color-contrast sequences compared with no-contrast ones. However, contrary to the prediction, looks to the target object in the anticipatory condition do not rise earlier than in the felicitous condition. Rather, looks to the target object start rising slightly later in the anticipatory than the felicitous condition.
Figure 4.2 shows the grand mean of looks to the target object for the four critical conditions. The figure shows that looks to the target object start rising at -221 ms, which is chosen as the beginning of the analysis window. 884 ms was chosen as the peak of the curve because the curve clearly falls after 884 ms. In addition, 884 ms is actually the highest data point.

The results from an omnibus GCA over an analysis window from -221 ms to 884 ms are shown in Table 4.5. Relevant for target fixations are the intercept and linear terms (see Section 2.10). The table shows a main effect of *Lexical Contrast* on the intercept term, but not on the linear term. The main effect on the intercept term confirms that looks to the target object rise earlier for color-contrast sequences than for no-contrast sequences. The absence of a reliable effect on the linear term suggests that looks to the target rise at an equal rate in the color-contrast compared to no-contrast sequences. Earlier looks to the target object in the color-contrast sequences (where the noun is repeated) than the no-contrast sequences.
(where the noun is novel) are expected because all conditions set up the expectation of a repeated noun.

Table 4.5 shows that there is no main effect of *Pitch Accent* or a *Lexical Contrast / Pitch Accent* interaction on the intercept or linear terms. Both results are not predicted. Rather, the conditions with two L+H* accents were predicted to elicit earlier looks to the color-contrast object than the conditions with a L+H* accent in the target instruction only. Thus, earlier looks to the target object were predicted for the *anticipatory* compared to the *felicitous* condition, and for the *garden-path* compared to the *anticipated garden-path* condition. This should have resulted in a reliable effect of *Pitch Accent* or a *Lexical Contrast / Pitch Accent* interaction on the intercept term.

Figures 4.3 and 4.4 show direct comparisons of the *felicitous* and *anticipatory* conditions and the *garden-path* and *anticipated garden-path* conditions, respectively. The graphs on the left show data points with error bars, and the graphs on the right show model fits for the curves predicted by the model. Figure 4.3 reveals that there is a small window, from –85 ms to 51 ms, for which there is an advantage for the *felicitous* over the *anticipatory* condition. Thus, contrary to prediction, there is a small window with reliably more looks to the target object if the context instruction carries a H* !H* pattern than if it carries a L+H* accent on the adjective.

Results from a GCA comparing the *felicitous* and *anticipatory* conditions are shown in Table 4.6. The analysis reveals a reliable effect on the intercept term, suggesting earlier looks to the target object in the *felicitous* condition compared to the *anticipatory* condition. Again, this result is the opposite of what was predicted. The *anticipatory* condition was expected to elicit earlier looks to the target object than the *felicitous* condition. It is possible that a L+H* accent can only evoke a contrast set with a preceding instruction, not with an upcoming instruction. The L+H* accent in the first instruction of the *anticipatory* condition may thus be somewhat infelicitous – since there is no preceding instruction with which a contrast set can be evoked. This may explain why looks to the target object are slower in
<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>20761.45</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>20696.07</td>
<td>-65.38</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>20695.57</td>
<td>-0.51</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>19802.29</td>
<td>-893.27</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>19801.61</td>
<td>-0.68</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 4.5: Experiment 2, data from EXPERIMENT A (see Figure 4.1): Target fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
Figure 4.3: Experiment 2, Color-contrast sequence from EXPERIMENT A: Comparison of the felicitous and anticipatory conditions presented in Table 4.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

the anticipatory condition compared to the felicitous condition. An additional GCA analysis (using an analysis window from -221 to 884) showed that the anticipatory condition does not differ reliably on the intercept or linear terms from the neutral condition (color-contrast sequence) of Experiment 1. Taking the neutral condition as a baseline, this suggests that the anticipatory condition presents neither a processing advantage nor delay. The felicitous condition, on the other hand, presents a processing advantage compared to the neutral condition (as was shown in Experiment 1).

Figure 4.4 shows no difference between the garden-path and anticipated garden-path conditions. The results from a GCA comparing the garden-path and anticipated garden-path conditions are shown in Table 4.7. The GCA reveals no reliable effects on the intercept or linear terms. This confirms the results from the omnibus GCA above, which found no main effect of Pitch Accent. An additional GCA (using an analysis window from -221 to 884) showed no reliable effects on the intercept or linear terms comparing the neutral
Table 4.6: Experiment 2, Color-contrast sequence from EXPERIMENT A (see Figure 4.1): Target fixation analysis comparing the felicitous and the anticipatory conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>9033.75</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>9029.62</td>
<td>−4.14</td>
<td>0.05</td>
</tr>
<tr>
<td>Linear</td>
<td>9029.46</td>
<td>−0.15</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>9029.41</td>
<td>−0.05</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>9009.98</td>
<td>−19.43</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 4.7: Experiment 2, No-contrast sequence from EXPERIMENT A (see Figure 4.1): Target fixation analysis comparing the garden-path and the anticipated garden-path conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8259.12</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>8257.38</td>
<td>−1.75</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>8255.65</td>
<td>−1.73</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>8253.84</td>
<td>−1.81</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>8233.12</td>
<td>−20.72</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

condition (no-contrast sequence) from Experiment 1 with both the garden-path condition and the anticipated garden-path condition. Taking the neutral condition as a baseline, this suggests that neither condition presents a processing advantage or delay.

Altogether the data suggest that the L+H* accent on the adjective of the first instruction of the anticipatory condition interfered with rather than strengthened the processing advantage found for the L+H* on the adjective of the second instruction. As a result, later looks to the target object were found in the anticipatory condition compared to the felicitous condition. The interference should not be due to the felicity of L+H* on the adjective in the first instruction, because the visual display includes more than one object consistent with the color adjective, and these objects are weakly familiar (Roberts, 2003) and could be referred to with a contrasting adjective. A possible explanation is that when a L+H* accent evokes a contrast set, generating the set and maintaining it in memory consumes processing resources. In sequences like the anticipatory condition, a contrast set is generated during the first and second instructions, while in sequences like the felicitous condition,
only one contrast set is generated (from the L+H* adjective during the second instruction). Referent resolution is less effortful in the felicitous condition, where the single contrast set combines with the most recently mentioned noun to generate anticipation of the repeated noun.

Data from Experiment B

Figure 4.5 shows looks to the target object cell in the four critical conditions of Experiment B (see Tables 4.2 and 4.4). The figure shows similar results as were found in Experiment A: Looks to the target object start to rise earliest in the felicitous condition. Unlike Experiment A, however, looks to the target in the garden-path condition rise earlier than looks to the target in the anticipated garden-path condition.

Figure 4.6 shows the grand mean of looks to the target object for the four critical conditions. The figure shows that looks to the target object start rising at -187 ms – last data
Figure 4.5: Experiment 2, data from EXPERIMENT B: Proportion of looks to the target object cell in the four critical conditions presented in Tables 4.2 and 4.4.
Figure 4.6: Experiment 2, data from EXPERIMENT A: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 4.2 and 4.4. The arrows indicate the beginning and end of the analysis window for the GCA analyses.

The results from an omnibus GCA, using an analysis window from -187 ms to 782 ms, are shown in Table 4.8. As for EXPERIMENT A, there is a main effect of Lexical Contrast on the intercept term, without a reliable effect on the linear term. This suggests a constant advantage for the color-contrast conditions over the no-contrast conditions. Unlike EXPERIMENT A, there is a main effect of Pitch Accent on the intercept term, and again no main effect of Pitch Accent on the linear term. This suggests a constant advantage for looks to the target in the conditions which have a H° !H° pattern in the first instruction (the felicitous and garden-path conditions) over the conditions which have a L+H° on the adjective of the first instruction (the anticipatory and anticipated garden-path conditions). This result goes counter to what was predicted: The anticipatory condition was predicted to elicit earlier
looks to the target object than the felicitous condition. There is no Lexical Contrast / Pitch Accent interaction on the intercept or linear terms.

Figure 4.7 shows a direct comparison of the felicitous and anticipatory conditions of EXPERIMENT B, as did Figure 4.7 for EXPERIMENT A. The Figure shows a small window with a reliable advantage for the felicitous condition over the anticipatory condition from about -17 ms to 102 ms. Results from a GCA comparing the two conditions are shown in Table 4.9 and reveal no reliable differences between the conditions on either the intercept or linear terms. These results differ from those found for EXPERIMENT A. Even though there is a small window with an advantage for the felicitous condition compared to the anticipatory condition in the data from both EXPERIMENTS, this advantage turns out to be reliable only for the data from EXPERIMENT A. Importantly though, data from both EXPERIMENTS A and B revealed that, contrary to prediction, the anticipatory condition did not elicit earlier looks to the target object than the felicitous condition.
<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td></td>
<td>16755.8</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>16681.81</td>
<td>-73.98</td>
<td>0.0001</td>
<td>15862.96</td>
</tr>
<tr>
<td>Linear</td>
<td>16681</td>
<td>-0.82</td>
<td>n.s.</td>
<td>15861.34</td>
</tr>
<tr>
<td>Quadratic</td>
<td>15878.39</td>
<td>-802.6</td>
<td>0.0001</td>
<td>15858.74</td>
</tr>
<tr>
<td>Cubic</td>
<td>15867.77</td>
<td>-10.62</td>
<td>0.01</td>
<td>15830.1</td>
</tr>
</tbody>
</table>

Table 4.8: Experiment 2, data from EXPERIMENT B (see Figure 4.5): Target fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
Table 4.9: Experiment 2, Color-contrast sequence from EXPERIMENT B (see Figure 4.7): Target fixation analysis comparing the felicitous and the anticipatory conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>6615.41</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>6612.95</td>
<td>-2.46</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>6612.78</td>
<td>-0.17</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>6584.69</td>
<td>-28.08</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>6545.2</td>
<td>-39.49</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 4.10: Experiment 2, No-contrast sequence from EXPERIMENT B (see Figure 4.8): Target fixation analysis comparing the garden-path and the anticipated garden-path conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>7096.9</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>7091.01</td>
<td>-5.89</td>
<td>0.05</td>
</tr>
<tr>
<td>Linear</td>
<td>7087.56</td>
<td>-3.45</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>7080.75</td>
<td>-6.8</td>
<td>0.01</td>
</tr>
<tr>
<td>Cubic</td>
<td>7076.22</td>
<td>-4.53</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Figure 4.8 shows a direct comparison of the garden-path and anticipated garden-path conditions of EXPERIMENT B, as did Figure 4.8 for EXPERIMENT A. The Figure reveals two windows with a reliable advantage for the garden-path over the anticipated garden-path condition. One window spans the time from 204 ms to 357 ms and the other one ranges from 595 ms to 697 ms. Results from a GCA comparing the two conditions are shown in Table 4.10. There is a reliable effect on the intercept, but not the linear term. This means that looks to the target object occured earlier in the garden-path condition than the anticipated garden-path condition. This result is consistent with the idea that the two successive L+H* accent generate two contrast sets, which increase the processing load and slow processing times in the anticipated garden-path condition as compared to the garden-path condition, which generated only one contrast set.

In summary, the results from analyses of looks to the targets in EXPERIMENTS A AND B reveal some similarities and some differences. Most importantly, data from neither EXPERIMENT showed the predicted earlier looks to the target object in the anticipatory condition.
compared to the felicitous condition. Rather, looks to the target object occur (numerically or reliably) later in the anticipated garden-path condition compared to the felicitous condition in both experiments. The data from the experiments differ in that only EXPERIMENT B, but not EXPERIMENT A, shows reliably later looks to the target object in the anticipated garden-path condition compared to the garden-path condition. Most likely as a result of this difference, only EXPERIMENT B, but not EXPERIMENT A, shows a main effect of Pitch Accent.

4.2.2 Color-Competitor Fixations

Data from EXPERIMENT A

Figure 4.9 provides an overview of the competitor fixation data from EXPERIMENT A. It shows looks to the color-competitor object cell in the four critical conditions of EXPERIMENT A (see Tables 4.1 and 4.3). As expected, the garden-path effects are larger for no-contrast sequences than color-contrast sequences. In addition, the figure shows that, as
Figure 4.9: Experiment 2, data from EXPERIMENT A: Proportion of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3.

predicted, the anticipated garden-path condition elicits the numerically largest garden-path effect.

Figure 4.10 shows the grand mean of looks to the color competitor object for the four critical conditions. The figure shows that looks to the color competitor object start rising at -153 ms, the beginning of the analysis window. 238 ms was chosen as the peak of the curve because it approximately cuts the mountain of the curve in half. It also happens to be the highest data point. To create an analysis window that is symmetrical around the peak, 629 ms was calculated as the end of the window.

The results from an omnibus GCA using the above calculated analysis window are shown in Table 4.11. Relevant for color-competitor fixations are the intercept and quadratic
Figure 4.10: Experiment 2, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3. The arrows indicate the beginning and end of the analysis window for the GCA analyses.

terms (see Section 2.10). The table shows a main effect of Lexical Contrast on the intercept and quadratic terms. The reliable effect on the quadratic term shows that the rise-fall pattern of looks is steeper in the no-contrast than the color-contrast sequences. The reliable effect on the intercept term shows that the area under the curve is larger for the no-contrast sequences than the color-contrast sequences. In this case, looks to the color-competitor do not differ across conditions at both ends of the analysis window. The steeper rise-fall pattern then results in a larger area under the curve in the no-contrast conditions compared to the color-contrast conditions. This pattern of results indicates a larger garden-path effect in the no-contrast sequences than the color-contrast sequences across the analysis window.

Table 4.11 further shows a reliable effect of Pitch Accent on the quadratic term. This reveals that the rise-fall pattern of the curves is reliably steeper in the conditions with two successive L+H* accents (the anticipatory and anticipated garden-path conditions) than in the conditions with a L+H* accent only in the target utterance (the felicitous and garden-path conditions).
However, there is no reliable effect of Pitch Accent on the intercept term, indicating no differences in area under the curve – i.e. no differences in size of garden-path effect – between the two pitch accent patterns. Finally, the GCA reveals no Lexical Contrast / Pitch Accent interaction.

Figure 4.11 shows a direct comparison of the garden-path and anticipated garden-path conditions. The figure shows reliably more looks to the color-competitor object in the anticipated garden-path condition over the garden-path condition for a small time window from 221 ms to 272 ms. Results from a GCA comparing the two conditions are shown in Table 4.12. The GCA shows a reliable effect on the quadratic term, but no effect on the intercept term. This means that the anticipated garden-path condition elicited a steeper rise-fall pattern than the garden-path condition across the time window of analysis. The absence of a difference in area under the curve between the two conditions suggests no difference in the size of the garden-path effect. Although it is tempting to interpret the numerically larger number of looks to the competitor object in the anticipated garden-path condition as an indication of more garden-path processing, this difference is limited to a brief (51 ms) time window, and showed a reliable effect on only the quadratic term, indicating differences in curve shape rather than overall size of the effect. Examination of the error bar display for the data in Figure 4.11 shows that functions for the two conditions are not reliably different throughout most of the analysis window. Thus, contrary to the predictions, the anticipated garden-path condition did not yield a reliably larger garden-path effect than the garden-path condition.

Figure 4.12 shows a direct comparison of the felicitous and anticipatory conditions. The figure shows no differences between the conditions across the analysis window. Results from a GCA comparing the two conditions are shown in Table 4.13. The GCA shows a reliable effect on the quadratic term. This means that the rise-fall pattern of the anticipatory condition is steeper than that of the felicitous condition. There is, however, no reliable effect on the intercept, suggesting that participants did not look more at the color-competitor in
<table>
<thead>
<tr>
<th>Model</th>
<th>Model fit</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
<td>-2LL</td>
</tr>
<tr>
<td>Base</td>
<td>14570.91</td>
<td>—</td>
<td>—</td>
<td>14072.15</td>
</tr>
<tr>
<td>Intercept</td>
<td>14540.4</td>
<td>−30.51</td>
<td>0.0001</td>
<td>14071.12</td>
</tr>
<tr>
<td>Linear</td>
<td>14536.61</td>
<td>−3.78</td>
<td>n.s.</td>
<td>14071.11</td>
</tr>
<tr>
<td>Quadratic</td>
<td>14182.85</td>
<td>−353.77</td>
<td>0.0001</td>
<td>14036.61</td>
</tr>
<tr>
<td>Cubic</td>
<td>14132.74</td>
<td>−50.11</td>
<td>0.0001</td>
<td>14034.19</td>
</tr>
<tr>
<td>Quartic</td>
<td>14072.15</td>
<td>−60.59</td>
<td>0.0001</td>
<td>14006.32</td>
</tr>
</tbody>
</table>

Table 4.11: Experiment 2, data from EXPERIMENT A (see Figure 4.1): Color-competitor fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.
Figure 4.11: Experiment 2, No-contrast sequence from EXPERIMENT A: Comparison of the garden-path and the anticipated garden-path conditions presented in Table 4.1. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits.

Table 4.12: Experiment 2, No-contrast sequence from EXPERIMENT A (see Figure 4.11): Competitor fixation analysis comparing the garden-path and the anticipated garden-path conditions.
Figure 4.12: Experiment 2, Color-contrast sequence from EXPERIMENT A: Comparison of the felicitous and the anticipatory conditions presented in Table 4.3. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits.

the anticipatory condition than the felicitous condition across the whole time window. These results are not predicted, but in line with the pattern of looks to the target object found in EXPERIMENT A (see Section 4.2.1). Earlier looks to the target object in any given condition may co-occur with a smaller garden-path effect in that condition. In EXPERIMENT A, the felicitous condition elicited reliably earlier looks to the target object and a numerically smaller garden-path effect than the anticipatory condition.

In summary, there is no evidence that the two Pitch Accent conditions (two successive L+H* accents vs. a L+H* accent only in the target utterance) elicited statistically significant differences in the size of the garden-path effect. That is, the presence of a H* !H* pattern on the NP (HHNP) vs. a L+H* accent on the adjective (LHA) of the context instruction had no effect on looks to the color-competitor object.
<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>5808.28</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>5806.66</td>
<td>-1.62</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>5806.64</td>
<td>-0.02</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>5793.87</td>
<td>-12.77</td>
<td>0.001</td>
</tr>
<tr>
<td>Cubic</td>
<td>5788.41</td>
<td>-5.46</td>
<td>0.05</td>
</tr>
<tr>
<td>Quartic</td>
<td>5785.13</td>
<td>-3.28</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 4.13: Experiment 2, Color-contrast sequence from EXPERIMENT A (see Figure 4.11): Competitor fixation analysis comparing the felicitous and the anticipatory conditions.

**Data from EXPERIMENT B**

Figure 4.13 presents an overview of looks to the color-competitor object in the four critical conditions of EXPERIMENT B (see Tables 4.2 and 4.4). As in EXPERIMENT A, the garden-path effects are larger for no-contrast sequences than color-contrast sequences. Unlike EXPERIMENT A, the anticipated garden-path condition does not elicit the numerically largest garden-path effect.

Figure 4.14 shows the grand mean of looks to the color competitor object for the four critical conditions. The figure shows that looks to the color competitor object start rising at -204 ms, the beginning of the analysis window. 170 ms was chosen as the peak of the curve because it approximately cuts the mountain of the curve in half. In this case, it is only the second highest data point. To create an analysis window that is symmetrical around the peak, 544 ms was calculated as the end of the analysis window.

The results from an omnibus GCA over the calculated analysis window are shown in Table 4.14. As in EXPERIMENT A, the table shows a main effect of Lexical Contrast on the intercept and quadratic terms. That is, the rise-fall pattern of looks is steeper and there are overall more looks to the color-competitor in the no-contrast than the color-contrast sequences. There is also a main effect of Lexical Contrast on the linear term, suggesting asymmetric differences between the Lexical Contrast conditions.
Figure 4.13: Experiment 2, data from EXPERIMENT B: Proportion of looks to the color competitor object cell in the four critical conditions presented in Tables 4.2 and 4.4.
Figure 4.14: Experiment 2, data from EXPERIMENT A: Grand mean of looks to the color competitor object cell in the four critical conditions presented in Tables 4.1 and 4.3. The arrows indicate the beginning and end of the analysis window for the GCA analyses.

Table 4.14 also shows a main effect of Pitch Accent on the quadratic term, with no reliable effect on the intercept term. This means that the two pitch accent patterns elicit looks that differ in rise-fall rate, but not in overall number of looks to the color-competitor across the window of analysis. That is, the two pitch patterns do not differ in the size of garden-path effect they elicit. The table shows no Lexical Contrast / Pitch Accent interaction.

Figure 4.15 shows a direct comparison of the garden-path and anticipated garden-path conditions. The figure shows no differences between the two prosodic conditions. Results from a GCA comparing the two conditions are shown in Table 4.12. The GCA shows a reliable effect on the quadratic term, but not the intercept term. This suggests that looks to the color-competitor rise (and fall) faster in the garden-path condition compared to the anticipated garden-path condition, but that both conditions elicit an equal amount of looks to the color-competitor object. Thus, importantly, the two conditions do not differ in the size of garden-path effect they elicit.
Table 4.14: Experiment 2, data from EXPERIMENT B (see Figure 4.5): Color-competitor fixation analysis: Main effects (Lexical Contrast and Pitch Accent) and interaction.

<table>
<thead>
<tr>
<th>Model</th>
<th>Lexical Contrast</th>
<th>Pitch Accent</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>12908.64</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>12849.45</td>
<td>-59.19</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>12844.60</td>
<td>-4.84</td>
<td>0.05</td>
</tr>
<tr>
<td>Quadratic</td>
<td>12491.97</td>
<td>-352.64</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>12453.33</td>
<td>-38.63</td>
<td>0.0001</td>
</tr>
<tr>
<td>Quartic</td>
<td>12436.44</td>
<td>-16.89</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
Figure 4.15: Experiment 2, No-contrast sequence from EXPERIMENT B: Comparison of the garden-path and the anticipated garden-path conditions presented in Table 4.3. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>5848.75</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>5848.39</td>
<td>−0.35</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>5848.32</td>
<td>−0.08</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>5828.05</td>
<td>−20.27</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>5807.63</td>
<td>−20.42</td>
<td>0.0001</td>
</tr>
<tr>
<td>Quartic</td>
<td>5807.56</td>
<td>−0.07</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 4.15: Experiment 2, No-contrast sequence from EXPERIMENT B (see Figure 4.15): Competitor fixation analysis comparing the garden-path and the anticipated garden-path conditions.
Figure 4.16 shows a direct comparison of the felicitous and anticipatory conditions. The figure shows reliably more looks to the color-competitor object in the felicitous over the anticipatory condition from the beginning of the analysis window at -204 ms to -68 ms and again from 17 ms to 51 ms. Results from a GCA comparing the two conditions are shown in Table 4.13. Unlike EXPERIMENT A, the GCA shows a reliable effect on the intercept and quadratic terms. This combination of effects suggests a larger garden-path effect in the felicitous over the anticipatory condition. The results also show a reliable effect on the linear term, suggesting asymmetric differences between the two curves.

Assuming that earlier looks to the target object may co-occur with a smaller garden-path effect, these results are predicted (since the anticipatory condition was predicted to elicit earlier looks to the target object than the felicitous condition), but not in line with the pattern of looks to the target object found in EXPERIMENT B (where numerically earlier looks were found for the felicitous compared to the anticipatory condition; see Section 4.2.1).

Notice, however, that we would expect a larger garden-path effect in the felicitous condition compared to the anticipatory condition to appear as more looks to the competitor object around the peak of the mountain. Rather, reliably more looks in the felicitous compared to the anticipatory condition occur during the rise of the curve. This suggests that the reliable effect on the intercept term may be the result of a difference in baseline at the beginning of the analysis window rather than a difference in the size of the garden-path effect.

In summary, the results from EXPERIMENT B are similar to those of EXPERIMENT A: All pairwise comparisons show a reliable effect on the quadratic term, and all but one comparison shows no effect on the intercept term. This suggests that the pitch accent conditions do not elicit garden-path effects of difference sizes.
Figure 4.16: Experiment 2, Color-contrast sequence from EXPERIMENT B: Comparison of the felicitous and the anticipatory conditions presented in Table 4.1. Left: Empirical logit of looks to the color-competitor object cell and error bars. Right: Model fits.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>4812.38</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>4804.5</td>
<td>-7.88</td>
<td>0.01</td>
</tr>
<tr>
<td>Linear</td>
<td>4797.62</td>
<td>-6.88</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>4792.62</td>
<td>-5.01</td>
<td>0.05</td>
</tr>
<tr>
<td>Cubic</td>
<td>4792.53</td>
<td>-0.08</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quartic</td>
<td>4792.51</td>
<td>-0.02</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 4.16: Experiment 2, Color-contrast sequence from EXPERIMENT B (see Figure 4.15): Competitor fixation analysis comparing the felicitous and the anticipatory conditions.
4.3 Discussion

This experiment tested whether the prosody of the first instruction affects the processing of the second instruction. In particular, it investigated whether L+H* accents on two successive color adjectives in a color-contrast sequence (YELLOW bottle ⇒ PURPLE bottle) lead to earlier looks to the second object compared to only a L+H* accent on the second color adjective (yellow bottle ⇒ PURPLE bottle). This would suggest that (in the absence of any previous discourse) a L+H* accent on YELLOW is interpreted as evoking an upcoming contrast set. Listeners would then anticipate a repeated noun as early as the end of the first instruction. The opposite would be expected in the no-contrast sequence: A L+H* accent on only the second color adjective (yellow scissors ⇒ PURPLE bottle) should elicit earlier looks to the target object than two successive L+H* accents (YELLOW scissors ⇒ PURPLE bottle). If the L+H* accent on YELLOW evoked an upcoming color contrast set, which is confirmed by the L+H* accent on PURPLE, listeners may take longer to recover from the assumption that the noun of the first instruction would be repeated than if only the L+H* on PURPLE suggested a color contrast set.

The results from Experiment 2, however, do not confirm this hypothesis. In Experiment A, the felicitous condition (yellow bottle ⇒ PURPLE bottle) elicits reliably earlier looks to the target object than the anticipatory condition (YELLOW bottle ⇒ PURPLE bottle), whereas the garden-path and anticipated garden-path conditions (yellow scissors ⇒ PURPLE bottle and YELLOW scissors ⇒ PURPLE bottle) do not differ reliably. In Experiment B, it is the other way around: The felicitous and anticipatory conditions do not differ reliably (even though looks to the target object are numerically earlier in the felicitous compared to the anticipatory condition), but the garden-path condition elicits earlier looks to the target object than the anticipated garden-path condition. Taken together, the data suggest that the conditions with only a L+H* accent in the target utterance (the felicitous and garden-path conditions) elicit earlier looks to the target object than the conditions with two successive
L+H* accents (the anticipatory and anticipated garden-path conditions) in both EXPERIMENTS. The omnibus GCAs revealed that this advantage is numeric only in EXPERIMENT A, but reliable in EXPERIMENT B.

EXPERIMENTS A and B were expected to yield the same results because critical instructions across the two EXPERIMENTS only differed in prosody (i.e. edge tone patterns) after the critical L+H* accent in the target instruction was heard. Thus, the differences in results were not predicted. The reason for these differences may after all lie in the differences in edge tone patterns across EXPERIMENTS. Participants in EXPERIMENT A only heard utterances ending in L-L% sentence-final intonation across the whole EXPERIMENT. Participants in EXPERIMENT B, on the other hand, heard utterances either ending in L-L% sentence-final intonation or a L-H% continuation rise. As a result, a L-L% pattern may have been more informative to participants of EXPERIMENT B than to participants of EXPERIMENT A. In particular, the L-L% pattern (compared to the L-H% pattern) may have indicated finality and separation of the current utterance from the one following it for listeners in EXPERIMENT B. But the L-L% pattern may have been more neutral for listeners in EXPERIMENT A since it was never contrasted with a L-H% pattern throughout the whole EXPERIMENT. Thus, even though the prosody of critical instructions was the same across EXPERIMENTS up to the target L+H* accent, the L-L% pattern of the first instruction may have been processed differently by participants of each EXPERIMENT.

With this in mind, let’s focus on the presence or absence of a L+H* accent in the first instruction. There are at least three possibilities for the interpretation of L+H* accents in instructions preceding the target instruction: (1) The L+H* accent evokes a contrast set with reference to the visual scene, (2) the L+H* accent evokes a contrast set with an upcoming utterance, or (3) the L+H* accent is ignored. If the L+H* accent in the initial instruction evokes a contrast set with the visual scene, it is felicitous and may be used immediately. In this case, looks to the target object should not have been delayed when the context instruction contained a L+H* accent as compared to when it did not. If the
L+H* accent evokes an expectation of an upcoming contrast set, this expectation may be stimied by the L-L% boundary that ends the first instruction. The L-L% finality contour separates rather than connects the initial utterance with the one following it, suggesting that the L+H* accent should not evoke an upcoming contrast set. If listeners heard only utterances ending in L-L%, the L-L% ending the first instruction may have interfered with the L+H* accent evoking an upcoming contrast set enough to cause a null-effect of Pitch Accent on the intercept term in the omnibus GCA (as in EXPERIMENT A). But if listeners heard utterances ending in L-L% or L-H% throughout the experiment, the L-L% ending the first instruction may have more actively suggested that the L+H* accent should not be interpreted as evoking an upcoming contrast set. Thus, in this case, the L-L% may have caused stronger interference. This would then result in reliably later looks to the target if the first instruction contained a L+H* accent than if it did not (as in EXPERIMENT B, where the omnibus GCA showed a reliable effect of Pitch Accent on the intercept term).

Finally, the L+H* accent of the first instruction could have been ignored. For example, prosodic information from the first instruction could have decayed from memory and not been available anymore when the second instruction was processed. In this case, however, data from both EXPERIMENTS A and B should have yielded a null-effect of Pitch Accent on the intercept term.

Notice that, as in Experiment 1, looks to the target object in the color-contrast sequence start rising about 100 ms earlier (around -225 ms compared to -125 ms) in EXPERIMENT B than in EXPERIMENT A. This is likely related to the duration differences of the adjectives, which are on average 80 ms shorter in EXPERIMENT A than EXPERIMENT B. With the graph aligned at the end of the adjective, lexical information from the adjective starts coming in a little earlier in EXPERIMENT B than EXPERIMENT A.

Unlike Experiment 1, the looks to the color-competitor reveal a consistent pattern. All pairwise comparisons for EXPERIMENTS A and B yielded a reliable effect on the quadratic term, suggesting a difference in slope between pitch accent conditions. Importantly, no
pairwise comparison yielded a reliable effect on the intercept term that could be interpreted as a difference in size of garden-path effect. The intercept term was reliable for only one of the comparisons and this effect was likely the result of a difference in baseline rather than a larger slope excursion. Thus, all pairwise comparisons yielded comparably large garden-path effects. Notice that the proportion of looks to the color-competitor remains under 35% throughout the analysis window in both Experiments 1 and 2. It is therefore possible that no reliable effects on the intercept term were detected because participants recovered so quickly from the garden-path.

Overall, the results from Experiment 2 do not confirm the hypothesis that two successive L+H* accents facilitate reference resolution in a color-contrast sequence. They do, however, suggest that a L+H* in an initial instruction could evoke a contrast set. When the initial instruction ends in L-L%, the finality contour may block the expectation of an upcoming contrast, and in the context of a visual scene with multiple objects of the same color, the contrast may refer felicitously to objects in the scene. When the initial instruction ends in a L-H% pattern, connecting the utterance to the one following it, the L+H* in the initial instruction could evoke the expectation of an upcoming contrast set. Further studies are needed to explore this possibility. The results also support the notion that L+H* accents are not interpreted in isolation. If this was the case, both Experiments A and B would have yielded null-effects. Rather, the prosody of the initial instruction does matter. This means that a representation of the interpretation of the prosodic pattern of the first instruction (including any generated contrast sets) must be kept in memory while the second instruction is processed. The results even allude to the possibility that information about distributional properties of prosody (in this case, edge tones) in the larger discourse is available when processing a L+H* in a given utterance.
Experiment 3a investigates how edge tone patterns interact with pitch accent and lexical information to affect reference resolution. The previous discussion of the results of Experiment 2 alluded to the possibility that listeners track the distribution of edge tone patterns throughout an entire discourse, and that this distributional information is used when interpreting pitch accent and lexical information. This would suggest that edge tones play a role for resolving reference. Experiment 3a investigates the role of edge tone patterns for reference resolution more directly. It examines whether the interpretation of pitch accent and lexical information is modulated by the frame of edge tones in which it occurs.

Pierrehumbert & Hirschberg (1990) proposed that a L-H% continuation rise indicates that the current utterance should be interpreted with respect to the one following it. A L-H% continuation rise compared to L-L% sentence-final intonation should thus link an utterance more closely to the one following it. As a result, a L-H% continuation rise compared to L-L% sentence-final intonation should facilitate resolving reference for the following instruction, as long as the sequence is otherwise felicitous. For example, listeners should show earlier eye movements to the target object, the purple bottle, in the
sequence in Example 5.1 if *YELLOW bottle* ends in a L-H% continuation rise rather than L-L% sentence-final intonation. In this example, *YELLOW bottle ⇒ PURPLE bottle* presents a felicitously accented color contrast. If a L-H% continuation rise on *YELLOW bottle* more closely connects this instruction to the one following it, listeners may have a stronger expectation of a repeated noun (*bottle*) occurring in the following instruction than if L-L% sentence-final intonation separates the two instructions. Such a result would suggest that listeners use a prosodic representation structured by edge tone information during discourse processing.

(5.1) *Click on the yellow scissors. Click on the YELLOW bottle. Click on the PURPLE bottle.*

Experiment 3a also explores the effect of the relative position in time of pitch accent information. Let’s consider again the sequence in Example 5.1. There is an object contrast between the first two instructions (*yellow scissors ⇒ YELLOW bottle*) and a color contrast between the last two instructions (*YELLOW bottle ⇒ PURPLE bottle*). The two L+H* accents are placed such that they felicitously mark the color contrast between the last two instructions. This placements, however, is infelicitous when considering the object contrast between the first two instructions. As a result, the L+H* accent on the repeated adjective *YELLOW* should elicit processing difficulties when it is first encountered (*Click on the yellow scissors. Click on the YELLOW...*). This expectation is based on previous results. Ito & Speer (2008) compared felicitous noun contrast sequences like *Hang the blue ball. Now hang the blue DRUM* with infelicitous noun contrast sequences like *Hang the blue ball. Now hang the BLUE drum*. They found later looks to the target object in the infelicitous noun contrast compared to the felicitous noun contrast sequences in the 600 ms to 900 ms after noun onset window. Notice that these processing difficulties occured late in processing the target utterance.

Now consider the sequence in Example 5.2, which presents a similar situation. Again, there are two lexical contrasts, a color contrast between the first two instruction (*green pencil
⇒ purple PENCIL) and an object contrast between the last two instructions (purple PENCIL ⇒ purple BOTTLE). Similar to the example above, the two L+H* accents are placed such that they felicitously mark the object contrast between the last two instructions, and are infelicitously placed with respect to the first two instructions. As a result, the L+H* accent on the repeated noun PENCIL should elicit similar processing difficulties when first encountered (Click on the green pencil. Click on the purple PENCIL...). Again, this expectation is based on previous results. Both Ito & Speer (2008) and Weber et al. (2006a) compared felicitous adjective contrast sequences like Hang the red drum. Now hang the BLUE drum or Click on the red scissors. Now click on the PINK scissors with infelicitous adjective contrast sequences like Hang the red drum. Now hang the blue DRUM or Click on the red scissors. Now click on the pink SCISSORS. Both found later looks to the target object in the infelicitous adjective contrast sequences compared to the felicitous adjective contrast sequences.

(5.2)  Click on the green pencil. Click on the purple PENCIL. Click on the purple BOTTLE.

The examples differ in that the expected processing difficulty related to the L+H* accent in the second instruction occurs relatively earlier in time in Example 5.1 than in Example 5.2. This also means that more time elapses between the expected processing difficulty and the onset of the upcoming third instruction in Example 5.1 than in Example 5.2. Experiment 3a explores how this difference in timing affects reference resolution in the third instruction. In particular, Experiment 3a explores the possibility that reference resolution for the third instruction is delayed in Example 5.2 compared to Example 5.1. In Example 5.2, the processing difficulty related to the infelicitously placed L+H* accent occurs immediately before the onset of the third instruction. Listeners may thus still be recovering from the processing difficulty when they encounter the third instruction. Remember that Ito & Speer (2008) found a late processing difficulty related to the infelicitously placed L+H* accent in the infelicitous noun contrast condition compared to the felicitous noun contrast condition. In the present EXPERIMENTS, instructions occurred in rapid succession.
late occurring processing difficulty may thus affect subsequent utterances, rather than the
target utterance. This would be consistent with results from Experiment 1, which found
no difference in looks to the target object in the neutral compared to the infelicitous condi-
tion. In Example 5.1, on the other hand, there is intervening sentence material between the
expected processing difficulty and the onset of the third instruction. Listeners may thus
be further in recovery or have already recovered from the processing difficulty when they
encounter the third instruction. Ongoing processing difficulty might result in late looks to
the target object of the third instruction. If this is the case, later looks to the target object of
Instruction 3 are expected for the sequence in Example 5.2 compared to Example 5.1.

As described above, a L-H\% pattern compared to a L-L\% pattern should facilitate ref-
ERENCE resolution for the sequence presented in Example 5.1. There are two possibilities
for how edge tones behave in the sequence presented in Example 5.2. A L-H\% pattern
compared to a L-L\% pattern ending the second instruction may help recover from the
L+H*-related processing difficulty by forging a closer connection of the second to the third
instruction. In this case, listeners should look at the target object of the third instruction
earlier if the second instruction ended in a L-H\% pattern than if it ended in a L-L\% pattern.
Alternatively, a L-H\% pattern compared to a L-L\% pattern ending the second instruction
may not modulate listeners’ responses. The proposed processing difficulty and the differ-
ence in edge tone pattern occur on the same lexical item (BOTTLE). Difficulty related to
processing the L+H* accent may thus interfere with considering edge tone information. In
this case, looks to the target object of the third instruction should not be affected by the
choice of edge tone pattern ending the second instruction.

To summarize, this experiment asks the following questions: Is reference of a target
object in a prosodically felicitous sequence resolved more easily if the preceding instruc-
tion ended in a L-H\% continuation rise compared to L-L\% sentence-final intonation? How
does the position in time of a preceding L+H*-related processing difficulty affect subse-
quent reference resolution and the consideration of edge tone information?
5.1 Experimental Conditions

Tables 5.1 and 5.2 show the critical conditions for Experiment 3a. The condition names in both tables refer to Instruction 2. The target instruction is always Instruction 3. Table 5.1 shows the *early L+H* conditions, where the L+H* accent of Instruction 2 occurs relatively early in time (i.e. on the adjective), whereas Table 5.2 shows the *late L+H* conditions, where the L+H* accent in the second instruction occurs relatively late in time (i.e. on the noun).

If the L+H* accent occurred early in the second instruction (as in the *early L+H* sequences), listeners should have recovered from any L+H*-related processing difficulty as
they processed the target instruction. If the L+H* accent occurred late in the second instruction (as in the late L+H* sequences), listeners may not yet have recovered from possible L+H* related processing difficulty as they processed the target instruction. As a result, looks to the target object may occur earlier in the early L+H* sequences compared to the late L+H* sequences. Notice that the target instruction presents a color contrast with the previous instruction in the early L+H* sequences, but an object contrast in the late L+H* sequences. In addition, the target instruction contains a felicitously placed L+H* accent in both the early L+H* and late L+H* sequences. Ito & Speer (2008) found earlier looks to the target object for felicitously accented color contrasts compared to felicitously accented object contrasts. Earlier looks to the target object in the early L+H* sequences compared to the late L+H* sequences could thus be related to differences in lexical contrast rather than the location of processing difficulties. This will be addressed in Experiment 3b.

In the finality conditions in Tables 5.1 and 5.2, Instruction 2 ends in L-L% sentence-final intonation, whereas it ends in a L-H% continuation rise in the continuation conditions. The finality conditions come from Experiment A, whereas the continuation conditions come from Experiment B. A L-H% continuation rise compared to L-L% sentence-final intonation should more closely link Instructions 2 and 3. This should lead to earlier eye movements to the target object of Instruction 3 in the early L+H* / continuation condition compared to the early L+H* / finality condition and possibly in the late L+H* / continuation compared to the late L+H* / finality condition. In the late L+H* sequences, the expected processing difficulty occurred close in time to (i.e. on the same lexical item as) the L-L% vs. L-H% patterns and may have interfered with the processing of the edge tone information. In particular, listeners may not have recovered from the processing difficulty when they encountered the edge tone pattern ending the second instruction. As a result, edge tone patterns may not modulate eye movement responses in the late L+H* sequences.

It needs to be emphasized that the edge tone manipulation compares data from Experiment A to data from Experiment B. Experiments 1 and 2 showed earlier looks to
target objects overall for data coming from EXPERIMENT B compared to data coming from EXPERIMENT A. It was pointed out that this may be related to individual differences in participants and/or differences in duration of adjectives – which are longer in EXPERIMENT B than EXPERIMENT A. Section 5.3 below addresses some of the concerns related to comparing data that came from different participants who responded to stimuli from different speakers.

5.2 Results

Unlike the previous two experiments, only looks to the target object will be reported. Looks to the color-competitor are not of interest because none of the conditions is expected to elicit a garden-path effect. In addition, in the late $L+H^*$ conditions, but not the early $L+H^*$ conditions, the color-competitor of the target instruction is also the target object of the preceding instruction. Participants are thus likely to still look at the color-competitor object at the beginning of the target instruction, rendering comparisons of looks to the color-competitor between early $L+H^*$ and late $L+H^*$ conditions meaningless.

Figure 5.1 provides an overview of looks to the target object cell in the four critical conditions of Experiment 3a (see Tables 5.1 and 5.2). The figure reveals the predicted pattern of responses. It shows earlier looks to the target object in the continuation conditions compared with finality conditions, suggesting that a L-H% continuation rise ending the previous instruction provided a closer connection to the target instruction than a L-L% pattern and facilitated resolving reference. In addition, early $L+H^*$ sequences elicited numerically earlier looks than late $L+H^*$ sequences, suggesting that the timing of the $L+H^*$-related processing difficulty affected reference resolution.

Figure 5.2 shows the grand mean of looks to the target object for the four critical conditions of Experiment 3a. The figure shows that looks to the target object start rising at -51
Figure 5.1: Experiment 3a, data from EXPERIMENTS A and B: Proportion of looks to the target object cell in the four critical conditions presented in Tables 5.1 and 5.2.
ms (the last data point lower than the previous one). 850 ms was chosen as the peak of the curve because the curve falls after 850 ms.

All GCA analyses for this experiment were performed over an analysis window from -51 ms to 850 ms. The results from an omnibus GCA are shown in Table 5.3. Since the *finality* and *continuation* conditions are between-subjects, the subjects fixed-effects had to be excluded from the analysis. The table reveals a main effect of *L+H* Position on the intercept and linear terms. The reliable effect on the intercept term suggests a constant advantage, i.e. earlier looks to the target object, for the *early L+H* conditions over the *late L+H* conditions. This result confirms the prediction that the relative position of a L+H* that caused a processing difficulty affected looks to the target object. In particular, if more sentence material (as in the *early L+H* conditions) compared to less sentence material (as in the *late L+H* conditions) intervened between the location of the L+H* accent in the second instruction and the noun phrase of the target instruction, listeners had had enough time to
recover from the processing difficulty when they encountered the target instruction noun phrase. The reliable effect on the linear term suggests a steeper slope of the curves in the late L+H* compared to the early L+H* conditions.

There is also a main effect of Edge Tone on both the intercept and linear terms. This suggests earlier looks, but a shallower rise in looks, to the target object in the continuation conditions compared to the finality conditions. The result confirms the prediction that a L-H% continuation rise compared to L-L% sentence-final intonation more closely linked an utterance to the one following it. This linkage facilitated resolving reference, and led to earlier eye movements to the target object. The analysis showed no L+H* Position / Edge Tone interaction.

Figure 5.3 shows a direct comparison of the two early L+H* (early L+H* / continuation and early L+H* / finality) conditions. The figure shows the predicted advantage for the early L+H* / continuation condition over the early L+H* / finality condition. There were reliably more looks to the target object in the early L+H* / continuation condition compared to the early L+H* / finality condition from the beginning of the analysis window until 323 ms. The advantage for the early L+H* / continuation condition over the early L+H* / finality condition begins as early as -255 ms.

Results from a GCA comparing the two early L+H* conditions are shown in Table 5.4 and reveal a reliable effect on the linear term, but not the intercept term. This pattern of results is novel. It did not occur in any of the target fixation analyses in Experiments 1 or 2. The reliable effect on the linear term indicates a difference in slope, such that looks to the target object rise faster in the early L+H* / finality condition over the early L+H* / continuation condition. The absense of a reliable effect on the intercept term suggests that the two lines cross near the middle of the analysis window (cf. Section 2.10). As a result, there are more looks to the target object in the early L+H* / continuation condition compared to the early L+H* / finality condition early in the time window, but fewer looks to the target object in the early L+H* / continuation condition compared to the early L+H* / finality condition late.
<table>
<thead>
<tr>
<th>Model</th>
<th>L+H* Position</th>
<th>Edge Tone</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>16637.74</td>
<td>16395.45</td>
<td>16347.46</td>
</tr>
<tr>
<td>Intercept</td>
<td>16613.77</td>
<td>16391.37</td>
<td>16347.09</td>
</tr>
<tr>
<td>Linear</td>
<td>16602.81</td>
<td>16377.16</td>
<td>16344.92</td>
</tr>
<tr>
<td>Quadratic</td>
<td>16536.62</td>
<td>16363.98</td>
<td>16341.35</td>
</tr>
<tr>
<td>Cubic</td>
<td>16395.45</td>
<td>16347.46</td>
<td>16341.19</td>
</tr>
</tbody>
</table>

Table 5.3: Experiment 3a, data from EXPERIMENTS A and B (see Figure 5.1): Target fixation analysis: Main effects ($L+H^*$ Position and Edge Tone) and interaction.
Figure 5.3: Experiment 3a, Early L+H* sequence from EXPERIMENTS A and B: Comparison of the early L+H* / finality and early L+H* / continuation conditions presented in Table 5.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

in the time window. Figure 5.3 shows that looks to the target object started rising earlier, but peaked lower, in the early L+H* / continuation condition compared to the early L+H* / finality condition. This result confirms the prediction that edge tones modulated the eye movement response to the target object. Here, looks to the target object were earlier, but with a more shallow slope, in the continuation condition than the finality condition. Thus, if listeners had recovered from the earlier processing difficulty, a L-H% linked the second instruction more closely to the third instruction and facilitated reference resolution.

Figure 5.4 shows a direct comparison of the two late L+H* conditions (late L+H* / continuation and late L+H* / finality). The figure shows a number of small windows where the lines for the late L+H* / continuation and late L+H* / finality conditions diverge enough to show significance. The largest of these windows includes four data points and ranges from -102 ms to -51 ms. The figure also shows that the two lines mainly diverge at the very beginning of the analysis window and before looks to the target object start rising, suggest-
Table 5.4: Experiment 3a, *Early L+H* sequence from EXPERIMENTS A and B (see Figure 5.3): Target fixation analysis comparing the *early L+H*/finality and *early L+H*/continuation conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>7855.41</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>7853.37</td>
<td>-2.04</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>7840.97</td>
<td>-12.39</td>
<td>0.001</td>
</tr>
<tr>
<td>Quadratic</td>
<td>7839.34</td>
<td>-1.63</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>7832.13</td>
<td>-7.21</td>
<td>0.01</td>
</tr>
</tbody>
</table>

The analysis reveals no statistically significant differences between the two conditions, with a slight difference in baseline. Thus, Figure 5.4 reveals no differences between the two conditions that could easily be related to cognitive processes. In particular, it is unlikely that the difference in baseline is related to the processing of the edge tone pattern of the second instruction. Differences in baseline are most likely to occur if, at the beginning of the current trial, listeners are still looking at the object they fixated towards the end of the previous trial (as in Figure 2.13 in Chapter 2, where the target object is repeated and listeners are still looking at the target object of the previous instruction at the beginning of the target instruction). A difference in baseline may thus be expected when comparing conditions that differ in whether an object is or is not mentioned in two successive instructions. This is not the case here.

Table 5.5 shows results from a GCA comparing the two *late L+H* conditions. Neither the intercept term nor the linear term are significant, suggesting neither a difference in area under the curve nor slope of the curves. This result reveals that the numeric advantage for the continuation conditions at the very beginning of the analysis window is not statistically reliable. The results suggest that if listeners had not had time to recover from the earlier processing difficulty, a L-H% did not link the second instruction more closely to the third instruction and did not facilitate reference resolution. This suggests that the L+H*-related processing difficulty interfered with the processing of the L-L% vs. L-H% pattern, and that a L-H% continuation rise only facilitated looks to the target object of an upcoming instruction if listeners had already recovered from the processing difficulty.
Figure 5.4: Experiment 3a, Late L+H* sequence from EXPERIMENTS A and B: Comparison of the late L+H* / finality and late L+H* / continuation conditions presented in Table 5.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8510.80</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>8509.13</td>
<td>-1.67</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>8506.1</td>
<td>-3.03</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>8491.86</td>
<td>-14.24</td>
<td>0.001</td>
</tr>
<tr>
<td>Cubic</td>
<td>8482.53</td>
<td>-9.33</td>
<td>0.01</td>
</tr>
</tbody>
</table>

Table 5.5: Experiment 3a, Late L+H* sequence from EXPERIMENTS A and B (see Figure 5.4): Target fixation analysis comparing the late L+H* / finality and late L+H* / continuation conditions.
Figure 5.5: Experiment 3a, Continuation sequence from EXPERIMENT B: Comparison of the early L+H* / continuation and late L+H* / continuation conditions presented in Tables 5.1 and 5.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Figures 5.5 and 5.6 show a direct comparison of the two continuation conditions (early L+H* / continuation and late L+H* / continuation) and the two finality conditions (early L+H* / finality and late L+H* / finality), respectively. Both figures show the predicted earlier looks to the target object in the early L+H* conditions over the late L+H* conditions. In both cases, there is a large window starting with the beginning of the analysis window and ending at 527 ms with an advantage for the early L+H* compared to the late L+H* conditions in both the continuation and finality sequences. This result is in line with the prediction that L+H*-related processing difficulty would interfere less with reference resolution if it occurred removed in time from when listeners encountered the target noun phrase (as in the early L+H* conditions) compared to close in time to when listeners encountered the target noun phrase (as in the late L+H* conditions).

Results from GCAs comparing the two continuation conditions and the two finality conditions are shown in Tables 5.6 and 5.7, respectively. Both tables show a reliable effect on
the intercept term, confirming the earlier looks to the target object in the early $L+H^*$ / finality conditions compared to the late $L+H^*$ conditions (see Figures 5.5 and 5.6). In addition, there is a reliable effect on the linear term comparing the two continuation conditions, but not comparing the two finality conditions. This suggests differences in slope between the early $L+H^*$ / continuation and late $L+H^*$ / continuation conditions, but not between the early $L+H^*$ / finality and late $L+H^*$ / finality conditions. In particular, looks to the target object in the early $L+H^*$ / continuation condition started rising much earlier than looks to the target object in the late $L+H^*$ / continuation condition, but peaked around the same time, resulting in a shallower slope for the early $L+H^*$ / continuation condition compared to the late $L+H^*$ / continuation condition. These results suggest that the early $L+H^*$ conditions yielded earlier looks to the target object than the late $L+H^*$ conditions, regardless of whether the second instruction ended in a $L-H^%$ continuation rise or $L-L^%$ sentence-final intonation.
Table 5.6: Experiment 3a, *Continuation* sequence from EXPERIMENT B (see Figure 5.5): Target fixation analysis comparing the *early* L+H* / *continuation* and *late* L+H* / *continuation* conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8394.9</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>8383.11</td>
<td>−11.79</td>
<td>0.001</td>
</tr>
<tr>
<td>Linear</td>
<td>8372.73</td>
<td>−10.37</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>8353.38</td>
<td>−19.35</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>8288.47</td>
<td>−64.91</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 5.7: Experiment 3a, *Finality* sequence from EXPERIMENT A (see Figure 5.6): Target fixation analysis comparing the *early* L+H* / *finality* and *late* L+H* / *finality* conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>8189.11</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>8175.89</td>
<td>−13.21</td>
<td>0.001</td>
</tr>
<tr>
<td>Linear</td>
<td>8173.55</td>
<td>−2.34</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>8122.46</td>
<td>−51.09</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>8045.33</td>
<td>−77.12</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

In summary, the results from Experiment 3a revealed a main effect of both L+H* Position and Edge Tone on the intercept and linear terms. This means that both L+H* Position and Edge Tone affected how early participants looked at the target object and how fast looks to the target object started rising. In particular, participants looked at the target earlier in *early* L+H* sequences than *late* L+H* sequences, and in *continuation* sequences compared to *finality* sequences. The GCA analysis comparing the two *late* L+H* sequences showed no reliable differences, suggesting that the main effect of Edge Tone is driven by differences in the *early* L+H* sequences.

5.3 Between-subject Comparisons

The results comparing the two *early* L+H* sequences and the two *late* L+H* sequences suggested that a L-H% continuation rise compared to L-L% sentence-final intonation con-
nected an instruction more closely to the one following it and facilitated resolving reference of the upcoming instruction if listeners had already recovered from an earlier L+H*-related processing difficulty. However, conclusions about the effects of a L-L% vs. L-H% pattern need to be made with caution because these comparisons were between-experiments (and as such between-subjects and between-speaker).

Experiments 1 and 2 revealed that looks to the target object in EXPERIMENT B tended to start rising about 100 ms earlier than in EXPERIMENT A. The auditory stimuli in EXPERIMENT B were longer in duration than in EXPERIMENT A, so listeners received lexical information from the adjective earlier in EXPERIMENT B than in EXPERIMENT A. In addition, participants in EXPERIMENT B may have been faster than participants in EXPERIMENT A. The earlier looks to the target in the *early L+H* / continuation condition (from EXPERIMENT B) compared to the *early L+H* / finality condition (from EXPERIMENT A) could have been a result of these differences in participants and/or auditory stimuli duration rather than a processing advantage resulting from the L-H% continuation rise.

If this is the case, i.e. if participants in EXPERIMENT B were overall faster than participants in EXPERIMENT A (be it due to individual differences or differences in stimuli durations), a pattern of earlier looks to a condition from EXPERIMENT B compared to one from EXPERIMENT A might simply reflect a null effect in the comparison of L-L% vs. L-H% edge tone patterns. Similarly, the finding of no differences in looks to the target object in conditions coming from both EXPERIMENTS would suggest a delay of looks to the condition from EXPERIMENT A. That is, the advantage for the *early L+H* / continuation condition (from EXPERIMENT B) and the *early L+H* / finality condition (from EXPERIMENT A) might in fact have been a null effect as far as the choice of L-L% vs. L-H% was concerned. Similarly, the null effect found for the comparison of the *late L+H* / continuation condition (from EXPERIMENT B) and the *late L+H* / finality condition (from EXPERIMENT A) might translate into a delay in looks to the target object in response to a preceding L-H% pattern compared to a L-L% pattern.
Importantly, differences among participants and auditory stimuli duration should have affected all conditions equally. We can therefore nevertheless conclude that the L-H% continuation rise modulated the timing of looks to the target object. However, there are two possible scenarios. If differences in looks to the target object are not related to individual differences in participants or auditory stimuli duration, we can conclude that the L-H% continuation rise facilitated looks to the target object compared to a L-L% pattern in the color-contrast conditions. That is, after having recovered from a L+H*-related processing difficulty, a L-H% pattern presented a closer connection to the target instruction than a L-L% pattern and facilitated reference resolution. But when listeners experienced processing difficulty as they encountered the edge tone pattern, a L-H% pattern compared to a L-L% did not facilitate reference resolution.

If differences in looks to the target object are related to individual differences in participants or auditory stimuli duration, the null effect found for the late L+H* conditions may actually have been a delay in looks to the target object for the late L+H* / continuation condition over the late L+H* / finality condition. However, this pattern of results would be difficult to explain.

Follow-up experiments where the two early L+H* and late L+H* sequences are within-subject and within-speaker comparisons would need to be conducted to resolve this issue. We can, however, get an idea as to whether differences in stimuli durations or individual participants may have affected the earliness of looks to the target object by comparing looks to the target object of Instruction 1. Instruction 1 provides a good comparison because it gives us a measure of how long it takes participants to find a target object when first presented with a visual display and asked to click on one object in the display. There are thus no confounding factors from differences in previous instructions.

Remember that the data is aligned at the end of the adjective. The longer stimuli durations of EXPERIMENT B compared to EXPERIMENT A affected how early lexical and pitch accent information from the adjective was heard. In particular, lexical and pitch accent
information from adjectives was heard earlier the longer the adjectives were. The longer adjective durations in EXPERIMENT B compared to EXPERIMENT A may have allowed participants to start searching for an object earlier, potentially leading to earlier looks to the target object. Such an advantage should appear when we compare looks to the target object of Instruction 1. Looking at Instruction 1 also has the advantage that prosodic patterns of Instruction 1 across EXPERIMENTS are identical.

Individual differences among participants should also appear when comparing looks to the target object of Instruction 1. If participants in EXPERIMENT B are overall faster at looking at the target object than participants in EXPERIMENT A, they should also find the target object of the very first instruction faster.

The following analyses compare looks to the target object of Instruction 1 across EXPERIMENTS. In particular, instructions with a L+H* accent on the adjective (LHA, e.g. Click on the PURPLE bottle) and a L+H* accent on the noun (LHN, e.g. Click on the purple BOTTLE) will be compared across EXPERIMENTS since these are the prosodic patterns of the target instructions in Experiment 3a. There will thus be the following comparisons: Instruction 1 is pronounced with a L+H* accent on the adjective and comes from EXPERIMENT A (Exp A: LHA), Instruction 1 is pronounced with a L+H* accent on the adjective and comes from EXPERIMENT B (Exp B: LHA), Instruction 1 is pronounced with a L+H* accent on the noun and comes from EXPERIMENT A (Exp A: LHN), or Instruction 1 is pronounced with a L+H* accent on the noun and comes from EXPERIMENT B (Exp B: LHN).

The comparison of looks in response to Instruction 3 with those in response to Instruction 1 is also convenient because average adjective durations for Instructions 1 and 3 are equivalent in both EXPERIMENTS A and B\(^1\). In particular, adjectives of Instructions 1 and 3 are approximately 80 ms longer in EXPERIMENT B than in EXPERIMENT A. Adjectives from LHA instructions have an average duration of 305 ms in EXPERIMENT A compared to 381

\(^1\)Due to the edge tone manipulation in EXPERIMENT B, average adjective durations for Instruction 2 differ from the average adjective durations for Instructions 1 and 3 in this EXPERIMENT.
Adjectives from LHN instructions have an average duration of 266 ms in EXPERIMENT A compared to 347 ms for EXPERIMENT B. Since Instructions 1 and 3 have the same adjective duration differences, response patterns due to these differences should appear equally in the latency of looks to Instruction 1 and Instruction 3.

Figure 5.7 provides an overview of looks to the target object cell in the four conditions. The figure reveals that there are small differences between the curves of the four conditions until about 550 ms. After that, the curves from EXPERIMENT A are steeper than those from EXPERIMENT B, resulting in a higher peak of the curves. It seems that the curves from EXPERIMENTS A and B differ not so much in how early looks to the target object start rising, but in the peak number of looks to the target object. The differences in peak height are most likely related to individual differences. Informal observations suggest that some participants used their peripheral vision, sometimes for several trials in a row, to find objects on the computer screen. That is, they clicked on the target object while looking at the cross in the middle of the screen. Some participants’ eye movements also sometimes “undershot” the target. That is, the eyes moved in the direction of the target object, but landed in a spot between the target object cell and the cross in the middle of the screen, often close to the target object cell.

Overall, Figure 5.7 suggests that the earlier looks in the early L+H* / continuation condition compared to the early L+H* / finality condition of Experiment 3a presents an advantage related to the choice of L-H% vs. L-L% pattern rather than one related to stimuli durations or individual differences. Notice that in EXPERIMENTS A and B, Instruction 1 is pronounced with a L+H* accent on the adjective and a L+H* accent on the noun in six experimental conditions, respectively (see Section 2.5). Therefore, each data point in Figure 5.7 is calculated from six times as many trials as each data point in Figures 5.3 through 5.6 of Experiment 3a. Comparing GCA pairwise comparisons from Experiment 3a with GCA pairwise comparisons run on the data in Figure 5.7 would therefore not be sensible. Due to the much larger number of data points in the data shown in Figure 5.7 compared
Figure 5.7: Instruction 1, data from EXPERIMENTS A and B: Proportion of looks to the target object cell in the four conditions described above.
to the data from Experiment 3a, a much smaller divergence of two curves would yield statistically reliable effects. For a better comparison between looks to the target object of Instruction 1 and looks to the target object in Experiment 3a, curves derived from similar numbers of data points should be compared. This would be achieved if only one of the six experimental conditions from EXPERIMENTS A and B, where Instruction 1 is pronounced with a L+H* on the adjective and noun, respectively, was compared. In this case, the number of data points going into the GCA analyses would be comparable to the number of data points that went into the GCA analyses of Experiment 3a.

There is no a priori way to decide which of the six LHA and six LHN conditions to choose for the GCA analyses. Therefore, for each EXPERIMENT and each pronunciation of Instruction 1 (LHA or LHN), the curves derived from an individual condition were compared to the curves derived from all six conditions (see Section 2.5). For example, the curve derived from all six conditions in EXPERIMENT A where Instruction 1 is pronounced with a L+H* accent on the adjective was compared to the curves derived from each individual condition in EXPERIMENT A where Instruction 1 is pronounced with a L+H* accent on the adjective. The condition whose curve most closely matched that of all six conditions was chosen for the GCA analyses. The best matching curves (as determined through visual inspection of relevant graphs) are shown in Figure 5.8.

Figure 5.9 shows the grand mean of looks to the target object for the four conditions chosen for the GCA analyses. The figure shows that looks to the target object start rising at -272 ms, then dip and start rising again at -204 ms. 204 ms was chosen as the beginning of the analysis window because it is the last data point that is lower than or level with the previous one. The peak of the curve ranges from 938 ms to 969 ms. 969 ms was chosen as the end of the analysis window because it is the beginning of the fall of the curve. The following GCA analyses will thus be conducted over an analysis window ranging from -204 ms to 969 ms.

The GCA analyses will focus on the two between-subject comparisons since they are
Figure 5.8: Instruction 1, data from EXPERIMENTS A and B: Comparison of the curves derived from all six conditions of EXPERIMENTS A and B where Instruction 1 is pronounced with either a L+H* on the adjective or the noun (Six Cond) with the best fitting curve derived from only one of those conditions (Best Fit).
the only comparisons of interest here. Figure 5.10 shows a direct comparison of the two LHA conditions from EXPERIMENTS A and B. That is, the figure compares looks to the target object of Instruction 1 pronounced with a L+H* accent on the adjective (LHA) across EXPERIMENTS. The figure shows no reliable differences between looks to the target object in EXPERIMENTS A and B across the whole analysis window. The figure suggests that the earlier looks to the target object in the early L+H*/continuation conditions compared to the early L+H*/finality conditions in Experiment 3a were not the result of differences in stimuli durations or individual participants. If this was the case, earlier looks to the target object of Instruction 1 should have been found for EXPERIMENT B compared to EXPERIMENT A.

Table 5.8 shows results from a GCA comparing Instruction 1 pronounced with a L+H* accent on the adjective in EXPERIMENT A and EXPERIMENT B. The table shows no reliable effects on the intercept or linear terms. This means that participants in EXPERIMENT B were not faster at finding an object on the computer screen than participants in EXPERIMENT A.
when the instruction describing the referred-to object was pronounced with a L+H* accent on the adjective (LHA). Rather, when presented with a computer screen containing twelve objects and instructed to click on one of the objects, participants in both EXPERIMENTS looked at the target object equally quickly. This also suggests that the longer adjective durations in EXPERIMENT B compared to EXPERIMENT A did not yield earlier looks to the target object.

Figure 5.11 shows a direct comparison of the two LHN conditions from EXPERIMENTS A and B, i.e. the two conditions where the first instruction is pronounced with a L+H* on the noun. The figure shows no reliable differences between looks to the target object in EXPERIMENTS A and B until 901 ms. That is, the curves diverge only at the very end of the analysis window. Overall, the figure suggests that the null effect of looks to the target object in the late L+H* / continuation conditions compared to the late L+H* / finality conditions in Experiment 3a was in fact a null effect and not the result of differences in stimuli.
durations or individual participants. If this was the case, earlier looks to the target object of Instruction 1 should have been found for EXPERIMENT B compared to EXPERIMENT A.

Table 5.9 shows results from a GCA comparing Instruction 1 pronounced with a L+H* accent on the noun (LHN) in EXPERIMENT A and EXPERIMENT B. The table shows no reliable effects on the intercept or linear terms. This confirms that participants in EXPERIMENTS A and B are equally fast at finding an object on the computer screen, regardless of whether the instruction is pronounced with a L+H* accent on the adjective (as in the comparison above) or with a L+H* accent on the noun (as in the current comparison).

The comparisons presented in Figures 5.10 and 5.11 and Tables 5.8 and 5.9 suggest that it is neither the case that the longer adjective durations in EXPERIMENT B compared to EXPERIMENT A allowed participants to find a target object on the computer screen faster nor that participants in EXPERIMENT B were generally faster than participants in EXPERIMENT A. This supports the validity of the above interpretation of the between-subject
comparisons of Experiment 3a. However, looking at participants’ eye movement behavior in response to hearing Instruction 1 to draw conclusions about participants’ response to hearing three instructions in a row is somewhat of a crutch. In addition, the null effects found for Instruction 1 do not explain why looks to the target object do start rising earlier (regardless of experimental condition) in the data from Experiment B compared to Experiment A in both Experiments 1 and 2. Ideally, Experiment 3a should be re-run in a future study with all factors within subjects.

5.4 Discussion

Experiment 3a investigated whether a L-H% continuation rise compared to L-L% sentence-final intonation can strengthen the expectation of an upcoming lexical contrast by more closely connecting an utterance with the one following it. In the early L+H* conditions,
the information from the early L+H* accent in the second instruction (YELLOW) was felicitously placed for the color contrast between the second and third instructions (yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle). When first encountered, the L+H* accent was predicted to cause processing difficulty because it evoked a contrast set of bottles of different colors, but the color adjective was repeated. Ending the second instruction with a L-H% pattern compared to a L-L% pattern provided a cue for the possibility of interpreting the L+H* accent as instead marking an upcoming color contrast. This led to earlier looks to the target object of the third instruction in the early L+H* / continuation compared to the early L+H* / finality condition.

In the late L+H* conditions, the information from the late L+H* accent in the second instruction (PENCIL) was felicitously placed for the object contrast between the second and third instructions (green pencil ⇒ purple PENCIL ⇒ purple BOTTLE). When first encountered, the L+H* accent was again predicted to cause processing difficulty because it evoked a contrast set of different kinds of purple objects, while the noun was repeated. In this case, ending the second instruction with a L-H% pattern compared to a L-L% pattern also provided a cue for the possibility of interpreting the L+H* accent as marking an upcoming contrast set. However, the L+H* that initially caused processing difficulty and the edge tone information occurred closer in time in the late L+H* conditions than in the early L+H* conditions. In particular, participants may not yet have recovered from the processing difficulty when they heard the edge tone information. As a result, edge tone information may have been ignored causing the null effect for looks to the target object of the third instruction in the late L+H* / continuation compared to the late L+H* / finality condition.

Together the data suggest that edge tones can play a role in resolving reference. In particular, the results are consistent with Pierrehumbert & Hirschberg’s (1990) notion that a L-H% continuation rise indicates that the current utterance should be interpreted with respect to the following one. A L-H% continuation rise compared to L-L% sentence-final
intonation facilitated reference resolution of the upcoming instruction, provided that listeners were not experiencing processing difficulties when they encountered the edge tones.

Experiment 3a also explored how the temporal distance between processing difficulties and the target noun phrase affected looks to the target object. In all sequences tested in this experiment, the first L+H* accent that listeners encountered was in an infelicitous sentence location with respect to what they had heard thus far. Encountering the L+H* accent was thus predicted to cause processing difficulties. More sentence material intervened between the L+H*-related processing difficulties and the target noun phrase in the early L+H* sequences than in the late L+H* sequences. As a result, listeners were more likely in the early L+H* sequences than in the late L+H* sequences to have recovered fully from the processing difficulties by the time they encountered the target noun phrase. If listeners were still experiencing processing difficulties when they encountered the target noun phrase in the late L+H* sequences, but not when they encountered the target noun phrase in the late L+H* sequences, later looks to the target object would be expected for the later L+H* compared to the early L+H* sequences. The experiment did indeed reveal this pattern of results.

At first sight, the data supports the idea that a late occurring processing difficulty compared to an early occurring processing difficulty contributed to the later looks in the late L+H* conditions compared to the early L+H* conditions. However, it is possible that this pattern of looks occurred without the presence of the first instruction, and thus without the L+H* accent being in an infelicitous sentence location. Ito & Speer (2008) found earlier looks for a felicitously accented adjective contrast (green drum ⇒ BLUE_{L+H*} drum_{no,accent}) than for a felicitously accented noun contrast (blue onion ⇒ blue_{H*} DRUM_{L+H*}). It is possible that felicitously accented adjective contrasts are processed faster than felicitously accented noun contrasts merely based on information from the target utterance. In particular, information from the felicitous L+H* accent of the target instruction, which evokes the contrast set, occurs earlier in a early L+H* sequence than a late L+H* sequence, and the
repeated, and thus most predictable, lexical item occurs later in an *early L+H* sequence than a *late L+H* sequence. In other words, the sequence *YELLOW bottle ⇒ PURPLE bottle* may elicit earlier looks to the purple bottle than *purple PENCIL ⇒ purple BOTTLE* merely based on the location of the *L+H* accent and the repeated lexical item in the target instruction. Experiment 3b explores the role that the presence of the first instruction played in this experiment. It compares *early L+H* and *late L+H* sequences with and without the context instruction.

Results from these experiments do, however, provide a cue that the later looks to the target object in the *late L+H* sequences compared to the *early L+H* sequences are related to the location of the processing difficulty. Notice that the presence of the first instruction created the processing difficulty related to the *L+H* accent in the second instruction. Without the first instruction, the *L+H* accent would not evoke a contrast set that can’t be reconciled with the previous instruction. If the earlier looks to the target object in the *early L+H* conditions over the *late L+H* conditions were unrelated to the location of the processing difficulty (created by the presence of the first instruction) and were merely a result of the position of the felicitous *L+H* and repeated lexical item in the target utterance, then the *L-H%* continuation rise in the second instruction should have affected looks to the target of the third instruction equally across conditions. That is, the *L-H%* pattern should either have led to earlier looks to the target object compared to a *L-L%* in both the *early L+H* and *late L+H* conditions, or it should not have affected either condition. A difference in when a processing difficulty occurs can explain why the *L-H%* only modulated looks in the *early L+H* conditions, but not in the *late L+H* conditions.

The results from this experiment have a number of implications. The experiment provided novel evidence that edge tone patterns between utterances could affect reference resolution. This suggests that listeners can use edge tone information to structure the discourse and that a *L-H%* continuation rise can link an utterance more closely to the one following it (cf. Pierrehumbert & Hirschberg, 1990). Furthermore, it seems that the *L-H%*
continuation rise could facilitate re-interpretation of the previous L+H* accent. In the early $L+H^*$ sequences, the L+H* accent first evoked a contrast set that was incompatible with the lexical content of the previous instruction and thus caused processing difficulties. The L-H% continuation rise then provided a cue that the L+H* accent marked a contrast set with the upcoming rather than the preceding instruction. Thus, the L-H% provided the necessary focus on the upcoming instruction to re-interpret the L+H* accent as evoking an upcoming rather than a preceding contrast set. As a result, the contrast set that the L+H* accent of the second instruction evoked could be integrated with the contrast set that the target instruction evoked. No such re-interpretation occurred if the second instruction ended in L-L% sentence-final intonation. As a result, the contrast set that the L+H* accent of the second instruction evoked could not be integrated with the contrast set that the target instruction evoked. Instead, two separate contrast sets were evoked, which increased the processing load and slowed processing times. Re-interpretation also did not occur if the processing difficulty overlapped with edge tone information. This suggests that pitch accents and edge tones interact to affect reference resolution. In particular, pitch accent and edge tone type may structure the discourse frame in which lexical information is interpreted. The interpretation of the prosodic structure is then not just maintained in the discourse representation, it is structuring the discourse representation, foregrounding certain aspects of the discourse over others, setting up expectations of upcoming material, and grouping together constituents.

The data again suggest that a L+H* accent can evoke an upcoming contrast set under certain conditions. In this experiment, the L+H* only evoked an upcoming contrast set if listeners considered information from a L-H% continuation rise that led to its re-interpretation. Compare this to the results from Experiment 2, where the L+H* accent of the initial instruction was never followed by a L-H% continuation rise. It was argued in Experiment 2 that ending the initial instruction in an L-L% pattern contributed to the initial L+H* accent not showing a facilitation effect. The fact that a L+H* accent could evoke
an upcoming contrast set in this experiment implies that there was no single interpretation of L+H* accents. The fact that this happened only in certain situations implies that L+H* accents were interpreted in their lexical and prosodic context rather than in isolation.

The data also suggest that L+H*-related processing difficulties interfered with interpreting immediately following seemingly conflicting prosodic information. The L+H* accent of the second instruction was initially taken to provide information about lexical contrast with the preceding instruction. Since the evoked contrast set did not match lexical information from the previous instruction, it caused processing difficulties (cf. Ito & Speer, 2008; Weber et al., 2006a). A following L-H% pattern provided a cue that the L+H* accent should instead be interpreted as evoking an upcoming contrast set. However, if both the L+H* accent and the L-H% pattern occurred successively on the same lexical item, the L-H% information was not considered (most likely because it occurred while listeners experienced processing difficulties). As a result, the L+H* accent was not re-interpreted as evoking an upcoming contrast set.

The data support the notion that prosody is processed immediately and in parallel with lexical information. In addition, listeners must build and maintain a detailed representation of the focus structure that they derived from the prosodic structure as they process language. This representation includes information about earlier pitch accents and edge tones. The pitch accent and edge tone information must further be kept in the discourse representation for long enough to affect upcoming utterances. Notice in particular that the early L+H* / continuation condition elicited earlier looks to the target object than the early L+H* / finality condition. These conditions only differed in the prosody of the instruction preceding the target instruction. That is, earlier looks to the target object were found in the early L+H* / continuation condition compared to the early L+H* / finality condition even though the target utterance had identical prosodic patterns in the two conditions. This suggests that listeners kept in their discourse representation not only lexical information, but also pitch accent and edge tone information, from preceding utterances.
Experiment 3a found later looks to the target object, the purple bottle, for late L+H* sequences (Click on the green pencil. Click on the purple PENCIL. Click on the purple BOTTLE) compared to early L+H* sequences (Click on the yellow scissors. Click on the YELLOW bottle. Click on the PURPLE bottle). The results from Experiment 3a suggested that this effect was due to the timing of earlier processing difficulties.

Experiment 3a suggested that listeners had fully recovered from the processing difficulties when they encountered the target noun phrase in the early L+H* sequences, but not the late L+H* sequences. As a result, the early L+H* sequences elicited earlier looks to the target object than the late L+H* sequences. However, Ito & Speer (2008) found earlier looks to felicitously accented color contrasts compared to felicitously accented object contrasts. The early L+H* sequences have a felicitously accented color contrast and the late L+H* sequences have a felicitously accented object contrast. The earlier looks to the target object in the early L+H* sequences compared to the late L+H* sequences may thus be solely due to the felicitously accented color contrast compared to a felicitously accented object contrast rather than the timing of earlier processing difficulties.

Experiment 3b is a follow-up experiment for Experiment 3a. It tests more directly the effect of the timing of earlier processing difficulties by comparing the early L+H* / finality and late L+H* / finality conditions from Experiment 3a with sequences lacking the initial instruction (these sequences have an instruction following the target instruction to ensure that all trials contained three instructions). If the initial instruction is missing, the L+H*
accent in the instruction preceding the target instruction cannot initially be taken to be in-
felicitously located and should not cause processing difficulties. Experiment 3a suggested
that in the early L+H* / finality condition, the processing difficulties were resolved by the
time listeners encountered the target noun phrase. The presence or absence of the first
instruction should therefore not affect looks to the target object. The processing difficul-
ties in the late L+H* / finality condition, on the other hand, were ongoing when listeners
encountered the target noun phrase. In this case, the sequence should elicit earlier looks
to the target object when the initial instruction is absent compared to when it is present.
Together these results would suggest that the later looks to the target in the late L+H* se-
quencies compared to the early L+H* sequences were (at least partially) due to processing
difficulties from a late L+H* accent.

Thus, Experiment 3b asks the following research question: Can the earlier looks to the
target object in the early L+H* sequences compared to late L+H* sequences in Experiment 3a
at least partially be attributed to the relative occurrence in time of L+H*-related processing
difficulties? It is predicted that the relative occurrence in time of processing difficulties
did indeed contribute to the earlier looks in the early L+H* sequences compared to the late
L+H* sequences of Experiment 3a.

### 6.1 Experimental Conditions

Tables 6.1 and 6.2 show the critical conditions for Experiment 3b. All the data for Ex-
periment 3b comes from EXPERIMENT A. The early L+H* / finality and late L+H* / finality
conditions from Experiment 3a also occur in this experiment. They are renamed as early
L+H* / target 3 and late L+H* / target 3. The finality conditions in Experiment 3a received
this name because they involved only final fall intonation and were compared with condi-
tions involving a continuation rise. Here, these same conditions receive the name target 3
because the third instruction is the target utterance and they are compared with conditions where the second instruction is the target utterance (these conditions are called target 2).

Table 6.1 shows the early L+H* conditions, whereas Table 6.2 shows the late L+H* conditions. As mentioned above, the early L+H* / target 3 and late L+H* / target 3 conditions also occurred in Experiment 3a. Here, the target utterance is Instruction 3. The target utterance for the early L+H* / target 2 and late L+H* / target 2 conditions is Instruction 2. The same lexical contrasts and prosodic patterns occurring in Instructions 1 and 2 in the target 2 sequences occurred in Instructions 2 and 3 in the target 3 sequences. The target 3 conditions contained a L+H* accent which should create processing difficulties, whereas the same L+H* accent in the target 2 conditions should not create processing difficulties.
6.2 Results

As in Experiment 3a, only looks to the target object will be reported here since no condition is expected to elicit a garden-path effect. Figure 6.1 provides an overview of the results. It shows looks to the target object cell in the four critical conditions of Experiment 3b (see Tables 6.1 and 6.2). The figure reveals the predicted pattern of responses. It shows numerically later looks to the target object in the late $L+H^*$, target 3 condition compared to the late $L+H^*$, target 2 condition (where the processing difficulties were suggested to not be resolved by the time the target noun phrase is encountered), but no difference in looks to the target object in the two early $L+H^*$ conditions (where the processing difficulties were suggested to be resolved by the time the target noun phrase is encountered). In addition, looks to the target object rise numerically earlier in the early $L+H^*$ sequences compared to the late $L+H^*$ sequences, confirming the results found in Ito & Speer (2008) that a felicitously accented color contrast elicits earlier looks to the target object than a felicitously accented object contrast.

Figure 6.2 shows the grand mean of looks to the target object for the four critical conditions of Experiment 3b. The figure shows that looks to the target object start rising at -85 ms (the last data point level with the previous one). 850 ms was chosen as the peak of the curve because the curve falls after 850 ms. Notice, however, that 782 ms represents the highest data point.

The results from an omnibus GCA are shown in Table 6.3. The table reveals a main effect of $L+H^*$ Position on the intercept and linear terms. This suggests earlier looks, but with a shallower rise in looks, to the target object for the early $L+H^*$ conditions over the late $L+H^*$ conditions. This result confirms data from Ito & Speer (2008) that felicitously accented color contrasts elicit earlier looks to the target object than felicitously accented object contrasts. There is also a main effect of Target Instruction on the intercept term. This suggests an advantage for the target 2 conditions over the target 3 conditions. Such an
Figure 6.1: Experiment 3b, data from EXPERIMENT A: Proportion of looks to the target object cell in the four critical conditions presented in Tables 6.1 and 6.2.
advantage would be expected if the initial instruction in the target 3 conditions caused processing difficulties related to an initially seemingly infelicitously placed L+H* accent. The analysis showed no L+H* Position / Target Instruction interaction.

Figures 6.3 and 6.4 show a direct comparison of the two target 3 (early L+H*, target 3 and late L+H*, target 3) and the two target 2 conditions (early L+H*, target 2 and late L+H*, target 2), respectively. Both figures show the predicted earlier looks to the target object in the early L+H* conditions over the late L+H* conditions. The comparison of the two target 3 conditions shown in Figure 6.3 is the same as that shown in Figure 5.6 of Experiment 3a. There are more looks to the target object of Instruction 3 in the early L+H* / target 3 condition over the late L+H* / target 3 condition for a window spanning from -68 ms to 527 ms (cf. Experiment 3a). Figure 6.4 shows more looks to the target object of Instruction 2 in the early L+H* / target 2 condition compared to the late L+H* / target 2 condition, but for a smaller window, ranging from 51 ms to 374 ms. The comparison between the two target
Table 6.3: Experiment 3b, data from EXPERIMENT A (see Figure 6.1): Target fixation analysis: Main effects ($L+H^*$ Position and Target Instruction) and interaction.

<table>
<thead>
<tr>
<th>Model</th>
<th>L+H* Position</th>
<th>Target Instruction</th>
<th>Interaction</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2LL</td>
<td>ΔD</td>
<td>p&lt;</td>
</tr>
<tr>
<td>Base</td>
<td>16065.4</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>16043.98</td>
<td>-21.41</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>16037.4</td>
<td>-6.58</td>
<td>0.05</td>
</tr>
<tr>
<td>Quadratic</td>
<td>15839.54</td>
<td>-197.86</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>15606.13</td>
<td>-233.41</td>
<td>0.0001</td>
</tr>
</tbody>
</table>
2 conditions confirms results from Ito & Speer (2008) that looks to the target object are indeed earlier for felicitously accented color contrasts than for felicitously accented object contrasts.

Results from GCAs comparing the two target 3 conditions and the two target 2 conditions are shown in Tables 6.4 and 6.5, respectively. Both tables show a reliable effect on the intercept term, confirming the earlier looks to the target object in the early $L+H^*$ conditions compared to the late $L+H^*$ conditions seen in Figures 6.3 and 6.4. In addition, there is a reliable effect on the linear term comparing the two target 2 conditions, but not comparing the two target 3 conditions. In particular, looks to the target object started rising later, but peaked slightly earlier and higher in the late $L+H^*$ / target 2 condition compared to the early $L+H^*$ / target 2 condition. The GCA in Table 6.4 compares the same conditions as the GCA in Table 5.7 of Experiment 3a. Notice that the log likelihood and change in deviance in the two tables differs. The GCA in Table 5.7 was run without the subjects fixed-effect because
some of the comparisons in Experiment 3a were between-subjects. In this experiment, all comparisons are within-subject, and the subjects fixed-effects were included, resulting in different log likelihoods and changes in deviance. Notice, however, that both GCAs yield the same results: a reliable effect on the intercept term, but not the linear term.

The results from the two GCAs suggest that felicitously accented color contrasts (as in the early L+H* conditions) did elicit earlier looks to the target object than felicitously accented object contrasts (as in the late L+H* conditions) (cf. Ito & Speer, 2008). This effect is possibly related to the position of the L+H* accent and the repeated lexical item in the early L+H* compared to the late L+H* conditions. In the early L+H* conditions, the target utterance received a L+H* accent on the adjective, which created the expectation of a repeated noun and led to early looks to the target object. In the late L+H* conditions, the target utterance received a L+H* accent on the noun, which felicitously marked the novel object type following a repeated adjective. If the felicitous L+H* accent is on the final lex-
Table 6.4: Experiment 3b, Target 3 sequence from EXPERIMENT A (see Figure 6.3): Target fixation analysis comparing the early L+H*, target 3 and late L+H*, target 3 conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>∆D</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>6859.26</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>6821.29</td>
<td>−37.97</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>6817.78</td>
<td>−3.51</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>6701.86</td>
<td>−115.91</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>6584.38</td>
<td>−117.49</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 6.5: Experiment 3b, Target 2 sequence from EXPERIMENT A (see Figure 6.4): Target fixation analysis comparing the early L+H*, target 2 and late L+H*, target 2 conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>∆D</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>7473.37</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>7466.96</td>
<td>−6.41</td>
<td>0.05</td>
</tr>
<tr>
<td>Linear</td>
<td>7461.17</td>
<td>−5.79</td>
<td>0.05</td>
</tr>
<tr>
<td>Quadratic</td>
<td>7346.29</td>
<td>−114.88</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>7190.3</td>
<td>−155.99</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

In the context of the utterance, any facilitation effect related to this accent should occur later than if the felicitous L+H* accent is on the penultimate lexical item. Thus, later looks were expected if the felicitous L+H* accent was on the last compared to the penultimate lexical item. The results suggest that the earlier looks to the target object in the early L+H* conditions compared to the late L+H* conditions were then at least partially due to the kind of felicitously accented lexical contrast (color contrast vs. object contrast). The following two comparisons will reveal if, as predicted, the timing of earlier processing difficulties were also involved in eliciting earlier looks in the early L+H* conditions compared to the late L+H* conditions.

Figures 6.5 and 6.6 show a direct comparison of the two early L+H* (early L+H*, target 3 and early L+H*, target 2) and the two late L+H* conditions (late L+H*, target 3 and late L+H*, target 2), respectively. Figure 6.5 shows no reliable differences between the two early L+H* conditions across the time window. Figure 6.6 shows a number of windows with reliably more looks to the target object in the late L+H*, target 2 than the late L+H*, target 2.
Figure 6.5: Experiment 3b, Early cue sequences from EXPERIMENT A: Comparison of the early $L+H^*$, target 3 and early $L+H^*$, target 2 conditions presented in Table 6.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

3 conditions. The windows span from the beginning of the analysis window at -85 ms to 51 ms, from 221 ms to 255 ms, from 425 ms to 510 ms, and from 663 ms to 816 ms. Thus, we do find more looks to the target object in the late $L+H^*$ / target 2 condition compared to the late $L+H^*$ / target 3 condition for a total of about half the analysis window. Unlike most previous comparisons, however, there is not one large region revealing the advantage for the late $L+H^*$ / target 2 condition over the late $L+H^*$ / target 3 condition. Rather, the advantage is seen in several smaller regions.

Results from a GCA comparing the two early $L+H^*$ conditions are shown in Table 6.6 and, as predicted, reveal no reliable differences between the two conditions. Listeners were predicted to not experience processing difficulties while processing the target noun phrase in both conditions. Therefore, no differences in eye movement patterns between the conditions were expected.

Table 6.7 shows results from a GCA comparing the two late $L+H^*$ conditions. The table
Figure 6.6: Experiment 3b, Late cue sequences from EXPERIMENT A: Comparison of the late L+H*, target 3 and late L+H*, target 2 conditions presented in Table 6.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

shows a reliable effect on the intercept term, but not the linear term, suggesting a difference in area under the curve, but no difference in slope of the curves. In particular, the area under the curve is larger in the late / target 2 condition than the late L+H* / target 3 condition, suggesting earlier looks to the target object in the late / target 2 compared to the late / target 3 condition.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>6855.07</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>6854.38</td>
<td>-0.69</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>6854.26</td>
<td>-0.12</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>6853.64</td>
<td>-0.62</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>6853.29</td>
<td>-0.34</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 6.6: Experiment 3b, Early L+H* sequences from EXPERIMENT A (see Figure 6.5): Target fixation analysis comparing the early L+H*, target 3 and early cye, target 2 conditions.
<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>6907</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>6892.21</td>
<td>−14.79</td>
<td>0.0001</td>
</tr>
<tr>
<td>Linear</td>
<td>6891.12</td>
<td>−1.09</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>6891.09</td>
<td>−0.04</td>
<td>n.s.</td>
</tr>
<tr>
<td>Cubic</td>
<td>6888.06</td>
<td>−3.02</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 6.7: Experiment 3b, Late L+H* sequences from EXPERIMENT A (see Figure 6.6): Target fixation analysis comparing the late L+H*, target 3 and late L+H*, target 2 conditions.

6.3 Discussion

Experiment 3b was a follow-up to Experiment 3a. Experiment 3a found earlier looks to the target object in the early L+H* conditions (yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle with either a L-L% or a L-H% pattern on bottle in Instruction 2) over the late L+H* conditions (green pencil ⇒ purple PENCIL ⇒ purple BOTTLE with either a L-L% or a L-H% pattern on PENCIL in Instruction 2). The data from Experiment 3a suggested that, depending on their occurrence in time, processing difficulties either interfered with target noun phrase processing (as in the late L+H* conditions) or were resolved when the target noun phrase was processed (as in the early L+H* conditions).

However, it is possible that the earlier looks to the target object in the early L+H* conditions compared to the late L+H* conditions were due solely to the lexical and L+H* information of Instruction 3 (and were unrelated to the occurrence in time of processing difficulties). While the effect of edge tone pattern (L-L% vs. L-H%) in Experiment 3a (which only modulated looks in the early L+H*, but not late L+H* conditions) supported the notion that the timing of processing difficulties in Instruction 2 affected looks to the target object of Instruction 3, these conditions were between-subjects and between-speaker. To avoid potential effects stemming from differences among subjects or stimuli durations, this experiment used within-subjects experimental conditions to support the idea that the pattern of looks found in Experiment 3a was related to when processing difficulties occurred.
in Instruction 2.

Experiment 3b compared the two finality conditions of Experiment 3a with and without the original first instruction (the context instruction). If the original context instruction was missing, the original second instruction did not have an infelicitously placed $L+H^*$ accent (since the $L+H^*$ accent was infelicitous with respect to the lexical information of the original context instruction) and did not cause processing difficulties. If the context instruction was present, the second instruction contained a seemingly infelicitous $L+H^*$ accent. This $L+H^*$ accent occurred early and was fully resolved when the target noun phrase was processed in the *early $L+H^*$ / finality* condition, but it occurred late and was not fully resolved when the target noun phrase was processed in the *late $L+H^*$ / finality* condition. Thus, in both *early $L+H^*$* conditions of Experiment 3b, no processing difficulties occurred while listeners encountered the target noun phrase, and no difference in looks was expected. But in both *late $L+H^*$* conditions of Experiment 3b, ongoing processing difficulties led to later looks to the target object in the *late $L+H^*$ / target 3* condition compared to the *late $L+H^*$ / target 2* conditions.

Experiment 3b showed the expected null-effect between the *early $L+H^*$* conditions. This suggests that neither condition is affected by $L+H^*$-related processing difficulties. The experiment also revealed, as predicted, reliably more looks to the target object in the *late $L+H^*$ / target 2* condition compared to the *late $L+H^*$ / target 3* condition. Since the conditions differed only in the presence or absence of the original first instruction, the differences in looks had to be related to the processing difficulties that the original first instruction set up.

The data from Experiment 3b also revealed that looks to the target object were indeed earlier across the board (i.e. independently of the presence or absence of the original first instruction) for felicitously-accented color contrasts than for felicitously-accented object contrasts. However, the window for which the advantage was reliable is smaller in the *target 2* conditions than the *target 3* conditions.
Overall, Experiment 3b supports the claims made in Experiment 3a. In particular, it is the case that looks to the target object were earlier across the board for felicitously-accented color contrasts compared to felicitously-accented object contrasts. In addition, the timing of infelicitous L+H* information in the original second instruction modulated responses.

Experiment 3b confirms that listeners built a detailed representation of prosodic structure that was maintained in a discourse representation long enough to structure expectations and anticipate the discourse referents that will appear in upcoming utterances. In the case of Experiments 3a and 3b, the combination of two preceding instructions affected processing of the target instruction. In particular, the effects found in Experiment 3b were based solely on the presence or absence of the original first instruction. This suggests that listeners maintain a discourse representation containing lexical and prosodic information from at least the two utterances preceding the current one. The data again suggest that pitch accents are not processed in isolation. Rather, relevant preceding lexical and prosodic information affects their interpretation.
Experiments 3a and 3b revealed that pitch accents and lexical contrast interacted with edge tones to influence reference resolution. Experiment 3a manipulated lexical contrast and pitch accent placement together: A change in lexical contrast always coincided with a change in pitch accent location to ensure that the pitch accent information always felicitously marked the lexical contrast between the last two instructions. Thus, only edge tones were manipulated independently. Experiment 3a showed that the *early* $L+H^*$ sequences, with accented adjectives, elicited earlier looks to the target object than the *late* $L+H^*$ sequences, with accents on nouns. It was proposed that the $L+H^*$ accent of the second instruction led to initial processing difficulties when first encountered in both the *early* $L+H^*$ and *late* $L+H^*$ sequences because it was infelicitous with respect to the lexical contrast between the first and second instructions (i.e. a repeated color adjective or object noun carried $L+H^*$; cf. Ito & Speer, 2008; Weber et al., 2006a). It was further proposed that these processing difficulties were resolved later for *late* $L+H^*$ than for *early* $L+H^*$ sequences, leading to reduced influence of information from the second instruction on processing of the third, target instruction. This was shown by the finding that a $L-H^\%$ continuation rise speeded reference resolution compared to $L-L^\%$ sentence-final intonation only in the *early* $L+H^*$ sequences.

Let’s consider the *early* $L+H^*$ / *continuation* condition from Experiment 3a (*yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle*). Here, the $L+H^*$ accent on the adjective in the second instruction felicitously marked the color contrast between the second and third instructions. In addition, the $L-H^\%$ continuation rise indicated that the second instruction
should be interpreted with respect to the following instruction. Notice that the edge tones are proposed guide the grouping of instructions for interpretation, suggesting that the second instruction should be interpreted with the third instruction. The edge tones do not, however, provide information as to how to interpret the following instruction. The L+H* accents provide this information. There are two possibilities for how the L-H% continuation rise of the early L+H*/continuation condition speeded reference resolution compared to the L-L% sentence-final intonation of the early L+H*/finality condition. (1) The L+H* accent on the repeated adjective of the second instruction conflicted with the first instruction (yellow scissors ⇒ YELLOW bottle). Both this initial conflict and the L-L% finality contour at the end of the first instruction might have signalled that the first two instructions should not be interpreted together. In addition, the L+H* on the adjective in the second instruction could be re-interpreted as providing information about an upcoming lexical contrast once the L-H% continuation rise (which groups the second and third instructions) was encountered. (2) Alternatively, the advantage for the early L+H*/continuation condition compared to the early L+H*/finality condition could be related solely to the L+H* accent of the third instruction. Experiment 1 as well as Ito & Speer (2008) and Weber et al. (2006a) found that a L+H* accent in the target instruction that felicitously marked a color contrast (as in yellow bottle ⇒ PURPLE bottle) speeded reference resolution compared to other prosodic patterns. Ending the instruction just prior to the target instruction with a L-H% continuation rise (as in the early L+H*/continuation condition) compared to L-L% sentence-final intonation (as in the early L+H*/finality condition) may have increased the facilitatory nature of the L+H* accent of the target instruction. That is, it could be the L+H* accent in the second or third instruction that listeners used to determine how Instructions 2 and 3 were related (if a L-H% pattern suggests that they should be interpreted together).

Experiment 4 explored whether the L+H* accent of the second instruction played a role in speeding looks to the target object in the early L+H*/continuation condition as compared to the early L+H*/finality condition of Experiment 3a. The early L+H* sequences from
Experiment 3a (yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle) are compared with sequences that differ from the early L+H* sequences only in that the L+H* accent in the second instruction was placed on the noun rather than the adjective (yellow scissors ⇒ yellow BOTTLE ⇒ PURPLE bottle). To avoid confusion with the late L+H* sequences of Experiment 3a, these sequences are called no early L+H* sequences. The second instruction in the no early L+H* sequences ends either in a L-L% pattern (the no early L+H* / finality condition) or a L-H% pattern (the no early L+H* / continuation condition).

According to our interpretation of the previous results, when the L+H* accent of the second instruction is on the adjective (as in the early L+H* sequences), it generates a contrast set. When the L-H% boundary is encountered, the pitch accent – edge tone combination creates the expectation of lexical contrast between the second and third instructions (even though in the context of the yellow scissors from the first instruction it may initially generate processing difficulty (see above and Experiment 3a)). That is, after an initial processing difficulty, the L+H* accent could be re-interpreted as providing information about the upcoming instruction when hearing a L-H% continuation rise compared to L-L% sentence-final intonation, ”announcing” the upcoming color contrast. The L+H* accent of the second instruction in combination with a following L-H% pattern should then provide a context that will speed reference resolution in the target utterance.

When the L+H* accent of the second instruction is on the noun (as in the no early L+H* sequences – yellow scissors ⇒ yellow BOTTLE ⇒ PURPLE bottle), it felicitously marks the lexical contrast between the first and second instructions. In this case, it should not cause processing difficulty when it is initially encountered (since it felicitously marks the lexical contrast between Instructions 1 and 2). However, in this case, the L+H* accent of the second instruction would be infelicitous if it was interpreted as suggesting an upcoming lexical contrast. In the example, it should generate a contrast set of objects of the previously mentioned color (yellow). This set would be incompatible with the contrast set generated by the L+H* on the adjective in the third instruction (in the example, PURPLE should gen-
erate a contrast set of differently colored bottles). That is, if it was the L+H* accent of the second instruction that drove the advantage for the early L+H* / continuation condition compared to the early L+H* / finality condition in Experiment 3a, then there should be no advantage for the no early L+H* / continuation condition over the no early L+H* / finality condition. But if it was the L+H* accent of the third instruction that drove the advantage for the early L+H* / continuation condition compared to the early L+H* / finality condition in Experiment 3a, then the no early L+H* / continuation condition should speed reference resolution compared to the no early L+H* / finality condition. Thus, Experiment 4 asks the questions: How much of the observed advantage of the early L+H* sequences in Experiment 3a was due to the L+H* pitch accent on the adjective of the second instruction felicitously marking the color contrast between the second and third instructions?

Earlier looks to the target object (in the third instruction) are predicted if the L+H* accent of the second instruction is on the adjective (and felicitously marks the upcoming lexical contrast) and the second instruction ends in a L-H% pattern as compared to conditions where either only one is the case (i.e. where either the L+H* accent of the second instruction is on the adjective or the second instruction ends in a L-H% pattern, but not both) or where neither is the case (i.e. where the second instruction neither has a L+H* accent on the adjective nor a L-H% continuation rise). Such a result would suggest that the L+H* accent in the second instruction interacted with edge tone patterns to affect reference resolution.

7.1 Experimental Conditions

Tables 7.1 and 7.2 show the experimental conditions for Experiment 4. Table 7.1 repeats the early L+H* sequences from Experiment 3a (early L+H* / finality and early L+H* / continuation). Here, the second instruction has a L+H* accent on the adjective, which felicitously
<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1: Context</th>
<th>Instruction 2: Object Contrast</th>
<th>Instruction 3: Color Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>early L+H* / finality</td>
<td>Click on the yellow scissors ↓</td>
<td>Click on the YELLOW bottle ↓</td>
<td>Click on the PURPLE bottle ↓</td>
</tr>
<tr>
<td>early L+H* / continuation</td>
<td>Click on the yellow scissors ↓</td>
<td>Click on the YELLOW bottle ↑</td>
<td>Click on the PURPLE bottle ↓</td>
</tr>
</tbody>
</table>

Table 7.1: Experiment 4, Early L+H* sequence from EXPERIMENTS A and B: CAPS indicate L+H* accents; Italics on an adjective indicate H* accents; Italics on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. Instructions ending in a L-L% pattern are shown with a ↓), whereas instructions ending in a L-H% pattern are shown with a ↑).

marks the upcoming color contrast. In addition, the L-H% pattern ending the second instruction in the *early L+H* / *continuation* condition is expected to more closely link the last two utterances than the L-L% pattern ending the second instruction in the *early L+H* / *finality* sequences. Experiment 3a showed that looks to the target object of the third instruction were speeded if the second instruction ended in a L-H% continuation rise compared to L-L% sentence-final intonation.

Table 7.2 shows sequences where the second instruction has a L+H* accent on the noun, felicitously marking the object contrast between the first and second instructions. As mentioned above, these sequences are called *no early L+H* acet Here, the L+H* accent of the second instruction felicitously marks the preceding lexical contrast. Edge tone patterns are not predicted to modulate responses if the facilitatory effect of the L-H% in the *early L+H* sequences of Experiment 3a was related to the position of the L+H* accent of the second instruction. However, edge tone patterns should modulate responses if the facilitatory effect of the L-H% in the *early L+H* sequences of Experiment 3a was related to the position of the L+H* accent of the third instruction.

---

1This name was chosen because the second instruction contains no L+H* accent on the adjective. Rather, the L+H* accent occurs later, on the noun, where it felicitously marks the lexical contrast between the first and second instructions. The sequences were not called *late L+H* only to not be confused with the *late L+H* sequences in Experiment 3a.
Table 7.2: Experiment 4, No early L+H* sequence from Experiments A and B: CAPS indicate L+H* accents; *Italics* on an adjective indicate H* accents; *Italics* on a noun indicate !H* accents. The instruction in **bold face** is the target instruction. Instructions ending in a L-L% pattern are shown with a \(\downarrow\), whereas instructions ending in a L-H% pattern are shown with a \(\uparrow\).

<table>
<thead>
<tr>
<th>Condition Name</th>
<th>Instruction 1: Context</th>
<th>Instruction 2: Object Contrast</th>
<th>Instruction 3: Color Contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>no cue / finality</td>
<td>Click on the yellow scissors (\downarrow)</td>
<td>Click on the yellow BOTTLE (\downarrow)</td>
<td>Click on the PURPLE bottle (\downarrow)</td>
</tr>
<tr>
<td>no cue / continuation</td>
<td>Click on the yellow scissors (\downarrow)</td>
<td>Click on the yellow BOTTLE (\uparrow)</td>
<td>Click on the PURPLE bottle (\downarrow)</td>
</tr>
</tbody>
</table>

7.2 Results

As in Experiments 3a and 3b, only looks to the target object will be reported here since no condition is expected to elicit a garden-path effect. Figure 7.1 provides an overview of the results. It shows looks to the target object cell in the four critical conditions of Experiment 4 (see Tables 7.1 and 7.2). The figure shows the predicted pattern of responses in that looks to the target object rise earliest in the *early L+H* / *continuation* condition.

Figure 7.2 shows the grand mean of looks to the target object for the four critical conditions of Experiment 4. The figure shows that looks to the target object start rising at -204 ms (the last data point level with the previous one). 782 ms was chosen as the peak of the curve because the curve falls after 782 ms. In addition, 782 ms represents the highest data point. The analysis window for the following GCA will thus range from -204 ms to 782 ms.

The results from an omnibus GCA are shown in Table 7.3. The table reveals a main effect of *Edge Tone* on the intercept and linear terms. This suggests earlier looks to the target object, but with a shallower slope, in the *continuation* compared to the *finality* conditions. There is neither a main effect of *Accent Location* nor an *Edge Tone* / *Accent Location* interaction on the intercept or linear terms. Contrary to expectations, this result suggests that the facilitation effect found for the *early L+H* / *continuation* condition compared to the
Figure 7.1: Experiment 4, data from EXPERIMENTS A and B: Proportion of looks to the target object cell in the four critical conditions presented in Tables 7.1 and 7.2.
Figure 7.2: Experiment 4, data from EXPERIMENTS A and B: Grand mean of looks to the target object cell in the four critical conditions presented in Tables 7.1 and 7.2. The arrows indicate the beginning and end of the analysis window for the GCA.

*early L+H* / finality condition of Experiment 3a was related to the L+H* accent of the *third*, rather than the *second* instruction. If the facilitation effect was due to the L+H* accent of the *second* instruction, a main effect of both *Edge Tone* and *Accent Location* as well as an *Edge Tone / Accent Location* interaction were expected (since the *early L+H* / continuation condition should have elicited earlier looks to the target object than all other conditions).

Figures 7.3 and 7.4 show a direct comparison of the two *finality* conditions (*early L+H* / finality and *no early L+H* / finality) and the two *continuation* conditions (*early L+H* / continuation and *no early L+H* / continuation), respectively. The figures show no reliable differences between the two *finality* and the two *continuation* conditions across the window of analysis, confirming the null effect of *Accent Location* found in the omnibus GCA in Table 7.3. This result suggests that, contrary to predictions, the position of the L+H* accent in the second instruction did not affect looks to the target object. Notice, however, that the curves for the two *finality* conditions differed from those for the two *continuation* conditions. While there
<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>18292.51</td>
<td>—</td>
<td>—</td>
<td>18064.73</td>
<td>—</td>
<td>—</td>
<td>18050.44</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>18285.14</td>
<td>-7.37</td>
<td>0.01</td>
<td>18064.21</td>
<td>-0.53</td>
<td>n.s.</td>
<td>18049.49</td>
<td>-0.95</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>18268.62</td>
<td>-16.52</td>
<td>0.0001</td>
<td>18064.18</td>
<td>-0.02</td>
<td>n.s.</td>
<td>18049.49</td>
<td>-0</td>
<td>n.s.</td>
</tr>
<tr>
<td>Quadratic</td>
<td>18144.32</td>
<td>-124.3</td>
<td>0.0001</td>
<td>18053.36</td>
<td>-10.82</td>
<td>0.01</td>
<td>18042.59</td>
<td>-6.9</td>
<td>0.01</td>
</tr>
<tr>
<td>Cubic</td>
<td>18064.73</td>
<td>-79.58</td>
<td>0.0001</td>
<td>18050.44</td>
<td>-2.92</td>
<td>n.s.</td>
<td>18041.65</td>
<td>-0.94</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

Table 7.3: Experiment 4, data from EXPERIMENTS A and B (see Figure 7.1): Target fixation analysis: Main effects (Edge Tone and Accent Location) and interaction.
Figure 7.3: Experiment 4, Finality sequence from EXPERIMENT A: Comparison of the early \(L+H^* / \text{finality}\) and no early \(L+H^* / \text{finality}\) conditions presented in Tables 7.1 and 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

was no numeric advantage for the early \(L+H^* / \text{finality}\) condition compared to the no early \(L+H^* / \text{finality}\) condition, looks in the early \(L+H^* / \text{continuation}\) condition rose numerically earlier than in the no early \(L+H^* / \text{continuation}\) condition. This is consistent with the view that, in the early \(L+H^* / \text{continuation}\) condition, a single contrast set was generated by the presence of the \(L+H^*\) on the adjective of the second instruction and maintained across the \(L-H\%\) continuation rise, allowing for anticipation of an additional member of the contrast set in instruction 3.

The results from Experiment 2 suggested that two successive \(L+H^*\) accents marking a color contrast (as in the anticipatory condition of Experiment 2 and the early \(L+H^*\) conditions in this experiment) may need to be connected by a L-H% continuation rise to facilitate reference resolution. The numerically earlier looks to the target object in the early \(L+H^* / \text{continuation}\) condition compared to the no early \(L+H^* / \text{continuation}\) condition support this idea.
Figure 7.4: Experiment 4, Continuation sequence from EXPERIMENT B: Comparison of the early L+H* / continuation and no early L+H* / continuation conditions presented in Tables 7.1 and 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Results from GCAs comparing the two finality conditions and the two continuation conditions are shown in Tables 7.4 and 7.5, respectively. Neither table shows a reliable effect on the intercept or linear terms. This confirms that, contrary to prediction, the L+H* pitch accent on the adjective in the early L+H* conditions does not facilitate reference resolution compared to a L+H* accent on the noun in the no early L+H* conditions, even though it felicitously marks the upcoming lexical contrast. This also means that the early L+H* accent in the second instruction of Experiment 3a did not contribute to the earlier looks to the target object in the late L+H* compared to the early L+H* sequences. It is possible that a L+H* accent on the adjective of the second instruction followed by a L-H% pattern (as in the early L+H* / continuation) did not reliably speed looks to the target object compared to a L+H* accent on the noun of the second instruction followed by a L-H% pattern (no early L+H* / continuation) because of the proposed initial processing difficulty in the former case.

Figures 7.5 and 7.6 show a direct comparison of the two early L+H* conditions (early
L+H* / finality and early L+H* / continuation) and the two no early L+H* conditions (no early L+H* / finality and no early L+H* / continuation), respectively. Both figures show earlier looks to the target object in the continuation compared to the finality conditions. Figure 7.5 shows a window ranging from the beginning of the analysis window at -204 ms to 323 ms with a reliable advantage for the early L+H* / continuation condition over the early L+H* / finality condition. Figure 7.6 shows reliably more looks to the target object in the no early L+H* / continuation compared to the no early L+H* / finality condition for a slightly smaller time window – from the beginning of the analysis window at -204 ms to 170 ms (with the exception of the data point at -51 ms). In addition, the figure shows a late advantage (from 578 ms to 663 ms) for the no early L+H* / finality condition over the no early L+H* / continuation condition. As proposed in Experiment 3a, the late advantage is most likely due to individual differences in participants.

Results from GCAs comparing the two early L+H* conditions and the two no early L+H*
Figure 7.5: Experiment 4, Early L+H* sequence from EXPERIMENTS A and B: Comparison of the early L+H* / finality and early L+H* / continuation conditions presented in Table 7.1. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.

Figure 7.6: Experiment 4, No early L+H* sequence from EXPERIMENTS A and B: Comparison of the no early L+H* / finality and no early L+H* / continuation conditions presented in Table 7.2. Left: Empirical logit of looks to the target object cell and error bars. Right: Model fits.
Table 7.6: Experiment 4, Early L+H* sequence from EXPERIMENTS A and B (see Figure 7.5): Target fixation analysis comparing the early L+H* / finality and early L+H* / continuation conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>9113.17</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>9107.95</td>
<td>-5.22</td>
<td>0.05</td>
</tr>
<tr>
<td>Linear</td>
<td>9099.44</td>
<td>-8.51</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>9003.91</td>
<td>-95.54</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>8954.35</td>
<td>-49.56</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

conditions are shown in Tables 7.6 and 7.7, respectively. Table 7.6 shows a reliable effect on both the intercept and linear terms. This reveals that, as predicted, looks to the target object occurred earlier, but with a shallower slope, in the early L+H* / continuation condition compared to the early L+H* / finality condition (cf. Experiment 3a). Figure 7.5 shows that the curve for the early L+H* / continuation condition rises earlier, but peaks at a similar time, than the early L+H* / finality condition, resulting in both a difference in area under the curve and slope between the two curves. Thus, if there is a color contrast between the second and third instructions, ending the second instruction with a continuation rise compared to sentence-final intonation facilitates resolving reference for the following, i.e. third, instruction.

Table 7.7 shows a reliable effect on the linear term, but not the intercept term. This suggests a difference in slope between the no early L+H* / finality and no early L+H* / continuation conditions, but no difference in area under the curve. This pattern of results suggests that the curves cross towards the middle of the analysis window, resulting in an early advantage for one condition and a late advantage for the other. In this case, there is an early advantage for the no early L+H* / continuation condition over the no early L+H* / finality condition, but a late advantage for the no early L+H* / finality condition over the no early L+H* / continuation condition.

Together, the results suggest that a L-H% continuation rise created a greater advantage compared to L-L% sentence-final intonation in the early L+H* sequences compared to the
Table 7.7: Experiment 4, No early L+H* sequence from EXPERIMENTS A and B (see Figure 7.6): Target fixation analysis comparing the no early L+H* / finality and no early L+H* / continuation conditions.

<table>
<thead>
<tr>
<th>Model</th>
<th>-2LL</th>
<th>ΔD</th>
<th>p&lt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base</td>
<td>9162.83</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Intercept</td>
<td>9160.41</td>
<td>-2.42</td>
<td>n.s.</td>
</tr>
<tr>
<td>Linear</td>
<td>9152.39</td>
<td>-8.02</td>
<td>0.01</td>
</tr>
<tr>
<td>Quadratic</td>
<td>9116.14</td>
<td>-36.24</td>
<td>0.0001</td>
</tr>
<tr>
<td>Cubic</td>
<td>9084.81</td>
<td>-31.33</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

no early L+H* sequences. This difference in advantage was measurable in the different effects on the intercept term of the GCA in Tables 7.6 and 7.7. Both GCAs showed a reliable effect on the linear term, suggesting an early advantage for the continuation conditions over the finality conditions. The reliable effect on the intercept term only comparing the early L+H* / finality and early L+H* / continuation conditions, but not comparing the no early L+H* / finality and no early L+H* / continuation conditions, suggests that the L-H% advantage was greater (i.e. great enough to result in a larger area under the curve) in the early L+H* conditions than in the late L+H* conditions.

Unlike the results from the two previous GCAs (shown in Tables 7.4 and 7.5 above), these results do suggest a role for the L+H* accent on the adjective of the second instruction (as in the early L+H* sequences) compared to the L+H* accent on the noun of the second instruction (as in the no early L+H* sequences) in modulating eye movement responses. In particular, looks to the target object were speeded more if the L+H* accent felicitously marked the upcoming color contrast than if it felicitously marked the preceding object type contrast. This suggests that the L+H* accent on the adjective of the second instruction (after an initial processing difficulty) evoked an upcoming color contrast and led to earlier looks to the target object than if the L+H* accent was on the noun of the second instruction. In addition, looks to the target object are generally speeded in the continuation compared to the finality conditions, suggesting that the L+H* accent of the third instruction is also considered as providing lexical contrast information for the two utterances that the L-H%
continuation rise suggests should be interpreted together.

Notice that the GCA comparing the early L+H* / finality and early L+H* / continuation conditions in Table 7.6 was run on the same data as the GCA shown in Table 5.4 of Experiment 3a. However, the analysis windows for the two analyses differ. In particular, the analysis window for the comparison in Experiment 3a started and ended later than that for the comparison in this experiment. The analysis windows were calculated by looking at the grand mean data for each experiment. In Experiment 3a, the early L+H* sequences were compared to the late L+H* sequences, but in Experiment 4, the early L+H* sequences are compared to the no early L+H* sequences. Since the late L+H* sequences elicited later looks to the target object than the no early L+H* sequences, the curve for the grand mean of the data rose and peaked later in Experiment 3a than Experiment 4 – resulting in different beginnings and ends for the analysis windows. The difference in analysis window affected the results of the GCA. In Experiment 3a, the direct comparison of the early L+H* / finality and early L+H* / continuation conditions yielded a reliable effect on the linear term, but not the intercept term. But in Experiment 4, the same comparison yielded a reliable effect on both the intercept and linear terms. Thus, both analyses revealed a reliable effect on the linear term, suggesting a difference in slope between the curves. A look at Figure 7.5 or 5.3 reveals that the two curves crossed somewhere between 500 ms and 550 ms. So if the analysis window ranged from -34 ms to 883 ms, as in Experiment 3a, the two curves crossed near the middle of the analysis window, yielding no difference in area under the curve between the two conditions. But if the analysis window ranged from -204 ms to 782 ms, as in this experiment, the two curves crossed closer towards the end of the analysis window, yielding a reliable difference in area under the curve and thus a reliable effect on the intercept term between the two conditions. The differences in results between Experiment 3a and Experiment 4 for the same pairwise comparison underline the importance of choosing an appropriate analysis window.
7.3 Discussion

The results from Experiment 3a revealed that looks to the target object in the early $L+H^*$ sequences (Click on the yellow scissors. Click on the YELLOW bottle. Click on the PURPLE bottle) were modulated by the edge tone pattern ending the second instruction. In particular, a $L-H%$ pattern led to earlier looks to the target object than a $L-L%$ pattern. It was suggested that this effect could be due to the edge tones interacting with the $L+H^*$ accent either in the second or third instruction. This experiment investigated whether the sentence location of the $L+H^*$ accent of the second instruction affected the earliness of looks to the target object. In particular, it was suggested that the $L-L%$ and $L-H%$ edge tone patterns separate or group successive utterances, and thus provide a frame for the interpretation of pitch accents. A $L-H%$ continuation rise, but not $L-L%$ sentence-final intonation, would then suggest that the second and third instructions should be interpreted together. If the second and third instructions were interpreted together, so that $L+H^*$ accents in common sentence positions referred to members of the same contrast set ($YELLOW$ bottle $\Rightarrow$ PURPLE bottle), as proposed for the continuation sequences, earlier looks to the target object should be found for the early $L+H^*/$continuation as compared to the no early $L+H^*/$continuation condition. In particular, a $L+H^*$ accent on the adjective of the second instruction could evoke a contrast set, suggesting an upcoming color contrast and leading to earlier looks to the target object in the early $L+H^*/$continuation as compared to the no early $L+H^*/$continuation condition. The results of Experiment 4 revealed that the advantage for the early $L+H^*/$continuation compared to the no early $L+H^*/$continuation condition was not reliable. There was, however, as numeric advantage for the early $L+H^*/$continuation compared to the no early $L+H^*/$continuation condition. It is possible that the $L+H^*$ accent on the adjective of the early $L+H^*/$continuation condition compared to the $L+H^*$ accent on the noun of the no early $L+H^*/$continuation condition showed only a numeric advantage because the $L+H^*$ accent on the adjective, but not that on the noun, caused an initial processing
difficulty.

Experiment 4 (just as Experiment 3a) involved between-experiment conditions. Again, these conditions need to be compared with caution, even though the arguments for the validity of the results that were made in Section 5.3 of Experiment 3a apply here too. In this experiment, the two between-experiment comparisons revealed interesting differences in results. Looks to the target object were speeded relatively more by a L-H% compared to a L-L% pattern if the second instruction had a L+H* accent on the adjective than if it had a L+H* accent on the noun. This suggests that a L-H% continuation rise provided a greater advantage compared to L-L% sentence-final intonation if the sentence location of the L+H* accent of the second instruction was consistent with an upcoming color contrast. This result does suggest that listeners considered information from the L+H* accent of the second instruction (possibly in addition to L+H* information from the target instruction) when the edge tone information suggested that Instructions 2 and 3 should be interpreted together.

The results from Experiment 4 are thus inconclusive. The between-experiment comparisons confirmed the predictions, but the within-experiment comparisons showed the predicted effect only numerically. Further research is therefore needed to establish the role of the L+H* accent of the second instruction in contributing the the advantage for the L-H% compared to the L-L% pattern ending the second instruction.

Together, Experiments 3a, 3b, and 4 revealed that lexical contrast types, and edge tone types, and (to a lesser extent) accentuation contributed to how quickly listeners could resolve reference. The experiments revealed that reference was resolved faster if an object presented a color contrast with the previously mentioned object (as in the early L+H* sequences in Experiments 3a and 3b) compared to an object contrast (as in the late L+H* sequences in Experiments 3a and 3b).

In addition, the experiments confirmed the notion that a L+H* accent can evoke an upcoming contrast set only with supporting evidence, in particular, if the L+H* accent was
in an instruction that ended in a L-H% continuation rise compared to L-L% sentence-final intonation. A comparison of the early L+H* sequences with the no early L+H* sequences of Experiment 4 hinted at the possibility that a L+H* accent on the adjective (marking the upcoming lexical contrast) followed by a L-H% continuation rise facilitated reference resolution more than a L+H* accent on the noun (marking the preceding lexical contrast) followed by a L-H% continuation rise.

The data further suggest a complex interplay of edge tone patterns with pitch accent and lexical contrast information. Notice that Experiment 3a showed earlier looks to the target object if an early L+H* sequence was connected by a L-H% compared to a L-L%, but no difference between a L-H% and L-L% pattern in a late L+H* sequence. Similarly, Experiment 4 showed that a no early L+H* sequence received less of an advantage from a L-H% pattern compared to a L-L% pattern than an early L+H* sequence. This suggests that a L-H% sequence compared to a L-L% facilitated looks to the target object most if the instruction it ended had a L+H* accent on the adjective that marked an upcoming lexical contrast. Such a result makes sense if edge tones tell the listener which utterances to interpret together and pitch accents reveal how utterances to be interpreted together are related.

The data from Experiments 3a, 3b and 4 suggest a complex interplay of lexical contrast type, accentuation, and edge tones. In particular, the data suggest that some cues are too weak to affect reference resolution by themselves. A L+H* accent only evoked an upcoming contrast set in combination with a L-H% pattern. A L-H% pattern compared to a L-L% pattern only facilitated reference resolution in a color-contrast sequence (the no early L+H* sequence in Experiment 4), but not an object-contrast sequence (the late L+H* sequence in Experiment 3a), even though the L+H* accent felicitously marked the upcoming lexical contrast only in the latter sequence.
The primary goal of this dissertation was to provide a more detailed picture of the abstract intonational structure that listeners use as they comprehend spoken language in visual situations in real time. In particular, the experiments examined the impact of aspects of intonational structure (pitch accents and edge tones) on how listeners determined the visual referents of definite noun phrases, given the visual and discourse context in which they occurred. Much of the previous work on prosodic processing in a discourse context has focused on pitch accents, especially the “contrastive” use of L+H*. The current research confirmed previous findings suggesting that the human language processor can use pitch accent information to anticipate upcoming discourse referents. It demonstrated both facilitation and interference (e.g. when misleading pitch accent information conflicted with lexical information available from spoken words). This dissertation also presented two primary novel findings. First, pitch accents were not interpreted in isolation, as their effects depended on the lexical and prosodic content of the utterances that preceded them. Second, preceding pitch accents and boundary tones structured the interpretation of spoken discourse, interacting with the interpretation of a particular local pitch accent.

The experiments presented in this dissertation had two foci. Experiments 1 and 2 expanded on work done by Ito & Speer (2008) and Weber et al. (2006a) and explored how L+H* accents interacted with other relevant pitch accents in the discourse to affect referent selection. Experiments 3 and 4 presented novel evidence that edge tone patterns played a role in structuring the discourse, affecting the ease with which listeners could resolve reference. While previous work focused on how the prosody of a target instruction
influenced referent selection, this dissertation explored how the lexical and prosodic information present in the utterance(s) preceding the target utterance influenced target referent selection. In particular, it was hypothesized that a L+H* accent could evoke an upcoming, future lexical contrast, if presented in an appropriate discourse, intonational, and visual context. Results presented here suggest that pitch accents in utterances connected by a L-H% continuation rise tended to be interpreted together, while those separated by L-L% finality intonation were not. Thus one L+H* may foreshadow another in the upcoming instruction when the two instructions are connected by a continuation rise. The experiments also investigated how much supporting evidence listeners needed to anticipate an upcoming lexical contrast, and how listeners processed conflicting information.

Experiment 1 tested whether previous findings of robust facilitation effects could be obtained with the current task and experimental design. The experiment compared color-contrast sequences (yellow bottle ⇒ purple bottle) and no-contrast sequences (green pencil ⇒ purple bottle), where the preceding context instruction was pronounced with a H* !H* pattern on the NP and the target instruction was pronounced either with a L+H* accent on the adjective, a L+H* accent on the noun, or a H* !H* pattern. Most importantly, the experiment showed a facilitation effect when the target instruction of a color-contrast sequence was pronounced with a L+H* accent on the adjective (as in Click on the yellow bottle. Click on the PURPLE bottle of the felicitous condition) compared to a L+H* accent on the noun (as in Click on the yellow bottle. Click on the purple BOTTLE of the infelicitous condition) or a H* !H* pattern on the NP (as in Click on the yellow bottle. Click on the purple bottle of the neutral condition). In addition, the experiment confirmed the prosodic "garden-path" effect found in Ito & Speer (2008). Most looks (either numerically or reliably more than in other conditions) to the competitor object (the purple pencil) occurred in the no-contrast sequences if the target instruction was pronounced with a L+H* accent on the adjective (as in Click on the purple pencil. Click on the PURPLE bottle of the garden-path condition). This finding confirmed results from Ito & Speer (2008) and Weber et al. (2006a) and suggested
that a L+H* accent on the adjective evoked the representation of a contrast set of different colored items (e.g. bottles) such that listeners anticipated the upcoming noun to be a repetition of the immediately preceding one. Experiment 1 demonstrated that this effect can be triggered solely by the presence of the L+H* accent on the adjective of the target instruction.

Experiment 2 tested whether a L+H* accent on the adjective of an initial instruction could evoke the expectation of a color contrast with an upcoming instruction (cf. Weber et al., 2006a). The experiment again compared color-contrast sequences (yellow bottle ⇒ purple bottle) and no-contrast sequences (green pencil ⇒ purple bottle). Here, the target instruction was always pronounced with a L+H* accent on the adjective, and the prosody of the context instruction was manipulated to have either a H* !H* pattern on the NP or a L+H* on the adjective. If a L+H* accent on the adjective of the context instruction could foreshadow an upcoming color contrast, looks to the target object should be earlier when both adjectives in a color-contrast sequence carried a L+H* accent (as in Click on the YELLOW bottle. Click on the PURPLE bottle of the anticipatory condition) than when only the target adjective did (as in Click on the yellow bottle. Click on the PURPLE bottle of the felicitous condition). Contrary to this prediction, Experiment 2 showed no reliable differences between the relevant conditions (data from Experiment A) or a delay in looks to the target object if the context instruction was pronounced with a L+H* accent on the adjective (data from Experiment B). These results suggested that the L+H* accent in the context instruction did evoke an upcoming color contrast set, but that the L-L% sentence-final pattern ending the first instruction hindered integration of the contrast set with that evoked by the L+H* accent on the adjective of the target instruction.

Experiment 3a tested how the occurrence in time of L+H*-related processing difficulties affected reference resolution. The experiment compared responses to the final instruction of three-instruction sequences, where the middle instruction presented a processing difficulty either when encountering an early L+H* accent on the adjective or a late L+H* accent on the noun. If the processing difficulty occurred late in time (as in Click on the green pen-
cil. Click on the purple PENCIL. Click on the purple BOTTLE of the late L+H* conditions), it was resolved relatively later than if the processing difficulty occurred early in time (as in Click on the yellow scissors. Click on the YELLOW bottle. Click on the PURPLE bottle of the early L+H* conditions). Late-occurring processing difficulties compared to early-occurring processing difficulties were more likely to remain unresolve at the time listeners began processing information from the target instruction, thereby delaying looks to the target object. In addition, the middle instruction in Experiment 3a ended either in L-L% sentence-final intonation or a L-H% continuation rise. Earlier looks to the target object were expected if an early compared to a late L+H* accent felicitously marked an upcoming color contrast and/or if a L-H% pattern compared to a L-L% pattern ended the middle instruction. As predicted, the results revealed earlier looks to the target for early L+H* sequences compared to late L+H* sequences. A L-H% continuation rise compared to L-L% sentence-final intonation only facilitated reference resolution in the early L+H* sequence. These results suggested that connecting the second and third instructions with a L-H% continuation rise (compared to separating them L-L% sentence-final intonation) facilitated reference resolution, but only if the earlier processing difficulties had been fully resolved by the time the edge tones of the second instruction were encountered. Experiment 3a presented novel evidence that a L-H% continuation rise could facilitate reference resolution compared to L-L% sentence-final intonation. Although post-hoc analyses presented in Section 5.3 suggested that between-experiment comparisons were appropriate for this data set, the critical comparisons between L-L% and L-H% conditions may suffer from potential confounds.

Experiment 3b provided a follow-up experiment to Experiment 3a using only within-experiment comparisons (from EXPERIMENT A). The experiment compared looks to the target object in the early L+H* and late L+H* conditions of Experiment 3a with and without the first instruction. If the first instruction was not included in the experimental trial, the first L+H* accent that listeners encountered presented no processing difficulties (as in Click on the YELLOW bottle. Click on the PURPLE bottle or Click on the purple PENCIL. Click
on the purple BOTTLE). If the first instruction was included in the experimental trial (as in the sequences from Experiment 3a in the previous paragraph), the first L+H* accent that listeners encountered did present processing difficulties. The results revealed no difference in looks to the target object in the early L+H* conditions. That is, looks to the target object were equally fast, regardless of whether or not the first L+H* accent caused processing difficulties (as in the early L+H*/target 3 and early L+H*/target 2 conditions, respectively). These results suggest that when the processing difficulties occurred early (as in the early L+H*/target 3 condition), they were fully resolved by the time the target instruction was encountered. The results also revealed later looks to the target object in the late L+H*/target 3 condition, where the first L+H* accent caused processing difficulties, compared to the late L+H*/target 2 condition, where the first L+H* accent did not cause processing difficulties. These results suggest that when the processing difficulties occurred late (as in the late L+H*/target 3 condition), they were ongoing when the target instruction was encountered.

Experiment 4 tested if the sentence location of the L+H* accent in the middle instruction in the early L+H* sequences of Experiment 3a, which felicitously marked the upcoming contrast, contributed to the earlier looks found in Experiment 3a for the early L+H*/continuation compared to the early L+H*/finality conditions. Since edge tones provide information about which utterances should be interpreted together, but not how such utterances are related, the advantage for the L-H% compared to the L-L% in the early L+H* sequences of Experiment 3a must have involved information from one of the L+H* accents – either that of the second instruction, the third instruction, or both. The experiment tested whether the sentence location of the L+H* accent in the middle instruction contributed to the advantage for the early L+H*/continuation compared to the early L+H*/finality conditions. The experiment compared the late L+H* sequences of Experiment 3a (yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle, with the middle instruction ending either in L-L% or L-H%) with sequences differing from the late L+H* sequences only in that the L+H* accent of the middle instruction was placed on the noun rather than on the adjective (yellow scis-
\textit{sors} ⇒ \textit{yellow BOTTLE} ⇒ \textit{PURPLE bottle}, with the middle instruction ending either in L-L\% or L-H\%. Such sequences (named \textit{no early L+H\*}) differed from the \textit{late L+H\*} sequences in two ways: They were not expected to cause processing difficulties when the first L+H\* accent was encountered, and the first L+H\* accent was not in a sentence position where it could felicitously mark the upcoming color contrast. Looks to the target object were expected to rise earliest in the \textit{early L+H\* / continuation} condition, i.e. if the L+H\* accent of the middle instruction was on the adjective (where it felicitously marked the upcoming contrast) and if the middle instruction ended in a L-H\% pattern (which suggested to interpret the current instruction with the following, target instruction). The results revealed that looks to the target in the \textit{early L+H\* / continuation} condition did indeed rise earliest. While this advantage was reliable in the between-experiment comparison with the \textit{early L+H\* / finality} condition, it was only numeric in the within-experiment comparison with the \textit{no early L+H\* / continuation} condition. The results thus provide some tentative evidence that the L+H\* accent on the adjective of the middle instruction in the \textit{early L+H\* / continuation} condition contributed to the speeded looks to the target object.

### 8.1 Pitch Accents

Overall, the data suggest that L+H\* cues from the target instruction had more influence on referent resolution than L+H\* cues from immediately previous instructions. Experiment 1 showed that a L+H\* accent felicitously placed on the adjective of the \textit{target} instruction in a color-contrast sequence (\textit{yellow bottle} ⇒ \textit{PURPLE bottle}) speeded reference resolution compared to target instructions with other prosodic patterns (\textit{yellow bottle} ⇒ \textit{purple bottle} or \textit{yellow bottle} ⇒ \textit{purple BOTTLE}). It was proposed that the L+H\* accent on the adjective of the target instruction evoked a contrast set of objects of the same type than that referred to in the context instruction, but in different colors. This created the expectation of a repeated noun. That is, the L+H\* information was used to predict lexical information of the upcom-
ing noun (before it was heard) based on information from the context instruction and the adjective of the target instruction. Experiment 2 showed that a L+H* accent felicitously placed on the adjective of the context instruction in a color-contrast sequence (YELLOW bottle ⇒ PURPLE bottle) did not have an equivalent effect. In this location, the L+H* accent also evoked a color contrast with (in this case) the upcoming instruction. However, the context instruction ended in L-L% sentence-final intonation, which prevented integration of the contrast set evoked during the context instruction with the contrast set evoked during the target instruction. As a result, looks to the target object were equally fast or slower when both adjectives carried a L+H* accent compared to only the adjective in the target instruction. Experiments 3a and 4 confirmed that the unexpected results of Experiment 2 were related to the L-L% pattern ending the context instruction. Both experiments suggested that L+H* accents on successive adjectives evoked color contrast sets, and that the two contrast sets were only integrated the instructions carrying the L+H* accents were connected by a L-H% continuation rise.

Taken together, the results from Experiments 1, 2, 3a and 4 confirmed that a L+H* accent felicitously placed on a color adjective evoked a color contrast and created the expectation of a repeated noun. In addition, the data suggest that a L+H* accent could evoke a color contrast either with the preceding instruction or with an upcoming instruction. Due to the amount of information that was available and that needed to be anticipated at the time of the L+H* accent, a preceding contrast set was evoked more easily than an upcoming contrast set. In particular, an upcoming contrast set may need to be reinforced, for example, by a L-H% pattern connecting the instruction that carried the relevant L+H* accent to the target instruction.

These results suggest that listeners considered cumulative evidence across utterances when anticipating an upcoming referent. In Experiment 1, the relevant L+H* accent was on the target utterance. That is, listeners had already heard Click on the yellow bottle. Click on the... when the L+H* accent on PURPLE evoked the relevant contrast and triggered the
expectation of a repeated noun. In Experiment 2, the relevant L+H* accent was on the context utterance, that is, listeners had only heard Click on the... when they encountered the L+H* accent on YELLOW. These situations differ in several ways.

If the relevant L+H* accent was in the target utterance (as in Experiment 1), there was already lexical evidence for a possible color contrast, only one lexical item was anticipated, and the L+H* accent occurred immediately before the anticipated item. Lexical evidence for a possible color contrast came from the context and target instructions having different color adjectives (yellow and purple). This lexical information was available when the L+H* accent on PURPLE was processed. In addition, the relevant repeated noun occurred immediately after the L+H* cue. This means that the lexical item whose accentual information triggered the anticipation and the lexical item that was anticipated occurred in succession. There was thus no intervening sentence material during which L+H* information could have decayed or additional lexical and prosodic information could have distracted from the evoked color contrast.

If the relevant L+H* accent was in the context instruction (as in Experiment 2), there was no lexical evidence for a possible color contrast, a more complex structure needed to be anticipated, and there was intervening sentence material between the relevant L+H* accent and the anticipated items. The absence of lexical evidence for a possible color contrast stemmed from the L+H* accent being on the first-mentioned color adjective. Having the L+H* accent in this discourse location, when the first-mentioned noun had not been heard, also means that what needed to be anticipated was more complex and abstract. In particular, if the L+H* accent evoked a color contrast, listeners would have to anticipate that the noun which immediately followed the L+H*-accented adjective was identical to the noun of the following instruction, and that the following instruction had an adjective that was lexically different from yellow, and potentially, that the following adjective also carried a L+H* accent (cf. Weber et al., 2006a). As a result, there may have been too little supporting evidence, too much information to anticipate, too much time for the L+H* cue to decay,
too much intervening lexical information, and/or too much demand on memory in order for the L+H* accent on the first-mentioned adjective to measurably facilitate reference resolution – in addition to the facilitation that was coming from the L+H* accent on the target instruction.

The results from Experiment 4 suggested that a felicitous L+H* accent on the adjective of the instruction immediately preceding the target instruction did evoke a color contrast, but that information reinforcing the L+H* cue as needed for it to have an effect on reference resolution. In Experiment 4, listeners had heard *Click on the yellow scissors. Click on the YELLOW...* when they encountered the relevant L+H* accent. This situation was similar to that of Experiment 2. The L+H* accent is heard in a discourse location where it could not evoke a color contrast with the previous instruction (since *yellow* is the adjective in both instructions). Instead it should evoke a color contrast with an upcoming instruction. For such a lexical contrast, the L+H* accent would be on the first-mentioned adjective, and listeners would need to anticipate the same complex information as in Experiment 2. If listeners heard *bottle* ending in a L-H% continuation rise (compared to a L-L% pattern) immediately after the L+H* accent on *YELLOW*, one aspect of the L+H* cue was reinforced: The L-H% continuation rise reinforced the focus on the upcoming instruction (i.e. that the L+H* accent evoked a color contrast with the upcoming, rather than the preceding instruction) and allowed integration of the evoked contrast set with that evoked later by the L+H* accent of the target instruction. At the same time, listeners heard lexical information from the noun. Thus, *bottle* ending in a L-H% pattern provided lexical information as to the noun being repeated and reinforced the expectation of an upcoming color contrast (before there was much time for the L+H* information to decay). At the time the L-H% pattern was heard, there were thus two (successive) cues to the upcoming color contrast, the identity of the noun was known, and only the repetition of the known noun and non-repetition of the adjective needed to be anticipated. If the L+H* accent was reinforced this way, it did speed up reference resolution in response to the target instruction.
The experiments present ample evidence that a L+H* accent was not interpreted in isolation, but in its discourse context. The nature of the evoked contrast set depended on the pitch accent, edge tone, and lexical context in which the L+H* accent was situated. Lexical information, information from relevant neighboring pitch accents, and edge tone pattern information were considered when deciding whether a L+H* accent created a contrast set with previous or upcoming material. The results suggested that any model of prosodic and discourse processing needs to take the global prosodic context of utterances as well as across utterances into account and cannot be based purely on local prosodic events (cf. Carlson et al., 2001, for a similar proposal for models of sentence processing).

8.2 Edge Tones

The data from Experiments 3a and 4 presented novel evidence that edge tone patterns affected reference resolution. In particular, the data suggested that a L-H% continuation rise can contribute to speeding reference resolution compared to L-L% sentence-final intonation. Pierrehumbert & Hirschberg (1990) proposed that a L-H% pattern indicates that an utterance should be interpreted with respect to the following utterance, whereas a L-L% pattern indicates finality and reduced connection to the following utterance. Listeners should thus have shifted their attention to interpreting the current instruction with reference to information in the upcoming instruction when hearing the current instruction end in a L-H% continuation rise compared to L-L% sentence-final intonation. Edge tones thus provide a frame in which pitch accents are interpreted. Pitch accents in instructions connected by a L-H% pattern should be interpreted together and contrast sets evoked by L+H* accents in instructions connected by a L-H% pattern should be easily integrated. In contrast, pitch accents in instructions separated by a L-L% pattern should be interpreted separately and L+H* accents in instructions separated by a L-L% pattern should evoke two separate, unintegrated contrast sets.
Experiment 3a showed that a L-H% pattern compared to a L-L% pattern speeded reference resolution only if a L+H* accent that caused processing difficulties was removed enough in time from the L-H% or L-L% information to not interfere with its processing. In the *early* L+H* sequences of Experiments 3a and 4 (*yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle*), the middle instruction carried a seemingly infelicitous L+H* accent on the adjective. This L+H* accent evoked a color-contrast set, which initially caused processing difficulties since the color adjective of the first and second instructions is repeated. The L+H* accent on the adjective of the target instruction also evoked color-contrast. A L-H% continuation rise as compared to L-L% sentence-final intonation allowed integration of the two evoked contrast sets (*YELLOW bottle* and *PURPLE bottle* are members of the same set), thus speeding reference resolution.

In Experiment 3a, the *early* L+H* sequences were compared to the *late* L+H* sequences (*green pencil ⇒ purple PENCIL ⇒ purple BOTTLE*), where the middle instruction carried a seemingly infelicitous L+H* accent on the noun. As in the *early* L+H* sequences, the L+H* accent evoked a contrast set that caused initial processing difficulties. In this case, a contrast set of different purple objects was evoked, but the noun *pencil* was repeated. Unlike the *early* L+H* sequences, these processing difficulties were not yet resolved when the edge tones were encountered and a L-H% compared to L-L% did not facilitate integration of the contrast set evoked by the preceding L+H* accent with that evoked by the L+H* accent in the target instruction. As a result, the L-H% pattern did not lead to earlier looks to the target object than the L-L% pattern.

In Experiment 4, the *early* L+H* sequences were compared to the *no early* L+H* sequences (*yellow scissors ⇒ yellow BOTTLE ⇒ PURPLE bottle*), where the middle instruction carried a L+H* accent on the noun, which felicitously marked the preceding object contrast (and thus did not present processing difficulties). Here, ending the middle instruction with a L-H% compared to a L-L% pattern did result in earlier looks to the target object. However, the size of the effect was smaller for the *no early* L+H* sequences than for the *early*
Notice that a L-H% pattern indicated to listeners to interpret the instruction that just ended with respect to the following one. Unlike a L+H* accent, it did not evoke any lexical contrasts, but grouped instructions. As a result, a L-H% continuation rise did not provide any information about lexical or prosodic aspects of the following instruction. It therefore makes sense that a L-H% pattern speeded reference resolution only in connection with appropriately placed L+H* accents.

Data from Experiment 4 further showed that the edge tone pattern of the middle instruction could interact with the L+H* accent of the target instruction. As mentioned above, edge tone patterns cannot provide information about upcoming lexical content. Their effect must therefore be linked to an aspect of the utterance that can, in this case a L+H* accent. In Experiment 4, the no early L+H* sequences elicited earlier looks to the target object if the middle instruction ended in a L-H% pattern compared to a L-L% pattern. However, the middle instruction provided no cue to the identity of the target instruction: The L+H* accent felicitously marked the preceding lexical contrast, not the upcoming lexical contrast. The facilitation effect must therefore be related to the L+H* accent in the target instruction, which felicitously marked the color contrast between the middle and target instructions. This result suggests that a L-H% continuation rise compared to L-L% sentence-final intonation can strengthen the effect of a future L+H* accent. In particular, the L-H% continuation rise indicated to listeners that the current instruction was closely linked to the following one, and the L+H* accent then revealed how the instructions were linked (by sharing the noun in the case of the no early L+H* sequences).

Data from Experiment 2 also alluded to the possibility that listeners tracked the frequency of edge tone patterns throughout the discourse. The target fixation analyses from Experiment 2 showed a main effect of Pitch Accent for the data from EXPERIMENT B, but not for the data from EXPERIMENT A. This suggested that conditions with two successive L+H* accents (as in YELLOW bottle ⇒ PURPLE bottle/pencil) elicited later looks to the target
object compared to conditions with a L+H* accent only in the target instruction (as in yellow bottle ⇒ PURPLE bottle/pencil) in EXPERIMENT B, but not EXPERIMENT A. Listeners in EXPERIMENT B heard instructions ending in a L-L% or L-H% pattern throughout the experiment, whereas listeners in EXPERIMENT A heard only instruction ending in a L-L% pattern. As a result, a L-L% pattern may have been more informative to listeners in EXPERIMENT B, where it was contrasted with a L-H% pattern throughout the discourse, than listeners in EXPERIMENT A, where L-L% sentence-final intonation ended all instructions. A L-L% pattern indicated finality and separation of the current instruction from the one following it, and a L+H* accent in an initial instruction may evoke an upcoming contrast set. Together, these present competing cues. If listeners considered the L-L% pattern to be informative (as in EXPERIMENT B), the competing cues may result in a processing delay for sequences with two successive L+H* accents compared to a L+H* accent only in the target instruction. But if listeners paid less attention to edge tone information (as suggested for EXPERIMENT A), the L-L% pattern ending the initial instruction may compete less with the L+H* accent and cause less processing difficulty. Further studies are needed to explore this possibility.

8.3 Lexical Information

Experiments 1 and 2 included color contrast (yellow bottle ⇒ purple bottle) and no-contrast (yellow bottle ⇒ purple pencil) sequences. The experiments showed a general preference for color contrast over no-contrast sequences, regardless of how instructions were pronounced. This result is likely related to a bias from the experimental design.

Experiments 3a, 3b and 4 included color contrast (yellow bottle ⇒ purple bottle) and object contrast (purple pencil ⇒ purple bottle) sequences. Experiments 3a and 3b showed earlier looks to the target object if the target instruction formed a felicitously-accented
color contrast compared to a felicitously-accented object contrast with the previous instruction. This confirms results from Ito & Speer (2008), who also found an advantage for felicitously-accented color contrasts over felicitously-accented object contrasts. Ito & Speer (2008) considered that the advantage was due to the organization of objects in the visual display. Their display contained all objects of the same object type within the same cell of the visual display. For example, all balls were in the same cell of the display, but all green objects were in different cells. If listeners heard *Hang the green ball. Now, hang the YELLOW...* and anticipated the repetition of the noun *ball*, they could direct their eyes to the location of the ball-cell. But if listeners heard *Hang the green ball. Now, hang the green...*, they could not direct their eyes to an appropriate cell since all green objects were in different cells. Thus, even if listeners used prosodic information to expect a felicitously-accented object contrast (e.g. if they were sensitive to pre-accent attenuation), the display did not allow them to guess the location of the object before lexical information from the noun had come in.

Compare this to the display here, where each object received its own cell. If listeners heard *Click on the yellow bottle. Click on the PURPLE...* and anticipated the repetition of the noun, they could direct their eyes to the cell containing the purple bottle. If listeners heard *Click on the purple pencil. Click on the purple...* and used pre-accent attenuation to anticipate an object contrast (i.e. a different noun), they could direct their eyes to the cell containing the other purple object in most trials (in some trials, there were two additional purple objects). Nevertheless, listeners were reliably faster at finding the target object if they heard a felicitously-accented color-contrast sequence compared to a felicitously-accented object-contrast sequence. This suggests that pre-accent attenuation is a weaker cue for an object contrast than a L+H* accent is for a color contrast. To investigate whether pre-accent attenuation can be used predictively, a display containing each object color and object type exactly twice may be needed.

Experiment 3b additionally revealed that listeners considered lexical information from
two preceding instructions when processing the target instruction. In the late $L+H^*$ sequences of Experiment 3b the presence or absence of the first instruction affected looks to the target object in the third instruction, suggesting that listeners keep in memory a fair amount of the preceding discourse and use it to make predictions. This result underlines that reference resolution is affected not only by the current utterance, but by the discourse context in which the utterance is situated.

8.4 Implications for Theories of Focus Structure

Semantic theories of focus structure link pitch accent placement to a sentence’s focus structure. That is, such theories ignore the two aspects of prosody investigated here: pitch accent type and edge tone information. The results from this dissertation as well as from Ito & Speer (2008) and Weber et al. (2006a) suggest the need for integrating both pitch accent type and edge tone information into semantic theories of focus structure.

The stimuli used in Weber et al. (2006a) provide a good illustration for the need to incorporate pitch accent type into theories of focus structure. Pronouncing the instruction *Klicke die lila Schere an* ("Click on the purple scissors") with only a $H^*$ accent on the noun *Schere* yields a typical utterance with broad focus. On the other hand, pronouncing the instruction with only a $L+H^*$ accent on the noun yields an utterance with narrow focus on the noun. That is, only if the noun is pronounced with a $H^*$ accent, but not if it is pronounced with a $L+H^*$ accent, can focus project to the whole sentence. Semantic theories that ignore pitch accent type would incorrectly predict the same set of possible focus structures for the two pronunciations. Féry & Samek-Lodovici’s (2006) constraint-based approach provides a step in the right direction. They distinguish two related STRESS-FOCUS (SF) constraints, $SF_{\text{contrast}}$ and $SF_{\text{new}}$, with $SF_{\text{contrast}}$ being ranked higher than $SF_{\text{new}}$. $SF_{\text{contrast}}$ requires contrastive focus to receive the highest prosodic prominence, and $SF_{\text{new}}$ requires new information focus to receive the highest prosodic prominence.
These constraints correctly predict that contrastive information receives higher prominence (usually by means of a L+H* accent compared to a H* accent) than new information if a sentence contains both new and contrastive information. In Weber et al.’s (2006a) stimuli, however, instructions contain either new or contrastive information. If *lila Schere* provides new information, the highest prominence is on *Schere* (since the whole noun phrase can be new even if only *Schere* receives the H* accent). If *Schere* provides “contrastive” information, the highest prominence is again on *Schere*. Thus, reference to pitch accent type is needed in order to capture the differences between the possible focus structures that correspond to the two pronunciations of *Klicke die lila Schere an*.

In addition, the experiments from this dissertation provided evidence that edge tones provided a frame for the interpretation of pitch accents and that reference resolution was affected by the choice of edge tone patterns. Neither of the semantic theories introduced in Section 1.3 consider edge tone information or the possibility that the prosodic pattern of preceding utterances may affect how pitch accents in the current utterance are interpreted.

### 8.5 General Conclusions

The experiments presented in this dissertation revealed a complex interplay of L+H*, edge tone, and lexical information. They confirmed that prosody plays an important role in reference resolution and needs to be considered in models of sentence and discourse processing. Furthermore, the experiments suggested that pitch accents and edge tones were not interpreted uniformly or in isolation. For example, L+H* accents seemed to more easily evoke a contrast set with a lexical item occurring earlier in the discourse. But given an appropriate discourse context, they could evoke a contrast set with a lexical item occurring in the future discourse. The experiments revealed one such situation. A L+H* accent could evoke an upcoming color contrast set if it was on the adjective of an utterance that ended...
in a L-H% continuation rise\textsuperscript{1}. The experiments presented here thus present evidence that pitch accents and edge tones are interpreted in a particular discourse context, and that changing the discourse context may change the interpretation of the pitch accents and/or edge tones.

Further studies are needed to achieve a more complete picture of how prosodic and lexical aspects of the speech stream interact to affect how easily reference is resolved. First, there are some loose threads. The results of Experiment 2 were unexpected. An anticipatory sequence like YELLOW bottle $\Rightarrow$ PURPLE bottle did not elicit earlier looks to the purple bottle than a felicitous sequence like yellow bottle $\Rightarrow$ PURPLE bottle, even though the anticipatory sequence contained two felicitously placed L+H* accents that evoked the color contrast set whereas the felicitous sequence contained only one such L+H* accent. The detailed results from Experiments 2 and 3a suggested that the anticipatory condition did not elicit earlier looks to the target object compared to the felicitous condition because the instructions were not connected by a L-H% continuation rise. Ending the first instruction with a L-H% continuation rise may lead to the current instruction to be interpreted with respect to the following one and cause the L+H* accent of the first instruction to evoke an upcoming contrast set. An experiment comparing the felicitous and anticipatory conditions of Experiment 2 with conditions differing from the these conditions only in that the first instruction ends in a L-H% continuation rise can reveal if the L-L% sentence-final intonation that ended the first instruction caused the unexpected results found in Experiment 2.

Experiment 2 also alluded to the possibility that listeners tracked the occurrence of prosodic patterns across discourse. The presence, absence, or frequency of prosodic patterns across a large discourse segment may then influence interpretation. In particular,

\textsuperscript{1}Since the direct evidence for this claim comes from between-experiment data (Experiment 4), it is somewhat tentative. However, data from Experiment 2 also suggested that a L+H* accent could evoke an upcoming contrast, even if this did not happen in the discourse context presented in Experiment 2. Notice also that non-local prosody did affect reference resolution in Experiment 2, albeit not as predicted.

251
Experiment 2 suggested that listeners tracked the occurrence of edge tone patterns and that these patterns became more informative as they were contrasted with other patterns. A reasonable hypothesis assumes that common edge tone patterns receive less attention and may be interpreted as a ‘default’, whereas less common patterns may attract attention. Follow-up studies comparing felicitously accented color and object contrasts and using three-instruction sequences with different numbers of utterances ending in a L-L% vs. L-H% pattern can explore this possibility. For example, earlier looks to the target object should be found for the early L+H* / continuation sequence of Experiment 3a (yellow scissors ⇒ YELLOW bottle ⇒ PURPLE bottle) if most instructions throughout the experiment ended in a L-L% pattern than if instructions ended equally frequently in a L-L% or L-H% pattern. Similarly, ending the first instruction of the felicitous and anticipatory conditions of Experiment 2 with a L-H% pattern should result in earlier looks to the target object if a number of instructions throughout the experiment ended in a L-L% pattern compared to all instructions ending in a L-H% pattern. Such results would suggest that reference resolution in a given experimental trial was not only affected by the discourse situation created in that particular trial, but by information from other trials of the experiment. Such a result would suggest that a much larger portion of the preceding discourse than suggested in the experiments presented here was considered when interpreting utterances. In addition, this may provide novel evidence that listeners tracked distributional properties of prosody in discourse.

Finally, further research can explore why felicitously-accented color contrasts (... ⇒ YELLOW bottle ⇒ PURPLE bottle) elicited earlier looks to the target object in Experiments 3a and 3b than felicitously-accented object contrasts (... ⇒ purple PENCIL ⇒ purple BOT- TLE). The advantage for felicitously-accented color contrasts may be related to L+H* accents on non-final lexical items evoking contrast sets more easily than those in sentence-final position or to L+H* accents on adjectival modifiers evoking contrast sets more easily than those on nouns. Instructions like Click on the purple bottle cannot separate possible
sentence-position and syntactic category effects since the adjective is always in non-final sentence position. An experiment using instructions like Click on the bottle that’s purple could provide insight into the mechanisms behind the effect\(^2\). If the sentence position of the L+H\(^*\) accent drove the effect, felicitously-accented object contrasts (Click on the pencil that is purple ⇒ Click on the BOTTLE that is purple) should elicit earlier looks to the target object than felicitously-accented color contrasts (Click on the bottle that is yellow ⇒ Click on the bottle that is PURPLE). If syntactic category drives the effect, felicitously-accented color contrasts should still elicit earlier looks to the target object than felicitously-accented object contrasts.

\(^2\)Thank you to Sharon Ross for suggesting this follow-up experiment.


The Audacity Team (2006). Audacity 1.2.6 [open source recording software].
URL http://audacity.sourceforge.net

The GIMP Development Team (2008). Gimp 2.4 [open source graphics software].
URL http://www.gimp.org


