Diachronic Adverbial Morphosyntax:
A Minimalist Study of Lexicalization and Grammaticalization

by

James Andrew Berry

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2011 by the
Graduate Supervisory Committee:

Elly van Gelderen, Chair
Karen Adams
Robert Mailhammer

ARIZONA STATE UNIVERSITY
May 2011
ABSTRACT

The historical study of sentence adverbs has, before now, been based mostly on models that emphasize the pragmatic and discourse-based motivations of processes of grammaticalization. This dissertation breaks from such tradition by exploring diachronic adverb development through syntactic and morphological lenses. A generative, feature-based approach is used that incorporates the cartographic architecture developed by Cinque and combines it with a more phenomenological approach to both grammaticalization and lexicalization.

Cinque's hierarchy of speech-act, evaluative, evidential, and epistemic adverbs is analyzed. It is determined (through corpus data) that these subcategories have grown in use primarily during the Modern English era, and particularly during the nineteenth and twentieth centuries. These four subcategories can be divided into two groups that are more general: speech-act adverbs, which arise from a (conditional) speech-act clause that undergoes ellipsis, and the other three types, which all arise from copula clauses. Each of these two groups is considered, and different methods of reanalysis by speakers are proposed for each.

In addition, a revised model for categorizing adverbs is proposed. This model is based on morphological lexicalization (or univerbation) processes, thus accounting for the wide variety of adverbial source materials. Such lexicalization offers a pattern for sentence adverbial formation. Finally, Standard Chinese adverbials are briefly examined, with results indicating that they show very similar signs of lexicalization (within the limits of the writing system).
ACKNOWLEDGMENTS

Although it is impossible with limited space to thank all those who have helped me complete this dissertation, I would like to single out a few people whose contributions have stood out.

First, I must give a warm and earnest “thank you” to my advisor and mentor, Elly van Gelderen. From the beginning of my academic career, through coursework, my MA, and my PhD research, Elly has been steadfastly supportive, offering useful feedback, thorough instruction, and sound advice at every turn. In addition, her cheerful, kind, and fun-loving manner has made my experiences at Arizona State uniquely enjoyable: thanks in large part to Elly, I have been able to spend the last several years doing something I love and learning far more than I ever thought I would. Thank you, my friend.

The other members of my doctoral committee have provided invaluable expertise and support as well. Karen Adams has had a tremendous influence on my academic life at ASU. She is a wonderful teacher and a great sounding board for various “important ideas” and often random thoughts I have had. In addition, she always brings a (much-welcomed) highly critical eye to my work that has challenged—and forced me to explain better—my assumptions more than once. I am very grateful to Rob Mailhammer for his expertise in the history of English, for his high level of scholarship, and for his willingness to join my committee rather late in the game while he was a new faculty member at ASU.

Other faculty members at ASU who have played important roles in my academic development include Dawn Bates, Carrie Gillon, Roy Major, Don
Nilsen, and Claire Renaud. I have learned important lessons from all of them and have enjoyed getting to know them all.

A significant forum for my continued education and for trying out new ideas has been the Syntax Reading Group coordinated by Elly at ASU every semester. Many thanks go to my friends and peers from the Reading Group, in particular to Hui-Ling (Ivy) Yang, Olena Tsurska, Mariana Bahtchevanova, Lynn Sims, Yi-Ting Chen, Uthairat Rogers, John Ryan, Victor Parra-Guinaldo, and Mohammed Al-Rashed.

Several of the ideas in this dissertation were first tried out in front of the audiences of LASSO XXXVIII in Provo and XXXIX in Las Cruces, and at the Workshop on Continuing Trends in Grammaticalization Research in Groningen. I appreciate the critical yet supportive comments and helpful questions I received at those conferences.

Outside academia, I have been surrounded by a wonderful group of friends and family, all of whom I have imposed upon, and to all of whom I owe a huge debt of gratitude. My wonderful and very patient parents have been incredibly supportive through this entire endeavor, and my two brothers and their wives and children have been tremendous sources of inspiration and advice, as well as fun. I have many friends who have all been cheering me on, and I am so glad to have them in my life.

Finally, I want to express my gratitude to my partner Philip, without whose constant love and encouragement I would never have started that first class, let alone reached this point. I owe you a few years of laundry duty.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>LIST OF TABLES</th>
<th>xi</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2 Research Questions</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1 What, exactly, are the criteria for adverbs in English?</td>
<td>2</td>
</tr>
<tr>
<td>1.2.2 How do sentence adverbs arise in the history of English?</td>
<td>3</td>
</tr>
<tr>
<td>1.2.3 Can such patterns be extended to languages other than English?</td>
<td>3</td>
</tr>
<tr>
<td>1.3 Methodology</td>
<td>3</td>
</tr>
<tr>
<td>1.4 Organization of this Dissertation</td>
<td>5</td>
</tr>
<tr>
<td>2 THEORETICAL FRAMEWORKS, PART I: HISTORICAL</td>
<td></td>
</tr>
<tr>
<td>LINGUISTICS</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>7</td>
</tr>
<tr>
<td>2.1.1 Context</td>
<td>7</td>
</tr>
<tr>
<td>2.1.2 Chapter organization</td>
<td>7</td>
</tr>
<tr>
<td>2.2 Diachronic Linguistics—Background</td>
<td>8</td>
</tr>
<tr>
<td>2.3 Grammaticalization</td>
<td>9</td>
</tr>
<tr>
<td>2.3.1 Introduction</td>
<td>9</td>
</tr>
<tr>
<td>2.3.2 Current understanding</td>
<td>13</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>2.3.3 Grammaticalization Theory</td>
<td>17</td>
</tr>
<tr>
<td>2.3.4 Grammaticalization and generative syntax</td>
<td>19</td>
</tr>
<tr>
<td>2.3.5 (Inter)subjectification and semantic considerations</td>
<td>20</td>
</tr>
<tr>
<td>2.3.6 Criticisms</td>
<td>21</td>
</tr>
<tr>
<td>2.3.7 Grammaticalization in this work</td>
<td>25</td>
</tr>
<tr>
<td>2.4 Lexicalization</td>
<td>25</td>
</tr>
<tr>
<td>2.4.1 Introduction</td>
<td>25</td>
</tr>
<tr>
<td>2.4.2 The state of the art</td>
<td>26</td>
</tr>
<tr>
<td>2.4.3 Lexicalization’s place vis-à-vis Grammaticalization Theory</td>
<td>30</td>
</tr>
<tr>
<td>2.4.4 Lexicalization in this work</td>
<td>32</td>
</tr>
<tr>
<td>2.5 Parataxis and Hypotaxis</td>
<td>34</td>
</tr>
<tr>
<td>2.5.1 Introduction</td>
<td>34</td>
</tr>
<tr>
<td>2.5.2 Levels of clausal connection: parataxis</td>
<td>35</td>
</tr>
<tr>
<td>2.5.3 Levels of clausal connection: hypotaxis and a confusion of terms</td>
<td>38</td>
</tr>
<tr>
<td>2.5.4 Clause combining and grammaticalization</td>
<td>43</td>
</tr>
<tr>
<td>2.5.5 Parataxis and hypotaxis in this work</td>
<td>45</td>
</tr>
<tr>
<td>2.6 Pragmatics and Discourse of Predicational Adverbs</td>
<td>46</td>
</tr>
<tr>
<td>2.6.1 Introduction</td>
<td>46</td>
</tr>
<tr>
<td>2.6.2 Pragmatic models I: Politeness Theory</td>
<td>47</td>
</tr>
<tr>
<td>2.6.3 Pragmatic models II: Relevance theory</td>
<td>49</td>
</tr>
<tr>
<td>PAGE</td>
<td>CONTENT</td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>51</td>
<td>2.6.4 Pragmatic models III: stance</td>
</tr>
<tr>
<td>53</td>
<td>2.6.5 Pragmatic markers</td>
</tr>
<tr>
<td>56</td>
<td>2.6.6 Sentence adverbs as pragmatic markers</td>
</tr>
<tr>
<td>57</td>
<td>2.7 Chapter Summary</td>
</tr>
<tr>
<td>58</td>
<td>3 ADVERBS AS A CATEGORY</td>
</tr>
<tr>
<td>58</td>
<td>3.1 Introduction</td>
</tr>
<tr>
<td>58</td>
<td>3.1.1 Adverbs are confusing</td>
</tr>
<tr>
<td>59</td>
<td>3.1.2 Chapter organization</td>
</tr>
<tr>
<td>60</td>
<td>3.2 Background</td>
</tr>
<tr>
<td>60</td>
<td>3.2.1 Adverbs and adverbials</td>
</tr>
<tr>
<td>64</td>
<td>3.2.2 Early methods of categorizing adverbs</td>
</tr>
<tr>
<td>67</td>
<td>3.2.3 Further semantic considerations</td>
</tr>
<tr>
<td>68</td>
<td>3.2.4 Meeting semantic conditions</td>
</tr>
<tr>
<td>74</td>
<td>3.2.5 Section summary</td>
</tr>
<tr>
<td>74</td>
<td>3.3 Supplementing Synchrony with Diachrony</td>
</tr>
<tr>
<td>75</td>
<td>3.3.1 Grammaticalization and pragmatization</td>
</tr>
<tr>
<td>77</td>
<td>3.3.2 Lexicalization</td>
</tr>
<tr>
<td>79</td>
<td>3.4 Adverbial Lexicalization</td>
</tr>
<tr>
<td>80</td>
<td>3.4.1 Fused adverbs</td>
</tr>
<tr>
<td>83</td>
<td>3.4.2 -ly adverbs</td>
</tr>
<tr>
<td>84</td>
<td>3.4.3 -like adverbs</td>
</tr>
<tr>
<td>87</td>
<td>3.4.4 A contrast: -like adjectives</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>3.4.5 Romance -ment(e)</td>
<td>90</td>
</tr>
<tr>
<td>3.5 Other Adverbs as Sites for Lexicalization</td>
<td>92</td>
</tr>
<tr>
<td>3.6 Chapter Summary</td>
<td>94</td>
</tr>
<tr>
<td>4 THEORETICAL FRAMEWORKS, PART II: MORPHOSYNTAX</td>
<td>96</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>96</td>
</tr>
<tr>
<td>4.1.1 Why generative grammar?</td>
<td>96</td>
</tr>
<tr>
<td>4.1.2 Chapter organization</td>
<td>98</td>
</tr>
<tr>
<td>4.2 Generative Grammar</td>
<td>99</td>
</tr>
<tr>
<td>4.2.1 Introduction: origins, history, and development</td>
<td>99</td>
</tr>
<tr>
<td>4.2.2 Principles and Parameters; the Minimalist Program</td>
<td>103</td>
</tr>
<tr>
<td>4.2.3 Distributed Morphology</td>
<td>114</td>
</tr>
<tr>
<td>4.3 Theoretical Models of Adverbs</td>
<td>118</td>
</tr>
<tr>
<td>4.3.1 Introduction: general difficulties with adverbs</td>
<td>118</td>
</tr>
<tr>
<td>4.3.2 Antisymmetry and cartography</td>
<td>120</td>
</tr>
<tr>
<td>4.3.3 Semantic-driven adjunction</td>
<td>134</td>
</tr>
<tr>
<td>4.3.4 Adverbial syntactic model for this study</td>
<td>137</td>
</tr>
<tr>
<td>4.4 Combining Generative Grammar and Historical Linguistics</td>
<td>138</td>
</tr>
<tr>
<td>4.4.1 Introduction: the history of diachronic syntax</td>
<td>138</td>
</tr>
<tr>
<td>4.4.2 Minimalist approaches to grammaticalization: Roberts and Roussou (2003)</td>
<td>140</td>
</tr>
<tr>
<td>4.4.3 Minimalist approaches to grammaticalization: van Gelderen (2004, 2008)</td>
<td>143</td>
</tr>
</tbody>
</table>
CHAPTER 4.5 Generative Grammar and Lexicalization ...................... 146

4.6 Chapter Summary ...................................................... 147

5 SPEECH-ACT ADVERBS .............................................. 149

5.1 Introduction ............................................................ 149

  5.1.1 Focus of the chapter ........................................... 149

  5.1.2 Chapter organization .......................................... 149

5.2 Semantics of Speech-act Adverbs ................................. 150

5.3 Speech-act Adverbs as a Syntactic Category .................... 152

5.4 Discourse and Pragmatics .......................................... 153

  5.4.1 The syntax–pragmatics interface ............................ 153

  5.4.2 Pragmatic and discourse models ............................. 154

5.5 The Speech-act Adverb *frankly* .................................. 155

5.6 How Does *frankly* Become a Sentence Adverb? .............. 156

  5.6.1 Grammaticalization ............................................. 156

  5.6.2 Lexicalization ................................................... 160

5.7 Corpora and Layering of Usages .................................. 163

  5.7.1 Older diachronic corpora ....................................... 164

    5.7.1.1 Early English corpora ................................... 164

    5.7.1.2 Davies-BYU *OED* corpus ............................... 166

  5.7.2 Recent diachronic and synchronic corpora ................. 169

    5.7.2.1 Davies-BYU *TIME* Corpus of American English ......................... 169
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.7.2.2</td>
<td>Davies-BYU Corpus of Contemporary American English</td>
<td>172</td>
</tr>
<tr>
<td>5.7.2.3</td>
<td>Davies-BYU Corpus of Historical American English</td>
<td>175</td>
</tr>
<tr>
<td>5.7.2.4</td>
<td>Davies-BYU interface for the British National Corpus</td>
<td>178</td>
</tr>
<tr>
<td>5.8</td>
<td>A Generative Analysis of Lexicalization</td>
<td>179</td>
</tr>
<tr>
<td>5.8.1</td>
<td>Cinque and the cartographic approach</td>
<td>180</td>
</tr>
<tr>
<td>5.8.2</td>
<td>Application of cartographic and generative principles to lexicalization phenomena</td>
<td>181</td>
</tr>
<tr>
<td>5.9</td>
<td>Chapter Summary</td>
<td>190</td>
</tr>
<tr>
<td>6</td>
<td>COPULA-BASED ADVERBIALS: EVALUATIVE, EVIDENTIAL, AND EPISTEMIC</td>
<td>192</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>192</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Combining these three types of sentence adverbs</td>
<td>192</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Chapter organization</td>
<td>194</td>
</tr>
<tr>
<td>6.2</td>
<td>The Subcategories: Evaluative, Evidential, and Epistemic</td>
<td>194</td>
</tr>
<tr>
<td>6.2.1</td>
<td>The Old English model: primarily truth-intensifying adverbs</td>
<td>194</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Evaluative adverbs in English</td>
<td>198</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Evidential adverbs in English</td>
<td>204</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Epistemic adverbs in English</td>
<td>206</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>6.3</td>
<td>Diachronic Approaches to the EEE Subcategories</td>
<td>209</td>
</tr>
<tr>
<td>6.4</td>
<td>Sentence Adverbs and Predication: A Syntactic Examination</td>
<td>210</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Underlying structure</td>
<td>210</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Den Dikken (2006) and the RELATOR</td>
<td>211</td>
</tr>
<tr>
<td>6.4.3</td>
<td>A “lexicalized” RELATOR (-ly)</td>
<td>215</td>
</tr>
<tr>
<td>6.5</td>
<td>Usage Patterns</td>
<td>220</td>
</tr>
<tr>
<td>6.6</td>
<td>Chapter Summary</td>
<td>223</td>
</tr>
<tr>
<td>7</td>
<td>A BRIEF LOOK AT CHINESE ADVERBIALS</td>
<td>224</td>
</tr>
<tr>
<td>7.1</td>
<td>Introduction</td>
<td>224</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Background for investigating Chinese adverbs</td>
<td>224</td>
</tr>
<tr>
<td>7.1.2</td>
<td>Chapter organization</td>
<td>224</td>
</tr>
<tr>
<td>7.2</td>
<td>Modern Chinese</td>
<td>225</td>
</tr>
<tr>
<td>7.2.1</td>
<td>An outline of Chinese</td>
<td>225</td>
</tr>
<tr>
<td>7.2.2</td>
<td>The typology of Chinese</td>
<td>226</td>
</tr>
<tr>
<td>7.2.3</td>
<td>Chinese VP-adverbs</td>
<td>227</td>
</tr>
<tr>
<td>7.3</td>
<td>Chinese Speech-act Adverbs</td>
<td>232</td>
</tr>
<tr>
<td>7.4</td>
<td>Chinese EEE Adverbs</td>
<td>235</td>
</tr>
<tr>
<td>7.4.1</td>
<td>Evaluative adverbs</td>
<td>235</td>
</tr>
<tr>
<td>7.4.2</td>
<td>Evidential adverbs</td>
<td>240</td>
</tr>
<tr>
<td>7.4.3</td>
<td>Epistemic adverbs</td>
<td>241</td>
</tr>
<tr>
<td>7.5</td>
<td>Chapter Summary</td>
<td>242</td>
</tr>
<tr>
<td>8</td>
<td>CONCLUSIONS</td>
<td>244</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>Page</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>8.1 Chapter Summaries</td>
<td>244</td>
<td></td>
</tr>
<tr>
<td>8.2 Contributions to the Field</td>
<td>247</td>
<td></td>
</tr>
<tr>
<td>8.3 Suggestions for Future Research Directions</td>
<td>248</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td>249</td>
<td></td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>2.1.</td>
<td>Grammaticalization of OE <em>pe</em> to ModE <em>the</em></td>
<td>22</td>
</tr>
<tr>
<td>2.2.</td>
<td>Criteria for pragmatic markers</td>
<td>54</td>
</tr>
<tr>
<td>2.3.</td>
<td>Functions of pragmatic markers</td>
<td>55</td>
</tr>
<tr>
<td>5.1.</td>
<td>Instances of <em>frankly</em> in the <em>OED</em></td>
<td>167</td>
</tr>
<tr>
<td>5.2.</td>
<td>Instances of <em>frankly</em> in the <em>TIME</em> Corpus of American English</td>
<td>169</td>
</tr>
<tr>
<td>5.3.</td>
<td>Tokens of <em>frankly</em> in the COHA (Davies), by type of syntactic structure</td>
<td>176</td>
</tr>
<tr>
<td>6.1.</td>
<td>Historical use of adjective–adverb pairs in American English</td>
<td>219</td>
</tr>
</tbody>
</table>
Chapter 1

INTRODUCTION

1.1 Background

Adverbs and adverbials occupy an interesting, uncertain “middle ground” in linguistic studies. There is little agreement on whether they comprise a lexical category or a functional one or whether they are even an independent category at all. Jespersen (1969, p. 109), for example, in his theory of Rank, allots to adverbs the lowest rank available:

\[
\begin{array}{ll}
\alpha. \text{Word-Classes:} & \beta. \text{Ranks:} \\
\text{Substantives} & \text{Primaries (1)} \\
\text{Adjectives} & \text{Secondaries (2)} \\
\text{Adverbs} & \text{Tertiaries (3)} \\
\end{array}
\]

Although adverbs bring a large amount of semantic and pragmatic information to the utterance, there seems to be little that is consistent syntactically or morphologically among them. Recent syntactic studies by Cinque (1999) and Ernst (2002) have attempted to find consistency and structure in adverb use; each of these studies has advantages and disadvantages.

Most syntactic work on adverbs and adverbials has been synchronic in nature. At the same time, most diachronic work on adverbs (e.g. Hopper & Traugott, 2003) has largely ignored generative grammar and the associated frameworks that have emerged. There are a number of generative approaches to historical linguistics that have emerged in the past 30 years (e.g. Lightfoot, 1979; Roberts & Roussou, 2003; van Gelderen, 2004); however, none of these has expressly addressed the adverb through the various synchronic frameworks.
This dissertation is focused, in particular, on addressing this gap in the research. My goal is to bring a diachronically syntactic perspective to the study of the adverb in English, thereby synthesizing formal (generative) and functional (grammaticalization- and lexicalization-oriented) insights. Another, secondary, goal is to propose a theory about the category *adverb* in English that is based on morphosyntactic characteristics rather than on primarily semantic criteria.

In my investigation into English adverbial diachrony, a fair amount of terminological confusion has presented itself. I have chosen definitions for terms such as *lexicalization* that do not necessarily agree with the orthodoxy established by grammaticalization theorists; I address these and other concerns in Chapter 2.

In order to give a direction to this study, I have focused on three main research questions.

1.2 Research Questions

1.2.1 What, exactly, are the criteria for adverbs in English?

This question is largely addressed in Chapter 3, but it has ramifications for all the work that is found in this dissertation. Basically, the issues at hand in answering this question include the use of syntactic, morphological, and semantic information for categorizing; the diachronic combinatory nature of most English adverbs that shows, with more or less opacity, their phrasal and clausal origins; and the tendency towards renewal of these combinatory patterns.
1.2.2 How do sentence adverbs arise in the history of English?

This question forms the core inquiry of the dissertation, and it was the first question I asked when broaching this topic. Chapters 5 and 6 focus on the growth of the adverb subcategories *speech-act*, *evaluative*, *evidential*, and *epistemic*. This growth is primarily a phenomenon of the Modern English era (especially the twentieth century). Issues covered in these investigations include the roles of grammaticalization and lexicalization phenomena in these shifts as well as generative frameworks for analyzing such phenomena.

1.2.3 Can such patterns be extended to languages other than English?

This question is addressed by this study in a limited fashion, leaving room for additional study. In Chapter 7, I examine data on sentence adverbs in Standard Chinese, looking for similar patterns to those in English and using some of the criteria and frameworks established in earlier chapters. Elsewhere there are brief discussions of some characteristics of Romance languages. In general, however, this is a question that will bear further investigation.

1.3 Methodology

This dissertation is primarily focused on the synthesis of different frameworks (synchronic–diachronic; generative–functional) to provide an analysis of adverb development in English. In addition, however, there is a wide range of source materials that have been used to add support to the claims I make in this work.
The online version of the *Oxford English Dictionary* (hence *OED*) has been used a great deal: to determine the entry point of a word in the language, to determine the rise of certain usages, and to collect early examples. Other online sources, such as ESL word lists and, for Chapter 7, Chinese-English dictionaries, have also provided a great deal of information. Current usages have been gathered in many cases through Google searches.

In addition, I have been able to make use of several English corpora in order to trace changes in the adverbs I examine in my case studies (in Chapters 5 and 6). Early corpora from the Oxford University website (e.g. Shakespeare’s First Folio, Michigan Early Modern English Materials) are only moderately used in Chapter 5 (and not at all in Chapter 6), primarily because the phenomena I investigate are characteristic of the more recent period of Modern English.

Instead, I depend more heavily on the highly useful corpora compiled by Mark Davies and his research team at Brigham Young University. At various times there have been five different English corpora available for use, and I have made use of all five. These include the British National Corpus (BNC), the *TIME* Magazine Corpus, the Corpus of Contemporary American English (COCA), and the recently-added Corpus of Historical American English (COHA). In addition, for a period of time a corpus interface for the *OED* was available on Davies’s website; it was removed under threat of legal action.

All of Davies’s corpora have the same highly useful and flexible interface and provide a level of detail and a size (up to 410 million words) that have been previously unavailable. Use of these corpora slants my analysis towards American
English, however. An extension of this investigation into other international varieties of English is planned as a future project.

1.4 Organization of this Dissertation

Because of the largely synthetic mission of this dissertation (i.e. to combine different linguistic models), there are a total of eight chapters. Chapter 2 is an introduction to one of the frameworks, namely the modern subfield of diachronic linguistics and specifically that of grammaticalization studies. In addition, I examine such topics as lexicalization and pragmicalization. Beyond diachronics, I also use this chapter to examine a few models of inquiry into pragmatic and discourse analytical questions, mostly Politeness Theory (Brown & Levinson, 1987[1978]), Relevance theory (Sperber & Wilson, 1986) and models dealing with questions of stance.

Chapter 3 addresses the question of adverb as category. In this chapter I look at some previous criteria for the categorization of adverbs, going back to the study of Ancient Greek by Dionysios Thrax. Semantic models are largely found to be too broad, as are purely syntactic theories. Instead, a morphosyntactic model based on word-level etymologies is suggested. The category adverb is proposed to be a collection of lexicalized adverbials that have their origins in other categories.

Chapter 4 turns to the other major framework used in this dissertation: generative grammar. A history of generative grammar turns up a pattern of inattention to adverbs, inattention that largely continues until the rise of two competing theories: cartography (Cinque, 1999) and a highly articulated semantic
approach (Ernst, 2002). These two theories are examined in some detail, and the choice of a cartographic approach is explained.

In Chapter 5, I turn to speech-act adverbs in general, and *frankly* in particular. These adverbs are found at the farthest left periphery (discounting discourse particles and CP-elements) and comment on the manner with which the upcoming proposition is spoken. In this chapter, I propose a model of change that does not rely on grammaticalization but instead on lexicalization as the main engine. In support of this proposal, I examine data from various corpora (concentrating on the *TIME* Magazine Corpus) and offer tree structures to represent the changes.

Chapter 6 continues the study of English sentence adverbs by looking at Cinque’s categories of *evaluative*, *evidential*, and *epistemic* (which I term *EEE adverbs*). These three subcategories are combined because I propose that all three arise (in English) from copula clauses with an adjective as the predicate (e.g. *It is clear that he fell*). I use a syntactic framework from den Dikken (2006) to demonstrate the reanalysis of a verbal copula to a *RELATOR* in the form of the *-ly* suffix. This solution is also expressed graphically through syntax trees.

Chapter 7 extends, in a very brief format, the morphological aspects of the proposals put forth in this dissertation to Standard Chinese adverbials. I use data from Ji’s (2007) dissertation to support the contention that highly similar changes have happened in a language as morphologically marked as Chinese.

Finally, Chapter 8 provides a brief set of conclusions and suggestions for future research.
Chapter 2

THEORETICAL FRAMEWORKS, PART I: HISTORICAL LINGUISTICS

2.1 Introduction

2.1.1 Context.

One of my goals for this dissertation is to continue the work of those linguists who are seeking to unite insights from both “functional” and “formal” areas of linguistics. In this chapter and in Chapter 4, I will mostly maintain a separation between the two camps for purposes of reviewing existing literature, which has itself been kept largely separate. In the data chapters, however, I will endeavor to unite various aspects of both generative and functional linguistics.

An interaction between the sociocultural requirements of communication and the structural parameters of the human language faculty is, in my opinion, ultimately the cause of internal language change, and therefore the focus of this dissertation.

2.1.2 Chapter organization.

This chapter is organized as follows: section 2.2 is focused on background to a critical analysis of current theories in historical linguistics, followed by sections covering grammaticalization (2.3), lexicalization (2.4), and the shift from parataxis to hypotaxis in the history of English (2.5). In addition, in these sections I address the way in which I am applying these topics to the study of adverbs. Section 2.6 turns to pragmatic approaches to understanding adverbs, particularly sentence adverbs and what are referred to (variously) as discourse particles. As
adverbs appear on the left periphery of sentences, they are assigned several
significant pragmatic roles in discourse. In 2.6 these roles will be examined, using
a variety of pragmatic models. Section 2.7 sums up the chapter.

2.2 Diachronic Linguistics—Background

The modern field of diachronic linguistics has its origins in *philology*, or
the study of classical languages (mainly Indo-European) through the examination
of the literature, cultural artifacts, and known history associated with those
languages. Philology demands a great deal of highly specific knowledge that may
or may not be useful in a broader context. In the eighteenth and nineteenth
centuries, with the growth and expansion of the scientific method in the natural
sciences came a corresponding interest in broader, taxonomic approaches to
language studies. The recognition in the late eighteenth century by Sir William
Jones of similarities among classical Latin, Greek, and Sanskrit—and by
extension, Germanic, Celtic, and Persian—helped spur interest in the
Comparative Method. All modern linguistics, both synchronic and diachronic, can
be traced back to this pattern of scientific study.

The twentieth century, as is well known, saw the separation of these two
methods of examining language under the influence of scholars such as Saussure.
Synchronic linguistics took on the leading research role, as Saussure was followed
by structuralist scholars like Bloomfield and Jakobson. They in turn were
succeeded by Chomsky, whose generative models, in various incarnations, have
strongly influenced the synchronic discussions of the past 50 years. Many of the
important insights of this period are either directly or indirectly attributable to Chomsky, his supporters, and his detractors.

Diachronic linguistics also shifted focus during the twentieth century, particularly in the latter half. Although comparative linguistics continued to play a role, especially in the study of ancient and proto-languages, a stronger interest in historical morphosyntax was developed. This interest was focused on historical processes and the motivations for those processes. The term for this recently revived branch of study is grammaticalization, coined by Meillet in his 1912 study “L’évolution des formes grammaticales”. Grammaticalization, it can well be argued, has sparked renewed interest in historical linguistics. As we shall see, it is also the subject of lively debates and competing definitions.

In section 2.3, I will carefully examine grammaticalization, including both its appeal and the frustrations that can be associated with this framework. In addition, I will examine two apparently related fields, pragmatization and lexicalization, and determine how these fit into (or are swallowed up by) grammaticalization studies. I also examine a topic somewhat less controversial, the shift of English from a paratactic to a hypotactic clausal structure, and I explain how these various historical concerns will be treated in this work.

2.3 Grammaticalization

2.3.1 Introduction.

As I mentioned above, grammaticalization is now a highly fashionable subject of inquiry in diachronic linguistics that has seen a great deal of recent
growth. Modern work in this area was initiated by Kuryłowicz (1965), who gives us an early but still highly relevant definition:

Grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to a more grammatical status, e.g. from a derivative formant to an inflectional one. (p. 69)

To fully understand this explanation, it is important to explain various terms that are used, in particular terms such as *lexical* and *grammatical*. These give us a starting point for understanding what grammaticalization is by definition. A lexical word, for the purposes of this study, meets certain semantic and syntactic criteria:

Lexical meaning is that aspect of meaning which is codified in a lexicon or dictionary, can be semantically analyzed, and together with the grammatical elements of meaning (such as mood, tense, …) yields the whole meaning of a linguistic expression. Normally lexical meaning consists of an open class of elements, whereas grammatical meaning is restricted to a closed class of elements…. The distinction between the two types is not always clear. (Bussmann, 1996, p. 277)

Semantically, then, lexical words are *autosemantic*, or possessed of referent-based meanings. Grammatical words, on the other hand, are *synsemantic*, without clearly definable meanings in isolation (Bussmann, 1996). They are instead dependent on their environment to produce meaning.

Category-wise, lexical words belong to the *open* class. New members of this class are added regularly. Nouns, verbs, and adjectives are all open classes. Grammatical words, in contrast, belong to *closed* classes. It is rare and noteworthy for new members (or even new subclasses) to appear. Verbal
auxiliaries, determiners, and complementizers are some examples of closed classes.

One of the important questions (for purposes of this dissertation) concerns adverbial categorization. Adverbs seem to combine elements of lexical and grammatical words. For example, they are generally autosemantic and members of a mostly\(^1\) open class. Yet their meanings can be affected by syntactic position—there are considerations of scope that must be discussed and accounted for. We will examine this in more detail in Chapter 3, where it will be seen that adverbs themselves seem to be a largely morphological class, primarily (with the exception of opaque examples such as those in note 1 below) made up of fused phrases and/or the results of derivational morphology processes. We will also discuss scope and syntactic models of adverbs, where the interactions of semantics and syntax are significant and are the subject of heated debate, in Chapter 4.

Implicit in such a distinction and in a concept like grammaticalization is the belief that language originates (in the spheres of both ontogeny and phylogeny) with only the most distinctly lexical elements, and that grammatical structures and words/morphemes must be developed over the course of time to meet the needs of speakers (see e.g. Bickerton, 1990). Such a model is borne out,

---

\(^1\) Scholars who study adverbs from various models (diachronic vs. synchronic, functional vs. formal) have noted the distinction between currently productive, derived forms (e.g. \textit{-ly} in English, \textit{-ment(e)} in Romance) which are open, and older, more unitary forms (e.g. deictic \textit{here} and \textit{there} in English, in which the morphology has become opaque, or \textsc{Adv\,fast}, where it has disappeared), which are closed. See, for example, Alexiadou (1997), Haumann (2007) for synchronic and Brinton and Traugott (2005) for diachronic discussions.
of course, by language acquisition patterns in children, where often concrete, strongly lexical words with real-world referents (usually nouns) are first acquired. Mastery of grammatical structures and the correct use of grammatical elements come much later.

Yet this discussion can be thought of as somewhat fragile, dependent as it is on judgments of what is lexical and what is grammatical. As Bussmann (1996) notes in the excerpt above, the distinctions between lexical and grammatical are not always clear-cut. Nouns (especially those whose referents are tangible, real-world items) may rest at one (the “lexical”) extreme and tense markers at the other (“grammatical”), as hypothetical examples, but there is clearly a gradient of categories in between.

Perhaps the most difficult categories to place on a spectrum of grammatical to lexical are adpositions and adverbs. Most adpositions (i.e. English prepositions over, from, at), for example, clearly indicate spatial–relational meanings. There are complex prepositions (such as atop, between, or next to) that require the use of a formerly independent noun (top, twain) or adjective (next) and seem to be more lexical than shorter prepositions like on or up. Preposition is, however, considered to be a closed class in English and many other languages.

Similarly, adverbs (as something of a “grab-bag” category) combine elements of both (as mentioned above) and are likely somewhere in the center of the spectrum. Adverb is a “fuzzy” category, without the general consistency of lexical verbs or grammatical determiners. As questions of grammaticalization and lexicalization are applied to adverbs or adpositions, we have to be careful.
Lehmann (2008) notes, “Little is know *a fortiori*, about the factors underlying the primary categorization of concepts such as adverbs.” Here he compares adverbs to the more easily understood categories of nouns, adjectives, and verbs.

Adverbs are very much a mystery class, somewhat resistant to facile attempts at assignment to either lexical or grammatical type. The relative instability of a “line of demarcation” between the terms *grammatical* and *lexical* is an important concern to keep in mind as we attempt to defend terms like *grammaticalization* and *lexicalization* for the purposes of this study.

2.3.2 Current understanding.

Certainly since the 1960s, when Kuryłowicz first published the definition excerpted above, the field of grammaticalization studies has grown, even exploded. Leading contemporary scholars (those who approach grammaticalization through a traditional, functional framework) include Christian Lehmann, Bernd Heine, Paul Hopper, and Elizabeth Closs Traugott. In fact, Traugott is by far the most heavily involved in the field of Grammaticalization Theory. She has written or co-written scores of books, articles, and conference presentations on the subject, and she is widely acknowledged as a leading figure in the field.

As the study of grammaticalization has developed from the original observations of Meillet into a modern linguistic theory, there have been several concerns typical of a young field: terminology; criteria for judging whether a
particular change is or is not grammaticalization; and even the question of whether grammaticalization is a discrete, definable phenomenon at all.

By far, the largest number of grammaticalization theorists work within the confines of a functionalist (rather than a formalist/generative) framework, and so the language change they focus on is driven by considerations of speaker innovation and socio-pragmatic influences. This innovation need not be found in child language acquisition, but, as Hopper and Traugott (2003, pp. 43ff.) indicate, can be found in speaker changes in young adulthood and beyond. Intergenerational acquisition of new forms is not overtly required to constitute a measurable change to a language. This can be a concern for a morphosyntactic change like those associated with grammaticalization, but it is one we shall turn to in detail in Chapter 4, where we examine the intersection between generative linguistics and diachronic phenomena like grammaticalization.

Perhaps the most comprehensive introductory study of grammaticalization can be found in Hopper and Traugott’s (2003) straightforwardly-titled textbook *Grammaticalization*. The shift from phenomenon to theoretical framework that has characterized the past three decades in particular can be seen in Hopper and Traugott’s redefined assessment of the term:

As a term referring to a research framework, “grammaticalization” refers to that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions. (2003, p. 1)

---

2 The first edition of Hopper and Traugott’s *Grammaticalization* was published in 1993. This dissertation references the second revised edition (2003).
Grammaticalization is, in this definition, no longer simply a type of phenomenon where a lexical item begins to be used in a grammatical sense. In the view of researchers such as Hopper and Traugott, it has been elevated to theoretical status. As a theory, grammaticalization depends upon several elements. Just what these elements are, and just which ones are crucial to the theory, is sometimes unclear, as there are different criteria used by different scholars at different times. Among those who consider grammaticalization to be theoretical, there is emphasis on mechanisms on the one hand, and on (nearly) universal tendencies on the other.

One of the distinctions that can be made in the study of grammaticalization is whether to look at diachronic or synchronic data. Usually, the processes of grammaticalization are studied over time, diachronically, and the changes that happen in the use and understanding of the word or phrase in question can be traced in such fashion. This is the approach that Hopper and Traugott prefer. However, grammaticalization can also be studied synchronically, and one of the phenomena associated with this sort of synchronic study is known as layering. Layering is “the synchronic result of successive grammaticalization of forms which contribute to the same domain” (Hopper & Traugott, 2003, p. 125). Elsewhere, Hopper and Traugott more simply indicate that layering can be found when “older and newer forms co-exist for individual speakers as well as for communities over time” (2003, p. 49).

Layering can be conceptualized as a phenomenon that is similar to the results of syntactic Move (now called Internal Merge). In generative grammatical theory, constituents (subjects, verbs, wh-pronouns, quantifiers) leave their original
(externally merged) positions in the sentence and are attracted by and move to positions higher in the tree (closer to the left periphery of the clause). Older theories speculated that what was left behind in the original position was a trace of the constituent; current theories offer a copy and delete alternative, where the constituent appears in several places in the syntax but is only given phonetic expression in its final (“highest”) position. Mistakes can occur in speech to indicate that such “lower” copies may indeed exist, as they can occasionally be left in place, especially in longer clauses.

Layering can be metaphorically thought of, therefore, as the remnant traces or copies, and as earlier forms of the word or phrase in question. After grammaticalization has begun, those earlier forms will often co-exist with the newer form for some time. In fact, the two forms may have a fairly permanent co-existence (consider English perfect have and the more lexical, older form, meaning “to possess”). And not all grammaticalization processes will be complete, resulting in the elimination of the lower, older, more lexicalized form.

In the case of will (which went from meaning “desire” or “intend” to being a grammatical future marker in ModE), there was a borrowed Scandinavian form in want that eventually shifted from only meaning “need” or “lack” to take on the meanings “desire” and “intend” from will (OED online). But this may be an unusual occurrence, which nevertheless has taken centuries to develop, as will began to be used as a future marker in the Old English period. More common may be the situation with have, where layers co-exist in a fairly stable relationship.
2.3.3 Grammaticalization Theory.

In Hopper and Traugott (2003), the most significant prerequisites to grammaticalization are pragmatic, as their focus is on communicative requirements. New communicative needs, or those that need reinforcement, trigger in speakers the two major mechanisms of grammaticalization: reanalysis and analogy (p. 39). Reanalysis is by far the more important of these two mechanisms. From a production-to-comprehension standpoint, then, reanalysis occurs when two interlocutors engage in discourse and “the hearer understands a form to have a structure and a meaning that are different from those of the speaker” (p. 50). Analogy is a different type of process, overt where reanalysis is covert, and concerning “the attraction of extant forms to already existing constructions” (pp. 63-4). In analogy, the initial framework has already been created, and new elements pattern after an originator.

Crucial to this theory, and perhaps the most controversial aspect of it, is the concept of unidirectionality. Hopper and Traugott (2003) indicate that grammaticalization is “hypothesized to be prototypically a unidirectional phenomenon” (p. 99). Of note is the relative strength of the explanation of unidirectionality as prototypical. For theoretical reasons, the authors do not claim universality (as that would echo the generative position they eschew), but this is nevertheless a powerful statement. Hopper and Traugott are confident enough in this position to regard degrammaticalization counterexamples (such as those in Norde, 2009) as largely “sporadic” (p. 99). Their model for unidirectionality is the cline. An example is below:
A cline of adverbial grammaticalization (from Hopper & Traugott, 2003, p. 37):

clause-internal adverbial > sentence adverbial > discourse particle

We will return to this particular cline, important as it is to this study, below.

Although Heine and Kuteva (2002) agree to the theoretical aspects of grammaticalization, their focus on mechanisms differs somewhat from Hopper and Traugott’s. Significantly, they do not include reanalysis in their theoretical model; instead their focus is on four closely interrelated mechanisms. These mechanisms are largely agreed upon by most grammaticalization scholars:

Four grammaticalization mechanisms (Heine & Kuteva, 2002, p. 2)

a. desemanticization (or “semantic bleaching”)—loss in meaning content,

b. extension (or context generalization)—use in new contexts,

c. decategorialization—loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms, and

d. erosion (or “phonetic reduction”)—loss in phonetic substance.

All of these elements are also found in Hopper and Traugott’s model; however, they do not appear as explicitly interrelated aspects of a single process as they do
here. Extension must occur first for Heine and Kuteva, followed by the others: desemanticization, decategorialization, and finally erosion. In fact, for most theories, phonetic change is least important to the definition of grammaticalization.

2.3.4 Grammaticalization and generative syntax.

In the past decade, diachronic generative scholars have turned their attention to grammaticalization phenomena and have offered models based on Chomskyan grammar (both the Principles and Parameters and the minimalist frameworks). In Chapter 4, I will examine these more closely in conjunction with my discussion of generative syntax. However, it will suffice to say here that the two most important works covering generative models of grammaticalization are Roberts and Roussou (2003) and van Gelderen (2004).

In Roberts and Roussou (2003), the shifts indicated by grammaticalization phenomena are triggered by a speaker resetting her/his parameters, thereby reanalyzing the word. In van Gelderen (2004), the triggers are related to principles (notably types of Economy Principles) that are considered inherent to all living beings and their systems. In this way, grammaticalization is situated in not only current minimalism but also the biolinguistic frameworks researched by current generative scholars. See Chapter 4 for a more complete discussion.
2.3.5 (Inter)subjectification and semantic considerations.

In Traugott’s more recent work, she emphasizes a unidirectional semantic change that she identifies with grammaticalization. This is carefully developed in Traugott and Dasher (2002) and in Traugott’s later articles, and it is termed *subjectification/intersubjectification*. Traugott and Dasher, borrowing from cognitive linguists such as Langacker (e.g. 1990), distinguish among *objective* language, *subjective* language, and *intersubjective* language.

The first is primarily declarative, with minimal deixis. The second type makes the speaker’s deictic considerations significant, emphasizing her/his *attitude*, both toward the proposition and toward the structure of the discourse. Intersubjective language brings the *connection* between the speaker and the hearer to the fore, and is strongly deictic as a result. Grammaticalization Theory is thereby tied to a shift in the semantico-pragmatic patterns of language use, from less to more subjective.

In a unidirectional schema, the shift for Traugott and Dasher (2002) is from objective to subjective to intersubjective. This shift is toward *embodied* language use, where communication is a reflex of generalized cognitive principles rather than a specific language faculty, as is found in the generative model. The speaker begins to use language to signal her/his place in the world at the same time as a communicative goal is achieved. Significant to subjectification is the extension of a pattern from one environment to another. In Brinton and Traugott (2005), the authors give this example, borrowed from Langacker (1990):
(3) An earthquake is going to destroy that town.

(Brinton & Traugott, 2005, p. 29; from Langacker, 1990, p. 2)

In example (3), the grammaticalization of *be going to* has been extended from an animate, mobile subject to an inanimate event. It is difficult to analyze the actions of an earthquake as involving a verb of motion. The extension has been (to use theta-role terminology) from *agent* to *causer*. The other shift here is that the speech time of the utterance is the focus of *be going to*; it is purely the speaker’s perspective that is expressed (Brinton & Traugott, 2005, p. 29).

This is subjectification as explained in Traugott’s work. As we will see, the movement of adverbs along Hopper and Traugott’s cline (above in example 1) follows this shift exactly. There is little doubt that a great deal of semantic shift follows the various directional paths indicated by Traugott and Dasher (2002). Well-known and well-trod roads tend to lead from concrete meaning to abstract, from spatial to temporal, and so on. The questions are whether such semantic shifting should be considered definitive or even emblematic of grammaticalization phenomena, and whether intersubjectification is truly the next step.

### 2.3.6 Criticisms.

If we look at an adverb like *frankly* (see further discussion in Chapter 5), we can see that there is a shift in its use, from the manner adverb in “They gave *frankly* of their time to help the cause”, to a more subjective, commentary-style
use that signals the interlocutor in “Frankly, they gave a lot of time to the cause”.
There is an increase in subjectivity, as the meaning of “freely” or “openly” shifts
to a personal meaning of “I speak frankly in saying”, and in intersubjectivity, as
the adverb is directly addressed to the hearer to prepare her/him for the
proposition, as well as being used to explain the speaker’s attitude to what s/he is
saying. So far, so good.

However, can such a change be compared to a traditional example of
grammaticalization such as the shift from demonstrative to definite article in
English and other languages? (Compare Hopper and Traugott, 2003, pp. 135-6.)

There are several markers of grammaticalization in the change: extension
of use; desemanticization, where the distal meaning is bleached; and phonological
erosion. There does not seem to be an increase in subjectivity or intersubjectivity
in such a case, as there is actually a decrease in speaker-oriented deixis and
therefore an apparent increase in objectivity. See table 2.1 below.

| Table 2.1 Grammaticalization of OE *pe* to ModE *the* |
|-----------------|-----------------|-----------------|-----------------|
| **Semantics** | **Context** | **Morphosyntax** | **Phonetic value** |
| se/pe distal deixis | used only in specific circumstances to indicate speaker perspective | inflected for gender/number (SG.M) | SG.NEUT. þæt retained in Modern English as that with strong stress |
| the definiteness; meaning has shifted | used in broad circumstances, with most common nouns | not inflected | reduced vowel in consonant environments |

(data from OED)
Other signals of grammaticalization are present, as mentioned above Heine and Kuteva’s (2002) criteria. There is a bleaching of semantic meaning from e.g. *se man* “the man there” to *the man* “the man”. There is phonetic erosion, with the vowel at schwa value in most consonantal contexts; there was cliticization for a period in early Modern English (e.g. Shakespeare’s Sonnet 129: “Th ‘Expense of Spirit in a Waste of Shame”). There is an extension of contexts, from specific nouns that are distant from the speaker to all nouns; yet in this extension is the loss of a degree of subjectivity.

Is this a counterexample to the subjectivity cline as it seems to be, as it is an uncontroversial example of grammaticalization? Hopper and Traugott acknowledge that “the continua of grammaticalization are not exceptionless” (2003, p. 211).

In my opinion, however, this is the strongest concern with elevating grammaticalization to a theory, and with adding broad semantic criteria like subjectification to the definition. The question arises: How exactly, then, should we define grammaticalization? For instance, many linguists would regard the frankly example as actually being a type of pragmaticalization, a similar phenomenon in which words take on stronger discourse-oriented roles rather than grammatical ones. Perhaps the weakness is in the term grammaticalization and its reliance on various definitions of what exactly it is that makes a word or expression “grammatical”.

Similar criticisms are expressed by Campbell (2001), in which paper he asserts that what is called “grammaticalization” is merely a collocation of several
different linguistic changes. Campbell points out, for example, that making unidirectionality a prerequisite of a phenomenon that is defined by its unidirectionality is something of a tautology; that neither semantic “bleaching” nor phonological reduction are diagnostic of grammaticalization; and that Grammaticalization Theory is not predictive but is instead descriptive of observed phenomena.

To me, an additional critique to Traugott’s conception of Grammaticalization Theory regards its reliance on pragmatic considerations alone as the driving force behind a unified model. The weakness here is, I believe, the very element that many functionalists point out in their review of formalist theories: namely the differences among cultures and individual languages. The assumption is that all cultures utilize the same metaphorical or metonymical patterns when extending meaning.

Despite the care that Traugott and her co-authors take to establish clines, unidirectionality, and increased subjectivity as merely very strong tendencies, making grammaticalization dependent on a consistent human pattern of metaphor and metonymy seems to be dangerously close, again, to proposing a linguistic universal. As has been shown in criticisms of other pragmatic models (see critiques of e.g. Politeness Theory, Brown & Levinson 1987[1978], in Watts, 2003), not all cultures value all communicative aims in the same way.
2.3.7 Grammaticalization in this work.

For the purposes of this dissertation, I will use grammatization as a term that is not specifically theoretical, but instead relatively mechanistic and phenomenological. Despite the inadequacy of the term to cover the territory it is assigned by some scholars, I will accord it some leeway. As I briefly mentioned above, for example, some linguists use pragmaticalization as a term to indicate the change from lexical to discursive (rather than grammatical). I will acquiesce to most scholarship, however, and consider pragmaticalization to be a subset of grammatization (with some reservations).

Grammaticalization appears to me to be results-oriented (in other words, truly predictable only after the path has been followed), and therefore I will apply the term in this descriptive fashion. My focus will be on the latest syntactic position of a form (and its use in the sentence) and comparing these criteria to its previous position to determine whether a process of grammatization has indeed occurred.

2.4 Lexicalization

2.4.1 Introduction.

When I began to work on this project, my understanding of the term lexicalization was simple and direct. Straightforwardly, lexicalization was to me the combination of the elements of a phrase (or an even larger syntactic unit) into a single lexical item. This has been the working definition I have used until reading the various perspectives of current grammatization and lexicalization.
scholarship, in which there is little agreement. Terminology, as evidenced by the discussion in section 2.3, is a chief concern in this work. Therefore, I will use this space to explain my interpretation, including a resistance to radically changing or narrowing my definition of lexicalization as listed above; however I will also examine the various current ways of understanding both the term and the phenomenon.

2.4.2 The state of the art.

Brinton and Traugott (2005) provide the most comprehensive model of understanding lexicalization; this model is influenced and restricted by its position vis-à-vis grammaticalization (as canvassed in the previous section). As in the various discussions of grammaticalization, the definition of lexicalization is crucial to its comprehension and use. First in the defining process is distinguishing *lexicalization from degrammaticalization*.

There has been a great deal of confusion between these terms, but it is important to carefully separate the two. Degrammaticalization is a comparatively rare occurrence in which a formerly dependent, grammatical unit (either a bound morpheme or a grammatical word/phrase) becomes independent and takes on autosemantic (as discussed in section 2.3) meaning. Norde (2009, pp. 3-5) describes the three types of degrammaticalization: *degrammation*, where “a function word is reanalysed as a content item”; *deinflectionalization*, where “an inflectional affix becomes less bound, while at the same time gaining in semantic or functional substance”; and *debonding*, where bound morphemes become free.
A well-known example of degrammaticalization (specifically deinflectionalization) is the change in English and Swedish in the structure of the genitive case. In both these Germanic languages, the old form is one where all elements in a Determiner Phrase (Noun Phrase) are marked with genitive -s endings. In both languages, the genitive is currently marked by the same affix at the end of the full phrase. See (4) for a Swedish example:

(4) ens riks mans hws >
    a-MASC.SG.GEN rich-MASC.SG.GEN man-MASC.SG.GEN house

    en rik mans hus
    [a rich man]’s house

(Old Swedish > Modern Swedish; from Norde, 2009, p. 4)

By contrast, lexicalization is far less infrequent. In addition, the use of “lexical” in the term is not precisely the same as it appears in the definition for grammaticalization, where it means autosemantic. Where degrammaticalization is, relatively speaking, a phenomenon that can be considered to be the opposite of grammaticalization, in which a unit of grammar becomes a noun, verb, or adjective in its own right, lexicalization is initially the creation of a single new

---

3 Norde (2009) offers such a definition with the caution that “there are no examples of degrammaticalization ‘all the way up the cline’—a degrammaticalization chain from suffix all the way up to lexical item has not been attested” (p. 8).
member of the lexical inventory of a language. Consider Brinton and Traugott’s (2005) definition:

The change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the word may become more lexical. (p. 96)

This definition appears to be mainly morphosyntactic. The shift here is from a syntactic structure that (in generative terms) requires operations of Merge to a morphological structure where the merger is (at least in some theories, but cf. the discussion of Distributed Morphology in Chapter 4) not required before the word’s use. Another way to see this definition is that lexicalization is the historical process in which a syntactic complex becomes a simplex.

There is a word in this definition that throws a small monkey wrench into the works, however, and that word is “contentful”. Although not stated directly in this definition, a contentful word is one of the three major lexical classes: noun, verb, and adjective. So in Brinton and Traugott (2005), the development of complex prepositions (in English and other languages) discussed earlier is dismissed as “typical of grammaticalization” (p. 65), primarily because the end result is not a member of one of their lexical classes.

Adpositions and adverbs are often thought of as similar classes, in that they occupy the middle ground between undoubtedly lexical nouns and verbs and clearly grammatical determiners, for example. As a result, this too becomes an
element of the characterization of change in adverbs as generally being examples of grammaticalization, instead of lexicalization.

So, narrowly, lexicalization applies only to the shift from a syntactic complex of items (i.e. phrase or clause) into a simplex that is qualified as belonging to a major lexical class. It is not clear that the conflation of the two types of *lexical* (“single word unit” vs. “only noun/verb/adjective”) is the basis for Brinton and Traugott’s definition. The movement, even vacillation, between these two criteria seems to make this definition problematic, and it is evident that the balance here has not yet been agreed upon.

Brinton and Traugott’s definition is not the only one, however. Lehmann (2002) focuses on somewhat different criteria:

*Grammaticalization involves an analytic access to a unit, lexicalization involves a holistic access to a unit, a renunciation of its internal analysis. Both processes do not concern signs in isolation, but signs in their paradigmatic and syntagmatic relations... A consequence of this explication of the notion of lexicalization is that the coalescence of two grammatical morphemes must be called lexicalization... The adjective *lexical* has two meanings in linguistics, 1) belonging to the inventory, 2) having a specific, concrete meaning... Lexicalization is a process in which something becomes lexical in the first of the two senses. (pp. 16-17)*

This definition seems more forthright and in keeping with the spirit of lexicalization as I have been given to understand it. The decision by Brinton and Traugott to limit lexicalization to the formation of open-class, lexical words (as they define them) seems to be arbitrary and counter to the spirit of the term.

In lexicalization, the primary concern is historical word formation, and the prototype of such a process is *compounding*. Compounding may include the combination of two items of the same class or of different categories (compare
birdhouse NOUN-NOUN with bluebird ADJECTIVE-NOUN). In the next chapter, we look at some examples of compounding and similar processes in adverb formation. Other typical formation patterns include blends, coinages, affixing, and ellipsis (Brinton & Traugott, 2005, pp. 33ff.).

The shift from syntactic to morphological structure is called *de-syntacticization* or *univerbation*, and this is the crucial signal that lexicalization is occurring. Other processes often associated with lexicalization include *fusion* (morphological reduction) and *coalescence* (phonological reduction) (Brinton & Traugott, 2005, pp. 47ff.). In words where lexicalization has occurred in the more distant past, there is often sufficient fusion and coalescence as to make the origins opaque.

### 2.4.3 Lexicalization’s place vis-à-vis Grammaticalization Theory.

With the growth of Grammaticalization Theory over the past couple of decades, there has been an effort to give a limited definition to lexicalization (as Brinton and Traugott do, above) and to allow grammaticalization to incorporate many elements (i.e. univerbation, fusion) that were traditionally restricted to lexicalization.

One of the clearer delineations of the new order between these two phenomena is given by Himmelmann (2004), in which grammaticalization is defined as a type of context expansion across three levels (host-class, syntax, and semantics/pragmatics). Lexicalization, in contrast, does not experience host-class expansion. It must be regarded as idiosyncratic, and thereby limited to erratic
occurrences; in fact, any sort of analogous patterning must be regarded as indicating grammaticalization instead (Himmelmann, 2004, pp. 36-8).

What is not clear to Himmelmann, however, is how to categorize the emergence of derivational formatives like -ly in English. This is of course important to the study of adverbs. Brinton and Traugott (2005, pp. 132-6), in their discussion of -ly adverbs, postulate that the suffix is moving from derivational to inflectional. This is evidence to them that, whatever its origins (i.e. in nominal lexicalization; see Chapter 3), -ly is becoming a grammaticalizing affix. A third viewpoint is found in Joseph (2003), who considers such a development an example of the (perhaps less terminologically loaded) phenomenon of

\textit{morphologization}, which he defines thus:

a set of developments by which some element or elements in a language that are not a matter of morphology at one stage come to reside in a morphological component—or at least to become morphological in type—at a later stage. (Joseph, 2003, p. 473)

This seems to be a more general characterization that could be applied to the definitions of both lexicalization and grammaticalization. We will look further at -ly in Chapter 3.

In summary, then, lexicalization is primarily regarded by grammaticalization theorists as a combination of several fairly restricted processes. Two or more elements must be combined to create a new form that is not necessarily predictable from its origins. Phonological reduction is important here, where it is less so for grammaticalization. Crucial to the definition of
lexicalization among grammaticalization theorists is its limited, idiosyncratic nature. Yet the terms used in these definitions still have some knotty problems.

The use of “lexical” in lexicalization does not, for some, indicate that the target morphological construction is necessarily a content word like a noun, verb, or adjective. Yet this point is not quite clear; Traugott, as mentioned above, does not accept complex prepositions as examples of lexicalization (see Brinton & Traugott 2005, p. 65) but purely as grammaticalization, despite the univerbation, fusion, and coalescence of e.g. *beside* (as in “Stand beside the tree.”). For her this is not a “contentful” form, despite its continued phrasal meaning (*by the side of*).

Yet for Lehmann (2002), such a process that forms complex prepositions can only be interpreted as lexicalization. So the criteria and the definitions are somewhat fluid, and much depends on the analysis and the scholar. This is a concern, as it is important for a subject of study to be agreed upon by all who study it.

2.4.4 Lexicalization in this work.

This research project was begun as an examination of the grammaticalization of sentence adverbs in English, was then switched to a study of their lexicalization, and has been brought full circle at this point, where both issues are of importance. Both Lehmann (2002) and Himmelmann (2004) regard grammaticalization and lexicalization as being orthogonal rather than opposed phenomena, and I agree for the purposes of this dissertation. Therefore, the presence of lexicalization processes in the history of a particular word in this
study does not preclude the inclusion of grammaticalization, and vice versa. Lexicalization is not the same as degrammaticalization.

Indeed I see these two processes (lexicalization and grammaticalization) as different in orientation: where I regard grammaticalization as being phenomenological and results-oriented, I see lexicalization instead as mechanistic and primarily morphological, a combination of its three elements of univerbation, fusion, and coalescence. Grammaticalization produces a functional or (in its guise as pragmaticalization) a metatextual/discourse-related term. Lexicalization produces a new word (or lexical entry/Vocabulary Item) from a frequent collocation. Therefore, the two occupy quite different spheres and do not overlap. Diachronic discourse and grammar schemas alike often layer functional meaning atop existing or residual semantic meaning—a phenomenon known as persistence (Hopper & Traugott, 2003, p. 96). In fact, the two processes have even been known to occur at different points in a word’s history.

Each of the specific uses of these terms in this study is notably a reduced, streamlined version of the broader meanings. I have chosen not to take on some of the “baggage” of the more complex, more restricted, and more theoretical meanings, as these theories seem to be somewhat unclear, particularly in the definitions of grammatical and lexical.

Issues like subjectification (for grammaticalization) and highly restricted idiosyncracy (for lexicalization) are important elements of the definitions for Grammaticalization Theory, but they are not as significant, except as coincidental support, for this study. In addition, one of my focuses in this dissertation is on the
morphology and syntax of adverbs, albeit from a diachronic point of view.

Although pragmatic considerations and semantic shifts are significant to
diachronic syntax and to adverbs, they are not the only focus of my work here.

2.5 Parataxis and Hypotaxis

2.5.1 Introduction.

For the typological case studies in Chapters 5, 6, and 7, there is another
important trend in diachronic linguistics that plays a part in the examination of
data. This tendency is the shift in historical English (and in many other
languages—particularly Indo-European) from simple sentence structure to more
complex clausal combinations. This is a pattern that is more widely seen in
written language than in spoken, and it therefore is often associated with literacy.

The development of complex sentence structures in English, I contend,
plays an important role in the historical shift of adverbs to the left periphery of the
sentence. In this I follow scholars such as Swan (1988) and Fischer (2007). In
later chapters I will argue that sentence adverbs are the final result of a pattern of
clausal combination, that said adverbs are the remnants of clauses that have been
reduced, or desententialized.

Important to this discussion are, of course, elements of terminology, and
those terms (not surprisingly) have a variety of different definitions. This portion
of the chapter will be devoted to hashing out those definitions from a synchronic
point of view, and then applying an understanding of historical change (e.g.
grammaticalization) to the clause.
Sections 2.5.2–2.5.3 are concerned with terms such as *parataxis*, *hypotaxis*, *dependency*, *coordination*, and *subordination*. In section 2.5.4, I examine how the grammaticalization framework has been applied to the shift from simple to complex clause structure. In section 2.5.5, I define how these terms are used in this dissertation. I also explain how the change from paratactic clause to hypotactic clause is the initial grammaticalization process that then triggers the lexicalization of adverbial clauses.

**2.5.2 Levels of clausal connection: parataxis.**

The clause in language is dependent in its organization on the presence of a verb and its arguments. The grammatical expressions of these arguments include, for English, the traditional subject, direct object, and indirect object. Different verbs require differing numbers of arguments, but in English there is at least one of these positions filled in all clauses.

The simplest means of connecting two clauses in English is simply to juxtapose them. In example (5), the two clauses are related, but there is limited signaling of this fact:

(5) Comma splices are often found in student writing. Full sentences should be connected by semicolons, conjunctions, or periods.

The only signal of a relationship is the lexical subject matter of the sentences, which concerns writing and punctuation. A further level of connection can be
signaled by cross-clausal anaphora. This is commonly seen across languages.

Consider example (6):

(6) I love my dog Rufus. He is a boxer.

These are types of paratactic clausal structures. Parataxis is a phenomenon where the clauses involved are of equal independence. Paratactic clauses can be combined into a single sentence either through punctuation as noted in the content of the sentences in example (5)—by using semicolons—or through the use of coordinating conjunctions, by which mutual independence is mostly maintained.

See example (7):

(7) I walk to school and Philip rides his bike.

This is a strategy seen in many Old English texts, such as the Anglo-Saxon Chronicle. The Chronicle, however, also uses coordinators to connect temporally ordered information. Such usage shifts us away from pure parataxis toward hypotaxis, in which clauses demonstrate hierarchical structure. See the OE example in (8), from the Chronicle (Peterborough version):

(8) Old English:

An. M.LXVI. On þyssum geare man halgode þet mynster æt Westmynstre on Cyldamæsse dæg 7 se cyng Eadward forðferde on Twelfts mæsse æfen
7 hine mann beýrġede on Twelftæn mæssēдеg innan þære niwa halgodre circean on Westmyntrē⁴

*Modern English:*

1066. In this year the monastery at Westminster was hallowed on Childermas day (28 December). And king Eadward died on Twelfth-mass eve (5 January) and he was buried on Twelfth-mass day, in the newly hallowed church at Westminster.

(Thorpe, 1861, from van Gelderen, 2006)

Note that with the lack of punctuation, the translator Benjamin Thorpe is able to choose whether to use a single sentence or separate sentences in representing the “and” connection.

Another way to temporally connect clauses is through *asynthetic* coordination (cf. Lehmann, 1988), where the “and” relationship is indicated simply by clausal order—by juxtaposition, as above, but with real sequential meaning. The best known of these is a quote from Julius Caesar:

(9) Veni, vidi, vici Latin

I came, I saw, I conquered

⁴ The use of “7” represents *ond* (“and”) in Old English. This mark is known as a *Tironian note* and is descended from the same origin (Latin *et*) as the ampersand “&” in modern typography. It is very frequent in OE texts.
2.5.3 Levels of clausal connection: hypotaxis and a confusion of terms.

Once clauses are combined in ways that move beyond simple juxtaposition or coordination, the relationship between them is said to be \textit{hypotactic}. There are different types of hierarchical relationships among clauses, and different linguists use terms such as \textit{subordinate}, \textit{dependent}, \textit{bound}, and \textit{embedded} to express levels of engulfment by one clause of another. Certain strategies of hierarchical relationship are expressed by terms such as \textit{interlacing} and \textit{extraposition}.

Jendraschek (2007) examines a corpus of eight different descriptive linguistic articles and finds, unsurprisingly, that each author (even the same author in different articles) uses these terms in different ways in each article. For some, \textit{subordinate} clauses are less separable from the main (or \textit{matrix}) clause than are \textit{embedded} clauses; for others, the reverse is true. For Jendraschek, from a synchronic perspective (and from a diachronic one as well, as we will see), the key concept of judging the dependency of a clause is \textit{desententialization}. He defines this term by using grammaticalization as a source:

\begin{quote}
Desententialization is the syntactic change by which a construction loses sentence properties, \textit{moving down the cline} from independent sentence to a simple deverbal lexeme within the matrix clause. (Jendraschek, 2007, p. 10; my italics)
\end{quote}

Desententialization, then, is the mechanistic part of the process by which a complex (in this case, a clause) loses its syntax and, crucially, its verb. For the result to be a “simple deverbal lexeme”, the verb of the clause may become non-finite (losing its tense and mood) or may be wholly lost through ellipsis.
Another important criterion for Jendraschek (2007) in his overview of dependent clauses is governance. In the syntactic hierarchy of clauses, complements of heads are governed, while modifiers are not. In such a hierarchy, sentence adverbs/adverbials would not be considered to be governed. A final consideration is given to the concept of free vs. bound clauses. Free clauses have finite verbs, while bound ones have non-finite verb forms (Jendraschek, 2007, p. 14). Bound clauses would therefore be more desententialized than free clauses and would therefore be more likely to be embedded.

Sinnemäki (2006), analyzing the distinction between embedding and hypotaxis in M. A. K. Halliday’s work, distinguishes three functions for embedded clauses. Following Halliday (1994), they function at “group level” rather than “clause level”:

- head of a nominal group (e.g. *It is obvious that Jeff wrecked the car*)
- post-modifier in a nominal group (*I saw the car that Jeff wrecked*)
- post-modifier in an adverbial group (*He came earlier than we had expected*)

(Halliday, 1994, p. 242, as quoted in Sinnemäki, 2006, p. 376)

For Halliday and Sinnemäki, embedded clauses must be an inseparable element within the matrix clause, semantically crucial to the meaning of the sentence.

Hopper and Traugott (2003) also examine clausal structure, from a diachronic perspective of course, and they have a three-way distinction among paratactic, hypotactic, and subordinate clauses. These are the three large-scale categories that are determined by use of a binary system that makes use of the
features [+dependent] and [+embedded]. Dependence is determined by whether a specific clause is separable and able to stand alone; embeddedness is determined further by whether a dependent clause sits within “a constituent of the nucleus” (p. 177), or, in other words, within the verb-argument structure core of the matrix clause. With these features, Hopper and Traugott are able to lay out a spectrum or cline of clausal connection:

(10) \[
\begin{array}{ccc}
\text{parataxis} & > & \text{hypotaxis} & > & \text{subordination} \\
[-\text{dependent}] & & [+\text{dependent}] & & [+\text{dependent}] \\
[-\text{embedded}] & & [-\text{embedded}] & & [+\text{embedded}] \\
\end{array}
\]

(2003, p. 178)

Hypotactic clauses would consist of adjunct clauses (*Because he was sick*, he called in). Subordinate clauses on the other hand would be subject clauses (*That he is sick* is true) or relative clauses (*The guy who is sick* stayed home). Hopper and Traugott’s distinctions are the three definitions I will make use of in this work, because they are broad enough to allow for a tripartite structure under which most clause types (including all those I am examining) fit.

With a definition of paratactic and hypotactic clauses relatively clear, I turn to one last consideration that relates to clause typology in this study: the source clause for the sentence adverb. There is not just one type of clause that is desententialized to become a sentence adverb; rather there are two main types that follow the separation between Chapters 5 and 6. *Speech-act* adverbs have a source in a *conditional* clause that is based on a verb of speech (e.g. *If I may speak*
frankly/honestly/seriously/briefly…). They are different from the other types of sentence adverbs I am examining, however, as evidential, epistemic, and evaluative adverbs are all based instead on copula clauses—specifically, extraposed copula clauses.

In considering the conditional clauses that are the source for speech-act adverbs, it is notable that they occur at the left periphery of the main clause. Diessel (2005) examines marginal (outside the nucleus) adverbial clauses in 40 languages and distinguishes between initial and final adverbial clauses from a typological perspective. He finds that adverbial clauses are placed in either position as a result of tension among three competing linguistic needs: 1) the parsing and processing needs of the addressee are benefited if the ADV clause follows the matrix; 2) discourse and pragmatic motivations favor an initial occurrence of the ADV clause, as that position provides thematic grounding and orientation for subsequent clauses; and finally 3) clause type plays a role, as conditional clauses are generally initial, and temporal clauses are often initial, but causative clauses are usually in final position (Diessel, 2005, p. 465; my italics). See example (11) below for illustration.

Diessel singles out conditional adverbial clauses in his article, explaining that this type “denotes a future situation whose realization is construed as a sufficient condition for the realization of the main clause event” (p. 461). However, speech-act conditional clauses are slightly different: “a speech act conditional presents information at a different speech act level than the associated clause. [It] does not affect the semantic interpretation” (p. 463; my italics).
Typologically, then, speech-act adverbial clauses are further removed from the proposition of the clause (as below) than are most conditional clauses.

(11) [If I may speak frankly], they won [because they worked as a team].

\[\text{conditional clause} \quad \text{causative clause}\]

(positioning of conditional and causative clauses following Diessel, 2005)

The other three major sentence adverbs bunched on the left periphery—evidential, epistemic, and evaluative—arise from a structure that is different from that of speech-act adverbs but is consistent among the three. The mechanism that affects these three and places them at the initial position in the sentence is known as \textit{IT-extraposition}. Other terms used to describe this phenomenon include \textit{expletive-IT, pleonastic-IT, anticipatory-IT, and preparatory-IT}.

Calude (2008) compares \textit{IT}-cleft sentences and \textit{IT}-extrapositions; where \textit{IT}-clefts are used for focus, \textit{IT}-extraposition is used to shift complex subjects to a later, less fronted, position. See example (12) for a contrast.

(12) a. \textit{IT}-cleft

\[
\text{It is on the playground where the children like to play } t, \text{ not in the backyard.}
\]

(Compare: The children like to play on the playground, not in the backyard.)
b. **IT-extraposition**

*It* is wonderful *that dogs now live on Mars*.

(Compare: That dogs now live on Mars is wonderful.)

Significantly, **IT-extraposition** is used to indicate speaker position regarding the shifted subject.

I will use “**IT-extraposition**” as a shorthand for the phenomenon I describe here, although this is now considered a somewhat antiquated term. My syntactic analysis (as will be seen in Chapter 4) is largely cartographic, based on antisymmetry (Kayne, 1994), and extraposition has been considered by some to be problematic for this model.⁵ In Chapter 6, however, I offer some solutions for understanding extraposition based on uniting cartography and more orthodox forms of minimalist syntax.

### 2.5.4 Clause combining and grammaticalization.

The movement from parataxis to hypotaxis and subordination is considered to be a type of grammaticalization and, as noted in section 2.5.3, these three types of sentence form the three stages of a cline of clausal structure (Hopper & Traugott, 2003, p. 178). The chapter in *Grammaticalization* (Chapter 7) that is devoted to this subject is somewhat broad in coverage, discussing the grammaticalization of clausal linkage words and combining this phenomenon

---

⁵ For other generative work on the syntax of extraposition, see e.g. Vicente (2003), who argues in favor of rightward movement, disallowed under antisymmetry.

Nevertheless, there are some important steps that Hopper and Traugott describe that correspond to the various stages of their clausal cline. As expected from these authors, there is a large amount of complex terminology involved. For them, there is a minimal process that involves the *unification* and *bonding* of two clauses. Bonding is accompanied (in the case of subordination) by a *hierarchical downgrading* of the subordinate clause and the process of desententialization (as above, following Lehmann, 2002). Such desententialization comprises *decategorialization* (loss of finite verb) and the movement of the clause to a marginal position. The reduction of the verb may be matched by *interlacing* (again following Lehmann) or sharing of the subject, as in raising constructions. See (13) for an example of a raising construction with a shared subject.

\[(13) \quad \text{It seems that they like the movie.} \quad \rightarrow \quad \text{They seem to like the movie.} \]
\[
\begin{align*}
\text{\quad (two subjects, two finite verbs)} & \quad \rightarrow \quad \text{\quad (one subject, one finite/non-finite verb)}
\end{align*}
\]

For the purposes of this dissertation, I am going to adopt a somewhat hybrid understanding of historical clause combining and reduction that follows Hopper and Traugott’s (2003) cline (example 10 above) at the macro level but
van Gelderen’s (2004) more fine-grained clausal strategy, shown in (14) below, for the micro level:

\[
\text{(14) } \begin{array}{c}
\text{clause} + \text{clause} \\
\text{subordinate strategy} \quad \text{coordinate strategy}
\end{array}
\]

\[
\begin{array}{l}
\text{CP} \quad \text{[clause [that/wh clause]]} \\
\text{IP/VP} \quad \text{[clause [Subj clause]]} \\
\text{IP/VP} \quad \text{[clause [(C)(DP) non-fin. clause]]} \\
\text{IP/VP} \quad \text{[clause Aux V]} \\
\text{IP/VP} \quad \text{[clause affix-V]}
\end{array}
\]/ \quad \text{[clause and/then [clause]}} \\
\text{[clause and VP]} \\
\text{[clause V+V]} \\
\text{(p. 5)}
\]

What this does is focus attention on the reduction of clauses once they are subordinated or coordinated. In addition, the portion of the clause that is affected is indicated in the far left column.

In this work, I look at the subordinate strategy in particular. The placement of speech-act and evidential/epistemic/evaluative adverbials on this more elaborated cline will be examined to determine whether it holds for the phenomena I investigate. It will be seen in Chapters 5 and 6 that there is a mix of grammaticalization and another process, that I consider to be a type of lexicalization, at work.

2.5.5 Parataxis and hypotaxis in this work.

To summarize the above, this study will utilize terminology from Hopper and Traugott (2003) to distinguish among parataxis, hypotaxis, and subordination, including the binary values for the characteristics dependent and embedded. This
tripartite structure will allow the examination of the two main types of clause combination found in the data: speech-act conditional clauses and extraposed IT-clauses. Since my study does not focus on the grammaticalization of clause linkers *per se*, I will adopt a hybrid understanding of the shifts that occur during and after hypotaxis, combining Hopper and Traugott (2003) and van Gelderen (2004).

### 2.6 Pragmatics and Discourse of Predicational Adverbs

#### 2.6.1 Introduction.

The use of sentence adverbs by speakers of English is an expression of the interaction of syntax and discourse-related communicative needs. This dissertation explores that interaction through examining adverbs, particularly sentence adverbs. The tension and interplay of these two aspects of language use (syntax and discourse) operate at an interface level that has its syntactic expression in the CP (complementizer phrase) at the left periphery of the clause in English. Study of this interface is not recent, and it can be expressed in a variety of ways, as in van Gelderen (2004), where the terminology is syntactic:

All of the [historical] changes show the interesting interaction, talked about in e.g. Jespersen (1921: Chapter 14, §6) as a ‘tug-of-war’, between economy and innovation. Economy eliminates specifiers; innovations reintroduce them (e.g. reinforcing *not* and *wh*-relatives); and prescriptive rules either stop their introduction (e.g. the ban on multiple negation) or try to stop their change to head (e.g. rendering *whether* phrasal by adding *or not*). (p. 12)
This section finishes Chapter 2 by exploring some of the different ways in which scholars who study pragmatics, discourse analysis, and sociolinguistics construe the use of sentence adverbs.

After this introduction, sections 2.6.2–2.6.4 examine some theoretical models of pragmatics and how they approach the communicative needs (innovations as in van Gelderen) of discourse cohesion. In section 2.6.5, we look at pragmatic markers, another area of study with an overabundance of competing terminology. I situate sentence adverbs with regard to pragmatic markers in 2.6.6 and explain the use of terminology for this study.

2.6.2 Pragmatic models I: Politeness Theory.

In this section I look briefly at three different models from discourse analysis. One of the best known is Politeness Theory (Brown & Levinson 1987 [1978]). In order to demonstrate their theory, Brown and Levinson postulate a Model Person (MP) who enters into verbal (or other language-based) interactions and, once the interaction is engaged, selects from an array of possible and predictable behaviors. The MP is rational: s/he recognizes the socioculturally defined means by which s/he can achieve certain social goals. In addition, the MP possesses face, a concept derived by Brown and Levinson from Goffman (1955 et seq.). As is well known, face is the public representation of an individual and, as such, is a significant property that must be preserved and protected.

For Brown and Levinson (1987[1978]), there are two sociocultural needs that must be met during interactions, as public identity is constructed: the need to
maintain independence and the need to be accepted and approved of by the community. These needs are expressed through negative and positive face, respectively. Negative face is widely considered to be competitive and hierarchical. Positive face comprises, among other qualities, concerns of collaboration and group membership.

In PT, all MPs have both negative and positive face-wants. Generally speaking, interactants in linguistic exchanges seek to maintain each other’s face. There are, however, communications (e.g. some speech acts; see Searle, 1969) that inherently challenge the face-wants of the interlocutor. These speech acts are known in PT as face-threatening acts (or FTAs); they can threaten either positive or negative face of either speaker (S) or hearer (H) and are largely mitigated through strategies of politeness.

When examining adverbs, particularly sentence adverbs, we can consider them to act as implicatures of politeness (or impoliteness) strategies in many cases. Brown and Levinson identify five strategies in response to FTAs. See example (15) for the four relevant ones:

(15) Four politeness strategies (Brown & Levinson 1987[1978], pp. 68-71):

**Threat is low: NO REDRESS**

Strategy 1. *FTA is performed baldly, without redress*

There is no fear of the hearer because

- the situation is urgent
- the danger to the hearer’s face is small
- the speaker is vastly superior to the hearer
Threat is moderate: REDRESS (counteracts FTA; implies no harm is meant)

Strategy 2. *FTA is performed with positive politeness*
The speaker wants to validate the hearer’s positive face-wants of belonging to the in-group or of friendship

Strategy 3. *FTA is performed with negative politeness*
- The speaker won’t interfere with the hearer’s freedom (uses self-effacement, formality, and restraint)
- The speaker uses apologies for interfering/transgressing, shows deference to hearer, hedges, uses mechanisms that impersonalize threat

Threat is high

Strategy 4. *FTA is performed off-record*
The speaker uses indirect or ambiguous language so that the FTA can be denied: vagueness, generalizations, ellipsis, metaphor, irony, rhetorical questions, understatement, tautologies, hints

Sentence adverbs, which signal discourse information to the interlocutor, can be used in any of these four strategies. Watts (2003, pp. 182-4) lists several types of adverbials used with illocutionary force. These include *hedges* (“rather”), *understaters/overstaters* (“briefly”/“absolutely”), *downtoners* (“perhaps”), *committers* (“certainly”), *forewarnings* (“frankly”), and *scope-staters* (“I’m afraid that…”).

**2.6.3 Pragmatic models II: Relevance theory.**

Jucker (1993), in his analysis of the discourse marker use of *well*, uses Relevance theory (Sperber & Wilson, 1987, et seq.) as a framework. Relevance theory is based around the third of Grice’s maxims, the Maxim of Relation (also known as *Be relevant*). Such a framework is closely related to Politeness Theory
in its assumption that there is an optimal communicative level. In Relevance theory, the goal of discourse is to reach a mutual, optimal level of relevance. The concepts of context, shared communicative environment, and continual negotiation are important to achieving such relevance.

Relevance theory focuses on the awareness and assumptions that interlocutors have and make about each other. Jucker (1993) summarizes the Relevance theoretical understanding of context, which comprises three main points:

First, every utterance comes with a guarantee of its own optimal relevance; second, the relevant context is established as part of the utterance interpretation; and third, discourse coherence is the outcome of negotiating relevant backgrounds. (p. 438)

Sentence adverbs, then, like more grammaticalized discourse markers, can act as “signposts” indicating the level of relevance of either preceding or upcoming discourse. A sentence adverb such as honestly can indicate the perceived inadequacy of current context and background assumptions. Consider the exchange in (16):

(16) A: Would you look at that?
    B: Honestly, that’s the ugliest surfboard I’ve ever seen.

The shift from A’s implicature of disapproval to B’s explicit condemnation of the surfboard is signaled by the signpost of the sentence adverb. The context of the utterance exchange is thus shifted from ostensibly polite (indirect) to openly
contemptuous (direct). This is a reaction by B to the relevance of the communicative context and a renegotiation to a more relevant state.

2.6.4 Pragmatic models III: stance.

Where the primary concern of both Politeness Theory and Relevance theory is the flow of interactive conversation across turns and among interlocutors, the concept of stance is more fully centered on the speaker. According to Johnstone (forthcoming), “Stance is generally understood to have to do with the methods, linguistic and other, by which interactants create and signal relationships with the propositions they utter and with the people they interact with.” Biber (2004) further limits stance to be “epistemic or attitudinal comments on proposition information” (p. 107).

Variations in stance are closely related to the speaker’s social environment, both at a micro and at a macro level. Eckert (1989) focuses, for example, on the different types of stance expressed by individuals in a high school setting. Stance is in some ways like face as envisioned by Brown and Levinson in PT, the public expression of the private self. Similarly, stance covers concerns such as group membership: “Since alignment or disalignment with another social actor could be accomplished through membership categorization moves, social identity claims and ascriptions also fall under the rubric of stancetaking” (Johnstone, forthcoming, p. 5). However, stance moves beyond the level of in-group marking to include, at the higher levels, the language and social choices that make up registers or other types of varieties (ibid, p. 7).
Although Biber (2004) indicates that several grammatical mechanisms are available for the expression of stance (e.g. modals, semi-modals, and complement clauses) one of the better-known mechanisms is the stance adverbial (p. 110). He distinguishes among them four classes, as in (17):

(17) Four classes of stance adverbials (Biber, 2004):

- **attitudinal**: surprisingly, hopefully, wisely
- **non-factive**: frankly, mainly, truthfully
- **factive**: undoubtedly, obviously, certainly
- **likelihood**: evidently, predictably, roughly

Attitudinal adverbials seem to be identical to the semantico-syntactic category of evaluative adverbs. Non-factives comprise speech-act adverbs (although *mainly* is an interesting case; it seems more epistemic). Factives can also be called epistemics, and likelihood adverbs are evidential under a different name.

As we will see in Chapter 4, these classes map quite neatly onto a syntactic structure such as that found in Cinque (1999). According to Biber, stance adverbials have notably increased in use over the past three centuries. The reasons for such a change, he indicates, may include 1) changes in social norms; 2) a change in the grammatical system for the expression of stance; and 3) a diversification among various registers for the patterns of use (p. 110).
2.6.5 Pragmatic markers.

Within the disciplinary constraints of discourse and pragmatic studies, the terminology of these elements (indicating social connectivity and/or discourse transitioning) is somewhat fluid. Fraser (1999), for example, offers a list of several of the terms given to these words (p. 932):

(18) cue phrases         pragmatic formatives              
discourse connectives   pragmatic markers              
discourse operators     pragmatic operators              
discourse particles     pragmatic particles              
discourse signaling devices semantic conjuncts 
phatic connectives      sentence connectives             
pragmatic connectives   pragmatic expressions

There are some distinctions, however. Currently, many scholars use the term *pragmatic marker* for the more general element that helps establish context for the utterance. Mosegaard Hansen and Rossari (2005), following Fraser, define *pragmatic marker* as “any signal that has an effect on the communicative, as opposed to the strictly propositional, level” (p. 178). *Discourse marker*, on the other hand, is more specifically used to illustrate those portions of an utterance that are “specialized in giving indications about intra-discursive relations” (Mosegaard Hansen & Rossari, 2005, p. 178). In other words, they specify how the nucleus clause relates to the foregoing discourse. Speech-act adverbs, for
example, belong more to the larger class of pragmatic markers than to discourse markers.

Models for understanding disjuncts such as speech-act (and other types of predicational) adverbs, as well as other CP-layer elements, owe a debt to the work of Schiffrin (1987), who calls all such elements “discourse markers”. For Schiffrin, such markers operate “independently” of syntax: “The distribution of other elements, e.g. the marker firstly as well as sentence adverbs such as frankly, can be constrained only by discourse and pragmatic facts” (p. 37) (her italics).

Although such a severe syntax–pragmatics disconnect is no longer widely proposed (and is in fact antithetical to the aims of this work), Brinton (1996) nevertheless uses criteria that are similar to Schiffrin’s in her discussion of the grammaticalization of pragmatic markers:

Table 2.2
Criteria for pragmatic markers (Brinton, 1996)

<table>
<thead>
<tr>
<th>Predominantly oral rather than written</th>
</tr>
</thead>
<tbody>
<tr>
<td>Used with high frequency in oral discourse – sometimes more than one in a single sentence</td>
</tr>
<tr>
<td>Stylistically stigmatized, negatively evaluated</td>
</tr>
<tr>
<td>Phonologically short – often reduced or unstressed</td>
</tr>
<tr>
<td>Of a separate tone group, often falling &gt; rising or rising intonation</td>
</tr>
<tr>
<td>Predominantly restricted to sentence-initial position (may occur elsewhere)</td>
</tr>
<tr>
<td>Little or no propositional meaning</td>
</tr>
<tr>
<td>“Outside” the syntactic structure or loosely attached to it – therefore with no clear grammatical meaning</td>
</tr>
<tr>
<td>Optional rather than obligatory (displaying a relation rather than creating it)</td>
</tr>
<tr>
<td>Marginal and derived from a variety of traditional categories</td>
</tr>
<tr>
<td>Multifunctional, both within and outside the sentence</td>
</tr>
</tbody>
</table>
Brinton indicates that such pragmatic markers are highly significant to the modern English sentence. Without such markers, discourse is grammatically acceptable but pragmatically difficult—“‘unnatural’, ‘awkward’, ‘disjointed’, ‘impolite’, ‘unfriendly’, or ‘dogmatic’” (Brinton, 1996, pp. 35-6).

The functions of pragmatic markers named by Brinton tend to fall into different modes, relating to “textual” issues or “interpersonal” ones. Textual issues often focus on the discourse itself, while interpersonal issues focus on the relation of the discourse participants. Of the below list, the first seven functions of pragmatic markers can be classified as textual, while the last two are interpersonal.

Of these functions, speech-act adverbs can be seen to effectuate initiating discourse, serving as a filler or delaying tactic for holding the floor, marking a discourse boundary, and bringing about increased intimacy between speaker and addressee.

<p>| Table 2.3 |</p>
<table>
<thead>
<tr>
<th>Functions of pragmatic markers (Brinton, 1996)</th>
</tr>
</thead>
<tbody>
<tr>
<td>to initiate discourse and close discourse</td>
</tr>
<tr>
<td>to aid the speaker in acquiring or relinquishing the floor</td>
</tr>
<tr>
<td>to serve as a filler or delaying tactic for sustaining discourse or holding the floor</td>
</tr>
<tr>
<td>to mark a discourse boundary (e.g. introduce a new topic; partially shift topics through correction, elaboration, specification, and expansion; or resume topics after an interruption)</td>
</tr>
<tr>
<td>to denote new or old information</td>
</tr>
<tr>
<td>to mark sequential dependence (e.g. the relation of the current utterance to previous utterances)</td>
</tr>
<tr>
<td>to repair discourse</td>
</tr>
<tr>
<td>to respond to preceding or anticipate future discourse</td>
</tr>
<tr>
<td>to effect cooperation, sharing, or intimacy between speaker and hearer</td>
</tr>
</tbody>
</table>
2.6.6 Sentence adverbs as pragmatic markers.

Sentence adverbs, the focus of this study, are often considered to belong to the larger class of pragmatic markers as defined in section 2.6.5. Yet, in a way that is similar to other debates in e.g. grammaticalization, lexicalization, etc., some scholars do not agree that sentence adverbs are a type of pragmatic marker.

This discussion generally follows the tripartite division that traditional grammar allots to adverbs: *adjuncts*, that are connected to verbal expressions (i.e. VP, vP, TP); *conjuncts*, that act as connectors between clauses (e.g. however, therefore); and *disjuncts*, adverbs that are not part of the proposition, such as sentence adverbs. An example of the confusion among these positions can be found in e.g. Aijmer (2002):

Discourse particles include elements that “look like” conjunctions (*however*), like main clauses (*I think*), sentence adverbials (frankly), imperatives (look) or interjections (oh). A distinction is sometimes made between conjuncts which are members of the class of discourse markers (Fraser 1990) and disjuncts such as frankly (*Frankly we are lost*) which are “commentary markers” with the procedural meaning to comment on some aspect of the basic message (Fraser 1990: 179). (Aijmer, 2002, p. 18)

Traugott marks the three types (adjunct, disjunct, conjunct) as the three stages of her adverbial cline, repeated here:

(19) Cline of adverbial grammaticalization (Traugott, 1995; also Hopper & Traugott, 2003):

adverbial > sentence adverbial > discourse particle
In this study, I will follow the distinction between sentence adverb and pragmatic marker; yet it is obvious that there are several similarities between the two.

### 2.7 Chapter Summary

This chapter examines the diachronic and pragmatic concerns that make up a large part of this dissertation. Issues such as grammaticalization, pragmatization, and lexicalization are examined from both a theoretical and a phenomenological standpoint, with the decision that this study will remain descriptive with regard to the phenomena. I therefore stick with a narrow definition of grammaticalization and with a less restrictive definition of lexicalization. I follow Hopper and Traugott (2003) in distinguishing among parataxis, hypotaxis, and subordination by using binary features; however, for grammaticalization in these three clause types I use van Gelderen’s (2004) model.

With regard to pragmatics and discourse analysis, there are three models (Politeness Theory, Relevance theory, and stance) that offer different insights to inform the innovative and communicative dimension of the syntax–discourse interface that this study investigates. Although there is some disagreement about sentence adverbs’ position vis-à-vis pragmatic markers (i.e. are they a type of marker?), this work will maintain a distinction between them.

In Chapter 3, we turn to questions of the broader category of adverb (including all three types: adjuncts, disjuncts, and conjuncts). A morphological approach that ties in with the definition of lexicalization used in this dissertation will be used to define a category that often evades definition.
Chapter 3

ADVERBS AS A CATEGORY

3.1 Introduction

3.1.1 Adverbs are confusing.

It is well known among linguists that no one knows quite what to do with the adverb. As a category, adverbs are ill defined: too broad semantically, too derivative morphologically, seemingly subject to free word order, and technically capable of modifying nearly anything else. In addition, the difference between adverbs (words) and adverbials (functions) is often confusing.

Adverbials can be realized by typical manner adverbs, by nouns (which are then reanalyzed as adverbs when acting as adverbials), by prepositional phrases, by clauses, and (in some languages and varieties) by adjectives. Adverbials can modify verbs, adjectives, other adverbs, nouns (in certain circumstances), prepositions, and clauses. Not all adverbs can be interchanged so that they modify all of the above, however. In fact, there are several restrictions on the types of modification allowed by different environments. Some adverbials can only act as degree adverbials, some only as verb phrase adverbials, and some only as sentence adverbials.

Because of all these concerns, there are many proposals for understanding, accounting for, and categorizing adverbs that have been suggested by various scholars. Nevertheless, adverbs in particular and modifiers in general have often been considered to be of minor significance. Generative grammatical theories, for
example, have largely ignored them until recently and have mostly concentrated on argument structure and verbal organization instead.

However, Chomsky (2004) notes, “There has never, to my knowledge, been a really satisfactory theory of adjunction, and to construct one is no slight task” (p. 117). *Adjunct* is, broadly, another term for a sentence-internal adverb. Current syntactic theorists have been more successful at explaining other types of adverbs, such as *conjuncts* (e.g. however, therefore) and *disjuncts* (e.g. frankly, fortunately). These forms are more marked in the sentence, because they generally appear on the left periphery and are discourse-related.

Early linguists determined the category of adverb by the absence of inflection, a useful technique when examining morphologically rich languages such as ancient Greek, but less so when studying more analytic languages such as English and Mandarin, where the lack of inflection is widespread. In comparison with modern syntacticians and ancient language scholars, many contemporary diachronic linguists focus instead on issues of grammaticalization and lexicalization. This seems to be a fruitful line of inquiry (as we saw in Chapter 2), one that I take up again here. This chapter is not based on a particular theoretical framework, although I utilize insights from both formalism and functionalism.

### 3.1.2 Chapter organization.

In section 3.2, I examine evidence from English in order to explore the heterogeneity of adverbs and adverbials, semantically and morphosyntactically. Then, in section 3.3, I concentrate more specifically on modern English adverbial
morphology, on both transparent and opaque examples, and suggest that the processes of grammaticalization and/or lexicalization have played a significant role in the development of the various elements categorized as adverbs.

Finally, in section 3.4, I propose that the *adverb* in English is a category that has a minimal number of (mostly deictic) examples that can be considered primitives; that the great bulk of its membership is compositional; and that this evidence supports the theory that the adverb is a derived lexical category. Many adverbs have an internal or covert syntax (at the interpretation interface, or LF in generative grammar), likely a remnant of grammaticalization and lexicalization whose implicit meaning has been retained as part of the semantic makeup of the word; that is, they are the result of a fused or elided phrase or clause—a reduced adverbial.

3.2 Background

3.2.1 Adverbs and adverbials.

In general, lexical categories are unified by various common linguistic markers. *Noun*, as a canonical example of a major category, is fairly unified semantically, morphologically, and syntactically. Semantically, nouns can be considered to satisfy some sort of indexing or existential quality; children are often taught that they are persons, places, things, and ideas. Time stability also helps to define the noun in several theories (e.g. Givón, 1984, pp. 51-6; Croft, 1991).
Morphologically, nouns are often marked for $\varphi$-features (person, number, gender) and structural case (nominative, accusative, dative, etc.). Syntactically, nouns trigger case marking by acting as subjects and objects. They also serve as antecedents to pronouns and as arguments/0-roles (agent, theme, experiencer, etc.). There are, of course, more and less semantically prototypical nouns, and various languages solve morphosyntactic problems by different means. However, most speakers have a sense of what a noun is—there is a certain “nouniness” shared by all.

Adverbs are much less well defined than nouns, and as a result they are only considered a major lexical word class by some scholars. Semantically, adverb(ial)s are often qualified by the $wh$-questions they answer: How? Why? When? Where? Yet there are several different words/phrases/clauses that are considered adverbials but do not answer these four questions; instead they explain, for example, the state of mind of the speaker or the subject. In example (1), there are some of the different types of adverb(ial)s found in English:

(1) Some English adverbial types (following Ernst, 2002, p. 9)

a. Predicational (gradable):
   
   *Honesty*, I thought Bam-Bam was cheating. (speaker-oriented or disjunct)
   *Joyfully*, Dino turned in his last exam. (subject-oriented)
   *Similarly*, the Rubbles lived in a stone house with a mammoth’s trunk for a shower. (exocomparative)
   Fred walked *rapidly* away from the sinkhole. (event-internal)

b. Domain:
   
   *Geographically*, the cities of Bedrock and New Rock City are very close; *politically* they are far apart.
c. Participant:
Fred drove down the street through the courtesy of his two feet.
Barney ran outside.
d. Functional (not gradable):
Wilma wanted to pick Pebbles up from college today. (time)
He ordered brontosaurus burgers again. (quantificational)
Mr. Slate was quite sure that Fred was sleeping at work. (focusing)
Stone-slab houses are not found outside Bedrock. (negative)
They didn’t go on vacation because the pterodactyls were on strike. (clausal relation or conjunct)

In most of the above examples, the adverbial is realized by a single adverb-ish word.

A participant adverbial is often difficult to express as a single word, as it generally replaces a noun in oblique case (e.g. instrumental, locative), and therefore is expressed as a prepositional phrase in English. An exception is a word like outside, where the prepositional phrase has fused; this sort of phenomenon is discussed in detail below. In the same way, today could be considered a noun (Today is the first day of summer), although it would likely not be considered a prototypical noun.

Ernst’s functional subcategory is evidently a hodgepodge of mixed adverbs. He concedes their heterogeneity (2002, p. 9) and unites them primarily because of the fact that they are non-gradable. His difficulty is not atypical of a semantic approach; the earliest linguists and teachers had to bring together a wide range of largely unrelated items as well (see below).

In addition to the different roles that adverbials can play in the sentence, there are a great number of forms that can be used by speakers to express an
adverbial function. Adverbials in English can of course be expressed by the set of traditional adverbs (as in 2),

(2) He arrived suddenly.

but also by certain nouns/noun phrases (3),

(3) He arrived this evening.

by prepositional phrases (4),

(4) He arrived with three friends/by horse-drawn carriage/at the break of dawn.

and by clauses (5).

(5) He arrived because he remembered the address this time.

In some varieties, adverbials are frequently realized by adjectives.

(6) He dances good.

Some adverbs are well known to share forms with prepositions.
(7) He drove around.

Even some verbal forms can act as adverbials.

(8) He arrived limping.

The picture of adverbial forms is therefore as muddled as the one of adverbial functions. Adverbs and adverbials often share word form and structure with several other categories and/or phrase types. In English, the only productive modern form adverbs can claim is the -ly affix. Even this is not originally a purely adverbial affix but is shared with adjectives in Old English (e.g. such adjectives as neighborly or friendly).

The original Old English adverb affix is -e, which disappears during the Middle English period. In the early forms, adjectives and adverbs were distinguished by -lic (ADJ) versus -lice (ADV), but this distinction disappears during Middle English and the two forms merge. When this happens, the affix becomes purely adverbial in meaning. Word pairs where the distinction is only through -e (e.g. fast and faste) merge as well, and such words take on both adjectival and adverbial meanings.

### 3.2.2 Early methods of categorizing adverbs.

The question of how to distinguish the category adverb is therefore not an easy one to answer. The first recorded solution is proposed by Dionysios Thrax, a
Greek from Alexandria who lived and taught in Rome circa 100 BCE. He wrote a *Grammar* that was used as a textbook by Roman students of Greek.

Focusing on what we now call Classical Greek, he describes adverbs as “indeclinable” words that modify a verb, taking either simple or compound forms; the lack of inflectional morphology is therefore an important marker for his definition. Adjectives may be marked for gender and number (as in modern Romance) as well as for case (as in Old English) through agreement or *concord* with the nouns they modify. In many morphologically rich languages, this is a useful means of identifying adjectives. Adverbs, on the other hand, show no such agreement.

When Dionysios gives a typology of Greek adverbs, he uses some general and some highly specific semantic definitions, several of which are still used today in English and other languages (the ellipses indicate Greek examples omitted for simplicity):

Some are indicative of *time*, …: to these we must subordinate as species those that connote particular times or seasons, … Some indicate *manner*, …; some, *quality*, …; some, *quantity*, …; some, *number*, …; some, *place*, …—of these there are three kinds, those signifying *in* a place, those signifying *to* a place, and those signifying *from* a place, … Some Adverbs signify a *wish*, …; some express *horror*, …; some, *denial or negation*, …; some, *agreement*, …; some, *prohibition*, …; some, *comparison or similarity*, …; some, *surprise*, …; some, *probability*, …; some, *order*, …; some, *congregation*, …; some, *command*, …; some, *comparison*, …; some, *interrogation*, …; some, *vehemence*, …; some, *coincidence*, …; some are *deprecative*, …; some are *asseverative*, …; some are *positive*, …; some express *ratification*, …; and some *enthusiasm*, … (Dionysios Thrax, trans. Davidson, 1874[100 BCE], pp. 14-15) (italics from translation; bold mine)
Here Dionysios is something of a “splitter” rather than a “lumper”; this is still a challenge for many semantics-based theories of categorization, as can be seen in example (1) above. Nevertheless, his semantic categories have had a long-lived influence. Teachers of language arts have used very similar semantic designations well into the modern era. We can compare Dionysios’s list to that of Roswell Smith (1882[1832]), found nearly two thousand years later in an American grammar book used extensively throughout the nineteenth century:

*Note* – Adverbs, though very numerous, may nevertheless be reduced to a few classes. You will now read with attention the following list, and I will then ask you some questions respecting each class.

2. Of *order*: as, ‘First, secondly, thirdly, fourthly, fifthly, lastly, finally’, &c.
3. Of *place*: as, ‘Here, there, where, elsewhere, anywhere, somewhere, nowhere, herein, whither, thither, upward, downward, forward, backward, whence, hence, thence, whithersoever’, &c.
4. Of *time*.
   - Of *time present*: as, ‘Now, to-day’, &c.
   - Of *time past*: as, ‘Already, before, lately, yesterday, heretofore, hitherto, long since, long ago’, &c.
   - Of *time to come*: as, ‘To-morrow, not yet, hereafter, henceforth, henceforward, by and by, instantly, presently, immediately, straightaways’, &c.
5. Of *quantity*: as, ‘Much, little, sufficiently, how much, how great, enough, abundantly’, &c.
6. Of *manner* or *quality*: as, ‘Wisely, foolishly, justly, unjustly, quickly, slowly’, &c. Adverbs of quality are the most numerous kind; and they are generally formed by adding the termination *ly* to an adjective or participle, or changing *le* into *ly*: as, ‘Bad, badly; cheerful, cheerfully; able, ably; admirable, admirably’.
7. Of *doubt*: as, ‘Perhaps, peradventure, possibly, perchance’.
11. Of *comparison*: as, ‘More, most, better, best, worse, worst, less, least, very, almost, little, alike’, &c.

When a preposition suffers no change, but becomes an adverb merely by its application: as, when we say, ‘He rides about’; ‘He was near falling’; ‘But do not after lay the blame on me’.

There are also some adverbs, which are composed of nouns, and the letter *a* used instead of *at, on, &c.*: as, ‘Aside, athirst, afoot, ahead, asleep, aboard, ashore, abed, aground, afloat’. (Smith, 1882[1832], p. 34; italics original)

It is evident from this excerpt that many of Dionysios’s diverse criteria are still in use well into the Modern English era.

These semantic definitions spur some questions. Where should speakers of a language (or students of language) draw the lines? Should Dionysios’s “enthusiasm” criterion (not mimicked by Smith) be a subset of his “vehemence” category, or “asseverative” be placed on a gradient opposite “congregation”? How many members should a subcategory possess to be considered separate from others? How useful are semantic subcategories in general?

These are questions that are still debated, as contemporary linguists such as Cinque (1999) and Ernst (2002) (as in example 1 above) have developed highly intricate syntactic hierarchies that are ultimately based on similar semantic distinctions.

### 3.2.3 Further semantic considerations.

Following Cinque and Ernst, Haumann (2007) notes that “[a]dverbs may assume a number of structural positions, but there is no single distributional property shared by all adverbs” (p. 3). The syntactico-semantic interpretation
plays the primary role in determining where adverbs may or may not be placed (in generative terms, merged) in the sentence. Consider (9) and (10) below:

(9)  
   a. He walked *(un)happily* up the stairs. [i.e. he walked in an *(un)happy* manner]  
   
   b. *He walked *(un)fortunately* up the stairs. [i.e. he walked in *an *(un)fortunate* manner]

(10)  
   a. *Luckily*, the car stopped on its own.  
   
   b. *Long*, the car stopped on its own.  
   
   c. *Thoughtfully*, the car stopped on its own.

Certain adverbs in English cannot operate in the verbal structure of the sentence (such as *(un)fortunately*), while certain ones cannot be on the left (in sentence-modifying position) without some fancy reasoning when parsing the sentence. In addition, considerations such as animacy (as in 10c), as found in other lexical items, must be met, particularly in adverbs derived from adjectives.

3.2.4 Meeting syntactic conditions.

However, while semantics is significant, in the end it is the interface between semantics and syntax that determines, both for adverbs and for other categories, how a word will be categorized by its users (and, indeed, how that
categorization can change). Lehmann (2008) distinguishes between final (phrasal) and primary (root) categorization of a sign. Final categorization is determined by the syntactic function it has to fulfill in the sentence. That in turn is determined by the propositional operation … to be performed on it. … Sometimes he [the speaker; sic] has to adapt the means that the system offers him. That is to say, the use of a certain expression in a certain category may necessitate some prior operation of recategorization on it. (Lehmann, 2008, p. 4; italics mine)

On the other hand, primary categorization “is essentially determined by universal cognitive principles” (Lehmann, 2008). Following Croft (1991), Lehmann focuses on time-stability as the most understandable and useful of those principles. Nouns have a high level of time-stability, while verbs have a low level of time-stability. Adjectives, more arbitrarily, have a medium level. Consider examples of primary and final categorization in (11):

(11) Primary category: Final category:

<table>
<thead>
<tr>
<th>Primary category</th>
<th>Final category</th>
</tr>
</thead>
<tbody>
<tr>
<td>table (N)</td>
<td>They <em>tabled</em> the motion. (V)</td>
</tr>
<tr>
<td>destroy (V)</td>
<td>The <em>destruction</em> of the village was total. (N)</td>
</tr>
<tr>
<td>bright (ADJ)</td>
<td>She <em>brightened</em> the painting by cleaning it. (V)</td>
</tr>
</tbody>
</table>

However, “[l]ittle is known, *a fortiori*, about the factors underlying the primary categorization of concepts as adverbs” (Lehmann, 2008, p. 5). This seems to be because there is no primary categorization for adverbs, indicating that they may not, in fact, be a primitive category of word in English. As demonstrated above,
many adverbs in English arise from adjectives through the use of -ly morphology; yet others are derivable from a variety of different sources.

For Lehmann, then, there is a compositionality to word forms that seems to be nearly generative. This stands against what is widely known as the *Lexicalist Hypothesis* (Chomsky, 1970), in which all forms of words are stored in a mental lexicon and morphology is static. Such an active characterization of morphology is largely paralleled in the theory of generative grammar known as Distributed Morphology (Halle & Marantz 1993, et seq.). Proponents of Distributed Morphology, or DM, also advocate a compositional approach to word formation in which there are primitive and final forms of words. (See Chapter 4 for a more detailed discussion.)

To give a highly simplified description, such a constructional framework for morphology is based on the cognitive use by the speaker of semantic roots in the “narrow syntax” operations of Merge and Move. Such roots are lexical (as opposed to functional) but, unlike Lehmann’s model, there is no base category for the root. Category is determined by the syntactic needs of the clause.

*Contra* the Lexicalist Hypothesis, where items are pulled based on preconceived speaker needs and are therefore dropped into syntax whole, in DM there is *underspecification* of the roots. Derivational and, in particular, inflectional morphology (and the attendant phonological reflexes) are inserted late in the process, based on syntactic role and on language-internal agreement marking. Morphology is thus a response to the needs of the syntax–semantics interface, with components corresponding to each.
Haumann (2007) approaches the question of adverb categorization from a somewhat different perspective. She indicates that “the resultant picture is a compartmentalized one: the traditional word class Adv does not come together as one discrete category, but is scattered across a large set of distributionally and semantically homogeneous subclasses” (p. 8). She examines three major category theories: adverb as adjective, as intransitive preposition, and as noun or NP. (Note that these—ADJ, P, and N—are three categories that have significant overlap with adverbs.) Each is useful but incomplete, breaking down in some fashion (Haumann, 2007, p. 8).

Let us examine one of these theories. Baker (2003), for one, espouses the “adverb as adjective” solution. First, he defines adjectives as lexical words that act as the elsewhere solution, appearing wherever nouns and verbs are not appropriate. Such a characterization is opposed to traditional frameworks, which assign adjectives the semantic quality of “property” or “state”. Instead adjectives are indicated by a binary semantic feature categorization, following Chomsky (1970), where the categories are N (noun) and V (verb). Baker’s (2003) nouns are {+N, −V}, verbs are {−N, +V}, and adjectives are {−N, −V} (there is no category that is {+N, +V}). This is different from e.g. Chomsky (1986), where adjectives are characterized as [+N, +V] and adpositions are [−N, −V] (p. 2).

There are at least a couple of concerns with Baker’s method. First, stating that a noun is {+N, −V} seems tautological and not very useful. Secondly, this process serves to make ADJ something of a grab-bag category, unified by what it is not—apparently noun and verb—rather than what it is.
Baker then states that adjectives (counter to traditional grammatical thought) can merge with verb phrases, tense phrases, and other adjective phrases by appearing as adverbs (2003, pp. 230-1). *Adverb*, then, is a subset of *adjective* (or the lexical elsewhere case) that is syntactically triggered to take on the *-ly* affix.

He goes on to list several shared properties between adjectives and adverbs: that the *-ly* affix is essentially inflectional rather than derivational (allowing adjectives to modify non-nouns); that *ADJ* and *ADV* can take the same degree modifiers (*very* quick/*very* quickly); that neither *ADJ* nor *ADV* can take complements in pre-head positions (*a proud [*of his daughter] father/*he proudly [*of his daughter] showed photos); that the modification of a verb by an adverb is mirrored in the modification of a verbal noun by an adjective (*completely destroyed/*complete destruction); and that the ordering of multiple adjectives and adverbs is the same (Baker, 2003, pp. 232-3).

Baker then suggests that *-ly* (and *-ment(e)* in Romance) should be analyzed as nouns, because they arise from nominal forms historically, because adverbs in the form *X*-ly can be substituted for by the PP “in a X manner”, and because in Romance, *-ment(e)* attaches to the feminine (marked) form of the adjective (2003, p. 234). Because *-ly* and *-ment(e)* are nouns, they are also syntactic heads and can take *θ*-roles, thus allowing them to appear in “adverbial” environments (Baker, 2003, p. 234).

This is an interesting analysis for adjective-based manner adverbs. However, it ignores other types of adverbs and adverbials as discussed above in
examples (3)-(8) or in the semantic lists compiled by Dionysios and Smith in their grammars excerpted above. Where Dionysios can be thought of as a “splitter”, basing subcategorization on seemingly random criteria, Baker is very much a “lumper”, veering instead in the other direction. He eschews semantics, which is an understandable response to the somewhat dizzying fine-grained variation of adverbs. Yet he seems too hasty in his simplification of the category as being itself a subcategory of adjective, and his solution glosses over some important considerations.

Considering some of the specifics of his theory, the idea of even adjunctive noun phrases receiving θ-roles is somewhat controversial, as they are outside the argument structure of the clause. The argument that -ment(e) attaches to the feminine form of the adjective in Romance, and that this is a marked form that is unusual, loses its strength when considering the fact that the noun mente in late Latin and modern Romance is also feminine (it is descended from the ablative singular form of L. mēns). This likely triggers a remnant concord relation between the adjective and the affix (originally a noun, as Baker indicates), and the morphology of the adjective (reanalyzed as the root of the word) inflects as feminine to reflect this historical agreement process. The -ment(e) affix will be considered in more detail below.

It can be seen from this brief discussion that finding a unified solution to understanding adverbs (and adverbials) is not yet feasible. Baker’s (2003) solution, for example, does not take into consideration several adverbs in English that make up an important minority of the cases—those, for example, that are
lexicalized or grammaticalized from other phrases and clauses, as will be discussed below.

3.2.5 Section summary.

Unlike the categories noun (reference), verb (predication), or adjective (property/state), adverbs have no unifying underlying cognitive/semantic quality. Adverbs have final categorizations (in the meaning of Lehmann, 2008) that are very specific and dependent on syntactico-semantic features; however, they have few primary categorizations in the sense of other lexical categories. Instead their primary categorizations are mostly borrowed from other categories. Morphology thus plays an all-important role in the construction of this category.

3.3 Supplementing Synchrony with Diachrony

Most authors who study adverbs as a category approach them from a synchronic viewpoint. Until recently, in fact, linguists from most of the generative frameworks mostly avoided historical linguistics, treating this branch of study as unhelpful in the least. The work of Lightfoot (1979), van Gelderen (2004), and Roberts and Roussou (2003), among others, has reversed this, however, and there is now a vigorous intersection of generative diachronic linguistics. The biggest issues in contemporary works in historical linguistics, whether studied from a generative or a functionalist model, are grammaticalization and similar phenomena such as pragmaticalization and lexicalization. These were discussed in detail in Chapter 2, but a quick reminder is helpful here.
3.3.1 Grammaticalization and pragmaticalization.

There are some historical principles of organization that may be useful to our consideration of the category of adverb. The best known of these is grammaticalization (e.g. Hopper & Traugott, 2003). This is a co-occurrence of historical processes (i.e. analogy, reanalysis) that tends to be unidirectional, from less to more grammatical. Open-class, lexical words are re-understood and re-utilized (in some environments) by speakers as closed-class, functional words that forward grammatical processes for the language. These new meanings are then acquired by new speakers of the language and thereby fixed as a part of the new linguistic structure.

One of the best-known English examples is the shift in “be going to” from a lexical expression indicating movement and direction:

(12) *I am going to* meet my brother in London. (I am walking to London to meet him.)

into a grammatical expression that indicates future tense (and implies irrealis mood):

(13) *I’m gonna* tell you all about it as soon as I finish this story. (I will tell you.)
Grammaticalization has been developed into a theory in recent years, and it has come to be applied to a large number of language-internal changes. The unidirectionality of change has become requisite to the theory rather than correlative to the phenomena, and Traugott in particular has focused in her work (Brinton & Traugott, 2005; Traugott & Dasher, 2002; Traugott & Trousdale, 2010) on the concept of subjectification as a unifying characteristic to what otherwise seems to be a broad and varied array of changes.

These theoretical developments have met with some challengers (e.g. Campbell 2001, Newmeyer 2001), who largely assert that grammaticalization is a collection of phenomena that can also be found in other types of language change, and that grammaticalization itself is a convenient label for these co-occurring phenomena, but not a driving force in itself.

A close cousin to grammaticalization, pragmaticalization, is essentially the same collection of operations, but with the result that the analyzed element becomes an element known as a discourse marker, generally found on the left periphery of the sentence. Discourse markers, very simply, act to connect upcoming speech with what has been said before. A result of this process would be present-day English well (which has shifted in meaning from a positive response “Very well”, to a pragmatic forewarning and hesitater “Well, I’m not sure…”). Most grammaticalization scholars argue that pragmaticalization is simply another name for the former.
3.3.2 Lexicalization.

A third historical operation, *lexicalization*, is what we turn to now. Unlike grammaticalization and pragmatization, the process of lexicalization involves the creation of simplex structures from complex ones. The major work on lexicalization is by Brinton and Traugott (2005), and their definition of this process (repeated from Chapter 2) is

the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and syntactic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the word may become more lexical. (p. 96)

In older works, this process was considered the antithesis of grammaticalization, based on the opposition of *lexical* words (generally the source for grammaticalization processes) and *grammatical* ones. But current understanding of the two concepts (in e.g. Himmelmann, 2004) is that they work in different and sometimes complementary ways.

Although Himmelmann does not expressly indicate this, the phenomena associated with *grammaticalization* seem to be more indicative of a typological shift towards analyticity in some languages. When morphological inflections lose their strength, new functional words are created to carry out necessary grammatical operations that are no longer being effectively carried out by the eroded inflections. It is usually considered more impactful, as new (closed) functional categories are created or added to. In the creation of such new categories, analogy is significant as a means of adding new members. Similar
logical or metaphorical extensions happen to a range of related elements, and a
new class is formed.

On the other hand, *lexicalization* as a process is focused more on
morphology, and it adds to open lexical categories. However, it generally is not
indicative (or even necessarily reflective) of a typological shift towards a more
synthetic character in languages. Here the focus is more on the initial, more
transparent morphological processes rather than on opaque, later-stage
morphology (as in cliticization and inflection) often associated with synthetic
languages. Lexicalization is often categorized with other word-formation
processes and is not considered to be a basis for analogous extension. Therefore,
lexicalization is considered more idiosyncratic.

Phrases or even clauses can be the origins for lexicalized words, which are
reduced through the process of *de-syntagmatization* (or collapse of the syntactic
structure into morphological structure) (Brinton & Traugott, 2005, p. 48). A
typical example of lexicalization is the name of the flower known in English as
the *forget-me-not*. In this case an entire predicate VP has been de-syntagmatized
and the resultant orthography marks the fusion or *univerbation* of the elements
(verb, direct object, negation) through hyphenation (Brinton & Traugott, 2005, p.
49).

The highly articulated system of compound and complex words in English
can be thought of as indicating the shift from more syntactic (separate) to more
morphological (fused). To consider one example, the *New York Times* (which is
fairly conservative with regard to prescriptive rules and largely slow to change),
in its style *Manual*, spells the sentence adverb incarnation of *overall* as over all, leaving the former as the spelling of the attributive adjective or the clothing item (Siegal & Connolly, 1999, p. 249). Hyphenated forms often appear as collocated items are regarded as unitary, but before complete orthographical fusion occurs. For example, we can consider the spellings of “to-day” and “to-morrow” (now archaic) in the excerpt from Smith’s (1882[1832]) *Grammar* above.

Other examples of lexicalization follow normal word formation patterns, including blending, coinages, acronyms, compounding, and ellipsis. In addition to morphological reduction (fusion or univerbation), there is often phonological reduction (coalescence) as a result of the process (Brinton & Traugott, 2005, p. 47).

### 3.4 Adverbial Lexicalization

Lexicalization can be found in every content-based word class (N, V, ADJ, ADV, P) in English, as well as in other languages. Within the category we will, for lack of a better choice, continue to call adverbs, most scholars fail to mention the most obviously lexicalized examples. Baker (2003), as noted above, ignores all English adverbs but those based on adjectives and ending in -ly. As another example, Alexiadou (1997), writing about English and Greek, explains

adverbs can be distinguished, from a morphological point of view, into a) non-derived ones, e.g. adverbs like *often*, *well* or *simera* “today”, *kthes* “yesterday” in Gr[ee]k], and b) derived one e.g. adverbs like *carefully*. These are formed by adding the suffix, -a or -os in Gr, -ly in English, -ment in French, -mente in Italian, and -weise in German, to the adjectival stem. (pp. 2-3)
The *-ly* suffix is, as has been noted several times, the prototypical adverb formant in English. We will discuss it more specifically below. First, however, let’s look at a group of adverbs not considered by Baker or by Alexiadou, except perhaps as a “non-derived” type: those with very evident traces of lexicalization showing.

### 3.4.1 Fused adverbs.

Word lists aren’t as plentiful, exhaustive, or accessible as one might expect. However, there are several online English-instruction sites that have partial adverb lists. One of the more detailed is the site of Paul and Bernice Noll (www.paulnoll.com), with a list of 1,200 common English adverbs. Of those 1,200, only a small fraction (64, or 5.33%) is made up of non-*ly* adverbs. Example (14) includes those from their list, plus a few more that I compiled from other sources:

(14) Non-*ly* adverbs (from www.paulnoll.com)

<table>
<thead>
<tr>
<th>about</th>
<th>also</th>
<th>around</th>
<th>below</th>
</tr>
</thead>
<tbody>
<tr>
<td>abroad</td>
<td>(al)together</td>
<td>askew</td>
<td>beneath</td>
</tr>
<tr>
<td>afterward(s)</td>
<td>always</td>
<td>aslant</td>
<td>beside(s)</td>
</tr>
<tr>
<td>again</td>
<td>anew</td>
<td>away</td>
<td>between</td>
</tr>
<tr>
<td>ahead</td>
<td>anymore</td>
<td>awry</td>
<td>beyond</td>
</tr>
<tr>
<td>almost</td>
<td>anyway(s)</td>
<td>backward(s)</td>
<td>elsewhere</td>
</tr>
<tr>
<td>aloud</td>
<td>anywhere</td>
<td>before(hand)</td>
<td>even</td>
</tr>
<tr>
<td>already</td>
<td>apart</td>
<td>behind</td>
<td>evermore</td>
</tr>
</tbody>
</table>
There are several short adverbs on this list (e.g. *far, fast, on, out*), many of which double as adjectives or prepositions. Other short adverbs are deictic (*here/there/where, nigh/near/next*), and most of these date to the Old English period or before.

However, many of the others are obvious fusions of e.g. preposition + noun (*behind, perhaps*), quantifier + noun (*sometimes*), modal + verb (*maybe*),...
quantifier + adjective (*already*). Most of these enter the language during the Middle English to Early Modern English period, and there is a mixture of Germanic and Latinate words (though far more of the former) found in these lexicalizations.

Interestingly, many of these lexicalizations show up at the left periphery of the clause in their first incarnations, at least according to the *Oxford English Dictionary* (*OED*):

(15) *nevertheless* (1382)

> Neuer þe lese withinne þoru vertu is all.  
> (Wycliffe Bible)

(16) *maybe* (1400) (probably a calque from Fr. *peut-être*)

> May be sum goost awey him ledde.  
> (*Cursor Mundi*)

(17) *perhaps* (1520) (coincidental with VP position)

> But herist not one word beware thow tell Nothing of the weddingys lest perhappis therby Her sorrow shold increace.  
> (translation of Terence *Andria*)

(18) *nonetheless* (1533)

> Nontheles he was 3eit reddy to except the said office…  
> (*Aberdeen Burgh Rec.*)
The source material for these adverbs has already been largely canvassed in examples (3)-(8): noun phrases (nonetheless, somewhat), prepositional phrases (ahead, below, indeed), reduced clauses with remnant inflection/verb phrases (maybe, notwithstanding). These are the adverbs from the list that are transparent; there are others (not, adverbs in -way(s) and -wise) that are opaque in Modern English but are also combinatory forms from Old English.

### 3.4.2 -ly adverbs.

By far the most common method of creating an adverb (particularly of manner) is to attach the -ly suffix to nearly any adjective. This is generally considered derivational morphology. There is some debate, however. Brinton and Traugott (2005) hold that -ly is a grammaticalizing suffix and therefore not fully derivational in Modern English, while (as mentioned above) Baker (2003), coming from a generative formalist framework, argues that -ly is still nominal. Yet he categorizes adverbs as a type of adjective; so despite its nominal status, the -ly affix must be regarded as a type of inflection that allows adjectives to merge in otherwise unavailable positions.

The *OED* indicates that the -ly suffix originates in proto-Germanic (“Teutonic” in their terminology) as the noun likom (body, appearance, form). It has cognates across the West Germanic languages and becomes in Old English a productive suffix for deriving adjectives from nouns (manly, neighborly, kingly) and adverbs from adjectives (quickly, hotly, brightly). In addition, it becomes (as an independent word) the noun, verb, and preposition (and multi-use word) like.

83
In its earliest form, the suffix was probably the root of a noun-noun or adjective-noun compound, as are extremely common in English (e.g. doghouse, flowerbed; lightweight, blackboard). So manly would have been “man-form” or “man-body”, and brightly would have been “bright-form” or “bright-appearance”. This sort of compounding can be considered lexicalization in its earlier morphological form, although it has since lost its transparency.

3.4.3 -like adverbs.

In addition to the Old English [noun + -ly = adjective] construction, there currently exists a fairly productive pattern in Modern English that seems to retain the older form -like. This is found in [noun + -like] constructions (childlike, snakelike). There is also a less productive adverbial construction with -like. The OED cautions us, however, about both of these:

suffix, forming adjs. and advs. In strictness, the words containing this suffix are compounds of LIKE a. and adv., in the senses in which these words govern a dative or are followed by an adj. (see LIKE a. 1b, LIKE adv. 1, 3). The compounds so formed not unfrequently resemble in sense the derivatives formed with -lik(e), ME. dial. form of -LY1, -LY2, but the two formations are entirely distinct: thus ME. gredilike adv. (= greedily) is not the same word as the mod. Sc. greedy-like. (OED)

and

Forming advs. With the sense ‘like one who is —’. Obs. exc. in Sc., where the sense of the advs. is rather ‘so as to appear —’. (OED)

Although the OED indicates that the two constructions (-ly and -like) are separate and should not be commingled, there are some similarities between the two that seem to indicate the possibility of a limited cycle of renewal. This is interesting,
as lexicalization processes are expected to be idiosyncratic (as noted above). Such idiosyncrasy would exclude the likelihood of any sort of cyclic behavior. Yet there does not seem to be any evidence of a broader pattern of grammaticalization at this point in the modern history of -like. This is an area that needs further study.

Examples (19)-(22) are based on a Google search for -like. The search was based on my own experience/knowledge of the -like form in American (particularly American Southern) English.

(19) quick-like

• Manipulate Windows Quick-Like with WinMover.
• Get organized quick-like with a wiki.

(20) easy-like

• Hand it over, slow and easy-like!
• If that’s the way the birds get up, sort of slow and easy-like, …
• Then I sift in some flour and my spices, slow and easy like, and stir like a Mardi Gras demon.
• I tend to get upset real easy-like.
• Did it just come out of the bearing real easy like?

(21) crappy-like

• Oh yeah, I can get my car running real crappy like, and after doing that I smell my tail pipes…
• I think it ended real crappy like.

(22)  *smart-like*

• So he sidles up to this other fish … and real smart-like, he says, …

Because Google’s search engine is not currently as sensitive as could be desired, I forced the issue after a point by searching for the phrase “real X-like”. The use of *real* as a degree adverb (rather than *really*) helps to limit the dialect/register to a largely self-conscious, even humorous, rural/Southern U.S. one.

The construction itself tends to reflect such an informal, even non-standard, dialect as well. The *OED* identification of the modern -like form with Scots may also fit with what I believe to be its distribution in the U.S., as many Southern and rural dialects descend from Scots-Irish dialects that were brought to America during the colonial and post-colonial periods.

Nevertheless, my reading of examples (19)-(22) is that the -like affix is being used in these varieties as a straightforward adverbial ending, synonymous with -ly. It is informal, but the meaning of, for example, “Get organized quick-like with a wiki” appears to be no different from that of “Get organized quickly with a wiki”.

This phenomenon is not widespread across regions or fully productive, however, and that may be why the -like affix works better with less formal adjectives like *crappy* (the form *crappily* can be found online but sounds somewhat odd to my ears). A Google search for e.g. *frank-like* (= *frankly*)
returned no adverbial hits, further indicating that not all manner adverbs can be substituted for with a -\textit{like} form.

\textbf{3.4.4 A contrast: -\textit{like} adjectives.}

This situation contrasts with that of the affix -\textit{like} when it is used in conjunction with nouns to make adjectives. There are no register or dialect restrictions in such a case. In fact, that denizen of correct and standard eighteenth/nineteenth century Modern English, Jane Austen, uses the -\textit{like} affix in her most famous heroine’s quite famous put-down of her most famous hero, when he first proposes:

(23) You are mistaken, Mr. Darcy, if you suppose that the mode of you declaration affected me in any other way, than as it spared the concern which I might have felt in refusing you, had you behaved in a more \textit{gentlemanlike} manner.

(\textit{Pride and Prejudice}, pp. 140-1; my italics)

It is interesting to note not only Austen’s use of \textit{gentlemanlike} when \textit{gentlemanly} had been in the language since the fifteenth century (as both an adjective and adverb) (\textit{OED}), but also her use of the periphrastic (and with this affix, the only possible) adverbial form, “in a more gentlemanlike manner”. This type of prepositional phrase is usually given as the restatement of -\textit{ly} adverbs (i.e. \textit{happily} = in a happy manner).

Yet there is no possibility for *\textit{gentlemanlikely}, just as there would be none for *\textit{gentlemanlily} or *\textit{gentlemanlike-like} (although, as noted, \textit{gentlemanly}
could be an adverb as well—and was so used late into the nineteenth century).

Adjectives constructed with -like are unusual in that they cannot take -ly to create adverbs (consider the below):

(24) They were able to close the futon-like convertible top quite easily.
      They were able *futon-likely to close the convertible top.
      \begin{em}
      BUT
      \end{em}
      They were able, futon-like, to close the convertible top.

This may be because of the nonce quality of some of these constructions, but even more commonly found examples (e.g. snakelike, womanlike) cannot take -ly (*snakelikely, *womanlikely). Yet, as a counterexample, the word likely (in which the relationship between like and -ly has become opaque) is unquestionably standard. This suggests the recognition on the part of modern speakers that -ly and -like are quite similar, perhaps interchangeable, and that -like and √like (root) are not the same.

In the online English-language corpora collected by Mark Davies of Brigham Young University, adverbs in -like are also evident (only) in the American corpora, though not terribly common. There also appears to be a counterexample to my assertion about -like and -ly co-occurring:
(25) Corpus of Contemporary American English (COCA) (Davies):

- friendly-like (8 tokens)
- quiet-like (8)

(26) Corpus of Historical American English (COHA) (Davies):

- quiet-like (27 tokens)
- sudden-like (20)
- easy-like (17)
- friendly-like (15)
- quick-like (12)
- careless-like (10)
- sad-like (10)
- wild-like (10)
- gentle-like (8)
- natural-like (8)
- peaceful-like (8)
- slow-like (8)
- solemn-like (8)
- soft-like (8)
- crazy-like (7)
- innocent-like (6)
- proud-like (6)
- polite-like (6)

(27) British National Corpus (BYU-BNC) (Davies): no tokens

We can note one final point of interest: the appearance in the two American corpora of the form friendly-like.

In the discussion of example (24) I did not consider layering -ly before -like. However, the old-to-new sequence appears to make sense. On the other hand, the meaning of friendly may have shifted (from “appearance of a friend” to “convivial”) to the extent that it is no longer easily subject to semantic/
morphological fission into parts. In such a case, *friendly-like* may be more of an approximant (“sort of friendly”) than a true manner adverb.

Indications therefore are that both *-ly* and *-like* can be considered lexicalizing affixes, particularly if their common origins are taken into account. The noun *like* (fossilized in e.g. “I’ve never seen the like of him before!”) is the source of both of these affixes, and Baker (2003) seems to be correct in his assertion that the affix(es) still retain traces of their original nominal status.

Both of these affixes are *predicational* in Ernst’s (2002) sense—they both indicate manner and, to me, they seem synonymous. Particularly in the case of *-ly*, lexicalization has become morphology—whether it is considered derivation or inflection. The *-like* affix is more obviously lexicalized because of its transparency, although its incompatibility with *-ly* raises additional questions about cyclicity.

### 3.4.5 Romance *-ment(e).*

Lexicalization is, of course, not only found in English adverb constructions. The widespread adverbial affix in Romance languages is *-ment(e).* There are differences among the various languages, however. As is well known, Spanish adverbs ending (as most do) in *-mente* appear to be in a transitional state (certainly when compared to similar constructions in sister languages like French or Italian), where it is possible to coordinate the stems of adverbs without needing to repeat the affix:
According to Torner (2005), no other Spanish affixes can be coordinated in this fashion, although some endocentric compounds can:

(29) datos tanto macro como microeconómicos
     data both macro with microeconomic-M-PL
     “both macro- and microeconomic data”

(Torner, 2005, p. 117)

His suggestion is that the -mente affix in Spanish is a hybrid called a phrasal affix. Such affixes behave in atypical ways, combining bound (affixal) and free (word-like) properties (Torner, 2005, pp. 126ff.).

Since mente and its cognates are also nouns in many Romance languages (a significant exception being French), the Spanish example provides us with a case study of the shift alluded to in the English examples—from lexicalization of a phrase (in this case, ADJ-N, just as in English) to more opaque morphology. As mentioned above, mente as a noun is feminine and is descended from Latin mēns (“mind”). Its continued combination with an adjective marked for feminine is therefore understandable and based on the history of the form.
Taking into account this behavior and Torner’s proposed solution, we may consider some of the entries in example (20) (repeated here as 20’) again:

(20’) *easy-like*

- Hand it over, **slow and easy-like**!
- If that’s the way the birds get up, sort of **slow and easy-like**, …
- Then I sift in some flour and my spices, **slow and easy like**, and stir like a Mardi Gras demon.

Interestingly, the *-like* affix coordinates much like Spanish *-mente*. This is a possibility (that *-like* is a phrasal affix) worth investigating further, although the coordination is fairly specific and quite possibly idiomatic to this collocation. What is clear is that, like *-mente* in Spanish, adverbial *-like* in English is far freer than e.g. *-ment* in French or English *-ly*.

### 3.5 Other Adverbs as Sites for Lexicalization

In Berry (2009), it is proposed that predicational sentence adverbs (specifically, speech-act adverbs) lexicalize from clausal structures by means of ellipsis. Independent clauses in the early stage of the change (30a), separated perhaps by the interruption of turn-taking, shift from paratactic to hypotactic (30b). This allows the speaker to seek ostensible permission while maintaining the floor. The initial matrix clause then becomes dependent (30c) and reduces from being finite, with agreement and a merged subject, to non-finite (30d). In 30e, the
adverb moves to first position, and in 30f, the verbal form has disappeared, leaving only the adverb. This lexicalization paradigm will be re-examined and further discussed in Chapter 5.

(30) Possible steps in the lexicalization process of *frankly* (sentences from Davies, COCA):


b. I must speak frankly now and tell you that I see no reason for you…

c. If I might speak frankly, there’s others [sic] still aboard more sickish…

d. To speak frankly, Barbara, do you think that this can be accepted…

e. Frankly speaking, if you’re looking for good nutrition in a hot dog…

f. Frankly, in the spirit of free speech, that’s good.

Lexicalization is probably not the only process involved in this sequence. The adverb, which has shifted from being a manner adverb in a verbal structure to being a speech-act adverb at the left periphery, has probably also gone through a process of grammaticalization or pragmationalization.
It no longer is part of the predicate or even of the proposition; instead it has *subjectified* in the sense of Traugott and Dasher (2002) and has become an extrapropositional, metalinguistic comment on the rest of the sentence by the speaker.

Yet grammaticalization usually occurs as a result of reanalysis (in this case, it would result from something like the topicalization of *frankly* so that it is reanalyzed as a speech-act adverb). There is no indication that such a reanalysis has occurred. Therefore, I continue to call this process lexicalization for the time being.

### 3.6 Chapter Summary

The proposal at the heart of this chapter is that lexicalization is perhaps broader in effect than has been indicated by contemporary historical linguists. If lexicalization can be considered idiosyncratic yet categorial, diachronic evidence indicates that the category of *adverb* can be regarded as being primarily morphological, a result of several different lexicalization processes. These may occur at the VP-level or at higher levels in the clause. Such lexicalization may be based on nearly any root or complex.

The resulting adverbs may be completely transparent (as in *nevertheless*) or more opaque (as in *simply* or Fr. *lentement*). Almost universally, however, single-word adverbs show signs of having once been syntactic units or of being constructed on an analogous pattern of derivational morphology based on former syntax. Adverbs can be considered (at least in English and probably in Romance)
to be a category that is essentially compositional, with only a very few deictic primitive forms.

Modern adverbs can have origins in noun phrases (the -ly and -like adverbs), in prepositional phrases (e.g. perhaps, besides), and in clauses that have been elided and/or fused (e.g. maybe, frankly). With so many historical sources for adverbs, the lack of semantic and syntactic cohesion among adverbs is more understandable. Unlike nouns, adverbs do not have a unique semantic “adverbiness” unifying them as a category. Their variety is the fossil record of their diverse origins.
Chapter 4

THEORETICAL FRAMEWORKS, PART II: MORPHOSYNTAX

4.1 Introduction

4.1.1 Why generative grammar?

As mentioned in the introduction to Chapter 2, one of the stated goals of this study is to continue the tradition that combines diachronic linguistics with generative theories of syntax and morphology. This chapter therefore has a structure that builds upon existing Chomskyan models of autonomous syntax (and, by extension, morphology); that specifically examines the treatment of adverbs and adverbials in two generative models (championed by Cinque and Ernst); and that reviews various generative approaches to historical linguistics.

Before we turn to an examination of the various literatures, I would like to discuss the use of generative grammar in diachronic studies such as this one. By a large margin, the majority of linguists working in historical linguistic studies use a functionalist/descriptivist framework of one type or another. The pressure on anyone working in diachronic studies, therefore, is to conform to such a model. However, there has been a movement over at least the past 30 years to combine formal and functional approaches, as each has different strengths.

One of the strengths that generative syntax offers that particularly appeals to me is the interest that Noam Chomsky and other formal syntacticians have in moving beyond descriptive adequacy in their work. Functional linguists, working in syntax and other fields, show great strengths in their ability to describe and organize observed facts of language use, and to hypothesize how certain
language-specific phenomena occur. But their determination to qualify, study, and understand language as a purely cultural tool limits their explanatory ability, in my opinion. I believe that generative grammar, with its autonomous syntax that is predicated on the operation Merge, offers instead an investigation that moves beyond the question of whether something linguistic happens to ask how and why it happens. This is a shift from descriptive adequacy to explanatory adequacy, and Chomsky (2004) urges the investigation of language to proceed even farther.

No model of language is perfect, and generative syntax has stumbled at times when it has been applied to some languages; yet the goal of Minimalism or Principles and Parameters is nevertheless admirable. The centrality of syntax in the Chomskyan models is crucial; the ability to merge or combine concepts freely is what, I believe, separates humans from other intelligent creatures. Other elements of linguistic structure and meaning are, of course, hugely important. But there are, ultimately, “work-arounds” for phonology (i.e. sign language or writing) and morphology (e.g. isolating languages like Chinese have very little). And in his famous and oft-quoted sentence pair (see example 1), Chomsky effectively argued against the primacy of semantics in grammar:

(1) Colorless green ideas sleep furiously.

*Furiously sleep ideas green colorless.

(Chomsky, 1957, p. 15; my asterisk)
This study therefore, as stated elsewhere, is an endeavor to continue a tradition and to tread the line between formalist and functionalist models of language. In it I hope to add understanding of motivation to phenomena of language change and thereby add to a growing body of scholarship.

**4.1.2 Chapter organization.**

The organization of this chapter is as follows: In section 4.2, I examine generative grammar in sum, starting with a quick look at its history and development. Then I examine the two major branches of recent theory: the Principles and Parameters model and the Minimalist Program. Finally, I give a brief background to the Distributed Morphology model of generative morphology. Section 4.3 is devoted to generative theories about adverbs and is centered around the two major theories: cartography, as advocated by Cinque (1999) and Alexiadou (1997); and semantic-based adjunction, as in Ernst (2002). In section 4.4 I examine in detail the generative tradition in diachronic linguistics, focusing mostly on models of grammaticalization by Roberts and Roussou (2003) and by van Gelderen (2004). Section 4.5 then applies such frameworks to not only adverbial grammaticalization but also lexicalization, as I use them in this study. Finally, section 4.6 summarizes the chapter.
4.2 Generative Grammar

4.2.1 Introduction: origins, history, and development.

What became known as transformational grammar first came to the attention of academia on a large scale with the publication of Noam Chomsky’s (1957) *Syntactic Structures*. This slim (117-page) volume set up the foundation for Chomsky’s theories to follow, first by separating syntax from other so-called “levels” of linguistic study, particularly from semantics and from statistical, empirical models:

Despite the undeniable interest and importance of semantic and statistical studies of language, they appear to have no direct relevance to the problem of determining or characterizing the set of grammatical utterances. I think that we are forced to conclude that grammar is autonomous and independent of meaning, and that probabilistic models give no particular insight into some of the basic problems of syntactic structure. (Chomsky, 1957, p. 17; my italics)

He goes on to set up and reject theories based purely on linear (left-to-right) or sequential models (p. 24) and on simple constituent analysis (p. 32), although the latter does play an important role in his model. In delineating the significance of phrase structure, he offers an early tree:

(2) Kernel sentence

```
(2) Kernel sentence

Sentence
  /   
NP   VP
  /   
T    N
the  man

  /   
Verb
  /   
T    N
hit  the

  /   
NP
  /   
T    N
the  ball

(Chomsky, 1957, p. 27)
```
It is important for him that a “diagram” like this indicates that $NP = T + N$ and that $VP = V + NP$. A collocation like “man hit” cannot be a constituent (p. 27).

Such a sentence as shown in example (2) is a kernel, and to achieve a passive (*The ball was hit by the man*) or other non-basic form such as question or negation, we must apply transformational rules. These rules apply to all sentences, including kernel sentences, but are distinguished between obligatory (kernel) transformations and optional (more complex) ones (p. 61).

A descriptive grammar, therefore, is a tripartite collection of rules. *Morphophonemic* rules are necessary for the expression of the elements of the grammar, and *phrase structure* rules are necessary for the construction of constituents. These are elementary rules. The higher-level rules are the *transformational* rules, by which more complex sentence types are created (p. 107). Chomsky reiterates in his conclusion that “[g]rammar is best formulated as a self-contained study independent of semantics. In particular, the notion of grammaticalness cannot be identified with meaningfulness” (p. 106).

One point that is notable in *Syntactic Structures*, as in most of Chomsky’s other works, is the concentration on the basic skeleton of the sentence, what we now think of (in semantic terms) as the *argument structure*. The core of his interests is there: in his early syntactic tree are the seeds of Merge ($VP = V + NP$) and his optional transformations are signals of Move-$\alpha$, which plays an important role in Principles and Parameters/Government and Binding theory. He shows no interest in and makes no mention of modifiers in the NP (i.e. adjectives) or the VP (adverbs).
By 1965’s *Aspects of the Theory of Syntax*, the transformational grammar model is focused and defined. However, the lack of interest in adjunction in general and in adverbials in particular is as pronounced as it had been in *Syntactic Structures*. In *Aspects*, Chomsky expands and clarifies the positions he introduced in 1957.

First among the expansion of early transformational ideas is the return of semantics to the model. Chomsky does not consider semantics to be generative in the same way that syntax is; for him, *subcategorization rules* (rules of grammaticality of expressions) are far more important than *selectional rules* (rules that use semantic features, such as \([±\text{Abstract}]\) or \([±\text{Animate}]\)) (1965, pp. 148ff.). One of the few roles that adverbials play in the model is determining subcategorization of verbs. Here Chomsky distinguishes among adverbial types:

…Verbs are subcategorized with respect to the Prepositional Phrases introduced by (52iii) [i.e. Manner Adverbials] but not with respect to those introduced by (52ii)—namely the Place and Time Adverbials that are associated with the full Predicate-Phrase, and that might, in fact, be in part more closely associated with the Auxiliary (cf. note 23) or with Sentence Adverbials which form a “pre-Sentence” unit in the underlying structure. (1965, p. 102)

Chomsky notes, for example, that manner adverbials are highly restricted with so-called “middle” verbs (such as *cost*, *resemble*, or *marry*). These verbs do not allow a passive transformation, so he elaborates the role of manner adverbials in determining verbal subcategorization by postulating a “dummy element” in their construction, “signifying that the passive transformation must obligatorily apply” (Chomsky, 1965, p. 103). This is exceptional; adverbs and adverbials seem to play no other major role in what is called *Standard Theory*.  

101
Another major shift found in *Aspects* is the further delineation of the ideal linguistic subject: namely the foregrounding of linguistic *competence* over *performance*. These are later re-named *I (internal)-language* and *E (external)-language*. Generative syntax takes I-language as its subject. A clear picture of Chomsky’s aim can be found in the first section of the book:

A grammar of a language purports to be a description of the ideal speaker-hearer’s intrinsic competence. If the grammar is, furthermore, perfectly explicit—in other words, if it does not rely on the intelligence of the understanding reader but rather provides an explicit analysis of his contribution—we may (somewhat redundantly) call it a *generative grammar*. (Chomsky, 1965, p. 4; italics original)

Further significant changes found in *Aspects* include, very briefly, a terminological shift from *kernel* to *base* as an indication of the initial, primitive state of a sentence (although he maintains the use of “kernel sentence” as a particular type, this term does not persist); the base form is represented in *deep structure*, the semantic interface that is resistant to the changes seen in *surface structure*, which in turn is the phonological string that expresses the result of grammatical transformations (Chomsky, 1965, pp. 15-18). Although the term “Universal Grammar” is not explicitly used in *Aspects*, the skeleton of this important element of generative theory is also briefly discussed in the section on formal universals (pp. 27-30).

The Standard Theory as set forth in *Aspects*, therefore, introduces the core of arguments that builds upon the initial forms in *Syntactic Structures* and codifies them. Syntax is an autonomous component of the language ability of humans; only syntax is capable of generating linguistic output, which it does while...
interfacing with phonological and semantic components. Meaning is based on the pre-transformational forms of deep structure, while surface structure expresses the transformational effects of, for example, passive voice.

This structure is only moderately changed (i.e. mainly through attributing some semantic interface properties to surface structure as well in Extended Standard Theory; see e.g. Chomsky, 1979, p. 163) until the 1980s, when the Principles and Parameters/Government and Binding/X-bar theoretical model becomes preeminent.

4.2.2 Principles and Parameters; the Minimalist Program.

The Principles and Parameters (hence P&P) model that emerged from Extended Standard Theory is a systematic model for separating Universal Grammar elements (principles) from language-specific parameters. In such a model, children learning language follow the parameters set forth in the language they are exposed to. All speakers, on the other hand, are biologically compelled to follow the universal principles in order to use language successfully.

Conflated with the P&P model is a model that pre-dates it, namely X-bar theory. This approach to constituent structure has its origins in Chomsky (1970) and is refined over the following decades so that it is broadly used by the time of the rise of the P&P model. In X-bar theory, there are three important levels to the projection of each lexical (and, eventually, functional) category.

The first is the minimal projection, the $X^0$ or head level, which is the main, label-giving element; moving higher on the tree, it is followed by the $X'$ (or
X-bar) level, an intermediate level that allows for complementation that follows the head (and for adjunction); and the maximum projection is the XP (for X-phrase) or X" level, which allows for a specifier to precede the head in the phrase. Despite later attempts to do away with X-bar theory, it still remains the most useful means of graphically showing the structure of phrases and clauses.

Over the past 20 years, there has been a shift away from some aspects of the P&P model. Part of the difficulty seems to have been the task of collecting and sorting through a list of agreed-upon universal principles (other than e.g. the phrase projection principle from X-bar theory) and parameters (such as pro-drop, phrase headedness, and the Extended Phrase Principle or EPP that attracts subjects/other phrases to higher spec positions in some languages).

For example, Radford (1988), in his very clear, widely-used, and well-regarded textbook, lists three major principles: the Endocentricity Constraint (only one head per phrase); the Modifier Maximality Constraint (all modifiers must be maximal projections) and the Category Neutrality Constraint (phrase structure rules apply equally to all types of phrases) (pp. 258ff.).

Nevertheless, P&P is still influential today, and when it was the predominant model it spawned several other theories. One such theory is found in Chomsky (1986), where the Government and Binding (hence GB) theory is very clearly explained. GB is a major component of P&P, and it makes use of several elements that come from X-bar theory as well, such as c-command and m-command. Significant to the purposes of our study is the theory of movement
This theory is often abbreviated as \textit{Move-\(\alpha\)}. There are two
types of movement that Chomsky recognizes. The first is \textit{substitution}:

\begin{enumerate}
\item Theory of movement (Substitution)
\begin{enumerate}
\item There is no movement to complement position.
\item Only \(X^0\) can move to the head position.
\item Only a maximal projection can move to the specifier position.
\item Only minimal and maximal projections (\(X^0\) and \(X"\)) are “visible” for the
rule Move-\(\alpha\).
\end{enumerate}
\end{enumerate}

Such rules within the theory prohibit rightward movement, restrict head
movement to existing heads, restrict specifiers to full phrases, and prohibit partial
phrase (\(X'\)) movement. There is another option within the theory of movement,
\textit{namely adjunction}:

\begin{enumerate}
\item Theory of movement (Adjunction)
\begin{enumerate}
\item Adjunction is possible only to a maximal projection that is a nonargument.
\end{enumerate}
\end{enumerate}
This limits adjunction to VP, rather than to DP (NP) or to CP when either is acting as an argument in the clause.

Although Chomsky does not explicitly do so, Radford (1988, pp. 197ff.) distinguishes between *Adjunct* and *Attribute*, so that the two would appear in this fashion:

(5)  

(a) Adjunct (based on Radford, 1988)

```
XP
  ├── Specifier
  │    ├── X'
  │    │    ├── X'
  │    │    │    └── Adjunct
  │    │    └── X^0
  └── X^0
      └── Complement
```

(b) Attribute (ibid)

```
XP
  ├── Specifier
  │    ├── X'
  │    │    ├── Attribute
  │    │    │    └── X'
  │    │    └── X^0
  └── X^0
      └── Complement
```

As we will see when examining Kayne (1994) and the cartographic model that emerges from his theories, both of these models for adjunction/attribution are considered somewhat controversial and are revised in order to simplify the structure.
The above models (examples 5a and 5b) from Radford (1988) are, however, useful to help demonstrate the important structural concepts of $c$-

\textit{command} and $m$-\textit{command}, as explained in Chomsky (1986):

(6) $c$-command

\[ \alpha \text{ c-commands } \beta \text{ iff } \alpha \text{ does not dominate } \beta \text{ and every } \gamma \text{ that dominates } \alpha \text{ dominates } \beta. \]

(7) $m$-command

Where $\gamma$ is restricted to maximal projections … we will say that $\alpha$ $m$-commands $\beta$.

In Chomsky (1986), $m$-command is important for government relations, but it also plays a role in Kayne’s (1994) antisymmetry (particularly as he employs various techniques to avoid using it). What $m$-command allows is for the head of a phrase in a specifier position to take a hierarchical role over the head and complement of the main phrase in question. While $c$-command has lasted, $m$-command is only used in limited fashion today.

Chomsky’s Minimalist Program (hence MP) (1995 et seq.) and the work it has inspired are a break from earlier syntactic work, significantly from GB and P&P. As befits the name, the MP seeks to \textit{simplify} the syntactic component of
language. Chomsky (2007) fits syntactic study into a larger biolinguistic framework of which there are three components: a genetic endowment of a language center (sometimes known as an “organ”) in the brain (Universal Grammar or UG); external language data to which speakers are exposed; and other properties of organic beings (which may or may not be purely human in nature). The MP continues to take as its focus the study of UG, though the focus is on limiting UG.

The background of such a study remains the consistency of properties across human languages, which Chomsky and minimalists continue to emphasize. Ultimately, all human language is used with the same goals of communication; the differences in approaching these goals are important but are not the aims of minimalist syntax. In Chomsky (2001), the continuing attitude toward human language is the Uniformity Hypothesis:

(8) Uniformity Hypothesis (from Chomsky, 2001, p. 2)

In the absence of compelling evidence to the contrary, assume languages to be uniform, with variety reduced to easily detectable properties of utterances.

Primary to minimalism is still the process of narrow syntax, known (since at least P&P) as a derivation. Derivations begin with a selection of items (a lexical array) from the individual’s lexicon. Chomsky (2007) has more recently
backed away somewhat from a hard-line lexicalist position (where each lexical item (LI) enters the syntax fully inflected) and now allows the possibility of models such as Distributed Morphology (e.g. Halle & Marantz, 1993), where syntactic position drives morphology. (See section 4.2.3 below for a more detailed discussion.)

Once lexical items are selected, the most important operation in a derivation is Merge. There are two types of Merge; these are used to form an utterance in a “bottom-up” structure. External Merge is the first operation (and arguably the simplest), where two LIs are combined. One element is assigned the position of head, and the other is the complement. The head projects and provides the label for the merged pair. In the later Bare Syntax model (cf. Boeckx, 2008), the labels are abandoned as being extraneous; this is not significant for this study.

Internal Merge is the later operation. Both types of Merge require an Agreement of features; Agree is the result of a Probe with uninterpretable (unvalued in later theory) features seeking a Goal with interpretable/valued features that Match. Interpretation of features is accomplished at the semantic level. Once the Probe has found an appropriate Goal (within the most localized domain), the Merge operation occurs and the u-features are deleted from the derivation. Internal Merge, unlike External Merge, may require movement of the Goal to the specifier position of the Probe (Chomsky, 2001).

The features used in derivations can vary according to the parametric settings of the language in question (a retention of parameters from P&P, though these are still largely undefined). Common to most derivations are so-called $q$-
features (person, number, and gender) that are used for subject-verb agreement. These $ϕ$-features are uninterpretable or unvalued on the tensed verb but are interpretable/valued on the subject nominal, which overtly shows these values. The T-head probes for the subject NP (its Goal), and the uninterpretable/unvalued features are deleted from T. T also (in English) has an Extended Projection Principle (EPP) feature (now also known as OCC, or occurrence) that triggers the Move portion of Internal Merge. Finally, the subject NP may carry an unvalued Case feature that the T values.

Syntactic derivation, according to Chomsky (2001), operates in cycles (phases) that are required by organic properties—particularly by memory constraints. There are two major phases in the syntax of a complete utterance: $vP$ (the light-verb phrase or “$vP$ shell”, where argument structure and $θ$-roles are determined) and CP (the complementizer phrase, with two macro-layers, CP and TP (tense phrase); the TP gives grammatical information while the CP turns the utterance outward toward the discourse situation).

To use traditional terminology, any of the subject, full predicate, verb, and object may internally merge out of a lower phase, although any element deeper than a head must move intermediately to the edge or spec of a phase before it can be probed (Chomsky, 2004). A simplified example of phase-cyclic merge can be seen in (9):
In this example, the Merge operations occur in this order: 1) the lexical verb *(eat)* and the direct object *(the food)* merge to form the VP; 2) the subject *(the dog)* merges with the VP at the Spec,VP position; 3) the edge features of the *vP* phase draw the head and spec of VP to the head and spec, respectively, of *vP*; 4) the TP is then able to probe the subject and lexical verb, because they are at the edge of the first phase; 5) probing values the features and allows the interpretable ones to appear (i.e. past tense on the verb) while the uninterpretable ones are deleted.

\[\text{The dog ate the food.}\]

\(^6\) In this example, I continue to use X-bar terminology, but there is a debate as to whether these terms constitute a trespass of the Inclusiveness Condition (i.e. nothing should be added to a derivation outside of the combining elements and features) (Chomsky, 1995).

\(^6\)
(struck through); 6) the EPP feature allows the subject to move from its position at the edge of vP to the specifier of TP.

Once a phase is completed, it is transferred to the interfaces. The sensorimotor (or SM) (also articulatory-perceptual, A-P) interface linearizes the derived components and (perhaps after a Morphology operation) sends them to Spell-Out (or pronunciation/signing/writing). The conceptual-intentional (C-I) interface focuses on the semantic meaning of the utterance component; unlike SM, C-I does not linearize but organizes by simple hierarchies (Chomsky, 2004).

This model is very useful for argument structure, but it lacks a useful approach to the insertion of “non-essential” or adjunct elements into the derivation. Chomsky acknowledges such a lack.

There has never, to my knowledge, been a really satisfactory theory of adjunction, and to construct one is no slight task. … An adjunction construction is plainly not the projection of a head: for NP-adjuncts, for example, the constituent structure appears to be something like [NP XP]. The construction is crucially asymmetric: if \( \alpha \) is adjoined to \( \beta \), the construction behaves as if \( \alpha \) isn’t there apart from semantic interpretation, which is not that of standard X-bar-theoretic constructions; island properties differ as well. (Chomsky, 2004, p. 117)

According to Boeckx (2008), the null hypothesis in generative grammar is complementation. Adjunction is exceptional, marked (p. 98). For him, adjuncts merge with their hosts through the operation Match, which is different (simpler, less significant) from Agree in that there is no feature checking as happens in more structural types of Merge. Similarly, Chomsky distinguishes between pair Merge \(<\alpha, \beta>\) and set Merge \(\{\alpha, \beta\}\) (2004, p. 119). For him, adjunction is pair Merge:
In $\alpha, \beta$, $\alpha$ is spelled out where $\beta$ is.

Here, the directionality of Merge is somewhat weak. If two elements are merged, how do we know what projects? For example, projection in set Merge cannot be based on some kind of “semantic weight”: Determiner Phrases (DPs) are centered around the functional rather than the lexical categories. For both SM and C-I interfaces, however, we need to have a clearly linear/hierarchical structure. Merge alone cannot provide this structure, so others have suggested alternate solutions that add elements to UG, thus moving away from minimalism.

One of the most influential of these solutions (and one we explore below) is Kayne’s (1994) Linear Correspondence Axiom (LCA). The LCA depends on asymmetric c-command to give both linearization at SM and hierarchy at C-I. The LCA breaks from Chomsky’s bare phrase structure by depending on projection into categories, as in X-bar theory. For example, External Merge cannot exist between two items (bare head and bare complement), but between a head and a phrase acting as a complement (1994). In later works, Kayne offers a solution to some of the problems of projection of a complement by requiring Self-Merge as a first step (2008).

Although Kayne’s theory of the LCA inspires some objections from MP followers (notably Chomsky himself), it has helped to open the subfield of cartography in minimalist syntax. Another highly influential contribution that falls
under this category is Cinque’s (1999 et seq.) articulation of the TP in a cartographic fashion.

Based on a structure that is highly compatible with Kayne’s LCA architecture, Cinque’s work provides a myriad of landing places for adverbs from the lower end of the CP down through the VP. His work is based on thorough and remarkably consistent empirical data, but there are several concerns with his work as well—not least of which is the non-minimal feel of dozens of functional heads scattered through the tree. We will explore both Kayne’s and Cinque’s theoretical models and their relationship with minimalism below in section 4.3.2.

Chomsky’s MP solves several problems with argument structure and with offering an explanation of the processes involved in language, and it does so in a simple and elegant manner. There are several natural language phenomena that aren’t addressed by his architecture, however, and attempts to address these issues tend to diverge from Chomsky’s minimalist simplicity. One of these, as we shall investigate further below, is the syntactic study of adverbs and adverbials, all of which fall outside the core of the clause.

### 4.2.3 Distributed Morphology

One of the fundamental aspects of generativist syntactic theory for the past 40 years has been the concept of the *lexicon*. As Chomsky and others have developed and refined syntactic theory, adding and subtracting such concepts as principles, parameters, and post-movement traces, the concept of a lexicon from
which most words are chosen whole (i.e. fully inflected) has remained comparatively stable.

Chomsky’s (1970) work “Remarks on Nominalization” has been considered to be the initial point where the *lexicalist* position was developed and articulated. Chomsky’s early position, according to Julien (2002), differentiates inflectional morphology (initially seen as syntactic) from derivational morphology, which is part of the lexicon (p. 10). Later incarnations of the minimalist approach, which have included feature checking on finite verbs, have moved farther away from allowing a syntactic role for inflectional morphology. Finite verbs under this aspect of the theory enter the syntax as fully inflected forms that only need to check features with other elements and move if required (Julien, 2002, p. 11).

Responses to the lexicalist approach to word insertion have gained attention in the past 20 years. One of the best-known morphological theories is Distributed Morphology, first postulated by Halle and Marantz in 1993 and refined, expanded, and re-articulated in several ways since then. Distributed Morphology (hence DM) refutes not only the lexicalist view of selecting completely derived and inflected items from a fixed lexicon, but also phonological models such as Optimality Theory, which unifies phonology and morphology.

DM relies on three key elements: *Late Insertion*, which inserts vocabulary items after the syntactic structure is developed; *Underspecification*, in which morphemes (which correspond with sound/meaning-based Vocabulary Items in
the theory) possess only a subset of the features specified by a terminal node in the syntax—competition exists among various Vocabulary Items and only the “fittest” is inserted; and *Syntactic Structure All the Way Down*, where the linear structure of Vocabulary Items is constrained by syntactic locality conditions, which can be modified by the morphological actions that include *head-to-head movement* and *morphological merger* of adjacent items (Halle & Marantz, 1994, pp. 275-6).

Other important features of DM include *Impoverishment*, where the irregularity of morpheme choices found in many languages (both in stems and affixes) is explained; *Fusion*, where semantically complex (or “portmanteau”) morphemes are accounted for; and *Fission*, where morphemic information is split, for example into the combination of a prefix and a suffix that is known as a *circumfix* (Halle, 1997, pp. 429-32).

Among morphemes there is a separation between *f-morphemes* and *l-morphemes* that correspond “approximately to the conventional division between functional and lexical categories” (Harley & Noyer, 1999, p. 4). F-morphemes are those that are inserted without choice to satisfy a specific grammatical need (i.e. gender, number, case); their spell-out is considered “deterministic” (p. 4). L-morphemes, by contrast and definition, are subject to choice by the language user. Any number of “nouns”, for example, could be inserted in a terminal syntactic node that calls for a noun. In addition, l-morphemes of undetermined category (*roots*) can be inserted at a terminal node, and the syntax will motivate the f-
morpheme that engages in morphological merger with that l-morpheme (Harley & Noyer, 1999, p. 4).

Embick (2003) speaks of cycles—root cycles and outer cycles—to discuss the nature of derivational morphology. These cycles may be thought of as roughly analogous to syntactic phases (as discussed above). The root cycle is based on conditions of locality, and it revolves around the original root or stem of a word; when an affix is adjoined to the root, the resulting construction is reanalyzed as the root of a new cycle (an outer cycle) and movement and merger result in a new affixation (Embick, 2003, p. 150). In English, for example, this pattern can be repeated several times to create nouns, verbs, adjectives, and adverbs.

One final consideration for DM scholars is their rather hostile relationship with the term lexicalization, a significant component of this particular study. Because they deny the Lexicalist Hypothesis, DM morphologists assert the impossibility of “lexicalizing” formerly separated words. Marantz (1997) indicates that “there is no sharp divide between the special meanings of words and the special meanings of phrases, nor has there been any systematic attempt to argue otherwise” (p. 207). Harley and Noyer (1999) argue similarly:

The internal structure of expressions is demonstrably not always a product of syntactic operations. In DM structure is produced both in syntax and after syntax in the Morphology component… Nevertheless, because of “Syntactic Hierarchical Structure all the Way Down”, operations within Morphology still manipulate what are essentially syntactic structural relations. The syntactic component produces a representation whose terminal elements are morphosyntactic features, which is then subject to operations such as “Merger Under Adjacency”, “Fission”, or “Fusion”, accounting for non-isomorphic mappings from syntactic terminals to morphophonological constituents. (p. 4)
Morphological merger, referenced in the above quote as “Merger Under Adjacency”, is the major actor in “lexicalization”. As Embick and Noyer (2007) say, “the ‘word’ is not a privileged derivational object as far as the architecture of the grammar is concerned, since all complex objects, whether words and phrases, are treated as the output of the same generative system (the syntax)” (p. 1). The concept of a syntactic morphological component that is constructivist, such as DM, is appealing for this particular study, which seeks to understand (what I will continue to term) lexicalization from a syntactic perspective.

Having set up the basic structure of the generative grammatical models I am using in this study, I will now turn to the application of these models to adverbs in general (and sentence adverbs in particular). There are two competing generative models (cartographic and semantic) that have attempted to add adverbs to the tree in a principled manner, as we shall see.

4.3 Theoretical Models of Adverbs

4.3.1 Introduction: general difficulties with adverbs.

Adverbs, as we saw in the above sections, pose a problem in general for generative grammarians. From the time of Syntactic Structures in 1957, Chomsky’s focus and the focus of others who work in syntax has mostly been the “core” of the sentence—namely the Subject, Verb, and Object structures by which we typologically categorize languages. Semantically, these elements are referred to (with slightly differing meanings) as argument structure or θ-roles. Adverbs and adverbials can add meaning to the core structure of the sentence, but they do
not carry any of the weight of the structure. Such is the meaning of Chomsky’s pair Merge, as in example (10) above.

In contrast to their syntactic poverty, adverbs (and adverbials) possess a great deal of semantic richness. In Chapter 3, I argue that adverbs as a category are morphologically defined. Such levels of morphological merger and fusion (in both DM terms and diachronic terms) as are found in their final form allows adverbs in particular to act as a sort of conversational “shorthand”, a group of abbreviated clause- and phrase-words through which larger semantic and pragmatic meanings are transferred.

This is highly useful in conversation and in other types of discourse, where economical methods of conveying information are essential to effective communication. But by essentially relegating the domain of adverbs to morphology and semantics, linguists are saying, in effect, that the syntax of adverbs (and of adverbials by extension) is unimportant. To the contrary, recent research has indicated that adverbial syntax is, instead, highly articulated in concert both with semantics and, as we shall see, with the discourse requirements of speakers.

The remainder of section 4.3 will look at the two leading syntactic solutions to the “problem” of adverbs. First we will examine the rise of the cartographic enterprise regarding adverbs, under the aegis of Cinque (1999 et seq.) but also encompassing work by Haumann (2007) and Alexiadou (1997). This work on adverbs is situated within a larger cartographic framework that
stretches from the CP layer at the left periphery of the English sentence (Rizzi 1997, et seq.) to the vP layer around the core of the clause (e.g. Ramchand, 2008).

In section 4.3.3, we will discuss the main competitor for cartography among generative theories of adverbs: the semantic-based model advocated by Ernst (2002 et seq.). In this model, syntax acts at the direction of semantic needs, and there is more freedom of adverb placement in the sentence than is found in Cinque’s cartographic model. There are, of course, pros and cons to this sort of structure, and they will be discussed as well. Finally, in section 4.3.4, I discuss the model of adverbial syntax chosen for this study.

4.3.2 Antisymmetry and cartography.

The cartographic enterprise in generative syntax, as Shlonsky (2010) terms it, is an attempt by some generative linguists to add detail to the existing skeletal structure of the clause as laid out in, e.g. P&P (CP>TP>vP>VP). Each of the top three “layers” adds new levels of information to the base of the clause. The basic proposition in cartography is, as the name indicates, to map out within the confines of this skeleton the position of various functional and semantic/pragmatic elements of the sentence, thereby assigning syntactic dimensions to these elements. Each of the three layers has its own grammatical and semantic details that are more fully exposed through cartography, as Shlonsky (2010) indicates:

Closest to the verb, one finds aktionsart projections, which determine the event-type of the predicate (Vendlerian classes; see e.g. Ramchand 2008). In the Cinquean hierarchy…, the two tense heads (Past and Future) are
arrayed above all the aspectual heads. The heads encoding illocutionary force, evidentiality, and other “moods” are positioned higher than tense. In the higher, left-peripheral space, one finds scope and discourse/informational projections, including quantificational elements, interrogative particles, topics, and focus. (p. 425)

Indeed, it is considered an accepted doctrine in generative grammar that the VP layer is the locus of lexical (in the sense of not functional) information, the vP layer is the extension of argument structure and inner or “lexical” aspect considerations, the TP (including tense, mood, and aspect) is the grammatical layer, and the CP is the discourse layer that connects to previous information and also privileges certain semantic items. The cartographic enterprise has as its focus the regularized exponence of such qualities.

As noted above, the growth of cartography in the late 1990s was largely triggered by Kayne (1994), whose monograph *The Antisymmetry of Syntax* postulates a highly consistent phrase structure, one which does not allow for multiple specifiers (as the Attribute construction in 5b above) or for right adjunction (as in 5a). Instead, Kayne states that phrase structure is the determiner of linear structure at a macro (clausal) level (i.e. CP encompasses TP, which in turn encompasses vP, etc.), and that such linearity can be extended to the components of the phrase itself, making the phrases “locally linear” (Kayne, 1994, p. 5).

To achieve a locally linear structure, *c-command* is the important component, just as both c-command and *dominance* are significant for a macro structure. Dominance at the intra-phrase level is less a concern, because every node is dominated by the X"-level. However, c-command must be limited—must
be asymmetric—in order to achieve a consistent hierarchical structure that also is consistently linear. Kayne’s solution to this problem is the Linear Correspondence Axiom (or LCA):

(11) Linear Correspondence Axiom (Kayne, 1994, pp. 5-6)

To express the intuition that asymmetric c-command is closely matched to the linear order of terminals, let us, for a given phrase marker, consider the set $A$ of ordered pairs $<X_j, Y_j>$ such that for each $j$, $X_j$ asymmetrically c-commands $Y_j$. Let us further take $A$ to be the maximal such set; that is, $A$ contains all pairs of nonterminals such that the first asymmetrically c-commands the second. Then the central proposal I would like to make is the following (for a given phrase marker $P$, with $T$ the set of terminals and $A$ as just given):

*Linear Correspondence Axiom*

$d(A)$ is a linear ordering of $T$.

The LCA, as an enforceable rule, moves syntax beyond its previously-acknowledged quality of asymmetry (i.e. there is an order that is expressed phonologically as *linear* and semantically as *hierarchical*) to *antisymmetry*. In other words, under the LCA, the structure must follow a specific order; there is no longer a directionality (or headedness) parameter. The only order is specifier-head-complement (Kayne, 1994, pp. 35ff.).

With such a structure, adjunction is no longer available on the right of a head, unless it is in the specifier position of a phrase that is c-commanded by that head. To contrast the structures, consider example (5) above, repeated as (12):
(12)  a. Adjunct (based on Radford 1988)

```
            XP
          /    |
Specifier X'         X'  Adjunct
          /    |
            X'  Complement
```

b. Attribute (ibid)

```
            XP
          /    |
Specifier X'         X'  Attribute
          /    |
            X'  Complement
```

These are the typical adjunction patterns of P&P, and are often called Chomsky Adjunction. In contrast, the phrase structure of Kaynean antisymmetry is more consistent, but arguably less flexible because of the loss of multiple X' levels:

(13)  Antisymmetric adjunction (following Kayne, 1994):

```
            XP
          /    |
Specifier X'         X'  Complement (YP)
          /    |
            X'  Complement (YP)
            /    |
Specifier Y'         Y'  Complement (ZP), etc.
```

123
What this does is set up a continuing series of specifier-head pairs that cascades from the beginning of a clause to its end. Ultimately, it gives a mathematically consistent structure that is repeatable *ad infinitum* and that gives a series of “landing sites” for all types of Move-α (as in Chomsky, 1986). Heads can engage in head-to-head movement and have various motivations for doing so, while phrases can land in the specifier positions of various functional heads. In MP, of course, this behavior is now known as Internal Merge, is feature-based, and is free.

The first cartographic proposal to use Kayne’s antisymmetric structure is Rizzi’s (1997) study of the CP-layer or left periphery of the sentence. This is the discourse-oriented layer, and it is host to several different types of phrases and heads. One of the questions that continually recurs for both CP and TP cartographic positions is the motivation for Merge (whether Internal or External) at the landing sites involved. Rizzi (1997) addresses this topic directly:

> I will phrase such requirements [i.e. of heads] in the style of the Criteria … rather than as feature checking, the main reason for this choice being that such features have an interpretive import (Wh, Neg, Top, Foc,…): they determine the interpretation of the category bearing them and of its immediate constituents …, function as scope markers for phrases with the relevant quantificational force in a local configuration, etc., so that *their role cannot simply be to trigger movement and disappear from representations.* (p. 242; italics mine)

Here it is clear that Rizzi regards the features that trigger CP movement as different from those that trigger SVO movement in the TP or vP. Where the features that drive movement at lower levels of the clause are largely uninterpretable (this term is briefly covered in section 4.2.2)—in other words,
they are present on the grammatical head or Probe and, once satisfied by a Goal, are deleted from the derivation—CP features that induce movement are interpretable, and are not deleted as a result of Merge.

In his original (1997) cartographic model, Rizzi breaks the CP into five different layers, made up of two main subsystems: the Force-Finiteness system and the Topic-Focus system. Force is a phrase that can host question/declarative/exclamative markers, relative and comparative markers, and certain adverbials. Finiteness is a phrase to distinguish between finite and non-finite clauses (e.g. in English, finite ones take that and non-finite ones take for) (Rizzi, 1997, pp. 243-4). Topic and Focus phrases highlight topics (old information) and focuses (new information), respectively (pp. 245ff.). So initially, Rizzi’s cartographic hierarchy is that found in (14):

(14) Original extended CP\(^7\) (Rizzi, 1997, p. 257)

\[
\text{ForceP} > \text{TopP}^* > \text{FocusP} > \text{TopP}^* > \text{FinP} > \text{IP}
\]

More recently, he has added further functional heads to host, for example, certain interrogatives (IntP) and preposed adverbs (that are left of Cinque’s hierarchy as defined below) (ModP):

\(^7\) Multiple Topic phrases are allowed in Rizzi’s theory to accommodate languages such as Italian (1997, pp. 255-7). IP is the Inflection phrase, another (earlier) way of indicating the TP-layer (including tense, mood, aspect, voice).
Revised extended CP (Rizzi, 2004, p. 242)

| ForceP | TopP* | IntP | TopP* | FocusP* | ModP* | TopP* | FinP | IP |

The advantages to such a structure are clear for languages that allow a crowded CP. In addition, the presence of a designated “higher landing site” for preposed adverbs (ModP) allows for some of the freedom of movement that has been supposed to be lacking from this framework.

The concentration of the work in this study is on the left periphery and on the “sentence adverb” positions, so I will not be canvassing the cartographic efforts that have been exerted on behalf of the vP. Instead, the next focus for us will be the expanded TP, and the work that Cinque has done to map the position of a variety of higher adverbs.

Cinque (1999), in his preface, asserts two goals for his cartographic approach. First, he seeks “to motivate an analysis of adverb phrases (AdvPs) as the unique specifiers of distinct maximal projections, rather than as adjuncts.” A further goal is “to argue for the existence of a fixed universal hierarchy or clausal functional projections” (p. v).

For many scholars who have (or haven’t) read Cinque’s book, the first goal is more theoretical, and generally more difficult to accept. We can see, however, that his goal fits in with Kayne’s (1994) antisymmetric approach and Rizzi’s (2004) solution for expanding the CP. On the other hand, Cinque’s use of data from a large number of languages makes his second stated goal, that of
establishing a fixed order of adverbials in all human languages, more easily accepted. He depends heavily on Indo-European languages (in particular, Italian, French, and English) to demonstrate his hierarchies, yet he also makes detailed use of Hebrew, Chinese, Malagasy, Korean, and Turkish (among others) in a concerted effort to be cross-typological. His data is so extensive as a result that he clearly provides support to his theory of a universal hierarchy of adverbs.

Cinque’s first goal is, as mentioned above, perhaps more challenging, particularly to many generativists. He argues that the traditional model of adverbials as adjuncts is essentially anti-minimalist:

A system that countenances both specifiers and adjuncts is clearly less restrictive than a system that does away with one or the other (while still expressing all the correct empirical generalizations). … One desirable direct formal consequence of Kayne’s (1994) restrictive version of X-bar theory – more generally of his antisymmetric view of syntax – is precisely the availability of a single specifier per projection (or adjunct, their difference being in fact neutralized). (Cinque, 1999, p. 44)

He therefore seeks to eliminate one of these positions from the clausal architecture and, in line with Kayne (1994), who argues for asymmetric phrase structure, and with Rizzi (1997, 2004), retains the specifier position rather than the adjunct position.

Cinque (1999, p. 106) initially proposes some 30 levels of functional head/specifier positions (and suggests that there may be as many as 40), ranging from just below the CP to just above the VP—essentially adding a great deal of detail to what he calls the TP (or IP) “zone”. Adverbs are base-generated (or, in current generative terms, Externally Merged) in specifier positions that depend on the semantic features required by the functional heads of certain phrases.
The relative order of these functional phrases corresponds, though not strictly, to relations of semantic *scope* (represented syntactically by asymmetric c-command) among the heads. Apparent similarities among the interpretation of differently placed adverbs can be due to three reasons:

... whenever one AdvP seems to have exactly the same interpretation in two apparently distinct positions, either (1) it occupies the same position and something else has moved around it (as I would claim happens in *John probably has been sick* and *John has probably been sick*), (2) it has moved from one position to the other, retaining the interpretation associated with the position of the trace (as in *How cleverly has John worded the letter?*), or (3) it deceptively has exactly the same interpretation in two positions. (Cinque, 1999, p. 20)

To explain these three possibilities, in possibility (1), the head has moved to another functional position around the adverb, which has remained in place. In possibility (2), limited movement is available to adverbs, primarily *wh*-movement. And finally in (3), Cinque states that there are positions for (particularly) manner adverbs both before and after the VP (1999, p. 21). These are distinct, with *different* kinds of scopal meaning, yet for many adverbs the reading is seemingly *the same* in the two positions when the sentences are simple. Consider these examples from Cinque (1999, p. 19):

16   a. John has answered their questions cleverly.
    b. John cleverly has answered their questions.
    c. John has cleverly answered their questions.
The readings of *cleverly* between (16a) and (16b) are noticeably distinct. The first instance shows a manner reading (*how did John answer their questions?*), while the second is a subject-oriented adverb (*how do we assess John’s intelligence by his answering their questions*?).

The problem of ambiguity arises in (16c), which can have both interpretations. This ambiguity is an argument for two different meanings and two different positions between *John* and *answered*. If an additional auxiliary is inserted, the distinctions are clearer (Cinque, 1999, p. 19):

(17)  
   a. John has been cleverly answering their questions.  
   b. John has cleverly been answering their questions.

Here the ambiguity disappears. The higher position before the verb is quite definitely the subject-oriented position, while the lower position is the manner position (Cinque, 1999, p. 19).

The head positions are filled by elements that can engage in head-to-head movement (i.e. verbs). The specifier positions are fixed, while movement (or Internal Merge) can occur relatively freely among the functional heads. For Alexiadou (1997), whose argument is quite similar to Cinque’s, this distinctiveness for adverbs is the *Adverbial Licensing Principle*:
Adverbial Licensing Principle (Alexiadou, 1997, p. 41)

Adverbs are licensed either as Specifiers of Functional Projections or via incorporation into the verbal head by the relevant (semantic) feature associated with the head.

The “incorporation” that Alexiadou is referring to here is a case where the morphological boundary between the adverb and the verb has been eliminated, and the adverb has fused onto the verb. In a theory like DM (as discussed above), this can be explained as a type of morphological merger (by adjacency).

Below are the various functional projections as postulated by Cinque and the related hierarchical order that they follow:

(19) Adverbial hierarchy (Cinque, 1999, p. 106)

\[
\begin{align*}
\text{Mood}_{\text{speech act}} & > \text{Mood}_{\text{evaluative}} > \text{Mood}_{\text{evidential}} > \text{Mod}_{\text{epistemic}} > \text{T(Past)} > \\
\text{T(Future)} & > \text{Mood}_{\text{irrealis}} > \text{Mod}_{\text{necessity}} > \text{Mod}_{\text{possibility}} > \text{Asp}_{\text{habitual}} > \\
\text{Asp}_{\text{repetitive(I)}} & > \text{Asp}_{\text{frequentative(I)}} > \text{Mod}_{\text{volitional}} > \text{Mod}_{\text{obligation}} > \\
\text{Mod}_{\text{ability/permission}} & > \text{Asp}_{\text{CELERATIVE(I)}} > \text{T(ANTERIOR)} > \text{Asp}_{\text{terminative}} > \\
\text{Asp}_{\text{continuative}} & > \text{Asp}_{\text{perfect}} > \text{Asp}_{\text{retrospective}} > \text{Asp}_{\text{proximative}} > \text{Asp}_{\text{durative}} > \\
\text{Asp}_{\text{generic/progressive}} & > \text{Asp}_{\text{prospective}} > \text{Asp}_{\text{SgCompletive(I)}} > \text{Asp}_{\text{PlCompletive}} > \text{Voice} > \\
& > \text{Asp}_{\text{CELERATIVE(II)}} > \text{Asp}_{\text{repetitive(II)}} > \text{Asp}_{\text{frequentative(II)}} > \text{Asp}_{\text{SgCompletive(II)}}
\end{align*}
\]
As is clear from this model, Cinque takes as his base the various subdivisions already accepted by generative syntacticians for verbal grammar, namely tense, mood, aspect, and voice. What this hierarchy accomplishes, in addition to a fine-grained placement of adverbs within the clause, is also a careful delineation of the various types of verbal expression permitted in the cross-section of languages that he has investigated. Here terminology becomes crucial, and a careful application of semantic notions such as perfect aspect must be undertaken in order to use the model successfully.

In English, speakers do not make use of all of these functional heads in most sentences, but the auxiliary order that is allowed follows this hierarchy directly:

\[
\begin{array}{ccccccc}
\text{He} & \text{may} & \text{have} & \text{been} & \text{being} & \text{seen} \\
\text{Subj} & \text{Mod}_{\text{possibility}} & \text{Asp}_{\text{perfect}} & \text{Asp}_{\text{durative}} & \text{Voice} & \text{Verb} \\
\end{array}
\]

by the witness then

\[
\begin{array}{ccccccc}
\text{PP} & \text{AdvP} \\
\end{array}
\]

In addition to these highly-ordered adverbs, Cinque also briefly addresses VP-internal adverbs and adverbials, which he terms circumstantial. He notes that, in comparison to the fixed order laid out above, the VP-internal adverbials (of place, time, manner, etc.) follow the complement of the verb and are far less rigidly ordered than the pre-VP adverbs. In addition, he indicates that such VP-
internal adverbs cannot appear in pre-VP positions, except for what he calls the “adverb of setting” position (e.g. *Tomorrow, I’m leaving for the coast*), possibly in a CP-internal topic position (pp. 28-9).

One of the major criticisms that has been leveled at Cinque’s hierarchy is its sheer size and the claim that all functional phrases are present in all languages, at all times (1999, p. 107). Many generative syntacticians consider such an excess of detail to be counter to the aims of the Minimalist Program. Since these functional phrases are present, according to Cinque, in all languages and in all clauses, he indicates that most of the time they are “default” rather than “marked”.

Default mode means that, for example, the head known as Mood_{evaluative} would have a default-value spec in *fortunately* rather than *unfortunately*. The default would often, though not always, go unexpressed, with the specifier and head positions left unfilled. In other words, unless otherwise indicated, most sentences would be considered to have a default evaluation of *fortunately*. The marked value, on the other hand, would always be expressed (Cinque, 1999, pp. 128-9). Therefore, the presence of all functional heads in every sentence is explained through a markedness criterion that allows most specifier positions to remain unspelled-out. Although this proposal is found in Cinque (1999), it is not a theme that he takes up in the continued debate over cartography (e.g. Cinque, 2004). The size of the cartographic apparatus thus remains a problem.

Shlonsky (2010) acknowledges this problem as one of the most important facing the cartographic model and its proponents. He goes on to suggest a
different solution from Cinque’s “default vs. marked” one by making use of minimalist terminology regarding the interpretability of features:

Cartographic work is not very explicit as to how the hierarchy of heads is formally guaranteed. If it is derived by precise statements of selection (e.g., Asp27 selects Asp26), then it follows that the entire cartographic hierarchy must be merged in every full clause… Supposing that the full structure is thus present in the clause, some mechanism is needed to distinguish “active” from “inactive” projections, that is, those which receive an interpretation and those which are interpretively null. One way of achieving this is to extend Chomsky’s dichotomy of interpretable and uninterpretable features to this domain. Thus, the entire structure is and must be represented syntactically, but only certain features are activated and feed interface interpretation. (p. 426; italics mine)

Cinque and Rizzi (2008), on the other hand, do not consider cartography and minimalism to be opposite or in direct conflict, but instead to be orthogonal:

An impoverished computational mechanism does not imply the generation of an impoverished structure: a very simple recursive operation can give rise to a very rich and complex structure, as a function of the inventory of elements it operates on, and, first and foremost, of its very recursive nature…. The division of labor here is that Minimalism focuses on the generating devices, and cartography focuses on the fine details of the generated structures, two research topics which can be pursued in parallel in a fully consistent manner, and along lines which can fruitfully interact. (p. 49; italics mine)

Therefore, in positioning their model vis-à-vis the MP, Cinque and Rizzi adopt a complementary stance, where cartography “fills in the details” that are generated by Chomsky’s intentionally impoverished language faculty. Where Chomsky and researchers of minimalism seek simplicity in the mechanism (features + Agree + Merge), if the cartographic position is accepted, such a simplified mechanism results in a highly articulated structure.
4.3.3 Semantic-driven adjunction.

Ernst (2002 et seq.) and others like Haider (2000, 2004) reject the cartographic model and believe that adverbs are truly adjuncts rather than specifiers of functional heads. Semantic properties, such as scope, determine the placement of adverbs in a syntactic structure. Adverbs have *selectional properties* as part of their semantic makeup. They choose a particular type of clausal structure (event, proposition, etc.) based on the properties they innately possess.

In this argument, Ernst is following a certain tradition of generative syntacticians who have broken with the Chomskyan “gospel” of an autonomous syntax that is the only generative device in human language. Among these scholars are names such as Jackendoff and Culicover, two of Chomsky’s former students who have chosen a more semantically-rich model of language that is still generative in nature. Consider this statement from Jackendoff (2002), which harkens back to earlier generative models by separating adverbs from syntactic concerns:

> The use of these expressions [“sentential adverbials”] is governed only by rudimentary syntactic principles. *As long as the semantics is all right*, a phrase of any syntactic category can go in any of the major breakpoints of the sentence: the front, the end, or the break between the subject and the predicate…. *this freedom bespeaks a somewhat protosyntactic phenomenon*. The relation of each phrase to the sentence is determined more or less pragmatically, using the meaning of the noun [in PPs] as a guide. (p. 256; italics mine)

This, in a nutshell, is the core of the argument from this particular branch of generative linguistics. Some elements simply “appear” in the sentence in a “protosyntactic” manner. This removes adverbials from the concern of Merge and
gives their placement over to other, non-generative units, such as semantics or pragmatics. Ernst (2002) follows and further articulates this position.

In opposition to theories like Cinque’s (or, in fact, those of Distributed Morphology), Ernst believes that the semantic information that is fully encoded in the lexical entry of an adverb drives the positioning of the adverb, in a sort of “tucking-in” type of adjunction. This is a type of Chomsky Adjunction: “Any phrase combined with some X node but that is not the sister of X⁰ is defined as adjoined” (Ernst, 2002, p. 24). A purely Kaynean structure restricted to specifier-head-complement (and following the LCA) is not part of Ernst’s clause as a result (2002, pp. 27-9).

Ernst further differs from Cinque in allowing for relatively free adjunction of functional adverbs; such adjunction is not mapped out in advance:

I assume in this book that adjunction is in principle quite free, that is, that there are no major syntactic restrictions on either the category or the level of a phrase to which something adjoins (thus adjunction to arguments and to both X’ and XP levels of structure are possible). (Ernst, 2002, p. 13)

Cinque’s structure could be analyzed as an argument for homophony among various iterations of adverbs, while Ernst’s focus allows for a more polysemous reading of different positions.

Ernst’s (2002) structure requires more work at the interfaces, however, and gives syntax a far more minor role in the placement of adverbs. His structure requires a Fact-Event Object Calculus (FEO Calculus) at Logical Form (the semantic interface, also known as C-I) (pp. 36ff.) and Directionality Principles.
(i.e. headedness parameters) and *Weight theory* at Phonetic Form (also known as SM; see above) to rectify the hierarchy and linearization problems.

The FEO Calculus is a set of rules that are used for building *events* and *propositions*. In such a semantic structure, there is a hierarchy of types that “layer” (Ernst, 2002, pp. 49-50) from a core event to add information that is expressed through syntactic means (e.g. event > state through the addition of progressive aspect in *be reading a book*).

(21) FEO Calculus structure (following Ernst, 2002, p. 50)

![Layers diagram](image)

*Layers* can include adverbials, quantifiers, aspect, modality, etc.

Such a semantic structure then adds syntactic/grammatical information to the clausal structure depending on the needs of the lexical items used in the derivation. When applied to adverbs, for example, certain evaluative adverbs (*strangely, approximately, significantly*) allow manner readings, while others (*luckily, fortunately, amazingly*) do not. Therefore, when these adverbs are used, the speaker must select for formal features that correspond to the semantic features of these adverbs.
The upshot of this theory is that *semantics drives syntax*, which is then merely reflexive and thereby less autonomous for adverbs (as well as for other clausal components). Such a model does, however, preserve the minimalist goal of less structure. This goal is reached by the reduction of the syntax “engine” and a greater focus on the semantics-syntax interface, both at LF and at the lexical level. In addition, PF takes on a greater role in this theoretical structure, as the phonetic interface is responsible for a larger proportion of the linearization process than in the cartographic model.

### 4.3.4 Adverbial syntactic model for this study.

Because this is a study that is focused more fully on syntax and morphology, and because I take a fully autonomous syntactic mechanism as a theoretical given, the cartographic model will be used in this study. This does not mean there are no misgivings associated with the choice of this model. Despite the reassurances offered by Cinque and Rizzi, the complexity of the model does seem to be incompatible with minimalism. At the very least, more integrative work must be done to show how the two areas of inquiry mesh. Nevertheless, my attitude toward the syntax-semantics interface is that, as Shlonsky (2010) states, syntax “imposes the pattern and the seams which delimit meaning and use” (p. 426) on semantics.
4.4 Combining Generative Grammar and Historical Linguistics

4.4.1 Introduction: the history of diachronic syntax.

To many, the goals of diachronic linguistics and generative syntax seem unrelated. One seeks to explain historical change, frequently (as we saw in Chapter 2) through sociological, pragmatic, and semantic shifts. The other seeks to explain the language faculty at a more abstracted level, purposely and expressly divorced from external factors such as society (except for the input of a language-based grammar). Combining these two fields seems to be a stretch, yet there is a history of subfield combining that, according to Lightfoot (1979, pp. 22-3), goes back to the early 1960s (the early history of transformational grammar) and the (unpublished) dissertation of Edward Klima.

Lightfoot himself is widely regarded as a pioneer in combining generative and historical linguistic approaches. Counter to the “gradient and gradual” pace of language change proposed by Traugott and others who work in a more functionalist framework, Lightfoot (1979) argues for a more sudden, “cataclysmic” type of language change. As support for this position, he offers the

*Transparency Principle:*

\[(22) \text{Transparency Principle (Lightfoot, 1979, p. 121)}\]

Derivations are minimally complex and initial, underlying structures are “close” to their respective surface structures.
For him, such a principle (phrased in terms of the Extended Standard Theory, as above) requires under optimal conditions that the results of transformational rules be as close to their deep structure meanings as possible. “Radical reanalysis” occurs when the complexity of the surface structure is increased, requiring the use of exception markers. As those exception markers grow in number, the tension on the system increases until a breaking point is reached. This is when the “cataclysm” occurs and there is a wholesale restructure of the grammatical system, such as the addition of a new category (p. 122).

Although he does not expressly use the term, it is clear that Lightfoot is referring to change such as grammaticalization here. Reanalysis has long been considered an important component to grammaticalization, and Lightfoot styles such behavior as a type of abduction. This follows a syllogistic approach, where reanalysis is the final stage: after induction (All X are Y) and deduction (Z is a member of X) comes abduction (Therefore, Z is Y) (Lightfoot, 1979, p. 347).

Other important scholars that have combined generative and diachronic approaches include Kiparsky (e.g. 1996, in press), Fuß (e.g. 2005), Garrett (in press) and Fischer, who “straddles the fence” (e.g. 2007) between formal and functional approaches. The two studies that have focused most extensively on grammaticalization for the purposes of this dissertation, however, are Roberts and Roussou (2003), and van Gelderen (2004, 2008), who bring minimalist perspectives to their work.
4.4.2 Minimalist approaches to grammaticalization: Roberts and Roussou (2003).

There are two central claims to Roberts and Roussou (2003, p. 2): first, that what is known as grammaticalization is a case of *parameter shift* under a generative model, and secondly, that therefore, grammaticalization itself is not a “real” action but is *epiphenomenal*, in concord with other scholars such as Newmeyer (1998) or Joseph (2001) (See Chapter 2 for further discussion.)

As discussed above, parameters are UG settings that account for much of the variation found cross-linguistically, from a synchronic perspective. Roberts and Roussou assume a *lexicalist* position (i.e. counter to the DM *constructionist* position) and assign parametric settings to the Lexicon (2003, p. 10). Such parameters are learned by speakers through the process of language acquisition, in line with conventional generative theories. Yet the standard paradigm of error-free conveyance of grammar from parent to child is not in line with the realities of language change. As a result, Roberts and Roussou (hence R&R) consider acquisition from a somewhat novel (for generativists) perspective:

… language acquisition is deterministic to the extent that all parameters have to be set. Most of the time, convergence is successful in that children arrive at the same parameter values as their parents. A change occurs when the trigger experience for a parameter setting has become obscure or ambiguous. (2003, p. 12)

The “obscurity” and “ambiguity” in this passage are directly descended from Lightfoot’s (1979) *exception markers*; they are interferences between the underlying (deep structure or I-language) meaning of an expression and its
representation (as surface structure or E-language). The trigger is the input to the language learner’s UG, which in such a case is unclear:

(23) Trigger (R&R, 2003, p. 15)

A substring of the input text S is a trigger for parameter $p_j$ if S expresses $p_j$.

If the expression of the parameter is “robust”, then the parameter passes correctly between generations. If it is “ambiguous” in contrast, there may be parameter resetting (perhaps to a safety or default value) and language change (R&R, 2003, p. 15).

For R&R, the locus of this parameter shift (and the resulting language change) is in the functional categories of human language. Lexical categories by contrast are stable cross-linguistically. In the face of bare syntax and cartographic approaches, R&R adopt a middle ground that allows for a larger number of functional heads but requires that they have “lexico-semantic content” (2003, p. 26). Therefore a Cinque-Rizzi-style cartographic approach is too large (though nevertheless utilized by R&R), but a completely minimalist approach is too scanty.

So in a grammaticalization schema, the language learner takes an existing pattern and applies new parameters to it. For R&R, those parameters are determined by not only their functionality but also by their interface strength (at LF and PF). Where lexical categories are universally interpretable at both
interfaces, functional categories often show deficiencies, particularly at PF. An example would be the category C; in English it may be phonetically expressed (I hope [that] you visit again) or not (I hope [C] you visit again) (R&R, 2003, pp. 27-8).

R&R accept a unidirectionality hypothesis (termed a “path” that follows Cinque’s markers upwards), but more interestingly, they note the shift in freedom of an item as it changes from lexical to functional: namely it loses the ability to Move head-to-head (or Internal Merge) and can only External Merge. As they indicate, such a shift can be systemic and result in fixed word order (e.g. a shift from synthetic to analytic) or can affect only one category (e.g. modals) (R&R, 2003, pp. 71-2).

Finally, the interface deficiencies that arise from the shift from lexical to functional are the cause for associated phenomena that are considered elements of grammaticalization. First, semantic bleaching is an LF deficiency that occurs as a result of the re-parameterization. Content is lost because it is no longer necessary or appropriate (R&R, 2003, pp. 218ff.). Phonetic reduction is another interface condition, as functional elements are necessarily atomic (i.e. not syntactically or morphologically complex) (R&R, 2003, pp. 224ff.).

Roberts and Roussou’s (2003) approach to grammaticalization is carefully argued and convincing from a position of diachronic syntactic study. They utilize a moderately cartographic approach (after backing away from Cinque in the early part of the book, and with the codicil that his paradigm is probably too detailed).
They account for various grammaticalization phenomena through structural analyses (i.e. PF and LF deficiencies) and through constraints on movement.


In van Gelderen (2004), a different approach is used than in Roberts and Roussou (2003). The focus here is more specifically motivating the changes that occur in grammaticalization phenomena, rather than simply offering a mechanism that fits with generative syntax and is based, as are Lightfoot’s models, primarily on chance. Such motivation is to be found in Economy Principles, specifically two principles of Economy of Derivation (used for avoiding costly extra steps in derivations) (van Gelderen, 2004, p. 4). These Economy Principles are found not only in language; corresponding principles are found in all biological systems, thereby simplifying the specifically linguistic UG.

The first one is a principle that states that a minimal projection is preferable to a maximal projection; van Gelderen terms it the Head Preference or Spec to Head Principle:

(24) Head Preference or Spec to Head Principle (van Gelderen, 2004, p. 11)

Be a head, rather than a phrase.
The second principle is similarly brief and to the point; it is called the Late Merge Principle (but see below for a more recent version):

(25) Late Merge Principle (van Gelderen, 2004, p. 12)

Merge as late as possible.

In addition, one of the motivations available for grammaticalization is the language’s selection from a macroparameter, called the Layer Parameter. In this parameter, different languages emphasize different regions or layers (as described above in section 4.2.2). For example, English makes extensive use of the CP/IP (or TP) layer, while Mandarin Chinese makes more use of the VP layer (van Gelderen, 2004, p. 4).

An issue (aside from the Spec to Head Principle, as Cinque’s hierarchy places adverbs only in spec positions) with using van Gelderen’s work to combine Cinque’s cartographic approach to minimalism with Hopper and Traugott’s traditional approach to grammaticalization is that van Gelderen herself is unconvinced by Cinque’s structures and addresses them directly in her volume. This is indeed a concern, but not a completely unapproachable one. Van Gelderen’s concerns (rightly) center themselves on the lack of co-occurrence of higher adverbs like \textit{frankly} and \textit{fortunately} in corpus data (2004, p. 45).

This is true, but Cinque’s structure does not depend very much on corpora but rather on speaker intuition about possible co-occurrences. He arrives at his
ordered structure through transitive processes (i.e. \( Is \ X > Y? \) \( Is \ Y > Z? \) Then the order must be \( X > Y > Z \).) that are not necessarily borne out by data. His detailed model therefore depends far more on theoretical potential than on actual speaker practice, and he seems to avoid the question of extensive supporting intra-language data as a result, relying instead on cross-linguistic structures to support his argument for the universality of functional heads and their corresponding specifier positions for adverbs.

In addition, van Gelderen does allow for a spec-to-spec process of grammaticalization specifically in the case of VP adverbs. She argues that adverbs are merged late in the derivation and can thus be merged in the VP, the TP, or the CP (van Gelderen, 2004, p. 130). In essence, this is not unlike Cinque’s argument for “base generation” of adverbs in various delineated specifier positions.

In more recent work (van Gelderen, 2008), there is a shift away from the Late Merge Principle; Move has been reinterpreted as Internal Merge (see above), and as a result it is no longer considered costly. This change from Move to Merge also affects R&R (2003), as there is no longer a concern about moving from a lower base-generated position to a higher one, and therefore the loss of movement is not economical or advantageous to the system.

Van Gelderen (2008) replaces Late Merge with a Feature Economy Principle. In this principle, the semantic features of lower layers (VP) are replaced by interpretable features in the CP/TP zone and further grammaticalization is stimulated by the (continuing) Spec to Head Principle, with the end result being a C-head (assuming an expanded Rizzian CP) (van Gelderen, 2008, p. 297):
(26) Feature Economy (van Gelderen, 2008, p. 297)

Minimize the semantic and interpretable features in the derivation, e.g.:

<table>
<thead>
<tr>
<th>VP-Adverbial</th>
<th>CP-Adverbial</th>
<th>C-Head</th>
</tr>
</thead>
<tbody>
<tr>
<td>[semantic] &gt;</td>
<td>[uF] &gt;</td>
<td>[iF]</td>
</tr>
</tbody>
</table>

A feature-driven account is in keeping with current minimalism, and therefore I will use both van Gelderen (2004, 2008) and Roberts and Roussou (2003) as models for connecting diachronic adverbial linguistics with generative grammar.

4.5 Generative Grammar and Lexicalization

Lexicalization does not seem to fit very well into any of the abovementioned frameworks. Distributed Morphologists regard lexicalization as running positively counter to the heart of their theory. A process of assembling new lexical items for most non-DM generativists is semantic, rather than syntactic. And finally, diachronic syntacticians can find little in a lexicalization process that meshes with theories dependent on Economy Principles or Triggers.

Yet the reductive power of lexicalization, the combination of items into a semantic and morphosyntactic unit, is not terribly different in some ways from a grammaticalization process that prefers a head to a phrase. There is obviously an economy principle at work, even though it may not be the same type as is found in grammaticalization processes. It should be different, if we are to maintain that...
lexicalization is a different phenomenon. The interface pressure may be PF-centered, rather than LF-centered, for example.

Syntactically, there is likely a savings in the shift from (as I argue) CP to AdvP. Although Merge remains free (both External and Internal versions), it is not clear whether the cognitive aspects of phase cycles are equally wide-open. If a CP reduces to an AdvP, there is no longer a vP or a CP phase, and this savings may indeed be an Economy Principle in a Chomskyan sense.

4.6 Chapter Summary

In this chapter, we examined the various aspects and mechanisms of the generative approach to syntax and morphology. First, we looked at the history of the transformational model and moved on to the modern incarnations of generative grammar, namely Principles and Parameters and the Minimalist Program. In the next section, we examined a useful constructivist model of morphology that fits within minimalism, Distributed Morphology.

After looking at generative grammar in a broad sense, we narrowed our focus to models of adverbial syntax. In those sections we examined the cartographic model, primarily advocated by Rizzi (1997) and Cinque (1999), and the semantic-calculus model of Ernst (2002).

In order to incorporate diachronic concerns (as discussed in Chapter 2), we then turned to minimalist models of grammaticalization in Roberts and Roussou (2003) and van Gelderen (2004, 2008), each of which provided different yet
similar approaches. Finally, we took a very brief look at how the phenomena associated with lexicalization might be treated in a generative manner.

In Chapter 5, we will turn to a specific type of sentence adverb, namely the speech-act adverb, and examine an example (*frankly*) in closer detail.
Chapter 5

SPEECH-ACT ADVERBS

5.1 Introduction

5.1.1 Focus of chapter.

In previous chapters we have set up diachronic, pragmatic, and syntactic models, and have examined the category *adverb* from a morphosyntactic perspective. This chapter narrows the focus to a case study of the category of speech-act adverb, one of the types of sentence adverb more frequently found in modern English. In this chapter, I will examine a prototypical speech-act adverb, *frankly*, from a synchronic and a diachronic perspective.

5.1.2 Chapter organization.

Section 5.2 introduces the semantics of the speech-act adverb, while in section 5.3 I first examine its syntactic categorization. Section 5.4 discusses pragmatic and discourse issues related to the category, while section 5.5 gives a brief introduction to and etymology of the word *frankly*. In section 5.6 I discuss the historical processes of grammaticalization and lexicalization and begin to apply these to *frankly*. Section 5.7 is devoted to an examination of corpus data from both synchronic and diachronic corpora. Section 5.8 involves a discussion of the corpus data, grammaticalization and lexicalization, and current generative syntactic theory, showing how they interact. Finally, section 5.9 is a conclusion and discussion of further research opportunities.
5.2 Semantics of Speech-act Adverbs

Sentence adverbs can be broken down into several categories, based primarily on semantic criteria. Current theories generally reference Jackendoff (1972, pp. 56ff.) as determining the class of predicational adverbs that are thought of as adjoining to a full clause or sentence. Of these, there are a number of sub-classes that vary according to each theory and its nomenclature. Four of the most commonly agreed-upon sub-classes are speech-act, epistemic, evaluative, and subject-oriented adverbs. All of the first three sub-classes are speaker-oriented, but while epistemic and evaluative adverbs indicate the speaker’s set of attitudes or beliefs about the proposition, speech-act adverbs instead turn outward, toward the addressee.

Speech-act adverbs are unusual in that the illocutionary force can change direction based on the type of sentence the speaker uses. In declarative sentences, speech-act adverbs express the speaker’s rhetorical stance to the addressee:

(1) a. Honestly, I like the apartment.
   b. Seriously, John doesn’t eat peanuts.

In each of the sentences in (1), the speaker is explaining by use of the speech-act adverbs (honestly and seriously) the manner in which s/he is speaking—it is a commentary on the act of discourse that the speaker is engaged in.

In interrogative sentences, however, the illocutionary force can be reversed, and the speaker can instead demand the manner of the response:
(2) a. Honestly, do you like the apartment?

b. Seriously, does John not eat peanuts?

Note, however, the ambiguity of the adverbs here. The speaker may be asking an honest or serious question, be expecting an honest or serious response, or both.

Nevertheless, the focus of the speech-act adverb is on some quality of the discourse itself, rather than on the message carried by the main clause (or, in Hopper and Traugott’s (2003, p. 176) terms, the nucleus (cf. Longacre, 1985)) of the sentence. Rather, the speech-act adverb, as all sentence adverbs are, is found in the margins.

Swan (1988) lists several of the more common speech-act adverbs found in Present Day English:

(3) A partial list of modern speech-act adverbs:

bluntly, briefly, broadly, candidly, (more) cheerfully, confidently, crudely, (more) encouragingly, flatly, frankly, generally, honestly, incidentally, literally, metaphorically, personally, poetically, precisely, prosaically, roughly, seriously, simply, sincerely, soberly, strictly, truly, truthfully (p. 64)

Among those listed, however, there is a great deal of variation as to frequency. In addition, many (if not most) of the adverbs listed are not only found as speech-act adverbs but as manner adverbs in general use. Some can also be found as subject-
oriented adverbs. One of the most common (and arguably among the most prototypical) speech-act adverbs in Present Day English is the word *frankly*.

### 5.3 Speech-act Adverbs as a Syntactic Category

Following Ernst (2002), speech-act adverbs are unified in two ways, thus forming a class. First, all speech-act adverbs have meanings that relate to (unspoken) “predicates of communication”, affecting the propositional clause. As shown by examples (1) and (2) above, “any speech act is potentially an act of communication (speaker’s) and invites an act of communication (addressee’s)” (Ernst, 2002, p. 70).

Secondly, all speech-act adverbs function (when in sentence adverb position) as manner adverbs. Indeed, in their VP-position, they are all adverbs of manner. When acting as a sentence adverb, however, they all modify what Ernst refers to as a “covert predicate of expression” (Ernst, 2002, p. 70).

Therefore, there is a mismatch between the syntax of the sentence and the conceptual-intentional (C-I) interface (also known as SEM or Logical Form in previous generative theory). What is spelled out via the articulatory-perceptual (A-P) interface (PHON or Phonetic Form in earlier works) does not accurately reflect the intended meaning of the sentence, although the extended meaning is understood by all participants. Such a mismatch has been noted for some time in syntactic literature, beginning at the latest with the Performative Analysis (Schreiber, 1972). See example (8) below for further discussion.
5.4 Discourse and Pragmatics

5.4.1 The syntax–pragmatics interface.

In syntax, as we discussed in Chapter 4, the clause is generally divided into three layers—vP, TP, and CP. Each of these layers bears responsibility for hosting different elements of clausal structure. The vP (known as the vP-shell or light verb phrase), which also contains the lexical VP or verb phrase, is largely concerned with argument/event structure and/or θ-roles—in other words, issues which arise regarding the transitivity of the main verb and its arguments, which are frequently expressed as DPs (or determiner phrases). In addition, questions of causality are generally resolved in this layer.

The next highest layer is the TP or tense phrase (also known as the IP or inflection phrase). This layer is generally thought of as the site of grammatical information (verbal tense, mood, aspect), as well as the locus of subject–verb agreement in English and other languages.

The uppermost layer, known as the left periphery in Indo-European languages, is the CP or complementizer phrase. As the site of interaction between pragmatics/discourse and syntax, it is among the least understood regions of clausal structure in the syntactic literature, although recent work by Rizzi (1997, 2004), in particular, has helped to articulate this layer. (See Chapter 4 for a more detailed discussion.)

Speech-act adverbs are generally related by syntacticians to the left periphery of the clause. Cinque, for example, places the speech-act functional head at the far left of his hierarchy (1999, p. 106) (see below, section 5.8) and
indicates that this adverb may at times (through movement) be located in the CP-layer instead. Ernst (2002), as well, ranges the level of adjunction for speech-act adverbs higher than that of any other adverb type (p. 114).

5.4.2 Pragmatic and discourse models.

Of all sentence adverbs, speech-act adverbs are most closely related to other elements found in the CP-layer that possess illocutionary force, acting as a bridge from previous discourse or context to the current utterance. Other types of speaker-oriented adverbs (evaluative, evidential, and epistemic) reference the main clause that follows, offering speaker commentary on various aspects of the clause. Speech-act adverbs are instead self-referential, discussing the manner of delivery of the utterance and thereby reinforcing the social positioning (or discursive positioning) of the speaker.

Speech-act adverbs can be used for initiating discourse, serving as a filler or delaying tactic for holding the floor, marking a discourse boundary, and bringing about intimacy between speaker and addressee (Brinton, 1996).

If I use the Politeness Theory model (Brown & Levinson, 1987[1978]), a speech-act adverb such as *frankly* can be interpreted in a variety of ways. It can be used as a type of positive politeness strategy, indicating an in-group membership that allows for “frank” conversation (similar to the “intimacy” of Brinton’s model). Yet it can also be used as a “forewarner” in the sense of Watts (2003, p. 184), indicating the speaker is about to issue a statement that can be interpreted as a negative face threat to the listener.
Similarly, an adverb like *frankly* can be interpreted as a non-factive stance adverbial (Biber, 2004). In establishing a position in the discourse for the speaker, *frankly* indicates a personal sense of openness in speech (and perhaps at a broader level), but it may also indicate a willingness to negatively affect the listener in the name of “frankness”.

### 5.5 The Speech-act Adverb *frankly*

The word *frank*, according to the *Oxford English Dictionary (OED)*, first appeared in the Middle English period (first written circa 1300) as a borrowing from Norman French. The word’s origins, however, are not Latinate but Germanic, derived from the Germanic tribal name *Frank*, the same origin for the word *France*. In the sixth century, Roman Gaul was conquered by the Merovingian Franks. The word in French acquired the meaning *free* because “in Frankish Gaul full freedom was possessed only by those belonging to, or adopted into, the dominant people” (*OED*). In fact, its earliest use in English was often as half of a coordinated phrase, *frank and free* (a phrase also found, for example, in Dutch).

By the sixteenth century at the latest, the meaning of freedom associated with *frank* had begun to shift by metaphorical extension to speech, indicating candidness, outspokenness, or a lack of reserve (*OED*). Also by the sixteenth century, the borrowed word had been reanalyzed as native to English and had followed the typical English pattern of adverb creation in *-ly* (as opposed to typical French adverb creation involving *-ment(e)*). In 1541, *frankelie* appears in
King Henry VIII’s official Royal Acts (OED), suggesting that the word was probably in common use many years beforehand.

According to the OED, the first recorded use of *frankly* as a sentence-initial adverb is quite late, in 1847 in Edward Bulwer-Lytton’s *Lucretia*. As might be expected of an adverb that is becoming associated with the speech act, it is found in dialogue:

(4) “Frankly, if you can like my niece, win her.”

The absence of sentence-initial *frankly* in earlier texts may well be related to the formality of the written register when compared to spoken language (of which we have minimal records) or to the types of texts used by the OED; nevertheless, such usage is comparatively modern.

5.6 How Does *frankly* Become a Sentence Adverb?

5.6.1 Grammaticalization.

As stated previously, one of the best known and most widely studied diachronic phenomena is known as *grammaticalization*. Grammaticalization is a conglomerate of different levels of linguistic change—in phonology, in morphology, in syntax, in semantics, and in pragmatics—that accompany historical reanalysis of a word or group of words. According to Brinton and Traugott (2005), grammaticalization is
… the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become more grammatical by acquiring more grammatical functions and expanding its host classes. (p. 99)

Brinton and Traugott argue that adverbs are the result of grammaticalization, and that -ly (which of course is common in adverb morphology; see Chapter 3) is a grammaticalizing affix.

The adjunctive qualities of adverbs are more versatile than those of adjectives (no matter whether scholars agree with Cinque or Ernst or someone else, adverbs obviously are found in more locales than adjectives); as a result, such versatility is a marker of a more “grammatical” function (Brinton & Traugott, 2005, pp. 132ff.).

In their earlier, seminal work on the subject, Hopper and Traugott (2003) also argue for an adverbial cline of grammaticalization:

(5) clause-internal adverbial > sentence adverbial > discourse particle (p. 37)

At first glance, such a cline of historical change seems plausible for speech-act adverbs. The typical grammaticalization path carries words from lower, lexical positions to higher positions in the tree—the grammatical range of Tense-Mood-Aspect, and the discourse-oriented range of Force-Topic-Focus-Finiteness (if we use Rizzi’s 1997 model of the expanded CP). However, there are doubts about whether sentence adverbs like the speech-act adverb frankly are truly grammaticalized from lower VP-adverbs.
There are several tests to indicate grammaticalization has taken place. Commonly, there are changes to words when they grammaticalize that affect not only the syntax but also the phonological strength of the word and the semantic meaning. Morphology may also change. If we examine *frankly*, however, particularly in comparison to a widely accepted example like *have*, there is little evidence of grammaticalization:

(6)  Grammaticalization criteria: *
have*
    
    Phonological stress reduction: ✓
    
    Morphological change: ✓
    (may include cliticization)
    
    Syntactic reanalysis as grammatical: ✓
    
    Semantic bleaching: ✓

At the phonology–morphology interface, the reduction of *have* when it grammaticalizes is dramatic, and it is often found in cliticized form as 'd (past) and 's or 've (present) when acting as a marker of perfect aspect. In North American English, such cliticization does not happen when the lexical version of *have* is used. Syntactically, perfect *have* is reanalyzed from the VP to the AspP. It takes on a grammatical function and no longer has any relevance with regard to the lexical position of the original verb. And finally, *have* undergoes significant semantic bleaching when it is reanalyzed. There is little, if any, trace of the original meaning of possession in the aspect marker.
It is apparent that none of these phenomena occur with the analysis of *frankly* as a sentence adverb rather than a VP-adverb. Phonologically, the use of *frankly* at the beginning of the sentence requires comma intonation, or emphasis rather than reduction. As a result, there is no cliticization of the sentence adverb form. There is syntactic reanalysis, but whether we follow Cinque’s structure or Ernst’s, the shift is not from lexical to grammatical. In Cinque’s model there are specifier positions scattered throughout the tree, adjoined to various functional heads. In Ernst’s model, the speaker uses a sort of lexical calculus for adjoining the adverb to one of many possible positions. (See Chapter 4 for a more detailed discussion.) In neither model do we have a lexical/grammatical contrast.

And perhaps most strikingly, in shifting from lower position to sentence-initial position, there is no semantic bleaching. In fact, we can say that *frankly* has been semantically *enriched* in its uppermost position. Since at least Schreiber (1972), a theory known as the Performative Analysis has assumed a reduced clause in the Lexical Form of “style disjuncts” like *frankly*. An example of the syntactic and semantic forms is below:

(7) The Performative Analysis (examples from Schreiber, 1972, p. 321)

a. *syntax:* Frankly, Merlin is a genius.

b. *semantics:* I tell you frankly that Merlin is a genius.
If we accept this analysis, it means that the sentence-initial use of *frankly* (marked in Ernst, 2002 as “predicational”) encompasses an entire clause, justifying its comma intonation in Phonetic Form.

The problem here is clear: We should ask just how many different types of “grammaticalization” there are, and how broad should the definition of this term be. If we follow the expected pathways and types of change at the major linguistic levels, from phonology to semantics, then the change from verb-of-possession *have* to perfect aspect *have* fits the pattern, while the change from VP-adverb *frankly* to sentence adverb *frankly* does not.

### 5.6.2 Lexicalization.

So if grammaticalization phenomena provide a poor answer for the shift of the adverb *frankly* from VP-adverb to sentence adverb, is there a better answer? Although less well understood, *lexicalization* seems to point in a direction that is more suited to the unique qualities of sentence-initial *frankly*. As a reminder from the discussion in Chapter 4, according to Brinton and Traugott (2005), lexicalization is

> The change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and syntactic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the word may become more lexical. (p. 96)

Therefore, lexicalization involves a group of phenomena that results in word creation. Unlike grammaticalization, where a word changes from being a member
of an open lexical category to a member of (or even the first member of) a closed grammatical category, lexicalization creates new members of open categories like Noun and Verb (or, I argue, Adjective and Adverb) from larger structures.

Phrases or even clauses can be the origins for lexicalized words, which are reduced through the process of de-syntagmatization (or removal of the syntactic structure) (Brinton & Traugott, 2005, p. 48). A typical example of lexicalization is the flower known as the “forget-me-not”. Here an entire predicate VP has been de-syntagmatized and the orthography marks the fusion of the elements (Verb, Direct Object, Negation) through hyphenation (Brinton & Traugott, 2005, p. 49).

Other examples of lexicalization follow normal word formation patterns, including blending, coinages, acronyms, compounding, and affixing. There is often morphological reduction (fusion) and phonological reduction (coalescence) as a result of the process (Brinton & Traugott, 2005, p. 47).

Under the Performative Analysis, the adverb frankly possesses a full propositional meaning distinct from the clause that follows. As it is a speech-act adverb, it comments on the manner of speaking the utterance that follows (as seen in example (7) above). Therefore, it seems like the semantic qualities of lexicalization fit this example well.

But what of the word-formation rules? There is no fusion of elements in the change from “I tell you frankly” to “Frankly” in example (7). What there is instead is the remains of a full clause—the remnant disjunct phrase. Brinton and Traugott address this, however. They give examples of ellipsis and clipping, distinguishing the one from the other by noting that “clipping concerns the
deletion of one or more syllables from multisyllabic words, whereas ellipsis leads to the formal reduction of a complex word or phrase. In ellipsis the *semantics of the omitted element is absorbed into the remainder by metonymy*” (2005, p. 40) (italics mine). Since *frankly* is so closely associated with the act of speech, it metonymically contains the omitted interlocutors (the first- and second-person deictics) and the speech-act verb.

(8) Examples of lexicalization through ellipsis

(from Brinton & Traugott, 2005, p. 41)

a. *canary* < bird from the Canary Islands
b. *damask* < *Damask cloth* (cloth from Damascus)
c. *denim* < Fr. *serge de Nîmes*
d. *jeans* < *Gene (= Genoa) cloth*

The force of these examples is somewhat weakened by both their concentration on nouns and on locations as source material for the metonymy. However, an important quality shown by these examples is the *salience* in each of them. Lost are the redundant words like “bird” or “cloth”, which are self-evident. Significant are the modifying phrases (“from the Canary Islands”, “from Genoa”), which bring new, identifying information. Similarly, in the reduction of “I tell you frankly” to “frankly”, the identification of the speaker (first person) and the addressee (second person) as well as the act of speaking (“tell”) are all
redundant semantically. Only the manner of the telling (“frankly”) is salient and significant, and that is the only surviving element of the clause (in the fully reduced version).

The syntactic study of ellipsis is extensive and far beyond my aim with this work. Suffice it to say, however, that the admission of phrasal ellipsis as a means of lexicalization allows for the possibility, even the likelihood, that speech-act adverbs arise through the lexicalization of entire clauses.

5.7 Corpora and Layering of Usages

When examining corpora for evidence of the use of *frankly* in English, there are both diachronic and synchronic corpora to choose from. Historical linguists generally prize diachronic evidence over synchronic, as synchronic evidence of change, particularly older change, can be sketchy and limited to highly specific, fossilized forms.

Such fossilized forms are examples of *layering*, a phenomenon that frequently results from the processes that make up lexicalization or grammaticalization. Layering has been defined as “the synchronous result of successive grammaticalization of forms which contribute to the same domain” (Hopper & Traugott, 2003, p. 125). Elsewhere, Hopper and Traugott more simply indicate that layering can be found when “older and newer forms co-exist for individual speakers as well as for communities over time” (p. 49).

Synchronic layering, however, need not be considered secondary evidence. One factor in the usefulness of layering as evidence of change is
relative age; the more recent the beginning of the change, the more likely it is that older forms will be evident in current synchronic corpora. With greater age, in fact, older forms are often relegated to the highly specific and/or fossilized realms reserved, for example, for the verb *to wend*, whose use in modern English is essentially restricted to suppletive use in the verbal paradigm of *to go* (*went*) and in collocations with a genitive nominal and the noun *way* (*she wended her way home*).

On the other hand, layering can provide ample evidence of different stages of a change, as those stages are still represented in the current version of the language. The differences among the longevities of older examples of grammaticalized (and lexicalized) forms are highly idiosyncratic. Therefore it is often most useful to use both diachronic and synchronic corpora where useful.

### 5.7.1 Older diachronic corpora.

#### 5.7.1.1 Early English corpora.

Several diachronic corpora have been examined in the search for examples of *frankly*. Since the word first appears in written texts in the first half of the sixteenth century (1540, according to the *OED*), the earliest corpora are largely devoid of examples of the word. What little evidence these corpora give, however, points to the early association of *frankly* with verbs of speech, an important metaphorical shift that prefigures its later use as a speech-act adverb. See examples (9) through (11) below for the limited attestation in some of the more widely-known early English corpora.
(9) Corpus of Early English Correspondence Sampler (1418-1680, 450,000 words):

4 instances; 1 of these associated with verbs of speech

I have millions of thankes to rendar you, that so frankly told to Cary suche offers…

(Elizabeth I to James of Scotland, May 15, 1588)

(10) Michigan Early Modern English Materials (roughly 1480-1702; based on the entries for 36,000 modal verb entries for the Early Modern English Dictionary):

1 instance; associated with a verb of speech

… that here he sayd that he wold more frankly opyn this matter to me then he did…

(Wyat, Letters, 1537)

(11) Shakespeare, First Folio (1589-1613; pub. 1623):

10 instances; 3 associated with verbs of speech, 2 associated with psych verbs

Note that I read the verb in this instance as opine, as this is a possible spelling. The verb, however, may be open, in which case its verb of speech status is less clear (pointed out by Dr. Karen Adams).
But I haue spoke with one that saw him die: Who did report, that very
frankly hee Confess’d his Treasons, …

(Shakespeare, *Macbeth*)

I doe beseech your Grace, for charity If euer any malice in your heart
Were hid against me, now to forgiue me frankly.

(Shakespeare, *Henry VIII*)

Although these corpora show very few examples, importantly they do indicate the
early shift of meaning to speech-related verbs or to so-called “psych” verbs like
*forgive* (which may have a speech act reading) in Shakespeare’s *Henry VIII*.

In Wyat’s letter (in 10), we have an instance of the use of *frankly* from
1537, fully three years before the first OED entry for the word. Significantly, that
first chronological use may demonstrate the association of *frankly* with verbs of
speech, as Wyat apparently uses the adverb with the archaic verb *opine*.

5.7.1.2 Davies-BYU OED corpus.

Mark Davies of Brigham Young University has developed several corpora
with a shared interface that is highly user-friendly and effective in finding
collocations, syntactic relationships, and other data besides simply lexical
searches. In addition to a Spanish and a Portuguese corpus, he has several English
corpora including the *TIME* Corpus of American English, which is diachronic,
containing all issues of *TIME* magazine (covering 1923–2006 and consisting of
100 million words); the Corpus of Contemporary American English (400 million words); access to the British National Corpus (100 million words); and the recently-completed Corpus of Historical American English (410 million words; online since summer 2010). From 2005 until October 2009, users also had access to the corpus of references used by the *Oxford English Dictionary* (*OED*), until Davies was compelled to remove this corpus under threat of legal action by Oxford University Press.

By using Davies’s online interface and search engine for the *OED* before its removal, I was able to look at the differences in the use of *frankly* through the centuries in such a way as I was unable to do with the other historical corpora. I could therefore mark the development of the semantic change in the word that likely led to its reanalysis and use in a lexicalized fashion.

*Frankly* can be found in *OED* sentences 153 times (out of a 37 million-word corpus). This amounts to a token frequency of 4.14 instances per million. From the sixteenth through the twentieth centuries, the use of the word increases from 3 to 88 tokens (actually 85 unique uses, with 3 duplications). The association of the adverb with verbs of speech (and later with speech-act, usually sentence-initial position) increases throughout.

Table 5.1

<table>
<thead>
<tr>
<th>(from Davies, OED interface)</th>
<th>1500s</th>
<th>1600s</th>
<th>1700s</th>
<th>1800s</th>
<th>1900s</th>
</tr>
</thead>
<tbody>
<tr>
<td>total</td>
<td>3</td>
<td>13</td>
<td>7</td>
<td>40</td>
<td>85</td>
</tr>
<tr>
<td>V-speech</td>
<td>1</td>
<td>3</td>
<td>6</td>
<td>19</td>
<td>31</td>
</tr>
<tr>
<td>speech-act</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2</td>
<td>19</td>
</tr>
</tbody>
</table>
The *OED* corpus shows significant evidence of layering. By looking at table 5.1 above, it can be seen that the combination of instances of *frankly* either associated with verbs of speech (as in 12 below) or acting as predicational speech-act adverbs (as in 13) increases dramatically through the centuries, even with a comparatively small sample size. There is a shift from a more general manner meaning to manner of speech and, eventually (particularly in the twentieth century), to a predicational, sentence-initial usage.

(12) An instance of *frankly* modifying a verb of speech (from the *OED* corpus):

Venture to own frankly that you came to Cambridge to learn what you can.

(William Pitt, Earl of Chatham, Letters, 1754)

(13) An early instance of *frankly* as a speech-act adverb (from the *OED* corpus):

SHE: Can you wonder that I’m disinclined for amusement?

HE: Frankly, I do.

(Kipling, *Under the Deodars*, 1888)

---

9 Here I read *own* as “admit”.

168
5.7.2 Recent diachronic and synchronic corpora.

5.7.2.1 Davies-BYU TIME Corpus of American English.

In the significantly larger TIME magazine corpus, there is a far larger number of instances of *frankly*: 1,850. However, the ratio is quite similar to that found in the *OED* corpus at 4.63 instances per million words. In the earlier decades there is the usual mix of non-speech and speech related uses; by the last partial decade (the 2000s), 66 of 81 instances (81.5%) of *frankly* are predicational speech-act adverbs. Of the 14 others, only 5 (6.2%) do not modify a verb of speech. Table 5.4 shows the shift in usage of *frankly* from TIME’s beginnings in the 1920s to the beginning of the twenty-first century:

Table 5.2
Instances of *frankly* in the TIME Corpus of American English (Davies)

<table>
<thead>
<tr>
<th></th>
<th>Speech-act</th>
<th>V-speech</th>
<th>Other 10</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1920s</td>
<td>14 (10.5%)</td>
<td>56 (42.1%)</td>
<td>63 (47.4%)</td>
<td>133</td>
</tr>
<tr>
<td>1930s</td>
<td>30 (9.8%)</td>
<td>118 (38.4%)</td>
<td>159 (51.8%)</td>
<td>307</td>
</tr>
<tr>
<td>1940s</td>
<td>46 (13.6%)</td>
<td>162 (47.9%)</td>
<td>130 (38.5%)</td>
<td>338</td>
</tr>
<tr>
<td>1950s</td>
<td>70 (18.7%)</td>
<td>162 (43.3%)</td>
<td>142 (38.0%)</td>
<td>374</td>
</tr>
<tr>
<td>1960s</td>
<td>55 (20.2%)</td>
<td>98 (36.0%)</td>
<td>119 (43.8%)</td>
<td>272</td>
</tr>
<tr>
<td>1970s</td>
<td>46 (33.6%)</td>
<td>48 (35.0%)</td>
<td>43 (31.4%)</td>
<td>137</td>
</tr>
<tr>
<td>1980s</td>
<td>43 (47.2%)</td>
<td>26 (28.6%)</td>
<td>22 (24.2%)</td>
<td>91</td>
</tr>
<tr>
<td>1990s</td>
<td>79 (71.2%)</td>
<td>18 (16.2%)</td>
<td>14 (12.6%)</td>
<td>111</td>
</tr>
<tr>
<td>2000s</td>
<td>66 (81.5%)</td>
<td>10 (12.3%)</td>
<td>5 (6.2%)</td>
<td>81</td>
</tr>
</tbody>
</table>

Although this data is limited by its single source (bringing considerations such as authorial and/or house style into question), there is no doubt in the shift during the twentieth century (and through the beginning of the twenty-first) of the use of

10 “Other” indicates use of *frankly* as either a manner adverb for non-speech verbs or, more frequently, as an intensifier modifying an AdjP or a DP.
frankly. In the 1920s and 1930s, roughly half of the tokens found in the corpus involve use of frankly as an intensifier:

(14) Typical intensifier use of frankly in the TIME Corpus of American English

a. The plot is a frankly threadbare clothesline on which to pin the songs and dances.

\[(TIME, November 12, 1923)\]

b. Frankly a sentimentalist, one of his outstanding deeds in office was sponsoring…

\[(TIME, February 13, 1931)\]

In these cases, frankly is roughly synonymous to openly or unabashedly and is often used to modify AdjPs or DPs that address controversial topics. In the TIME corpus, there are several adjectives, for example, that frankly collocates with (100 total). Below are the most common of these:

(15) Adjectives frequently modified by frankly in the TIME corpus (5+ tokens)

skeptical, political, commercial
puzzled, worried, disappointed
partisan, sexual, sentimental
By the 1940s through the 1960s, however, such usage of *frankly* begins to drop, and it hovers around 40% during this 30-year period; it approximates 30% by the 1970s. The change then accelerates and becomes more dramatic, from the 1980s, when intensifier usage averages approximately 25% (manner use in non-speech verbs has become unknown by this time), to the 2000s, when it averages about 6%.

As the use of *frankly* as an intensifier weakens, its use first as a modifier of verbs of speech and finally as a sentence adverb both change significantly. Throughout the twentieth century, verb of speech modifier and “other” usage (mainly intensifier) remain close in number. In the 1940s, however, the most frequent use of the adverb is as a modifier of verbs of speech:

(16) He speaks his mind frankly and clearly on matters of public interest and importance.

(*TIME*, June 17, 1943)

Nearly half the usage in the corpus modifies such verbs. However, this trend begins to recede in the face of rapid growth in the use of *frankly* as a sentence (speech-act) adverb in the magazine. At first limited to merely 10% of usage, by the 1960s speech-act tokens have doubled in frequency. Over the next 36 years, the speech-act usage comes to dominate the token frequency in the corpus, with a percentage of more than 80% by the 2000s.
Interesting to note at the same time, however, is the decline of total frequency of the word *frankly* after the 1950s. From a high point of 372 total tokens during that decade, the frequency drops to 111 tokens during the last full decade, the 1990s. It is not clear why this has occurred. It may be tied to the decrease in space devoted to news articles in magazines such as *TIME* after the 1950s (as advertising increased in an attempt to increase revenues).

### 5.7.2.2 Davies-BYU Corpus of Contemporary American English.

A third corpus from Davies’s website, the 400-million word Corpus of Contemporary American English (COCA), covering 1990-2009 (now 425 million words covering 1990-2011), offers further evidence of the shift in use of *frankly*. As stated earlier in this section, where diachronic corpora can more directly trace the paths of change, synchronic corpora are nevertheless valuable because they can demonstrate the presence of older forms side by side with modern usages of a word.

The COCA is split into five separate text types, each roughly the same size: transcripts of speech from television, works of fiction, magazines, newspapers, and academic texts. As of 2009, there are 8,638 total tokens of *frankly* in its database, for a frequency of 21.6 instances per million words. This is far greater than the overall frequencies for most of the diachronic corpora, indicating perhaps increased usage in the modern age.

However, these tokens are not split evenly by text type. By far the highest frequency of *frankly* appears in the spoken language sub-corpus, with 5,534
tokens and a token frequency of 67.74 instances per million words. The lowest frequency is in the academic sub-corpus, with 343 tokens and a frequency of 4.33 instances per million. This split is appropriate to differences in usage; as we saw in the TIME corpus, modern usage of this adverb is highly correlated with the act of speaking.

Based on the trends observed in the TIME corpus, similar results are expected for the COCA. Overall, this turns out to be true. Davies’ website interface with his corpora allows the user to check collocations by part of speech and even by punctuation. Examining the COCA for the structure “frankly + ADJ” results in 179 of 8,638 instances, or 2.1%, intensifier usage. This is significantly lower even than the results from the last (partial) decade of the TIME corpus.

Similarly, there are 1,130 tokens where frankly appears either directly before or following a verb. This amounts to 13.1% of its total frequency in the COCA; however, there are no examples of manner modification outside of the semantic category of verbs of speech. The only verbs in the immediate environment of frankly are speech verbs, auxiliaries, and copula verbs.

One way to think of the process of lexicalization through ellipsis is to consider the possible steps that may have led to the use of frankly as a speech-act adverb, from its origin in the VP of a separate sentence to its final position(s). These steps follow the cline of parataxis > hypotaxis, widely thought to be the diachronic pattern of clausal structure relations found in English and other languages (see Chapter 2 for further discussion). The COCA offers evidence of such a shift in its various layers:
(17) Possible steps in the lexicalization process of *frankly*

(all from the Corpus of Contemporary American English)


b. I must speak frankly now and tell you that I see no reason for you…

c. If I might speak frankly, there’s others [sic] still aboard more sickish…

d. To speak frankly, Barbara, do you think that this can be accepted…

e. Frankly speaking, if you’re looking for good nutrition in a hot dog…

f. Frankly, in the spirit of free speech, that’s good.

Represented in example (17) by these entries from the corpus are some steps that may have been taken by speakers as they lexicalized *frankly* and thereby gave the adverb propositional meaning. Example (17a) shows a type of politeness strategy, where the first speaker asks permission from the addressee, waiting a turn before delivering a message thought to be both honest and potentially disfavored.

In example (17b) the permission-seeking is shifted into a coordinated clause structure. Here the speaker is likely concerned about ceding the floor when asking permission to speak frankly. To avoid losing the floor, the speaker
continues to use a modal to maintain a level of politeness through irrealis but changes the interrogative structure to declarative. The speech-act clause is conjoined to the main proposition by means of the coordinating conjunction and as well as through use of the parallel verbs speak and tell; the main proposition of the sentence is subordinated as a that-clause acting as the direct object under tell.

By sentence (17c), the main proposition of the utterance is now the matrix clause of the sentence, and the frankly clause has become a subordinated adverbial, although the (modal) verb is still finite.

In sentences (17d) and (17e), the subject is lost and the verb has become non-finite. By (17e), “frankly speaking” (which in the COCA outnumbers “speaking frankly” in this position by a ratio of 23:1) has moved the adverb to initial position, so that the loss of the verb in (17f) is unsurprising. In this final form the adverb has thus become propositional. The ellipsis of the other lexical items in the clause has shifted the meanings of those items to a covert level, understood by speaker and audience alike (i.e. available in Logical Form), but no longer spelled out (in Phonetic Form).

Each of these steps involves a fairly significant syntactic change. I will examine possible explanations for these changes in the case of frankly and other speech-act adverbs through the framework of generative syntax in section 5.8.

5.7.2.3 Davies-BYU Corpus of Historical American English.

Corpus-based support for the proposed ellipsis pattern from (17) is found in another of Mark Davies’s corpora, in this case the Corpus of Historical
American English (COHA), which Davies developed from a variety of text types (fiction, magazine, newspaper, other non-fiction). Like the COCA, this is a very large corpus, with 400 million words spread across a 200-year span of American English (1810-2009). And like Davies’s other corpora, the COHA has a user-friendly interface that allows for a wide range of search types.

Interestingly, like the COCA (and unlike the TIME and OED corpora), *frankly* has a fairly high occurrence rate in the COHA (7,333 tokens total), with an average of 18.33 tokens per million words. This rate is highly skewed toward the twentieth century (unsurprisingly considering the evidence from the TIME corpus); use of *frankly* in the nineteenth century is fairly rare until the 1860s.

Table 5.3 below traces the use of *frankly* through the period covered by the COHA.

<table>
<thead>
<tr>
<th>Decade</th>
<th>Total</th>
<th>2-Clause</th>
<th>2nd Per.</th>
<th>Imper.</th>
<th>1st Main</th>
<th>Infinitive</th>
<th>-ing</th>
<th>S-Adv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1810s</td>
<td>8</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>1820s</td>
<td>57</td>
<td>–</td>
<td>1</td>
<td>9</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>–</td>
</tr>
<tr>
<td>1830s</td>
<td>141</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>21</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1840s</td>
<td>197</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>23</td>
<td>10</td>
<td>3</td>
<td>–</td>
</tr>
<tr>
<td>1850s</td>
<td>182</td>
<td>6</td>
<td>4</td>
<td>3</td>
<td>23</td>
<td>–</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1860s</td>
<td>252</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>20</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1870s</td>
<td>348</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>32</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>1880s</td>
<td>370</td>
<td>12</td>
<td>6</td>
<td>3</td>
<td>33</td>
<td>3</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>1890s</td>
<td>339</td>
<td>8</td>
<td>3</td>
<td>4</td>
<td>28</td>
<td>–</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>Decade</td>
<td>Total</td>
<td>2-Clause</td>
<td>2nd Per.</td>
<td>Imper.</td>
<td>1st Main</td>
<td>Infinitive</td>
<td>-ing</td>
<td>S-Adv.</td>
</tr>
<tr>
<td>--------</td>
<td>-------</td>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>----------</td>
<td>------------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>1900s</td>
<td>568</td>
<td>3</td>
<td>5</td>
<td>–</td>
<td>24</td>
<td>4</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>1910s</td>
<td>731</td>
<td>3</td>
<td>5</td>
<td>6</td>
<td>34</td>
<td>1</td>
<td>3</td>
<td>41</td>
</tr>
<tr>
<td>1920s</td>
<td>711</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>20</td>
<td>5</td>
<td>–</td>
<td>63</td>
</tr>
<tr>
<td>1930s</td>
<td>640</td>
<td>11</td>
<td>2</td>
<td>–</td>
<td>21</td>
<td>5</td>
<td>4</td>
<td>83</td>
</tr>
<tr>
<td>1940s</td>
<td>474</td>
<td>4</td>
<td>3</td>
<td>–</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>111</td>
</tr>
<tr>
<td>1950s</td>
<td>535</td>
<td>6</td>
<td>3</td>
<td>1</td>
<td>15</td>
<td>3</td>
<td>1</td>
<td>190</td>
</tr>
<tr>
<td>1960s</td>
<td>409</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>15</td>
<td>3</td>
<td>–</td>
<td>173</td>
</tr>
<tr>
<td>1970s</td>
<td>364</td>
<td>3</td>
<td>4</td>
<td>–</td>
<td>9</td>
<td>1</td>
<td>–</td>
<td>185</td>
</tr>
<tr>
<td>1980s</td>
<td>376</td>
<td>10</td>
<td>1</td>
<td>–</td>
<td>12</td>
<td>2</td>
<td>2</td>
<td>231</td>
</tr>
<tr>
<td>1990s</td>
<td>291</td>
<td>1</td>
<td>–</td>
<td>–</td>
<td>4</td>
<td>1</td>
<td>–</td>
<td>200</td>
</tr>
<tr>
<td>2000s</td>
<td>340</td>
<td>3</td>
<td>–</td>
<td>–</td>
<td>1</td>
<td>1</td>
<td>–</td>
<td>233</td>
</tr>
</tbody>
</table>

The explanation for the column types is as follows. The first column lists the date range by decade, and the second lists the total tokens of *frankly* for that decade. The other columns separate the tokens that are significant for this study (excluding intensifier use or non-speech-act use).

The “2-Clause” column refers to structures, usually in reported dialogue, that correspond to (17a)—namely, those instances where there are multiple turns. The “2nd Per.” column refers to structures not directly discussed in (17), second-person imperative sentences that are speech-act-related (e.g. “Tell me frankly, what do you think?”). “Imper.” refers to impersonal speech-act structures (e.g. “It may be frankly admitted that...” from the 1840s). “1st Main” references the type of structures found in (17b,c), where *frankly* appears in the first-person main clause of the sentence, frequently in forms that include modals (“I will/must tell
you frankly that…” or lexical verbs of compulsion/confession (“I frankly admit/confess/acknowledge/own…”). It also includes the far less frequent “if” clauses (“If I may speak frankly,…”).

The last three columns (“Infinitive”, “-ing”, “S-Adv.”) are relatively self-explanatory and refer to (17d,e,f) respectively. The infinitive and -ing forms are found in quite small numbers, while the sentence adverb function grows significantly through the decades, so that by the 1990s, it alone accounts for two-thirds of the tokens.

The numbers show, aside from a strong increase in the sentence adverb use, a decline in clausal usage that corresponds with the layered synchronic data from (17). The numbers of non-finite clauses are consistently quite small, and there is no step-by-step progression through the levels postulated in (17). Yet the overall picture fits the proposed directional model quite well.

5.7.2.4 Davies-BYU interface for the British National Corpus.

Like the COCA, the British National Corpus (BNC) is a synchronic corpus, in this case of British English. It was compiled between the 1970s and 1993 of contemporary texts and consists of 100 million words. The number of tokens in the BNC is not as large as in the COCA, either in total number or in percentage (931, for a frequency of 9.31 instances per million words). This may be dialectal in origin (i.e. British vs. American English) or may relate to the earlier time frame of the data gathered.
In the BNC, there is a high token frequency for spoken texts (16.06 per million), although it is very far below the frequency for speech in the COCA (67.74 per million). The highest frequency is in fictional texts instead (18.61 per million, compared to 12.46 per million for the COCA). In other ways, however, it is very much like other more modern corpora. Intensifier use is quite small while speech-act and modifier of verb of speech usages predominate.

5.8 A Generative Analysis of Lexicalization

As mentioned before (particularly in Chapter 3), adverbs in English are considered by many to be quite messy as a category. They seem to be able to be adjoined almost anywhere, from the right to the left peripheries of the sentence. Semantically, they are a somewhat loose grouping of manner, time, and place, among VP adverbs, and subject-oriented, speech-act, epistemic, evidential, and evaluative, among sentence adverbs.

Rather heated debates exist, just within the generative community, as to whether adverbs are the expression of rigidly ordered functional heads or whether they are directed to the VP or left periphery by a semantic calculus that determines scope conditions (see Chapter 4). The very concept of “scope” is often debated, as is the gradient of semantic variation that occurs for many adverbs as they are adjoined in different positions.

As a reminder, in this study I am working within a cartographic framework, which is situated (at times uneasily) within the confines of generative syntax and, more particularly, the Minimalist Program (Chomsky, 1995, et seq.).
5.8.1 Cinque and the cartographic approach.

The basics of the cartographic approach are surprisingly straightforward. Cinque (1999) proposes some 32 levels of functional head/specifier positions (and suggests that there may be many more), ranging from within the CP to just above the Verb Phrase (or VP)—essentially adding a great deal of detail to what he calls the TP (or IP) “zone” (p. 106).

Adverbs are base-generated (or, in current generative terms, externally merged) in specifier positions (as we saw in Chapter 4) that depend on the semantic features required by the functional heads of certain phrases. The relative order of these functional phrases corresponds, though not strictly, to relations of semantic scope (represented syntactically by asymmetric c-command) among the heads. The head positions are then filled by verbs, verb phrases, or possibly by entire clauses. The specifier positions are fixed (outside of topicalization), while movement (or Internal Merge) can occur from head to head.

Below are the various functional projections (repeated from earlier) as postulated by Cinque and the related hierarchical order that they follow:

(18) Cinque’s hierarchy of adverbial structure

\[
\begin{align*}
\text{Mood}_{\text{speech act}} & > \text{Mood}_{\text{evaluative}} & > \text{Mood}_{\text{evidential}} & > \text{Mod}_{\text{epistemic}} & > \text{T(Past)} & > \\
\text{T(Future)} & > \text{Mood}_{\text{irrealis}} & > \text{Mod}_{\text{necessity}} & > \text{Mod}_{\text{possibility}} & > \text{Asp}_{\text{habitual}} & > \\
\text{Asp}_{\text{repetitive(I)}} & > \text{Asp}_{\text{frequentative(I)}} & > \text{Mod}_{\text{volitional}} & > \text{Mod}_{\text{obligation}} & > \\
\text{Mod}_{\text{ability/permission}} & > \text{Asp}_{\text{celerative(I)}} & > \text{T(Anterior)} & > \text{Asp}_{\text{terminative}} & > \\
\text{Asp}_{\text{continuative}} & > \text{Asp}_{\text{perfect}} & > \text{Asp}_{\text{retrospective}} & > \text{Asp}_{\text{proximate}} & > \text{Asp}_{\text{durative}} & > 
\end{align*}
\]
Asp\textsubscript{generic/progressive} > Asp\textsubscript{prospective} > Asp\textsubscript{SgCompleitive(I)} > Asp\textsubscript{PtCompleitive} > Voice

> Asp\textsubscript{celerative(II)} > Asp\textsubscript{repetitive(II)} > Asp\textsubscript{frequentative(II)} > Asp\textsubscript{SgCompleitive(II)} (p. 106)

The speech-act adverb (in boldface) sits at the top of the hierarchy and takes scope over all other adverbs, as well as any other element in the clause.

5.8.2 Application of cartographic and generative principles to lexicalization phenomena.

Cinque’s cartographic structures may be somewhat controversial among generative and more functionalist linguists alike, but they provide useful scaffolding on which to demonstrate historical processes (or collections of processes) like grammaticalization and lexicalization. In the case of speech-act adverbs lexicalizing from full clause structures to the heads of AdvPs in the canonical position indicated in example (17) above, Cinque’s hierarchy provides a landing place for the newly imported adverb to rest.

How do we explain the “import” process, though? Two of the more influential works that discuss historical language change in terms of generative syntax are, as discussed in Chapter 4, Roberts and Roussou (2003) and van Gelderen (2004). If we adopt a Cinquean hierarchy, however, it seems largely incompatible with Roberts and Roussou (by their own admission). They describe language change in terms of parameter setting, which is appropriate for grammaticalization processes where words (and semantic categories) shift from lexical to functional type:

181
…[W]e assume that parameter change is an aspect of the process of parameter setting. A change is initiated when (a population of) learners converge on a grammatical system which differs in at least one parameter value from the system internalized by the speakers whose linguistic behaviour provides the input to the learners. As the younger generation replaces the older one, the change is carried through the speech community. (2003, p. 11)

Cinque’s hierarchy, however, is continually available to speakers. He addresses the question of parametric variation among functional heads:

How much variation should UG allow in the number and type of functional projections available to different languages? The strongest position would be that UG allows no variation at all…. The order appears cross-linguistically invariant. A related question is whether we should take the entire array of functional projections to be present in every sentence. I will suggest that this is the least costly assumption, once we recognize that each head comes with a marked and a default value. (1999, p. 127)

Cinque (1999) allows for no parametrization in functional categories, asserting that all languages have the structure in example (p. 18) in all sentences. So it seems that these two perspectives are incompatible.

When we look at van Gelderen’s work, as noted in Chapter 4, it is different from that of Roberts and Roussou. She concentrates on Principles of Economy (or “efficient computation” in Chomsky, 2007) that are present in all biological structures. Among these are Late Merge Principle (van Gelderen, 2004, p. 12) and the Head Preference Principle (van Gelderen, 2004, p. 18). In van Gelderen (2008), she presents a different version of the Late Merge Principle in terms compatible with current Minimalist Program feature requirements:
Minimize the semantic and interpretable features in the derivation, e.g.:

\[
\begin{array}{ccc}
VP-Adverbial & CP-Adverbial & C-Head \\
[\text{semantic}] > & [\text{iF}] > & [\text{uF}]
\end{array}
\]

This principle, of Feature Economy based on ideas of simplification and thereby on efficiency of computation, can be expanded to include interpretable features in order to meet the needs of lexicalized sentence adverbs (as I postulate in this dissertation).

Let us look again at the “stages” of lexicalization repeated from example (17) (as 17′ on the following page). Each of these steps requires a type of reanalysis by the speaker. Such reanalysis can be thought of in terms of feature revision.
Possible steps in the lexicalization process of *frankly*

(all from Davies, the Corpus of Contemporary American English)


b. I must speak frankly now and tell you that I see no reason for you…

c. If I might speak frankly, there’s others [sic] still aboard more sickish…

d. To speak frankly, Barbara, do you think that this can be accepted…

e. Frankly speaking, if you’re looking for good nutrition in a hot dog…

f. Frankly, in the spirit of free speech, that’s good.

As previously mentioned, the motivation to move from a discourse structure like (17’a) to (17’b) is primarily pragmatic, in order to hold the floor. Yet there is syntactic economy in making this change as well. Such a set of structures as are in (17’a) can be expressed graphically as in (20):
(20) Turn-taking, represented syntactically (from example 17'a)

---INTERLOCUTOR’S TURN---

The sentence in (17'b), where the speech act clause is the matrix clause, is similarly straightforward. The propositional that-clause is the direct object of the verb tell, and as a CP-object, it has features (interpretable φ-features, uninterpretable Case) that must be valued for the External Merge with tell to occur. In addition, there is an interesting coordination (parataxis) between speak
and tell. The speaker distinguishes between the intransitive speak while directly addressing the listener with the ditransitive tell.

In losing the permission-seeking quality of the first exchange, the speaker in (17′b) shifts from a lower functional head (Mod\textsubscript{ability/permission}) to a higher one (Mod\textsubscript{necessity}) (i.e. may > must). This fits in with a floor-holding strategy, where the speaker is “compelled” to continue talking. This offsets the negative face threat of the upcoming proposition. See (21) (following page) for a representation of this structure.
The shift from (17b) to (17c) is also quite significant, and perhaps the most
difficult to account for. I argue that the CP [If I might speak frankly] has been
reanalyzed from the matrix clause of a sentence like (17b) to a speech-act clause
with appropriate feature markings. In a subject or object position, a CP will have
interpretable third person singular, but in an adverbial position, these features are
superseded by other interpretable features necessary to the Mood_{speech-act} node.

These could incorporate (lower) irrealis mood features (present in the subjunctivity of the C-head and in the modal verb) and extra-propositional features that relate to the discourse situation (e.g. interpretable illocutionary features). See the simplified tree in (22) below.

(22) The shift from matrix (parataxis) to subordinate (hypotaxis)

```
(…)  Mood_{sp-a}P
      \              /  \\
     (Force)CP [+iF] /    \\
       (Force)C'    /     \\
         (Force)C^0 /      \\
          If         /       \\
             Mod_{a/p}P /        \\
                Mod_{a/p}' /         \\
                  vP (RELATOR^{11}) /           \\
                      there’s others aboard… /               \\
                          Mod_{a/p}^0 /                    \\
                             vP /                      \\
                                DP \\
                                  ↓
                                    I
                                      DP
                                        ↓
                                          Mod_{a/p}^0
                                            may
                                              DP
                                                ↓
                                                  v^0
                                                    speak
                                                      VP
                                                        speak frankly
```

The double-arrow line indicates the feature checking that is now taking place between the hypotactic CP and the speech-act head. There is an interpretable feature checking process, as the operation is External Merge.

^{11} “RELATOR” follows den Dikken (2006) to indicate, among other possibilities, the use of the copula in there’s. This will be discussed more fully in Chapter 6. In figure 5.3, it does not affect the analysis.
In (17’d) and (17’e), internal structural erosion has begun to take effect. Overt structures such as the C-head and the subject are redundant for discourse purposes. Feature checking within the Speech-act specifier node has been reduced to non-finite tense checking, which is reduced from overt subject-verb checking in English (including, in this case, no EPP), and the semantic feature checking of the verb with *frankly*. The non-finite verb forms cannot overtly show irrealis mood, but interpretable features are still evident. See (23) below for another simplified tree:

(23) Infinitive clause structure

By (17’f), full internal reduction has taken place, and the only features remaining are the interpretable features necessary for *frankly* to agree with the functional
head at the Speech-act node. See (24) below for a final simplified tree representing the positioning:

\[
(24) \quad \text{Frankly as the speech-act specifier}
\]

\[
(\ldots) \quad \text{Mood}_{sp\text{-}a}\text{P}
\]

\[
\text{AdvP} \quad [+iF]
\]

\[
\text{Frankly} \quad \text{Mood}_{sp\text{-}a}^{'}
\]

\[
[+iF] \quad (\ldots)
\]

I would therefore argue for a Feature Economy that prioritizes and/or deprioritizes features as necessary for the syntactic layer targeted by the lexical item.

5.9 Chapter Summary

Speech-act adverbs are a highly specific, discourse-oriented type of adverb. They have their origins in adverbs of manner that become closely associated with verbs of speech. In this chapter I have proposed that speech-act adverbs do not result from typical grammaticalization patterns but rather from a pragmatically-driven, discourse-based process of lexicalization by ellipsis.

The rise of speech-act adverbs can be tracked through corpus data that shows a rapid increase in the use of \textit{frankly} as such an adverb, particularly through the last 40 years. Finally, the lexicalization process can be analyzed as a type of Feature Economy in which discourse features supersede semantic and grammatical features in the left periphery of the sentence.
Additional research will add both depth and breadth to this part of the study. The exact nature of the interpretable features that the CP gains in becoming an adjunct needs to be thrashed out. An examination of the full spectrum of speech-act adverbs through corpus data would give the study further strength, while additional historical and dialectal data would add a broader focus.

In the next chapter (Chapter 6), we will switch from speech-act adverbs to look at three types fairly similar to one another: evaluative, epistemic, and evidential adverbs. These adverbs arise through different types of reanalyses, whose mechanisms will be seen to correspond more closely with grammaticalization phenomena.
Chapter 6

COPULA-BASED ADVERBIALS:
EVALUATIVE, EVIDENTIAL, AND EPISTEMIC

6.1 Introduction

6.1.1 Combining these three types of sentence adverbs.

Not all sentence adverbs are alike, of course. Speech-act adverbs, as we saw in the last chapter, are directly connected to the discourse situation. Using an adverb like *frankly* is a means of communicating various messages to the listener. These messages are indicative of the manner in which the speaker would like the proposition received/understood by the listener: as *frankly* or *honestly* or *seriously* or *briefly* presented via speech. They are openly deictic. These are the locutionary messages; they are bolstered by illocutionary meanings, such as a warning of a face threat or the expression of stance in the communicative environment.

Using evaluative adverbs (such as *naturally* and *unfortunately*), evidential adverbs (*obviously, apparently*), and epistemic adverbs (*possibly* and *definitely*), by contrast, appears to de-emphasize somewhat the speaker’s relationship with the listener. Instead, it is the speaker’s consideration of the propositional content of the sentence that is highlighted. On the surface, this primarily affects the locutionary content of the message; these adverbial types are nevertheless quite useful for establishing stance or conveying other illocutionary information.

These three adverbial types are connected to one another by their canonical position in the clause, near the far left periphery but (according to Cinque, 1999) staggered sequentially below the speech-act adverbs. I distinguish
the three types of adverb based on Cinque’s hierarchy, although there are several different nomenclatures among different scholars. See example (1) below for a reminder of how Cinque fits these adverbs (in boldface) in the syntactic structure of the clause:

(1) Cinque’s hierarchy of adverbial structure

Mood\textsubscript{speech act} > Mood\textsubscript{evaluative} > Mood\textsubscript{evidential} > Mod\textsubscript{epistemic} > T(Past) > T(Future) > Mood\textsubscript{irrealis} > Mod\textsubscript{necessity} > Mod\textsubscript{possibility} > Asp\textsubscript{habitual} > 
Asp\textsubscript{repetitive(I)} > Asp\textsubscript{frequentative(I)} > Mod\textsubscript{volitional} > Mod\textsubscript{obligation} > 
Mod\textsubscript{ability/permission} > Asp\textsubscript{CELERATIVE(I)} > T(Anterior) > Asp\textsubscript{terminative} > 
Asp\textsubscript{continuative} > Asp\textsubscript{perfect} > Asp\textsubscript{retrospective} > Asp\textsubscript{proximative} > Asp\textsubscript{durative} > 
Asp\textsubscript{generic/progressive} > Asp\textsubscript{prospective} > Asp\textsubscript{SGCompletable(I)} > Asp\textsubscript{PLCompletable} > Voice > Asp\textsubscript{CELERATIVE(II)} > Asp\textsubscript{repetitive(II)} > Asp\textsubscript{frequentative(II)} > Asp\textsubscript{SGCompletable(II)} (p. 106)

In addition, these three adverb types can largely be paraphrased by copula clausal structures of the type traditionally known as \textit{IT-extraposition}. This type of structure was briefly introduced in Chapter 2, and in example (2) there is a sample sentence:

(2) \textit{IT-extraposition}

\emph{It is wonderful/possible/obvious that dogs now live on Mars.}

(Compare: That dogs now live on Mars is wonderful/possible/obvious.)
This construction allows us to gain insight (as we will see below) both into the sentence adverb-form being investigated here and into the historical processes involved in the rise of such forms.

### 6.1.2 Chapter organization.

In section 6.2, I first examine the three subcategories in more detail, acknowledging the imperfections of the semantic subcategorization model used. Section 6.3 is focused on the orthodox grammaticalization approaches to such adverbs and some of the problems that arise as a result. Section 6.4 turns to the IT-extraposition case in syntactic literature. I contend that insights from this literature help inform an understanding of the diachronic processes that seem to be involved in the rise of these three sentence adverb types. In section 6.5, historical usage patterns are traced and examined in light of the previous discussions. Finally, section 6.6 concludes the chapter.

### 6.2 The Subcategories: Evaluative, Evidential, and Epistemic

#### 6.2.1 The Old English model: primarily truth-intensifying adverbs.

Thanks to limited textual evidence, relatively little is known about the extent to which sentence adverbs were used in Old English (OE). It is known, however, that there is a significant gain in number of sentence adverbs found in the Middle English (ME) period in comparison. This lack of knowledge about OE is likely attributable to the availability to present-day scholars of examples from
only a few genres (i.e. religious, legal, and historical texts—many of which are Latin translations—plus poetry).

Swan (1988), in her exhaustive examination of sentence adverb occurrence and use across the various historical English periods, notes that by far the most widely used sentential adverbs in OE are what she terms “truth intensifiers” (pp. 87ff.). Such adverbs are roughly comparable to Modern English (ModE) modal (evidential and epistemic) or speech-act adverbs. Example (3) gives a list of the most common truth-intensifying OE adverbs:

(3) Truth intensifiers (from Swan, 1988, p. 87)

<table>
<thead>
<tr>
<th>OE</th>
<th>ModE</th>
</tr>
</thead>
<tbody>
<tr>
<td>soðlice (truly)</td>
<td>untweogendlice (indubitably)</td>
</tr>
<tr>
<td>witodlice (truly, certainly)</td>
<td>openlice (openly, plainly)</td>
</tr>
<tr>
<td>sweotole (evidently, clearly)</td>
<td>nede (necessarily)</td>
</tr>
<tr>
<td>sweotolice (evidently, clearly)</td>
<td>uncuð (unknown, uncertainly)</td>
</tr>
<tr>
<td>gewislice (truly)</td>
<td>unleaslice (not falsely)</td>
</tr>
<tr>
<td>cuðlice (truly, certainly)</td>
<td>untwylice (undoubtedly)</td>
</tr>
</tbody>
</table>

The most common of these is soðlice (the stem of which is preserved in ModE words soothsayer and forsooth), which appears 135 times in Swan’s corpus. Second most is witodlice, which, although it appears 67 times, is the most often preposed (43 times vs. 30 for soðlice). The others are comparatively rare. See examples for soðlice and witodlice below.
Further modeling for the modern pattern of left-periphery adverbs can be seen in OE topicalization and temporal adverbs, both of which are generally found on the left of the TP (Pintzuk, 1993). An example of each is below:
(6) OE adverbial topicalization (from Pintzuk, 1993, p. 156):

\[
\text{swiðe eæðe mæg on smyltre sæ ungelaered} \\
\text{very easily can on smooth sea untaught}
\]

\[
\text{scipstiora genoh ryhte stieran} \\
\text{steersman enough straight steer}
\]

“An untaught steersman can very easily steer straight enough on a smooth sea.”

\[(Cura Pastoralis 58.2-3)\]

(7) OE temporal adverbial (from Pintzuk, 1993, p. 157):

\[
& \text{fullice .lxx. wintra syðdan on an was se} \\
\text{and fully 70 years afterwards continually was the}
\]

\[
\text{ðeodscype eall geðeowod under heora feonda gewealde} \\
\text{nation all enslaved under their enemies’ power}
\]

“And for fully 70 years afterwards, all the nation was continually enslaved under their enemies’ power.”

\[(Wulfstan’s Homilies 6.120-1)\]
This pattern of preposing adverbs, going back at least to OE, is evidence that later grammaticalization phenomena of sentence adverb formation is at least partially based on analogy.

In example (7) above, for example, there are temporal adverbs that may sit in the ModP specifier in Rizzi’s expanded CP (“for fully 70 years afterwards”), followed by an Asp\textsubscript{continuative} adverb (“continually”). Neither of these is one of the adverb types being studied, but this still stands as evidence that English speakers are comfortable using sentence adverbs. With a longstanding tradition of adverbs in the first position in English, the left periphery makes sense as a landing site for sentence-scope adverbs.

The availability of the disjunct positions (i.e. external to the main clause) for sentence adverbs (e.g. in examples 4 and 5) can therefore be traced in English from the earliest record. Pintzuk (1993) indicates that “[i]n particular, the position of sentential adverbs does not ‘count’ for verb-seconding” (p. 164). In her P&P-based characterization of OE as a “well-behaved West Germanic language” (i.e. with regard to V2 patterning), this is possible because sentence adverbs are adjoined to IP or to I\textsubscript{0} (p. 164).

6.2.2 Evaluative adverbs in English.

Evaluative adverbs are used to offer comments on the propositions that are under their scope. Unlike epistemic and evidential adverbs, there is not a concern with truth value in the use of evaluatives. Ernst (2002) gives a semantic definition to this subcategory (and the sub-subcategories within it):
Evaluatives represent the speaker’s evaluation of some state of affairs according to how good or bad it is (luckily, unfortunately), how (ab)normal it is (normally, strangely, curiously, surprisingly), its desirability (ideally, preferably), or a wide range of other criteria (e.g. for significantly, absurdly, conveniently, astonishingly, etc.). (p. 76)

Most of these adverbs (aside from the “desirability” category in the above quote) are factives; that is, they assume the truth of the proposition and seek to propose ways that the speaker and listener(s) should regard such a truth. Bonami and Godard (2006) explain the basic pragmatic use of evaluative adverbs:

In the case of simple assertion, the speaker asserting $p$ without an evaluative commits himself [sic] to the truth of $p$, at the same time as he asks the addressee(s) to evaluate $p$; when he asserts evaluative $p$, the same conversational moves are present, but, in addition, the speaker commits himself to the proposition associated with the adverb while withdrawing it from the addressee’s evaluation. (p.1)

In other words, the use of an evaluative adverb with a proposition is the speaker’s opportunity to structure the framework of how that proposition is to be regarded by the interlocutors.

Cinque (1999), in his hierarchy (see example 1 above), indicates that evaluative adverbs follow speech-act adverbs. There are very few sentences where both types of adverbs are used, at least in English. Nevertheless, with a Google search, it can be seen that some evaluatives do combine with speech-act adverbs, and generally along expected lines:

(8) The only conversation about immigration is, frankly, unfortunately, is political pandering, and it’s really sad…

(Mario Diaz-Balart, CQ Politics)
We had a very schedule in three days that frankly, unfortunately prevented us from being able to go to all the places that you might have wanted to visit during a trip to India…

(Ben Rhodes, NDTV.com)

The only counterexamples I found were false, where the evaluative is either at the end of a previous clause or the speech-act adverb is serving an unusual role.

Unfortunately, frankly speaking from the military is not appreciated…

(militaryphotos.net)

…and would never have been able to convey clearly what I wanted to a developer, unfortunately, frankly b/c I didn’t even know.

(webhostingtalk.com)

This is empirical support for Cinque’s hierarchy. It shows that there may indeed be a shift in direction from *frankly* to *unfortunately*, from looking outward to the listener to looking inward to the proposition, and that the order is likely based on this type of criteria.

Swan (1988) notes that the evaluative subcategory of adverb, seen in very small numbers in earlier periods of English, “expands enormously” during the twentieth century, in particular (p. 464). She proposes three subcategories to the larger class of evaluatives: a category that is [± fortunate], one that is [± strange], and one that is a miscellaneous class. These three semantic subcategories, though
perhaps rather unexpected, are found in the literature. The three can be seen below in examples (12)-(14).

(12) [± fortunate] evaluative adverbs (Swan, 1988, p. 464)
luckily mercifully sadly
(un)fortunately (un)happily

(13) [± strange] evaluative adverbs (Swan, 1988, pp. 468-9)
naturally understandably
absurdly astonishingly curiously
funnily incredibly inexplicably
magically miraculously oddly
paradoxically perplexingly remarkably
staggeringly strangely stupendously
surprisingly unbelievably unexpectedly
wonderfully

(14) Miscellaneous evaluative adverbs (Swan, 1988, p. 474)
amazingly appropriately complexly
disconcertingly hopefully importantly
inescapably inexplicably intriguingly
ironically notoriously predictably
In addition to categorizing these adverbs by their semantic content, Ernst (2002) separate evaluatives into pure evaluatives and dual evaluatives (p. 76). Pure evaluatives cannot take a more narrow-scope position; in other words, they cannot be used as manner adverbs in the VP. There are several of these, but the most common are adverbs such as *(un)fortunately* and *surprisingly*. See below:

(15) a. Unfortunately, the ship has sailed.
    b. *The ship has sailed unfortunately. (i.e. *It sank.*)

(16) a. Surprisingly, the hamster spoke French.
    b. *The hamster spoke French surprisingly. (i.e. “Alors!”)

Dual evaluatives, however, can act in both contexts, although they often have quite different readings in the different scopal positions.

(17) a. Curiously, the cat kept sneaking into the paper bag.
    b. The cat kept curiously sneaking into the paper bag.

In the first sentence, the speaker finds the cat’s behavior perplexing. In the second, the cat is exhibiting curiosity (as cats are wont to do). Such a shift when the adverb is found in a wide-scope position is not at all unusual. In addition, in
this example, we can get a narrow-scope reading with the first sentence if we treat
the preposed adverb as a type of topicalization.

Some evaluations, as is well known, are “more equal” than others. There
is a further subclass of evaluative adverbs that have been widely condemned by
language mavens, particularly in written use. The best known of these is hopefully
but there are others such as thankfully and even mercifully. The APA Publication
Manual, Fifth Edition, for example, has this to say about hopefully:

Another adverb often misused as an introductory or transitional word is
hopefully. Hopefully means “in a hopeful manner” or “full of hope”;
hopefully should not be used to mean “I hope” or “it is hoped.”

Incorrect:
Hopefully, this is not the case.

Correct:
I hope this is not the case.

(American Psychological Association, 2001, p. 54)

The OED puts the blame on the United States for both hopefully and thankfully:

Hopefully, adv.
2. It is hoped (that); let us hope. (Cf. German hoffentlich it is to be hoped.)
orig. U.S. (Avoided by many writers.)

Thankfully, adv.
II.4. Let us be thankful (that); one is thankful to say. orig. U.S.
This use as a sentence adverb, like hopefully adv. 2, is deprecated by
some writers.

It is not clear why the reaction to hopefully is so strong, but there are blog entries
and personal pet peeves devoted to this subject.

I suggest that the main resistance arises from a sense among many
language users of the underlying structure of most evaluative adverbs, namely IT-
extraposition, also known as “It is ADJ that”. For example, the sentence “Clearly,
she won” can be restated as “It is clear that she won”. As will be seen in section 6.4, there is a parallel structure between the copula and the -ly suffix.

For hopefully and thankfully, however, the adjective forms are hopeful and thankful—neither of which can be used in an IT-extraposition structure. (Cf. *“It is hopeful that she won”; *“It is thankful that she won”.) As a result, people who reject this usage may recognize the break in the pattern without understanding what exactly the pattern is.

6.2.3 Evidential adverbs in English.

Evidential adverbs are, generally, expressive of the manner in which something has been perceived by the speaker. The adverbs are frequently based on roots that indicate sensory perception. In ModE, often the use of evidential adverbs is only tangentially related to the report of how evidence/knowledge has been gathered by the speaker. There is a concern with truth (as there is in epistemic and in speech-act adverbs, as well) but the focus is on the speaker’s connection with and confidence in the manifestation of that truth.

As in the previous section, there is a contention in Cinque (1999) that evidential adverbs follow evaluatives. Let’s look at a few examples from Google.

(18) Samuel Eto’o, unfortunately, obviously isn’t an isolated case.

(gunnergirl.tumblr.com)
(19) Unfortunately, obviously, not many native French teachers know about it or would feel confident enough to enter a candidate.

(20) Obviously, unfortunately, the gentleman has passed away in the next couple of clips, because that’s what we’re trying to determine.

In the last example, where the two have swapped places, it seems that there is topicalization of *obviously*, as the speaker is discussing the unfortunate video footage of a man being run down by a bus. As a result, the visual aspect of *obviously* has increased in importance.

As with evaluative adverbs, Swan (1988) provides a fairly complete list of most current evidential adverbs:

(21) Evidential modal adverbs (Swan, 1988, p. 451)

<table>
<thead>
<tr>
<th>apparently</th>
<th>clearly</th>
<th>evidently</th>
</tr>
</thead>
<tbody>
<tr>
<td>manifestly</td>
<td>obviously</td>
<td>ostensibly</td>
</tr>
<tr>
<td>patently</td>
<td>plainly</td>
<td>recognizably</td>
</tr>
<tr>
<td>seemingly</td>
<td>unmistakably</td>
<td></td>
</tr>
</tbody>
</table>
In addition, she recognizes a set of “miscellaneous” adverbs that are related to evidentials (and that are often categorized with them). She places them into three different subcategories: true miscellaneous, performative, and distancing adverbs.

(22) Miscellaneous (evidential) modal adverbs (Swan, 1988, p. 455)

<table>
<thead>
<tr>
<th>MISCELLANEOUS</th>
<th>PERFORMATIVE</th>
<th>DISTANCING</th>
</tr>
</thead>
<tbody>
<tr>
<td>actually</td>
<td>admittedly</td>
<td>allegedly</td>
</tr>
<tr>
<td>indeed</td>
<td>assuredly</td>
<td>supposedly</td>
</tr>
<tr>
<td></td>
<td>(in)disputably</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(un)arguably</td>
<td></td>
</tr>
<tr>
<td></td>
<td>unquestionably</td>
<td></td>
</tr>
</tbody>
</table>

6.2.4 Epistemic adverbs in English.

The epistemic adverb is an adverb that marks the speaker’s confidence in the truth of a proposition. This, as we have seen in other areas, is a very important question that spills over into speech-act adverbs (honestly, frankly), evaluative adverbs (hopefully, incredibly), and evidential adverbs (actually, seemingly). Because these boundaries are fuzzy, there are several terms for the adverbs that are allocated to this category, such as epistemic, modal, non-factive, and truth conditional.
Once again, according to Cinque (1999), the epistemic adverb follows the evidential adverb in the hierarchy. Here are some Google search results to test his theory:

(23) Wave B obviously probably peaked yesterday, which means that SPX (S & P 500) is in a short term Wave C downcycle.

(tradethecycles.blogspot.com)

(24) I’m obviously probably not the first one to come up with it but I didn’t steal it from anyone haha I was just blazed and thought it up one…

(highideas.com/philosophy)

(25) You don’t know... she probably, obviously, has a phobia now, but you can’t know she had a phobia before the incident.

(www.huffingtonpost.com)

One clue that distinguishes (25) from the others with the “canonical” order is the use of commas. These seem to indicate that there has been movement of elements around this particular use of obviously. It also may be a self-correction, as obviously is a stronger word than probably.

It is not surprising that the terminology for this type of adverb, like the evidential adverb just discussed, is somewhat confused. Swan (1988) labels these as “medium-probability modal adverbs”, sharing the larger category (modal) with
evidentials (p. 441), while Ernst (2002) considers modal and evaluative to be two types of epistemic adverbs. Examples (26)-(28) show the semantic distinctions assigned by Swan’s (1988) typology:

(26) High-probability modal adverbs (Swan, 1988, p. 436)
   certainly  inevitably  doubtless
   surely  necessarily  indubitably
   undoubtedly  definitely  decidedly

(27) Medium-probability modal adverbs (Swan, 1988, p. 441)
   probably  presumably  likely

(28) Low-probability modal adverbs (Swan, 1988, p. 445)
   perhaps  maybe  possibly
   conceivably

The low-probability adverbs are generally thought to sit below the Tense heads, and Cinque places them in that position (i.e. in the specifier of the Mod possibilité head). The high-probability adverbs, on the other hand, appear to be categorized in the epistemic category.
6.3 Diachronic Approaches to the EEE Subcategories

As we saw in Chapter 2, the standard approach to grammaticalization is the cline of unidirectionality. This cline, as I discussed before, is not predictive but reactive to previous patterns of change. The cline for grammaticalization of adverbs (clause-internal adverbial > sentence adverbial > discourse particle) (Hopper & Traugott, 2003, p. 37) indicates straightforward upward mobility of the adverb, but as we saw in the case of the speech-act adverb *frankly*, it only tells us the facts (that *frankly* has changed from a VP-adverb to a sentence adverb) without supplying either the motivating factors or the mechanisms for change.

In the case of *frankly*, its position near the left periphery of the sentence is more likely due to the erosion of a main clause: first changing into a subordinate clause, then through ellipsis (suggested here to be a type of lexicalization) becoming a single-word adverb. This is a more finely-grained solution than a cline can offer, of course, and it shows the interplay of different types of change.

The situation with regard to the evaluative-evidential-epistemic adverbs (hence EEE adverbs) is, I argue, essentially the same (main clause > sentence adverb) from a macro perspective. In this case, though, the original main clause is far more closely connected with the propositional clause than in the speech-act adverbial situation. This group of adverbs shows far more direct signs of grammaticalization over lexicalization, as well.
6.4 Sentence Adverbs and Predication: A Syntactic Examination

6.4.1 Underlying structure.

Deciding upon an “underlying structure” for almost any expression can be daunting. The same is true for the EEE adverbs, but throughout the literature on these words there is noted a paraphrase for an adverb like unfortunately in the copula clause it is unfortunate that. The two structures, morphological and syntactic in nature, are obviously related. It also seems clear from the OED that adjectives precede adverbs in essentially all cases of borrowing from French and Latin (the sources for many EEE adverbs).

Swan (1988) specifically proposes a copula-adjective structure (along with intensifying adverbs) in OE as a possible source for evaluative sentence adverbial development in later periods. Below are two examples of such sentences:

(29) a. þæt is sarlic þætte…

“it is grievous that…” (Bede 96, from Swan 1988, p. 137)

b. Wæs eac wundorlic þæt…

“It was wonderful too that…” (Ælfric, Lives I 438, ibid)

These are clearly examples of it-extraposition (or the equivalent, with a demonstrative in (29a) and a null subject in (29b)), and they demonstrate how long this construction has been used in English.
In ME, the pattern seen in (29b) is retained—namely, the null subject lasts far longer in sentences that make use of expletive *it* than in any other sentence type (Fischer, van Kemenade, Koopman, and van der Wurff, 2000, p. 71). Below are two examples of late ME “it-less” sentences (from Fischer et al., 2000, p. 71):

(30) *to us surgiens aperteneth* that we do to every wight the beste that we kan

“It is our duty as surgeons to treat every person as well as we can”

(Chaucer, *Melibee*, 1011)

(31) hard is to knowe in al poyntis to holde the meene, lyght is hit to faille

“It is hard to know exactly how to rule society; it is easy to fail”

(Lydgate, *Secreta Secretorum*, 130/26)

Interestingly, in the Lydgate example, the use/non-use of *it* is mixed between the two clauses; this is a good example of the non-obligatory nature of the construction. According to Fischer et al. (2000), “In the course of the fifteenth century, this option is used less and less frequently and by 1500 the use of expletive *it* has become the rule” (p. 72). One of the well-known hallmarks of the ModE period, in fact, is the loss of the null subject option.

### 6.4.2 Den Dikken (2006) and the RELATOR.

Within the framework of the Minimalist Program, den Dikken (2006) has examined predicate constructions of all types, including copula constructions, but
also including adjectival and adverbial modification. He proposes a functional \textsc{relator} head that can be inserted in different kinds of structural environments, from \(v\)Ps to “of” PPs and manner AdvPs. The hierarchical structure within a \textsc{relator} phrase is not the same sort of structure as within other phrases: “predication relationships, while always hierarchically asymmetrical, are fundamentally nondirectional” (den Dikken, 2006, p. 3). See the simplified trees below for illustration:

(32) Nondirectionality of \textsc{relator} phrases (den Dikken, 2006, p. 3)

\[
\begin{array}{c}
\text{RP} \\
/ \ \\
\text{SUBJECT} & \text{R'} \\
| \ \\
\text{The flowers} & \text{RELATOR} & \text{PREDICATE} \\
| \ \\
| \ \\
\text{are} & \text{lovely} \\
\end{array}
\]

\[
\begin{array}{c}
\text{RP} \\
/ \ \\
\text{PREDICATE} & \text{R'} \\
| \ \\
\text{Lovely} & \text{RELATOR} & \text{SUBJECT} \\
| \ \\
| \ \\
\text{are} & \text{the flowers} \\
\end{array}
\]

Den Dikken (2006) argues that both of these structures can be base-generated by the syntax, and indeed, his focus is on motivating predicate-inversion structures. Such concerns of inversion are significant for the purposes of this study.
Additionally, den Dikken argues for the lack of θ-role assignment by RELATOR heads, and that is a further important consideration for our purposes (2006, pp. 14-15). Since lexical heads (i.e. V, P—note that, for den Dikken’s purposes, P is considered lexical) assign θ-roles (i.e. agent, theme, goal), he argues that RELATOR heads are functional instead (p. 22).

Copulas are not the only structures that produce predicate complements. Adjectival and adverbial structures, den Dikken states, do so as well (2006, p. 29). In particular, manner adverbs utilize the -ly suffix, he argues, as a RELATOR head. See the example in (33) (from den Dikken, 2006, p. 30):

(33)  a. Imogen dances beautifully.

\[
\text{[RP [Imogen dances] [RELATOR=-ly [beautiful]]]}
\]

b. Imogen dances like a beauty.

\[
\text{[RP [Imogen dances] [RELATOR=like [a beauty]]]}
\]

In these two examples, the subject of the RELATOR phrase is the TP Imogen dances. The predicates are beautiful in (33a) and a beauty in (33b). The relators are -ly and like, respectively. These parallel structures help with our understanding of the similarity between copula and adverbial structures in this study; however, they also specifically point out the diachronic connection between -ly and like.

In his discussion, den Dikken (2006) specifically refers to the -ly relator as a “lexicalization” of the RELATOR head:
By doing so, it elevates -ly to the status of a syntactically autonomous marker—a RELATOR of subject-predicate relations similar to like in [Xb]. Thus the analysis captures the historical relationship between -ly and like, by generating them in the same structural position and giving them the same syntactic function: that of relating a predicate to its subject. (p. 31)

If we express (33) above as graphic (tree) structures, they would look like (34) below:

(34)  Like and -ly as RELATOR heads (based on den Dikken, 2006, p. 31)

By using den Dikken’s (2006) model of RELATOR phrase structure, then, we can apply a syntactic framework to the shift that English speakers undergo, both synchronically and, I argue, diachronically, from matrix copula clause to disjunct adverbial phrase.
6.4.3 A “lexicalized” RELATOR (-ly).

Syntactically, then, there are two main shifts that need to occur in the reanalysis and grammaticalization of the copula constructions as they become EEE sentence adverb constructions. Neither of these, I argue, is the topicalization of a manner adverb to a fronted position. Instead, we have the shift, as we saw in speech-act adverbs, of a syntactically matrix clause to a subordinated position, and the shift from a verbal RELATOR to an adverbial RELATOR head, which leads to univerbation (which I term here a type of lexicalization) of the RELATOR-predicate structure.

It is easiest to see this change through the use of tree diagrams, so let us take a simple extraposed sentence with an evidential predicate:

(35) It is clear that he fell.

This, we will argue, is the surface structure (to use older terminology) of the sentence whose deep structure is in (36):

(36) That he fell is clear.

In order to get the structure in (35), a structure that is more common in English than the structure in (36), a couple of instances of Merge (transformations in older terminology) must occur. See (37) below.
In this example, the reciprocity of subject-predicate can be seen, as the base structure is a \( \text{vP}-\text{RELATOR} \) (no VP) and the adjective is the predicate \((p)\) while the clause is the subject \((s)\). The verb merges internally, moving in this situation to \( T^0 \), which here is also a RELATOR head. The CP does not move, and an expletive \((it)\) subject is required to enter the derivation via External Merge into the specifier position of the upper RELATOR head by the EPP feature.

Also in (37), the matrix (main) clause of this sentence is syntactically the copula clause; however, the lower CP (acting as the subject) and the AdjP (the predicate) are hooked up by the lower relator head, the vP. Before the rise of strict SVO order in later ME, the verb was able to move to the higher position without attracting the expletive pronoun \( it \) (as we saw in example 31 above). From the start of ModE to the present day, however, this is impossible.
So in order to move from this structure to the evidential adverb structure

(Clearly, he fell), as we said, a couple of processes need to be undergone by
speakers. Example (38) shows the adverb structure analyzed through Cinque and
den Dikken.

(38) Changing structure via Cinque (1999) and den Dikken (2006): Option 1

From this tree, we can see that in some ways den Dikken’s explanations are
useful, and in some ways they are not. There are two fairly major concerns to be
overcome.

First, if clear morphologically merges with -ly, which is a head (from
either the specifier or complement position), and then somehow moves to the
specifier of the Mood$_{evidential}$ head, then we have Head-Spec movement, which is a
major no-no. The second concern is moving the new main clause of the sentence out of den Dikken’s RELATOR structure (where it is a specifier, even if such a distinction is only linear rather than hierarchical) and into place in the main part of the tree.

However, the second concern is, at least, in line with the type of diachronic reanalysis that speakers must make; the adverb must be reanalyzed as disjunct from the main clause as well, rather than the predicate of that clause. The motivation for such a reanalysis is likely structural, in that English speakers frequently want to give clausal weight to the proposition. Another possible solution does not involve movement:

(39) Modern English evidential-adverb structure (Cinque, 1999; den Dikken, 2006): Option 2

```
CP
  C'
    C^0
    Mood_evidential^P (RELATOR)
      AdjP(p)
        clear
          Mood_evidential'
            Mood_evidential^0
              [+iF] -ly
              TP(s)
              he fell
```

In (39), the solution is another hybrid between Cinque and den Dikken, but one that is not based on an Internal Merge (movement) solution. In this case, the -ly is the RELATOR head that attracts clear as an adjective to the specifier (predicate)
position through a Match operation that matches interpretable features. The TP that is the remainder of the sentence is in a typical complement (subject) position that follows the Mood\textsubscript{evidential} phrase.

This solution undoubtedly fits the intersection of the two theoretical models better than the first attempt. Here we have the functional head (functional in both frameworks) that contains the predicating element (the adverbial suffix \textit{-ly}). Through a DM process such as morphological merger via adjacency, the two nodes combine in a manner that preserves asymmetry.

Such a proposed solution also observes the syntactic guidelines for grammaticalization, in that it simplifies structures. If we were still utilizing the Late Merge Principle (van Gelderen, 2004), this would be a strong argument on its behalf. Yet with Feature Economy (van Gelderen, 2008), that has replaced it, there is still a correlation.

Although the vP/TP \textsc{relator} heads are not lexical, there are nevertheless semantic qualities found in the verbal form of the \textsc{relator} (the copula) that cause the subject and predicate to pair up on either side. Essentially, they must have matching semantic identity. In the case of the External Merge to the Mood\textsubscript{evidential} head, on the other hand, the head element (\textit{-ly}) can be interpreted as a grammaticalizing suffix. Brinton and Traugott (2005) assert, for example, that “\textit{-ly} is becoming grammaticalized, and that individual adverbs formed with it have a tendency to grammaticalize” (p. 132).

The features that must match between \textit{clearly} and \textit{he fell} in the above examples and figures are more removed than, say, between \textit{he fell} and \textit{quickly}. In
“He fell quickly”, the manner adverb ending -\textit{ly} is acting to create a \textsc{relator} structure that is something like \texttt{[he fell] = [quick]}. In “Clearly, he fell”, it looks more like \texttt{[truth(he fell)] = [clear]}. The truth value of the statement distances the subject and predicate. The Match process, therefore, is not purely semantic but also grammatical.

Now that I have offered a possible analysis for the grammaticalization of EEE adverbs, I will examine some corpus data to trace, in quite broad lines, the shift from adjective-based clauses to clauses with sentence adverbials.

\section*{6.5 Usage Patterns}

Let us assume a process rather shorter but not completely unlike the one assumed for speech-act adverbs.

(40) Possible steps in the history of evidential adverbs (from Davies, Corpus of Contemporary American English)

\begin{quote}
a. What we say is clear.
That they are not is clear.

\begin{itemize}
\item \textbf{▼}
\end{itemize}

b. It is clear that things will never be the same.

\begin{itemize}
\item \textbf{▼}
\end{itemize}

c. But, clearly, they are still in charge.
\end{quote}
In the COCA, there are 47,750 occurrences of *clearly* in a 410-million-word corpus. This compares to 9,044 occurrences of *frankly*. Based on a careful examination of the occurrence patterns, approximately one-third of the tokens of clearly in the corpus are sentence adverbs (~16,000). This compares to 2,964 instances of *It is clear*, and an estimated number fewer than 100 (again based on occurrence patterns; in examining 500 sentences I found 2) instances of *CP is clear*. In the two cases of *CP is clear* that I found (24a), they are both notably small CPs that are acting as subjects.

The changes that have happened to evidential adverbs have largely been echoed in the other copula-based sentence adverb types. The frequency of adverb use, as is well known, has grown tremendously during the Modern English era. Table 6.1 below compares ADJ–ADV usage from the Corpus of Historical American English; the results across prototypical evaluative, evidential, and epistemic groups are fairly strong.

What can be seen in the table, which traces the adjectival and adverbial forms of words through the course of the COHA’s time span, is that in each of the three subcategories there is growth in the use of the adverb. There is also significant decline in the use of the corresponding adjective, although in the case of *obvious* that decline is stopped and even reversed somewhat. In contrast, the decline of *probable* is startling: where in the 1810s, *probable* had a frequency of more than 63 per million words, by the 2000s the frequency has plummeted to fewer than 5 per million. The decline for *unfortunate* is just about as severe.
In contrast, the adverb forms all show fairly strong growth during the same time. *Unfortunately* and *probably* each double in frequency; *probably* is, in addition, a word of very high frequency in comparison with the others.

The contrast between *probably* and *probable* (243.55:4.84) is extreme.

Where the adverb form had been approximately twice as frequent as the adjective in 1810, by the twenty-first century it is more than 50 times as frequent. A similar story can be told about *unfortunately* and *unfortunate*; where the adjective is originally four times as frequent, the adverb reaches the point of triple its frequency by the 2000s.

Table 6.1
Historical use of adjective–adverb pairs in American English
(Data from Davies, Corpus of Historical American English)
(Figures are occurrences per million words)

<table>
<thead>
<tr>
<th>Decade</th>
<th>Unfortunate</th>
<th>Unfortunately</th>
<th>Obvious</th>
<th>Obviously</th>
<th>Probable</th>
<th>Probably</th>
</tr>
</thead>
<tbody>
<tr>
<td>1810s</td>
<td>64.34</td>
<td>16.93</td>
<td>58.41</td>
<td>8.47</td>
<td>63.49</td>
<td>123.60</td>
</tr>
<tr>
<td>1820s</td>
<td>64.24</td>
<td>19.78</td>
<td>61.50</td>
<td>11.84</td>
<td>62.80</td>
<td>181.75</td>
</tr>
<tr>
<td>1830s</td>
<td>57.79</td>
<td>20.47</td>
<td>50.02</td>
<td>17.35</td>
<td>47.41</td>
<td>155.94</td>
</tr>
<tr>
<td>1840s</td>
<td>40.07</td>
<td>17.01</td>
<td>39.51</td>
<td>14.71</td>
<td>39.88</td>
<td>134.03</td>
</tr>
<tr>
<td>1850s</td>
<td>40.31</td>
<td>18.27</td>
<td>39.28</td>
<td>8.07</td>
<td>41.28</td>
<td>147.83</td>
</tr>
<tr>
<td>1860s</td>
<td>38.11</td>
<td>20.99</td>
<td>28.20</td>
<td>9.38</td>
<td>43.57</td>
<td>153.21</td>
</tr>
<tr>
<td>1870s</td>
<td>39.43</td>
<td>23.60</td>
<td>31.19</td>
<td>12.50</td>
<td>40.14</td>
<td>161.24</td>
</tr>
<tr>
<td>1880s</td>
<td>36.57</td>
<td>20.67</td>
<td>27.22</td>
<td>11.08</td>
<td>44.45</td>
<td>181.88</td>
</tr>
<tr>
<td>1890s</td>
<td>32.62</td>
<td>19.08</td>
<td>25.97</td>
<td>14.13</td>
<td>38.64</td>
<td>181.69</td>
</tr>
<tr>
<td>1900s</td>
<td>29.60</td>
<td>21.72</td>
<td>31.81</td>
<td>21.18</td>
<td>33.71</td>
<td>171.78</td>
</tr>
<tr>
<td>1910s</td>
<td>28.68</td>
<td>19.96</td>
<td>38.32</td>
<td>26.65</td>
<td>30.48</td>
<td>184.31</td>
</tr>
<tr>
<td>1920s</td>
<td>26.94</td>
<td>20.97</td>
<td>42.26</td>
<td>38.82</td>
<td>32.04</td>
<td>207.61</td>
</tr>
<tr>
<td>1930s</td>
<td>21.99</td>
<td>23.01</td>
<td>48.78</td>
<td>54.02</td>
<td>27.44</td>
<td>220.42</td>
</tr>
<tr>
<td>1940s</td>
<td>19.22</td>
<td>23.04</td>
<td>53.06</td>
<td>65.43</td>
<td>21.15</td>
<td>224.00</td>
</tr>
<tr>
<td>1950s</td>
<td>15.12</td>
<td>21.19</td>
<td>62.91</td>
<td>72.07</td>
<td>18.50</td>
<td>214.79</td>
</tr>
<tr>
<td>1960s</td>
<td>15.18</td>
<td>23.73</td>
<td>66.86</td>
<td>77.28</td>
<td>13.55</td>
<td>217.96</td>
</tr>
<tr>
<td>1970s</td>
<td>14.74</td>
<td>30.11</td>
<td>63.57</td>
<td>77.39</td>
<td>11.25</td>
<td>235.23</td>
</tr>
<tr>
<td>1980s</td>
<td>12.17</td>
<td>29.90</td>
<td>58.66</td>
<td>72.84</td>
<td>8.49</td>
<td>234.95</td>
</tr>
<tr>
<td>1990s</td>
<td>10.24</td>
<td>36.29</td>
<td>55.15</td>
<td>48.78</td>
<td>6.01</td>
<td>243.19</td>
</tr>
<tr>
<td>2000s</td>
<td>11.57</td>
<td>35.07</td>
<td>53.81</td>
<td>49.41</td>
<td>4.84</td>
<td>243.55</td>
</tr>
</tbody>
</table>
6.6 Chapter Summary

These frequency data points are not definitive with regard to the phenomenon of grammaticalization that is proposed to be happening; yet they indicate, along with the syntactic motivation and the patterns of analogy from Old English discussed above, that there is significant change between adjectives and adverbs in Modern English.

Unlike the pattern seen with speech-act adverbs, which I argue to be chiefly a type of lexicalization, the pattern found with the EEE adverbs is strongly aligned with the traditional definition of grammaticalization. The major shift seems to be twofold: first is the reanalysis of the RELATOR head in the extraposed IT clause. From being a traditional copula, it changes to the -ly affix that now connects the subject and predicate. For this new structure to be effective, though, the subordinate THAT-clause must be reanalyzed as the matrix.

Next we turn to Chapter 7 and to Standard Chinese. In this chapter we examine some data from Chinese and search for parallels in sentence adverb construction.
7.1 Introduction

7.1.1 Background for investigating Chinese adverbs.

This study is grounded in generative grammar (see Chapter 4), and as a result one of its aims is to suggest the universality (or at the very least, the broad nature) of its central proposals. These proposals are that adverbs (in general, and in particular sentence adverbs) frequently are morphologically reduced adverbial phrases and clauses; that speech-act adverbs are based on reduced speech commentary clauses; and that evaluative, evidential, and epistemic adverbs are based on reduced and reanalyzed copula clauses.

Although the focus of this dissertation is English, it is important to look outside the English/Germanic/Indo-European traditions for evidence in unrelated languages. The scope of such a search for this particular study is necessarily limited; however, the doctoral dissertation of Mingping Ji (2007) discusses many of the same kinds of topics, with somewhat different goals, about Standard (Mandarin) Chinese. This chapter makes use of some of her data to show similarities between English and Chinese in the combination and reduction of elements in modern adverbials that serve similar purposes.

7.1.2 Chapter organization.

This chapter is organized as follows: section 7.2 gives a very brief look at Chinese from a historical and typological point of view. Section 7.3 is focused on
Chinese speech-act adverbs, and section 7.4 turns to the EEE (evaluative, evidential, and epistemic) adverbs. Section 7.5 wraps up the chapter.

7.2. Modern Chinese

7.2.1 An outline of Chinese.

For purposes of this study, what we are calling “Chinese” is actually Standard Chinese, which is mostly based on the Beijing dialect of Mandarin. The Chinese dialects/languages are unified by a continuous cultural heritage (of more than 3,000 years) and a shared writing system that dates back nearly as far. Chinese is a somewhat distant cousin from Tibetan and Burmese languages and historically emerged from proto-Sino-Tibetan, but there has been some difficulty in reconstructing a proto-language because of the highly isolating nature of Chinese from the Old Chinese period (first millennium BCE) forward (Norman, 1988, pp. 12ff.).

Today, Standard Chinese is the language of instruction, of most culture, and of government both in the People’s Republic of China and in the Republic of China (Taiwan). Most residents of both nations have either native or a very high level of second-language competency in Mandarin Chinese. Many of the local dialects (as they are still called) remain in large numbers. As a result there is a large amount of bilingualism/bidialectalism across China and Taiwan.
7.2.2 The typology of Chinese.

There are several distinguishing characteristics about Chinese that make it an interesting subject of study. First, morphologically, Chinese is considered monosyllabic (Norman, 1988, p. 8). This is true of Chinese throughout its recorded history. In addition, consonant clustering is largely lost by the Middle Chinese period. Because there is a limited number of consonant-vowel combinations available to create these monosyllabic morphemes, Chinese is more dependent on phonological distinctions than other languages; like many Southeast Asian languages, Chinese is tonal (Norman, 1988, p. 9).

The monosyllabic morphological nature of Chinese allows for very little inflection, and the language is considered highly analytic or isolating, far more so than, for example, English. According to Norman (1988), this has been true throughout the history of the language (p. 10). In Modern Chinese, there are several polysyllabic words, but these are strings of monosyllabic morphemes that have been combined through compounding:

By far the greatest number of words in the dictionary are compounds… Compounding is the most productive process of word formation in the modern language. (1988, p. 156)

The word-phrase difference is rather difficult to distinguish as a result. This is not aided by the fact that the logographic writing system represents morphemes in a continuous visual stream with little indication of either syntactic or morphological structure (Norman, 1988, p. 156).

The organization of a Chinese sentence is SVO (similar to that of English) and the order of modifiers and heads is, for example, ADJ-N and ADV-V. This is
similar to Southeast Asian languages but in contrast to most Tibeto-Burman ones, which are SOV (as are Japanese and Korean). A further characteristic of Chinese that seems to be a regional, Sprachbund trait is the use of classifiers in determiner phrases, as in many Southeast Asian languages (Norman, 1988, p. 10).

In the next section, we will look at several examples of “lower adverbs” (i.e. VP-adverbs) to see whether tokens from a language such as Chinese, that is somewhat typologically marked, will support the proposal from Chapter 3: that “adverb” is, by and large, a morphological class of highly diverse source material that begins as other phrases (e.g. DP, CP, VP, AdjP) and lexicalizes into single-word units.

7.2.3 Chinese VP-adverbs.

Norman (1988) indicates that there are two broad super-categories of Chinese words (just as there are in Indo-European words): “full” words or shici (lexical categories) and “empty” words or xuci (functional categories). While nouns, verbs, and adjectives are shici, adverbs are xuci (p. 157).

However, the literature offers a distinction between regular adverbs and “adverbial adjuncts”, which are largely sourced from adjectives and the derivational suffix -de (p. 158). It is not clear whether scholars agree on the categorization of derived adverbial adjuncts—whether they remain “full” words because of their sources or whether their conversion to adverbs makes them “empty”. See example (1) below (all Chinese words are represented by romanized pinyin script):
Adverbial adjunct creation (from Norman, 1988, p. 158):

\[ man \text{ “slow” } \rightarrow manmanrde \text{ “slowly”} \]

The -de suffix, similar to the -ly suffix in English, does double duty—in this case by also indicating nominal “subordination”, including possession (Norman, 1988, p. 159). See (2) below.

(2) wode dianshiji

I-de television set

“my TV”

Ji (2007) gives examples of both VP-adverbs and sentence adverbs in her work. Although the adverb examples she gives are mostly made without the -de suffix, there are some patterns that emerge. In the below examples, there are a range of different Chinese VP-adverbs, separated according to traditional semantic type (glosses are mine and are based on the Chinese-English dictionary at www.mandarintools.com). The variety of phrase types that comprise these adverbs is quite similar to the English variety discussed in Chapter 3. (Please note that many of the adverbials were separated into two or more “words” in order to show the morphological components. In Ji’s original, the morphemes enclosed by the square brackets are written as a single word in pinyin.)
(3) Chinese adverbials of time:

a. [jin tian] “today”
   this/current day

b. shang [xing qi] “last week”
   last small time-period

(4) Chinese adverbials of place:

a. [zhe li] “here”
   this neighborhood

b. [jia li] “at home”
   home neighborhood

For the adverbs of place and time (3 and 4 above), there are elements that look quite familiar; namely, these are DPs acting as adverbials. As we saw in Chapter 3, nouns and adverbs can be closely related, particularly in these semantic sub-categories. In addition, these two types of adverbial phrase frequently use deictic elements (e.g. jin “this” in jin tian “today”; zhe “this” in zhe li “here”) to indicate the speaker’s relationship with the time/place discussed.
(5) Chinese adverbials showing aspect:

a. [gang gang] “just”
   just just

b. [yi jing] “already”
   already pass

(6) Chinese adverbials of frequency:

a. [jing chang] “often”
   pass often

b. [zong shi] “always”
   always be

c. [you shi] “sometimes”
   be time

Among the adverbials of aspect and frequency (Ji’s categories), there are three main patterns that emerge. First, there is the simple adverb (gang “just”) that is shown reduplicated, possibly for emphasis. Both of the other two patterns are verbal/clausal. One of the recognizable characteristics of Chinese is the lack of
morphological tense on verbs. Instead, there are often aspect particles and adverbials that are used to indicate speaker perspective.

The two verb types used in these adverbials are jing ("pass", "happen") and the copulas shi and you. Null-subject sentences are allowed in Chinese, so with jing, sentences can appear (in translation) like “It already happened that he saw the dog” or “It often happens that the dog jumps over the fence”. With the copula constructions, sentences like “It’s always that the sun rises in the east” or “It’s time (sometimes) he parks in the garage” are available.

(7) Chinese adverbials of manner:

a. [ren zhen] “carefully”
   recognize real/genuine

b. [qin fen] “diligently”
   diligent exert-oneself

For manner, again, these are lexicalized verb phrases. The interesting semantic extension is in (7a), where “recognize the genuine” is translated into “carefully”.

In addition to lexicalization/univerbation processes like the above, there are some adverbs (particularly degree modifiers) that have shown signs of grammaticalization. Chui (2000) discusses the case of the degree adverb hen ("very") and argues for a “morphologization” reading, as hen can be used in three
(layered) ways: 1) as an adverb modifying a predicate adjective (copulas are not necessary in predication constructions); 2) as a “phrasal clitic” that is bound to a larger phrasal component; and 3) as a constituent in a compound, with little of its original meaning remaining (p. 46).

The shift from phase 1 to phase 2 is signaled by the loss of prosodic stress, of autonomy, and of intensifier meaning. Yet there is a grammatical requirement that it remain in place modifying a predicate adjective (Chui, 2000, p. 50). The final step is integration into a compound as a morphological component, although the meaning is no longer based on its meaning at all. An example of this final step is the adverb *henshao* (“seldom”), where the two morphemes *hen* (“very”) and *shao* (“be less”) have merged, where *hen* cannot be replaced by another degree adverb (e.g. *feichang-shao* “strongly-be.less”), and where the meaning has become idiosyncratic (Chui, 2000, p. 52).

In the following sections we will move from VP and degree adverbs to look at Chinese sentence adverbs. Following the pattern established in the earlier chapters, we will first look at speech-act adverbs and then turn to the evaluative, evidential, and epistemic adverbs.

### 7.3 Chinese Speech-act Adverbs

Although I do not have access to Chinese corpora (in part this is because I am unable to read Chinese characters), my understanding is that speech-act adverbs in Chinese, just as in English, are the farthest leftward in the clause (Hui-ling Yang, p.c.). However, the focus of this chapter is less syntactic (outside of
categorization) and more morphological. So the question is: Do speech-act adverbs form in the same way in Chinese as in English? In Chapter 5, I suggest that speech-act adverbs are formed from full clauses (commenting on the act of speaking that is to come) that attach to the left periphery of the propositional clause. Once attached, they become more dependent through a process of ellipsis until the only remaining element from the original speech clause is the adverb.

In Ji (2007), she lists several speech-act adverbials that are analyzed as units. She adds a distinction among her adverbials that I did not, that between content speech-act adverbials (these comment on the truth value of the upcoming proposition) and format speech-act adverbials (these offer a promise to the listener that the proposition will be expressed in a particular structural manner). See example (8) below:

(8) Speech-act adverbials (from Ji, 2007, pp. 50-2)

a. Truth-oriented
   laoshi shuo       “honestly”
   shuo shihua      “honestly”
   tanbai shuo      “frankly”

b. Format-oriented
   jian er yan zhi  “briefly”
   juti di jiang/shuo “specifically”

233
All of these are based on verbal constructions, and all have evidently lexicalized from VPs based on speech-verbs. These verbs come from different periods in the history of the Chinese language.

(9) Verbs found in Chinese speech-act adverbs (from Ji, 2007)

\[
\begin{align*}
\text{shuo} & \quad \text{“say”—high frequency in Modern Chinese} \\
\text{yan} & \quad \text{“say”—high frequency in Old/Early Chinese} \\
\text{jiang} & \quad \text{“say”/”state”—found in formal registers}
\end{align*}
\]

The adverbials formed with these verbs combine in different ways with the elements that provide the semantic force. In some cases (10), the structure is “say” + DP. In others, there is an adverb as in English (11) and (12), or even a predicate adjective as in (13). Because this adjective is a predicate as well, it must be connected to the verb of speech by a coordinator.

(10) shuo shihua (“honestly”)

V N
say truth

(11) laoshi shuo (“honestly”)

ADV V
truly say
Very much like English speech-act adverbs, Ji (2007, p. 47) indicates that all of these can be combined with arguments (i.e. subjects) to form full clauses. Compare English “Frankly, you have to go” with “Frankly speaking, you have to go” with “I am telling you frankly that you have to go”. The Chinese adverbials are closer to the second English example, with V + ADV, rather than just ADV.

According to Ji, all of these phrases became sentence adverbials in the Modern Chinese era (2007, p. 52).

Next we’ll look at the evaluative, evidential, and epistemic adverbials in Standard Chinese.

7.4 Chinese EEE Adverbs

7.4.1 Evaluative adverbs.

As in English, evaluative adverbs are used by speakers to evaluate the quality of the proposition: whether it is a surprise to the speaker, whether the proposition is positive or negative, and whether it is strange or unusual. Cinque’s
(1999) hierarchy indicates that evaluative adverbs follow speech-act adverbs in the clause but precede all other mood/modal adverbs and the tense head nodes in the Chinese sentence. Hsieh (2005) offers support for Cinque’s order:

… future time cannot block the use of evaluative modals even though their function is to modify a proposition believed to be true. In fact, the truth of a proposition is not governed by the time expression appearing in the sentence. (p. 51)

In other words, the truth condition of the proposition takes scope over the expression of future time, even though a future adverbial implies an unknown state (irrealis). This can be seen in example (14) below:

(14) Nanguai ta mingnian yao huan gongzuo
no-wonder he next-year going-to change job

Xingkui ta mingnian yao huan gongzuo
Fortunately he next-year going-to change job

“No wonder } he is going to change his job next year.”
“Fortunately ]

(Hsieh, 2005, p. 51)

Ji (2007) follows Swan (1988) in separating evaluative adverbs that encompass commentaries about good vs. bad (generally phrased as the “luck” adverbs, as most of them are phrased in terms of luck or fortune) and the normal vs. strange
adverbs. In addition, she characterizes Chinese evaluatives based on syntactic flexibility:

(15) Evaluative adverbs in Chinese (from Ji, 2007):

a. Closed set:

LUCK

\[ xing+kui \text{ ("fortunately") = "lucky+deficit" } \]
\[ duo+kui \text{ ("fortunately") = "much+deficit" } \]
\[ hai+hao \text{ ("luckily") = "also+good" } \]
\[ hao+zai \text{ ("what’s good is…") = "good+what" } \]

NORMS

\[ jing+ran \text{ ("unexpectedly") = "unexpected+(ADV)" } \]

b. Open set:

LUCK

\[ bu+xing \text{ ("unfortunately") = "not+lucky" } \]
\[ xing+yun \text{ ("fortunately") = "lucky+fortune" } \]

NORMS

\[ bu+liao \text{ ("unexpectedly") = "not+know" } \]
\[ chu \text{ ren yi+liao ("unexpectedly") = "occur man idea+know" } \]
\[ mei xiang dao \text{ ("unexpectedly") = "not think-of" } \]
\[ qi+guai \text{ ("oddly") = "strange+queer" } \]
The difference between the sets is that the closed set does not allow for adding *de shi*, and no degree modifiers are allowed, while the open set allows for both:

(16) \[ \text{de shi} = (\text{adverbializer/nominalizer} + \text{“be”}) \]

(17) \[ \text{degree modifiers: } \text{hen (“very”)} \]
\[ \text{geng (“more”)} \]
\[ \text{zui (“most”)} \]

As elsewhere, there are different strategies in the lexicalization that have occurred among evaluative adverbials in Chinese. There is redundancy in combinations like *xingyun* (“lucky fortune”) and *qiguai* (“strange queerness”). Two of the open class have apparently lexicalized from VPs:

(18) \[ \text{mei xiang dao} \quad (\text{“unexpectedly”}) \]
\[ \text{Neg V} \]
\[ \text{not think-of} \]

(19) \[ \text{bu liao} \quad (\text{“unexpectedly”}) \]
\[ \text{Neg V} \]
\[ \text{not know} \]

238
For an adverbial such as *unexpectedly*, which is based on a psych verb (*expect*/
*know*), the VP origins make sense. Unlike a more static evaluative adverb like *
luckily*, *unexpectedly* indicates a sudden change of understanding (similar to that
in *surprisingly*). However, one aspect of these adverbials that is quite different
from similar adverbs in English is the relation of the adverbial to a copula clause.

In English, there is a progression of derivational morphology that allows a
word like “expect” to take on a series of affixes and verbal relators:

(20)  expect > (V)
      expect-ed > (ADJ PARTICLE)
      un-expect-ed > (ADJ PARTICLE)
      is-un-expect-ed > (COPULA-ADJ PARTICLE)
      un-expect-ed-ly (ADV)

This is not the case in Chinese, and as a result there are far fewer allowable steps:

(21)  liao > (V)
      bu-liao (V)
      bu-liao-de-shi (V-ADV PARTICLE-COPULA)

Therefore, although there is a parallel structure in that there are AdjPs, DPs, and
VPs acting as AdvPs in the Mood<sub>evaluative</sub> head of the Chinese sentence, these do
not convert to true “adverbs” as they do in English. Nevertheless, the parallel role of copulas in the two processes indicates a strong connection between them.

In the next section, we move on to evidential adverbs in Chinese.

### 7.4.2 Evidential adverbs.

Evidential adverbs in Chinese, as in English, are used by speakers to indicate what sort of support they have for the truth value of the proposition they are offering. Chinese evidentials have parallel semantic concerns to English ones:

(22) Evidential adverbs (from Ji, 2007, p. 81)

<table>
<thead>
<tr>
<th>Chinese</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>xianran</td>
<td>“evidently”</td>
</tr>
<tr>
<td>hen mingxian de</td>
<td>“quite clearly”</td>
</tr>
<tr>
<td>jushuo</td>
<td>“allegedly”</td>
</tr>
</tbody>
</table>

These adverbials show signs of being descended from adjectives and adjective-noun combinations, similar to those found in English.

(23) xian ran (“obviously”)

    obvious state

(24) hen ming xian de (“very obviously”) (must include degree mod.)

    very clear obvious ADV

240
Xian is, however, originally a verb, meaning “to bring glory/prestige”. In (25), the verb shuo, last seen in the speech-act adverbs, is found in the evidentials for knowledge that has been received through speech.

(25)  ju  shuo  (“reportedly”)  
      depend-on  say

7.4.3 Epistemic adverbs.

Epistemic adverbs essentially indicate probability of the proposition involved. All of the adverbs listed by Ji (2007) can be translated into English as “probably”. However, they arise from a number of origins, including spatial directionals, degree modifiers, and full verbs. See below:

(26)  Epistemic adverbs (from Ji, 2007)

       huoxu  “probably”
       yexu  “probably”
       dagai  “probably”
       kongpa  “probably”

(27)  huo  xu  (“probably”)  
      Coord (“or” EarlyC)  <zheyang (“this way”)  
      “or this way”
(28) \textit{ye xu} (‘probably’)
also this way

“also this way”

(29) \textit{da gai} (‘probably’)
<degree modifier (‘approximately’ with quantities EarlyC)

(30) \textit{kong pa} (‘probably’)
<V (‘fear’ EarlyC)

\textit{Kongpa} is the only epistemic adverb to allow a clausal reading with an argument, most likely because of its past form as a verb (Ji, 2007, p. 107).

7.5 Chapter Summary

Adverbials in Chinese are composed of essentially the same material as those found in English: Determiner Phrases, Adjective Phrases (with or without derivational morphology), and Verb Phrases. Verb Phrases are generally found in place of Complementizer Phrases (as are found in English), since as van Gelderen (2004, p. 4) notes, Chinese does not make a great deal of use of the CP, but uses the VP far more.

Because the combinatory actions in Chinese are far less fusional over time than in English and other Indo-European languages, the morphemes can still be easily pulled apart, as we did in the examples above. Many of these are in fact
written separately in pinyin (and word boundaries in general are somewhat unclear).

In spite of the differences, the evidence from Chinese generally supports the contentions forwarded in this dissertation. Such evidence supports not only the proposal that adverb is a morphological category, but also the idea that speech-act adverbs are reduced conditional speech-verb clauses, and (to a lesser extent) that evaluative, evidential, and epistemic adverbs are based on copula clauses.

In the final chapter, we will briefly revisit the various points that have been made in Chapters 2 through 7 and offer suggestions for future research that may stem from this work.
Chapter 8

CONCLUSIONS

8.1 Chapter Summaries

This dissertation has been primarily focused on answering three research questions I posed in Chapter 1. First I asked what, exactly, the criteria are for adverbs in English. The second question asked how sentence adverbs arise in the history of English. And the third question asked whether patterns found in my research can be extended to other languages. I believe these three questions have been addressed in the course of this study.

In Chapter 2, I examined, from a critical perspective, the theoretical and phenomenological aspects of grammaticalization studies. I established that, for this dissertation, I would adopt older, more basic definitions of both grammaticalization and lexicalization that are based primarily on morphology and syntax. In so doing, I expressed a sense of caution about Grammaticalization Theory and its codification of unidirectionality, about reduction of lexicalization phenomena to a minimal, major-category-only level, and about expansion of grammaticalization so that it has far wider reach than it did previously.

In addition, I used Chapter 2 to explore some of the pragmatic frameworks that work well with speech-act, evaluative, evidential, and epistemic adverbs. These included Politeness Theory, Relevance theory, and questions of stance.

Chapter 3 provided me with an opportunity to suggest a reunderstanding of the adverb as a category. In this chapter, I considered the semantic criteria for adverbs, dating from 100 BCE and Dionysios Thrax, and extending to the modern
era. I suggested that an approach that is diachronic in scope and that focuses on adverbs as *lexicalized adverbials* arising from a number of different source materials would be more effective than an approach that tries to explain the diversity of “adverbs” in other ways. Adverbs in this model are essentially the relics of former phrases and clauses that have fused, have been reduced by ellipsis, and/or have been constructed by a derivational morphology whose source is a reduced noun (-ly).

In Chapter 4, I turned more directly to models of syntax, centered on the frameworks from the generative grammar tradition. First I explored a brief history of generative grammar, primarily as developed by Noam Chomsky. After presenting a history and some terminology, I then turned to questions of adverbs and their role in a generative framework. Until recently, that role has been quite limited. Finally, I considered generative approaches to adverbs from Cinque (the cartographic approach) and Ernst (the semantic approach). I decided to limit my study to cartography (despite some concerns over the cumbersome nature of the structure) because of the primacy of syntax in this model.

Also in Chapter 4, I examined some generative frameworks for studying diachronic linguistics, primarily grammaticalization. I compared van Gelderen (2004) and Roberts and Roussou (2003) and found that van Gelderen’s focus of the motivation of syntactic change fit my research model more closely, although Roberts and Roussou offered important insights.

Chapter 5 then moved on to the study of speech-act adverbs, in the form of *frankly*. After looking at some preliminary information about the word, I used
corpus data (particularly from the *TIME* Corpus) to show the shift in its use
during the twentieth century. In the last section of the chapter, I proposed various
stages of reanalysis that would fit with a lexicalization–grammaticalization
combination and showed these stages through syntactic tree representations.

In Chapter 6, I treated evaluative, evidential, and epistemic adverbs in a
similar way. Because they all share an underlying copula structure, I suggested a
reanalysis schema that utilizes synchronic work by den Dikken on relator phrases.
The reanalysis involves a shift from a [copula+adjective] to an [adjective+-ly]
paradigm, where the adverbial suffix acts as a copula-like relator that allows for
predication similar to that in the underlying structure. Data from the COHA added
to the analysis by indicating a marked shift from adjectival to adverbial use of
each of the subcategories: *unfortunately*, *obviously*, and *probably*.

Finally, in Chapter 7, I extended the grammaticalization and lexicalization
models I have proposed, along with the concept of adverb as relic, to another
language (Chinese). By examining a language with no genetic link to English, I
was able to suggest (with a limited amount of data) that the patterns of complex-
to-simplex that play such a major role in English adverbs are also active in
Chinese. The main distinction is that Chinese, partly because of a writing system
that is morphemic, does not have fusion of morphemes to the same extent as is
found in English.
8.2 Contributions to the Field

This dissertation is not the first to notice a tendency for sentence adverbs to arise from larger syntactic structures. Swan (1988) proposed such a solution for speech-act adverbs (pp. 64ff.). Fischer (2007) is even more explicit in her discussion of such sentence adverbs:

From [the] historical and comparative details, it becomes clear that the wide-scope sentence adverbial/pragmatic marker must originally have been placed outside the main clause, in the form of a prepositional phrase, a reduced clause or a predicative clause followed by a “that”-complement. …These clauses function as sentence adverbials and have scope over the “that”-clause. When reduced to an adverbial, they would resemble other adverbials but retain their scope over the original “that”-clause which now becomes the main clause. (p. 296)

These are very much the same kinds of claims that this dissertation makes. However, one of the significant differences this study offers is its use of a synthesis of diachronic and generative models. Motivation for the type of change that Fischer alludes to must include interplay between the cross-purposes of economy and innovation that humans bring to any sort of endeavor. In this dissertation I propose an articulation of that interplay.

In addition, the embedding of lexicalization- and grammaticalization-related processes within the bounds of a category (the adverb in English) that seems peculiarly open to such phenomena, is an important contribution by this study to the science of linguistics. Few studies have suggested that a “lexical” category could arise almost entirely from historical processes. The fact that parallel phenomena are carried out in Chinese (to the extent allowable by this
language’s limited morphological component) indicates that this characterization of adverbs may in fact be evidence of a universal tendency.

8.3 Suggested Future Research Directions

Although in this dissertation I offer some possible explanations for the rise of certain sentence adverbs in English, there is a great deal more that can be researched in this area. It is important, for example, to examine other types of reanalysis models to see what insights can be gleaned. In addition, the models I have suggested should be tested in a wider range of languages (and varieties) to see what may or may not work, cross-linguistically. A sample of two languages is not sufficient.

Other phenomena can and should be examined, including other adverbs but also including adpositions, particularly complex (non-structural) ones. There is a similarity between many complex prepositions in English and many adverbs. Finally, I am very much interested in the category-based connection between adverbs and nouns and would like to explore that interface. It is clear that adverbs connect to many other word categories and can be thought of as central to the diachronic study of language.
REFERENCES


