Rewriting rules, derivations and underlying representations are enduring characteristics of generative phonology. In this book, John Coleman argues that they are unnecessary. The expressive resources of context-free unification grammars are sufficient to characterise phonological structures and alternations.

According to this view, all phonological forms and constraints are partial descriptions of surface representations. This framework, now called Declarative Phonology, is based on a detailed examination of the formalisms of feature theory, syllable theory and the leading varieties of non-linear phonology. Dr Coleman illustrates this with two extensive analyses of the phonological structure of words in English and Japanese. As Declarative Phonology is surface-based and highly restrictive, it is consistent with findings in cognitive psychology and amenable to straightforward computational implementation.
**In this series**

<table>
<thead>
<tr>
<th>Page</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>Michael S. Rochemont and Peter W. Culicover</td>
<td>English focus constructions and the theory of grammar</td>
</tr>
<tr>
<td>53</td>
<td>Philip Carr</td>
<td>Linguistic realities: an autonomist metatheory for the generative enterprise</td>
</tr>
<tr>
<td>54</td>
<td>Eve Sweetser</td>
<td>From etymology to pragmatics: metaphorical and cultural aspects of semantic structure</td>
</tr>
<tr>
<td>55</td>
<td>Regina Blass</td>
<td>Relevance relations in discourse: a study with special reference to Sissala</td>
</tr>
<tr>
<td>56</td>
<td>Andrew Chesterman</td>
<td>On definiteness: a study with special reference to English and Finnish</td>
</tr>
<tr>
<td>57</td>
<td>Allessandra Giorgi and Giuseppe Longobardi</td>
<td>The syntax of Noun Phrases: configuration, parameters and empty categories</td>
</tr>
<tr>
<td>58</td>
<td>Monik Charette</td>
<td>Conditions on phonological government</td>
</tr>
<tr>
<td>59</td>
<td>M. H. Klaiman</td>
<td>Grammatical voice</td>
</tr>
<tr>
<td>60</td>
<td>Sarah M. B. Fagan</td>
<td>The syntax and semantics of middle constructions: a study with special reference to German</td>
</tr>
<tr>
<td>61</td>
<td>Anjum P. Saleem</td>
<td>Universal grammar and language learnability</td>
</tr>
<tr>
<td>62</td>
<td>Stephen R. Anderson</td>
<td>A-Morphous morphology</td>
</tr>
<tr>
<td>63</td>
<td>Lesley Stirling</td>
<td>Switch reference and discourse representation</td>
</tr>
<tr>
<td>64</td>
<td>Henk J. Verkuyl</td>
<td>A theory of aspectuality: the interaction between temporal and atemporal structure</td>
</tr>
<tr>
<td>65</td>
<td>Eve V. Clark</td>
<td>The lexicon in acquisition</td>
</tr>
<tr>
<td>66</td>
<td>Anthony R. Warner</td>
<td>English auxiliaries: structure and history</td>
</tr>
<tr>
<td>67</td>
<td>P. H. Matthews</td>
<td>Grammatical theory in the United States from Bloomfield to Chomsky</td>
</tr>
<tr>
<td>68</td>
<td>Liljana Progovac</td>
<td>Negative and positive polarity: a binding approach</td>
</tr>
<tr>
<td>69</td>
<td>R. M. W. Dixon</td>
<td>Ergativity</td>
</tr>
<tr>
<td>70</td>
<td>Yan Huang</td>
<td>The syntax and pragmatics of anaphora</td>
</tr>
<tr>
<td>71</td>
<td>Knud Lambrecht</td>
<td>Information structure and sentence form: topic, focus, and the mental representations of discourse referents</td>
</tr>
<tr>
<td>72</td>
<td>Luigi Burzio</td>
<td>Principles of English stress</td>
</tr>
<tr>
<td>73</td>
<td>John A. Hawkins</td>
<td>A performance theory of order and constituency</td>
</tr>
<tr>
<td>74</td>
<td>Alice C. Harris and Lyle Campell</td>
<td>Historical syntax in cross-linguistic perspective</td>
</tr>
<tr>
<td>75</td>
<td>Liliane Haegeman</td>
<td>The syntax of negation</td>
</tr>
<tr>
<td>76</td>
<td>Paul Gorrell</td>
<td>Syntax and parsing</td>
</tr>
<tr>
<td>77</td>
<td>Guglielmo Cinque</td>
<td>Italian syntax and universal grammar</td>
</tr>
<tr>
<td>78</td>
<td>Henry Smith</td>
<td>Restrictiveness in case theory</td>
</tr>
<tr>
<td>79</td>
<td>D. Robert Ladd</td>
<td>Intonational phonology</td>
</tr>
<tr>
<td>80</td>
<td>Andrea Moro</td>
<td>The raising of predicates: predicative noun phrases and the theory of clause structure</td>
</tr>
<tr>
<td>81</td>
<td>Roger Lass</td>
<td>Historical linguistics and language change</td>
</tr>
<tr>
<td>82</td>
<td>John M. Anderson</td>
<td>A notional theory of syntactic categories</td>
</tr>
<tr>
<td>83</td>
<td>Bernd Heine</td>
<td>Possession: cognitive sources, forces and grammaticalization</td>
</tr>
<tr>
<td>84</td>
<td>Nomi Erteschkishir</td>
<td>The dynamics of focus structure</td>
</tr>
<tr>
<td>85</td>
<td>John Coleman</td>
<td>Phonological representations: their names, forms and powers</td>
</tr>
</tbody>
</table>

**Supplementary volumes**

<table>
<thead>
<tr>
<th>Page</th>
<th>Authors</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>76</td>
<td>Liliane Haegeman</td>
<td>Theory and description in generative syntax: A case study in West Flemish</td>
</tr>
<tr>
<td>77</td>
<td>A. E. Backhouse</td>
<td>The lexical field of taste: A semantic study of Japanese taste terms</td>
</tr>
<tr>
<td>78</td>
<td>Nikolaus Ritt</td>
<td>Quantity adjustment: Vowel lengthening and shortening in early Middle English</td>
</tr>
</tbody>
</table>

Earlier issues not listed are also available.
Phonological and morphological representations of the phrase

*phonological representations*
Contents

List of figures x
List of tables xiii
Preface and acknowledgements xv
Key to symbols xvii

1 Introduction 1
1.1 Rules and representations in linguistic theory 1
1.2 Rules and representations in phonological theory 2
1.3 Imperative, declarative and interrogative approaches to grammar 7
1.4 Linguistic representations and cognitive representations 11
1.5 Notation 12
1.6 Overview 15

2 Segmental representations and their phonetic interpretation 18
2.1 Introduction 18
2.2 Historical context of phonetic notations 19
2.3 Segmental phonetic notation 21
2.4 Notation and denotation 22
2.5 The syntax of IPA transcriptions 26
2.6 The semantics of IPA transcriptions 29
2.7 Another example: Sweet’s Organic Alphabet 38
2.8 Feature-matrix representation 39
2.9 Segmentation and temporal overlap 41
2.10 Superweak segmental interpretation 43
2.11 Conclusion 45

3 Segmental and transformational phonology 46
3.1 Introduction 46
3.2 Historical context 47
### List of contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3</td>
<td>Phonemic phonology</td>
<td>48</td>
</tr>
<tr>
<td>3.4</td>
<td>Transformational-generative phonology</td>
<td>66</td>
</tr>
<tr>
<td>3.5</td>
<td>The excessive power of transformational grammars</td>
<td>77</td>
</tr>
<tr>
<td>3.6</td>
<td>Phrase-markers and representations of derivations</td>
<td>92</td>
</tr>
<tr>
<td>3.7</td>
<td>Conclusion</td>
<td>108</td>
</tr>
<tr>
<td>4</td>
<td>Non-linear phonological representations in contemporary generative phonology</td>
<td>110</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>110</td>
</tr>
<tr>
<td>4.2</td>
<td>Origins of non-linear generative phonology</td>
<td>111</td>
</tr>
<tr>
<td>4.3</td>
<td>Non-linear phonological representation</td>
<td>113</td>
</tr>
<tr>
<td>4.4</td>
<td>Constituency, heads and modifiers in phonological theory</td>
<td>146</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusion</td>
<td>164</td>
</tr>
<tr>
<td>5</td>
<td>Phonological representations in Declarative Phonology</td>
<td>165</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>165</td>
</tr>
<tr>
<td>5.2</td>
<td>Historical context</td>
<td>166</td>
</tr>
<tr>
<td>5.3</td>
<td>A non-derivational approach to phonology</td>
<td>167</td>
</tr>
<tr>
<td>5.4</td>
<td>Phonological ‘processes’ in Declarative Phonology</td>
<td>177</td>
</tr>
<tr>
<td>5.5</td>
<td>A formalism for phonological representation</td>
<td>184</td>
</tr>
<tr>
<td>5.6</td>
<td>Phonetic interpretation, 1: temporal interpretation</td>
<td>216</td>
</tr>
<tr>
<td>5.7</td>
<td>Phonetic interpretation, 2: parametric interpretation</td>
<td>229</td>
</tr>
<tr>
<td>5.8</td>
<td>Conclusion</td>
<td>232</td>
</tr>
<tr>
<td>6</td>
<td>A declarative analysis of Japanese words</td>
<td>233</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>233</td>
</tr>
<tr>
<td>6.2</td>
<td>Japanese syllable-structure</td>
<td>234</td>
</tr>
<tr>
<td>6.3</td>
<td>Moras</td>
<td>235</td>
</tr>
<tr>
<td>6.4</td>
<td>Syllables and moras</td>
<td>236</td>
</tr>
<tr>
<td>6.5</td>
<td>Structure of CV moras</td>
<td>251</td>
</tr>
<tr>
<td>6.6</td>
<td>Regressive palatalisation</td>
<td>252</td>
</tr>
<tr>
<td>6.7</td>
<td>Phonetic interpretation of CV moras</td>
<td>258</td>
</tr>
<tr>
<td>6.8</td>
<td>Consonant ‘clusters’ in onset and coda</td>
<td>265</td>
</tr>
<tr>
<td>6.9</td>
<td>Concluding remark</td>
<td>267</td>
</tr>
<tr>
<td>7</td>
<td>A declarative analysis of English words</td>
<td>269</td>
</tr>
<tr>
<td>7.1</td>
<td>A grammar of English monosyllables</td>
<td>269</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>-----------</td>
<td>--------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>7.2</td>
<td>A grammar of English words</td>
<td>307</td>
</tr>
<tr>
<td>7.3</td>
<td>Lexical and postlexical phonology</td>
<td>323</td>
</tr>
<tr>
<td>7.4</td>
<td>Envoi</td>
<td>324</td>
</tr>
</tbody>
</table>

References 326
Index 344
Figures

Phonological and morphological representations of the phrase *phonological representations* frontispiece

2.1 Sets and relations in phonetic notation page 25
2.2 ‘Relaxed’ segmentation 37
2.3 ‘Strict’ segmentation 37
2.4 Strict segmental interpretation of a componential representation 42
2.5 Relaxed segmental interpretation of a componential representation 42
2.6 Coarticulation modelled by temporal overlap (coproduction) 44
3.1 Componential phonological representations 59
3.2 An ‘analytic transcription’ of *Joe took father’s shoe bench out* 60
3.3 Distinctive-feature transcription 61
3.4 A top–down, depth-first, left-to-right derivation 97
3.5 A bottom–up, height-first, right-to-left derivation 98
3.6 Derivation-graph of a$^3$b$^3$c$^3$ 101
4.1 A paddle-wheel graph 130
4.2 A cubic graph 133
4.3 Non-planar graphs 137
4.4 Two embeddings of a planar graph 138
4.5 A planar graph 139
4.6 A planar embedding of a paddle-wheel graph 142
4.7 Non-planar Autosegmental Phonological representations 143
4.8 $K_{3,3}$ subgraphs of figure 4.7 144
4.9 Necessarily non-planar Autosegmental Phonological representations 144
4.10 Representation of minimal syllable structure in Dependency Phonology 148
4.11 Representation of branching syllable structure in Dependency Phonology 148
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.12</td>
<td>A compound structure</td>
<td>150</td>
</tr>
<tr>
<td>4.13</td>
<td>'Germanic' word structure</td>
<td>150</td>
</tr>
<tr>
<td>4.14</td>
<td>'Latinate' morphological structure</td>
<td>150</td>
</tr>
<tr>
<td>4.15</td>
<td>'Latinate' metrical structure</td>
<td>151</td>
</tr>
<tr>
<td>4.16</td>
<td>Pierrehumbert and Beckman's proposed surface representation for a Japanese sentence</td>
<td>162</td>
</tr>
<tr>
<td>4.17</td>
<td>Figure 4.16 translated into a UPSG representation</td>
<td>163</td>
</tr>
<tr>
<td>5.1</td>
<td>Linguistic forms and interpretation</td>
<td>175</td>
</tr>
<tr>
<td>5.2</td>
<td>Declarative characterisation of allophony</td>
<td>178</td>
</tr>
<tr>
<td>5.3</td>
<td>A declarative view of assimilation</td>
<td>179</td>
</tr>
<tr>
<td>5.4</td>
<td>Declarative (coproduction) model of coarticulation</td>
<td>179</td>
</tr>
<tr>
<td>5.5</td>
<td>Declarative characterisation of epenthesis</td>
<td>181</td>
</tr>
<tr>
<td>5.6</td>
<td>Phonetic characterisation of elision</td>
<td>182</td>
</tr>
<tr>
<td>5.7</td>
<td>Declarative characterisation of metathesis</td>
<td>183</td>
</tr>
<tr>
<td>5.8</td>
<td>Hierarchical representation of syllable-structure</td>
<td>194</td>
</tr>
<tr>
<td>5.9</td>
<td>Ambisyllabicity represented by a re-entrant graph (butter)</td>
<td>196</td>
</tr>
<tr>
<td>5.10</td>
<td>Assimilation represented by a re-entrant graph (it put)</td>
<td>196</td>
</tr>
<tr>
<td>5.11</td>
<td>A place-of-articulation category</td>
<td>197</td>
</tr>
<tr>
<td>5.12</td>
<td>A category-valued feature, [consonantal], and its value</td>
<td>198</td>
</tr>
<tr>
<td>5.13</td>
<td>Graph of the category /h/</td>
<td>198</td>
</tr>
<tr>
<td>5.14</td>
<td>Graph of the coda /h/</td>
<td>201</td>
</tr>
<tr>
<td>5.15</td>
<td>Graph of the syllable full</td>
<td>202</td>
</tr>
<tr>
<td>5.16</td>
<td>A re-entrant dag</td>
<td>204</td>
</tr>
<tr>
<td>5.17</td>
<td>Two dags</td>
<td>205</td>
</tr>
<tr>
<td>5.18</td>
<td>Two trees joined by a constraint</td>
<td>205</td>
</tr>
<tr>
<td>5.19</td>
<td>Part of a non-segmental phonological representation of the word cousin</td>
<td>206</td>
</tr>
<tr>
<td>5.20</td>
<td>Feature propagation</td>
<td>207</td>
</tr>
<tr>
<td>5.21</td>
<td>Temporal structure of a context-free string of words</td>
<td>218</td>
</tr>
<tr>
<td>5.22</td>
<td>Concatenative interpretation (b) of syllable-structure (a)</td>
<td>221</td>
</tr>
<tr>
<td>5.23</td>
<td>Timing of $v_1cv_2$ transitions</td>
<td>222</td>
</tr>
<tr>
<td>5.24</td>
<td>(a) Structure and (b) temporal interpretation of nucleus and coda</td>
<td>223</td>
</tr>
<tr>
<td>5.25</td>
<td>Temporal interpretation (b) of rime structure (a)</td>
<td>223</td>
</tr>
<tr>
<td>5.26</td>
<td>Temporal interpretation (b) of syllable, rime and onset constituents (a)</td>
<td>224</td>
</tr>
<tr>
<td>5.27</td>
<td>Non-concatenative temporal interpretation of syllable constituents</td>
<td>224</td>
</tr>
<tr>
<td>5.28</td>
<td>Simpson's analogy</td>
<td>226</td>
</tr>
</tbody>
</table>
List of figures

5.29 Labiality and retroflexion of /r/ are present throughout /tr/ 226
5.30 Labiality and retroflexion of /r/ are present throughout /spr/ 226
5.31 Structure of obstruent + glide onsets 226
5.32 Possible onset structures: (a) doubly branching; (b) ternary branching 227
5.33 ‘Unreduced’ forms of prefixes 228
5.34 ‘Reduced’ forms of prefixes 229
5.35 Second-formant motion for /ou/ 231
5.36 Second-formant motion for /gou/ 231
7.1 Relationships between morphosyntactic, semantic and morphophonological information 275
## Tables

<table>
<thead>
<tr>
<th>3.1</th>
<th>‘CV’ and ‘CyV’ moras</th>
<th>page 51</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Phonetic variability of Japanese /i/ and /u/</td>
<td>52</td>
</tr>
<tr>
<td>3.3</td>
<td>The Chomsky hierarchy of generative grammars</td>
<td>79</td>
</tr>
<tr>
<td>7.1</td>
<td>Summary of the vowel categories</td>
<td>280</td>
</tr>
</tbody>
</table>
Preface and acknowledgements

I shall begin in antiquated style, if I may, with an apology to the reader for burdening them and the academic libraries of the world with another pound of paper and probably many thousand more words of text than the world really needs. The thesis on which this monograph is based was too long: I have tried to be more judicious in revising it. In order to explain my purpose, it is perhaps easiest to start with a little personal history.

In the first half of the 1980s, as I began graduate study, the Chomskyan linguistic mainstream was shaken by a revisionist, non-transformational approach, Generalised Phrase-Structure Grammar (GPSG). It was not the first plausible challenger to transformational grammar, but for many it was the most credible. One part of GPSG's approach was the articulation of a detailed theory of linguistic representations employing features and structures of features. How attractive, then, was the prospect of extending the approach to phonology! I was not the only graduate student who saw the research opportunity in that ambition: at the University of Edinburgh at that time, Ewan Klein and a number of his students, Jo Calder, Jim Scobbie, Steven Bird and Michael Broe, were working with a similar aim in mind. We met at times during our graduate years, and though we did not exactly work collaboratively, the exchange of ideas, papers and eventually theses led to the approach now called Declarative Phonology (Scobbie, Coleman and Bird 1996). As well as offering this work to you, I commend to you also Calder and te Lindert (1987), Scobbie (1991), Bird (1995) and Broe (1993).

In addition to the references and credits in the text, I would like to acknowledge with thanks the help and feedback I have received from many people. Special thanks are due to John Local, for the years of tuition and patient guidance that he gave me, and for the combative, iconoclastic and Cartesian (i.e. sceptical) attitudes that he passed on. I am also grateful to him for agreeing to the incorporation of large parts of a co-authored paper (Coleman and Local 1991) into chapter 4. John Kelly provided guidance and comments on phonetic theory and prosodic phonology. Steve Harlow
introduced me to formal semantics, gave useful comments on chapters 2 and 3, and lent me a copy of McCawley (1968) for six years. I am indebted to my Japanese informants Rika Shin and Mr and Mrs Nakai, and to my Japanese teacher Taeko Crump.

Most of the ideas and arguments in this thesis have appeared in papers, draft papers, conference abstracts, seminar handouts and postings on electronic mailing lists. For their comments on various sections of the draft, I am grateful to Steven Bird, Ted Briscoe, Erik Fudge, Patrick Griffiths, Dick Oehrle, Richard Ogden, Andrew Slater, Anthony Warner and Pete Whitelock. I am particularly grateful to two Cambridge University Press referees, and for the referees of those papers, parts of which have been incorporated into the text of this work. I am grateful to the Cambridge linguistics editors, Catherine Max, Joanna West, Judith Ayling and Hilary Gaskin, for their work in seeing this book through the press, and for the work of the others on the editorial side. This includes my colleague Rebecca Posner, one of the general editors, for her encouragement and useful advice, and Jenny Potts, the copy-editor.

I thank D. Reidel Publishing Company for their permission to reproduce parts of Coleman and Local (1991) in chapter 4.
# Key to symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>{ ... }</code></td>
<td>set braces; disjunction of phonological terms</td>
</tr>
<tr>
<td><code>[ ... ]</code></td>
<td>feature-value pair; sequence; phone; string of phones</td>
</tr>
<tr>
<td><code>[ ... ]</code></td>
<td>feature-value structure</td>
</tr>
<tr>
<td><code>/ ... /</code></td>
<td>phoneme; string of phonemes</td>
</tr>
<tr>
<td><code>[x]</code></td>
<td>denotation of x</td>
</tr>
<tr>
<td>`</td>
<td>x</td>
</tr>
<tr>
<td><code>⊠</code></td>
<td>concatenation operation</td>
</tr>
<tr>
<td><code>∩</code></td>
<td>intersection</td>
</tr>
<tr>
<td><code>&amp;</code></td>
<td>logical 'and'</td>
</tr>
<tr>
<td><code>∨</code></td>
<td>logical 'or'</td>
</tr>
<tr>
<td><code>⊕</code></td>
<td>logical exclusive 'or'</td>
</tr>
<tr>
<td><code>¬</code></td>
<td>logical 'not'</td>
</tr>
<tr>
<td><code>@</code></td>
<td>temporal 'AT'</td>
</tr>
<tr>
<td><code>○</code></td>
<td>(partial) temporal overlap</td>
</tr>
<tr>
<td><code>θ</code></td>
<td>non-commutative cocatenation</td>
</tr>
<tr>
<td><code>⊕⁺</code></td>
<td>commutative cocatenation</td>
</tr>
<tr>
<td><code>∪</code></td>
<td>Unification</td>
</tr>
<tr>
<td><code>≈</code></td>
<td>contemporaneous ('while')</td>
</tr>
<tr>
<td><code>≤</code></td>
<td>precedence relation</td>
</tr>
<tr>
<td><code>≤₀</code></td>
<td>temporal adjacency</td>
</tr>
<tr>
<td><code>&lt;</code></td>
<td>strict autosegmental precedence</td>
</tr>
<tr>
<td><code>≤</code></td>
<td>autosegmental precedence</td>
</tr>
<tr>
<td><code>□</code></td>
<td>inclusion (of intervals or events)</td>
</tr>
<tr>
<td><code>≤</code></td>
<td>subsumption</td>
</tr>
<tr>
<td><code>⁺</code></td>
<td>(superscript) transitive closure</td>
</tr>
<tr>
<td><code>∗</code></td>
<td>(superscript) reflexive transitive closure</td>
</tr>
<tr>
<td><code>→</code></td>
<td>(in string grammars) rewrites as; (in directed graphs) arc</td>
</tr>
<tr>
<td><code>←</code></td>
<td>mapping, function; (in directed graphs) arc</td>
</tr>
<tr>
<td><code>↔</code></td>
<td>pair of arcs in an equivalence relation; logical equivalence</td>
</tr>
<tr>
<td><code>⇒</code></td>
<td>directly derives</td>
</tr>
<tr>
<td><code>&lt;=$</code></td>
<td>graph ordering</td>
</tr>
<tr>
<td><code>■</code></td>
<td>end of a proof</td>
</tr>
<tr>
<td><code>⇒</code></td>
<td>logical implication</td>
</tr>
<tr>
<td><code>⊂</code></td>
<td>subset of</td>
</tr>
<tr>
<td><code>∈</code></td>
<td>element of</td>
</tr>
<tr>
<td><code>≡</code></td>
<td>non-distinctness</td>
</tr>
</tbody>
</table>
I Introduction

1.1 Rules and representations in linguistic theory

In generative grammar, a distinction is made between linguistic representations, which are formal or mental objects of linguistic theory, and rules, which are (usually derivational) relations between representations of different kinds. The idea of different kinds or levels of representation (morphemic, morphophonemic, phonemic, phonetic etc.) predates generative linguistics. Generative linguistics inherited the notion of 'levels of representation', and by making explicit the notion of linguistic rule, explored the relations between representations of various kinds. The distinction, articulated by Chomsky (1980), has remained so fundamental to generative grammar from its earliest days that until recently it was not questioned.

The distinction was explored at length in relation to theories of phonology by Anderson (1985), who regarded rules and representations as notions of grammars and languages respectively:

the notion of linguistic representations is one that arises as the central object of study in a theory concerned with languages (construed as sets of sentences, words, utterances, etc.); while the notion of rules is one that arises particularly in connection with the study of grammars. Many of the central concerns of the field at present are fundamentally questions about the basic properties of representations; but if this is a notion that pertains particularly to theories of languages, and we accept the proposal that the appropriate object of study in linguistics is actually grammars, then it follows that concerns about properties of representations must at minimum be raised anew. (Anderson 1985: 7)

In a recent group of closely related syntactic theories, however, collectively known as unification-based grammar formalisms (Shieber 1986), the distinction between rules and representations has been significantly undermined. In that approach, syntactic representations can be regarded as the combination of two kinds of partial representations: lexical representations, which represent lexically idiosyncratic information, and constraints,
which represent recurrent or general syntactic properties. (Though they are conceptually different, lexical representations and grammatical constraints are similar in kind in unification-based grammar formalisms, as lexical representations are constraints too, albeit rather specific ones.) A sentence is said to be grammatical or well formed if it can be assigned a well-formed syntactic representation. A syntactic representation is said to be well formed if it can be constructed out of lexical representations and constraints using the operations of concatenation and unification. Unification-based grammar formalisms developed from the desire to constrain the theory of generative grammar by eliminating transformational and context-sensitive rewriting rules from syntactic theory.

Canonical transformational rules are quite rare in phonological theory (Chomsky and Halle 1968: 427, Kenstowicz and Kisseberth 1979), but context-sensitive rewrite rules are commonplace. If it can be shown that they are unnecessary, however, they should be dropped, so that the rule component of phonological theory is constrained as much as possible.

In later chapters, I shall apply the techniques of unification-based grammar formalisms to phonological theory, to produce a new theory which has no transformational or context-sensitive rewriting rules (or any rules at all, in the traditional sense), only representations. In order to reach this goal, I shall first critically examine a variety of alternative phonological theories, namely phonemic phonology, transformational-generative phonology, Dependency Phonology, Autosegmental Phonology and Metrical Phonology. Drawing on techniques from formal language theory, I shall examine some of the equivalences and differences between those theories, paying special attention to the notation of phonological representations in earlier theories. Informed by this critical background, I shall then construct, motivate and exemplify the new framework.

1.2 Rules and representations in phonological theory

From the mid 1960s to the mid 1970s, phonological theory was dominated by the theory of transformational-generative phonology, or the ‘SPE model’, as it came to be known. According to that theory:

1. Phonological units were of two kinds: fully specified columns of distinctive features (segments), and various boundary symbols.
2. Phonological representations were strings of segments and boundary symbols. There were three levels or kinds of phonological
1.2 Phonological rules and representations

representation: lexical (systematic phonemic), surface (systematic phonetic) and intermediate.

Lexical representations were mapped onto surface representations by an ordered sequence of rewriting rules of the form $A \rightarrow B / C \rightarrow D$. Since each rule could be reapplied many times over, and each rule was applied one at a time, the mapping from lexical to surface representations proceeded via many intermediate representations.

With such a simple view of phonological units and representations, research in this framework concentrated almost exclusively on discovering the phonological rules of each language, the recurrent kinds of rules in the world's languages, and consequences of different ways of applying rules. In short, although there were three different kinds of representation, the research concentrated on rules. Anderson (1974), a typical work of the period of research immediately following publication of Chomsky and Halle (1968), illustrates this statement as well as any other work from then.

In the mid 1970s, phonological theory underwent a period of rapid change. Within the space of a few productive years, each of the three positions of the SPE model was challenged. The start of this change can be taken, I think, as Anderson and Jones (1974), who argued that:

1 phonological segments were graphical objects with a rich internal structure, and boundaries were dispensable because
2 phonological representations were also graphical objects, viz. structured strings.

Kahn (1976) took up the second of these ideas. Then, Goldsmith (1976), drawing on ideas raised by Leben (1973), argued that phonological units need not be segments – a single feature could act as a unit by itself.

In syntax, by 1981 attention had moved away from the search for language-specific rules towards principles of greater generality. Perlmutter (1980) argued that all movement transformations could be attributed to the interaction of a universal set of relation-changing operations (raising etc.), interacting with universal constraints on such operations, and language-particular linearisation rules. Kaplan and Bresnan (1982) and Gazdar (1982) argued in slightly different ways that transformations were completely unnecessary, attributing movements to semantic (rather than syntactic) correspondences between sentences, such as synonymy. Chomsky (1981) argued that the various kinds of transformational rule of Aspects-style
transformational grammar could be attributed to the interaction of various independent principles of Universal Grammar, such as binding theory, theta theory etc. (Chomsky 1981: 5). The differences between human languages and the child's ability to acquire any of them were attributed to different settings of a number of universal parameters, rather than to the acquisition of different grammatical rules.

This model was soon adopted by phonologists too. The 'Principles and Parameters' model did not spring fully fledged from the head of Chomsky, as it were, but developed over a number of years out of the search for general principles of rule application and form. Though they predate Chomsky (1981) by several years, Goldsmith's (1976) Well-formedness Condition and Liberman and Prince's (1977) Lexical Category Prominence Rule are more like principles in their expression than rewrite rules. In Metrical Phonology, Zubizaretta (1982) showed how the various patterns of tone and accent in nine different Japanese dialects could be attributed to the setting of three main parameters:

(1.1) 1. Direction of branching feet:
   - left-branching
   - right-branching

2. Type of foot:
   - both feet polarised (accent feet)
   - only initial foot polarised
   - only final foot polarised
   - both feet non-polarised (harmony feet)

3. Melody or percolating feature:
   - high
   - low

The Principles and Parameters model was adopted by Autosegmental Phonologists too. In the same volume as Zubizaretta (1982), Clements and Sezer (1982: 217) argued that

the autosegmental framework allows a substantial reduction in the abstractness and arbitrariness of conventional descriptions, by reducing the emphasis on the role of rules in favor of the specification of a small number of parameters along which individual languages make a selection (Clements (1981)). These parameters appear to include the following:

(2) a. The class of \( P \)-segments (melody units) which constitute the autosegmentally-represented harmony features;

   b. The class of \( P \)-bearing units (melody-bearing units) defined as the class of units to which \( P \)-segments are associated under the universal Well-Formedness Conditions;
1.2 Phonological rules and representations

c. The (possibly null) class of *opaque segments*, defined as those which are underlyingly associated with a P-segment;
d. The (possibly null) class of *transparent segments* which must be formally excluded from the class of P-bearing units;
e. The *domain* within which the Well-Formedness Conditions initially apply.

The new phonological theories did not define their representations by rewriting rules of the old type, but by a new, more declarative technique, which (put simply) can be expressed:

A well-formed phonological representation is anything which conforms to a certain set of language-universal constraints.

Although a small number of language-specific rules were still permitted by these new theories, language-universal constraints were to be preferred.

With this period of rapid developments in phonological theory came a fragmentation of the discipline, however, which was marked by an explosion of competing proposals for new phonological notations. For example, consider briefly how English word stress could be represented in the three theories:

(1.2) Dependency Phonology | Autosegmental Phonology | Metrical Phonology

<table>
<thead>
<tr>
<th>Main Stress:</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>application</td>
<td>application</td>
<td>application</td>
</tr>
<tr>
<td>Compound Stress:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>blackboard</td>
<td>blackboard</td>
<td>blackboard</td>
</tr>
</tbody>
</table>

There are obvious differences between the three kinds of representation, certainly. The Autosegmental representation has no single foot-level object. The Dependency representation is not labelled. In the Metrical tree representation, every non-terminal node must branch. Such differences were the subject of extensive debate; for instance, in defence of Dependency representation, Ewen (1982: 29) writes: 'Whereas prosodic structures in metrical phonology are assigned a constituency-based system of representation, in dependency phonology they are characterised in terms of dependency structures. Each construction, such as the syllable or foot, is assigned a
head, which is obligatory for that construction, and which may govern a modifier (or modifiers), which are then dependent on the head.' My reading of Ewen's argument is that Dependency Phonology is preferable to Metrical Phonology because it is more elaborated, and thus more restrictive: although both theories give syllables and feet a constituent structure, Dependency Phonology makes an additional claim, that one constituent of every domain is distinguished.

But this view is wrong. While Metrical Phonology uses a different notation and theoretical vocabulary from Dependency Phonology, in Metrical representations too, one constituent of every domain is distinguished: the node labelled 's'. If we pursue this further, it can be argued that Metrical Phonology is in fact more restrictive than Dependency Phonology, because every head (labelled 's') is the sister of exactly one modifier node labelled 'w', whereas in Dependency Phonology there may be any number of modifiers, or none at all. In Autosegmental Phonology, likewise, 'head' and 'modifier' are distinguished in foot structure (H and L) and in syllable structure (V vs. C).

I shall not attempt to argue for any one of those theories, however. My concern here is that disagreements about such areas of overlap between the theories could arise in the first place. It is important not to confuse notation with the linguistic, mental or social reality of the objects denoted by such notations. Whatever the form of phonological representations in people's mental or social reality, it seems most implausible that they consist of geometrical arrangements of lines and/or letters. Consequently, it cannot matter in cases such as this whether head constituents are indicated in examples and diagrams by H or s or * or l or thick lines or vertical alignment or whatever: such diagrams are, in any case, only symbols whose meaning derives from their interpretation, not their form.

Despite the fact that phonological representation has received considerable attention from linguists, very little effort has been expended on getting to grips with the formal properties of such notations and representation. This lack of rigour has too often led phonologists to confuse their diagrams for the phonological representations which they want to portray. I take up this issue in chapters 2–4, and present an examination of various aspects of the structure and interpretation of phonological notation and representation. In a few places in those chapters I let slip an alternative, non-derivative view of generative phonology, which will be discussed at greater length in chapters 5–7.
1.3 Imperative, declarative and interrogative approaches to grammar

The non-transformational approach advocated in later chapters has come to be known as Declarative Phonology (Scobbie 1993, Scobbie, Coleman and Bird 1996), though other terms, in particular Unification-based Phonology and Constraint-based Phonology have been used from time to time as near synonyms elsewhere, all meaning non-transformational, non-derivational, and working with a single level of phonological representations, without intermediate levels. In order to understand some of the terminology used in this approach and its motivation, let us take a look at the abstract computational architecture of generative phonology.

The computational mechanisms employed in transformational phonology such as SPE (Chomsky and Halle 1968) have persisted through to more recent work in non-linear phonology. For example, the following rules of syllabification in Harris (1983) are typical:

(1.3) **Onset Rule (final version)**
Construct a maximally binary branching tree of category O(nset) whose branches dominate [+consonantal] segments that are not adjacent on the universal sonority scale. (p. 21)

The effect of this statement is that the only onsets are single segments and obstruent-plus-liquid clusters. For example:

Input: petaca abogado triple frustrar

Output: petaca abogado triple frustrar

\[ \begin{array}{cccc}
\mid & \mid & \mid & \mid \\
O & O & O & O \\
\end{array} \]

(1.4) **Rime Rule R1**
Construct a maximally binary branching tree of category R(ime) whose obligatory left branch dominates [+syllabic, -consonantal] and whose optional right branch dominates [-syllabic]. (p. 24)

This rule constructs core rimes; for example:

Input: petaca abogado triple frustrar

Output: petaca abogado triple frustrar

\[ \begin{array}{cccc}
\mid & \mid & \mid & \mid \\
O & O & O & O \\
\end{array} \]
Introduction

(1.5) **Rime Rule R2**  
Adjoin a [-consonantal] segment to a rime. (p. 25)

For example:

<table>
<thead>
<tr>
<th>Input:</th>
<th>swave</th>
<th>vjuda</th>
<th>bwej</th>
<th>fjel</th>
</tr>
</thead>
<tbody>
<tr>
<td>R</td>
<td>V</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output:</th>
<th>swave</th>
<th>vjuda</th>
<th>bwej</th>
<th>fjel</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

(1.6) **Rime Rule R3**  
Adjoin the segment /s/ to the right of an existing rime. (p. 28)

Examples:

<table>
<thead>
<tr>
<th>awstral</th>
<th>vals</th>
<th>konsta</th>
</tr>
</thead>
<tbody>
<tr>
<td>V</td>
<td>V</td>
<td>V</td>
</tr>
<tr>
<td>R</td>
<td>R</td>
<td>R</td>
</tr>
</tbody>
</table>

Rules (1.3)–(1.6) are imperative in character. They are phrased as instructions employing the commands ‘construct (a tree)’ and ‘adjoin (a segment)’. In such a formulation, the question of rule ordering and/or parallel application arises naturally. Such an approach to syllabification, and to the computation of phonological structure more generally, is typical in generative phonology.

A declarative phonological analysis equivalent to Harris’s rules is the following set of statements:

(1.7) An onset is well formed if it consists of no more than two [+consonantal] segments that are not adjacent on the universal sonority scale.

(1.8) A core rhyme (R[+core]) is well formed if it consists of a [+syllabic, -consonantal] segment, and an optional [-syllabic] segment.

(1.9) An augmented rime (R[−core]) is well formed if it consists of an optional [-consonantal] segment and a core rime.

(1.10) An augmented rime is well formed if it consists of a core rime followed by /s/.

Because of their declarative expression, such a set of statements may not appear to be sufficient for constructing syllable structure. That impression, however, is incorrect. Although the statements, like Harris’s, are written in English, they may be more formally expressed as a context-free
1.3 Approaches to grammar

We take this opportunity to flesh out the consequences of the sonority constraint in branching onsets more concretely, using distinctive features:

\[(1.11) \ O \rightarrow \begin{array}{c}
+\text{consonantal} \\
-\text{sonorant} \\
-\text{continuant}
\end{array} \left( \begin{array}{c}
+\text{consonantal} \\
+\text{sonorant} \\
-\text{nasal}
\end{array} \right) \]

\[(1.12) \ R[+\text{core}] \rightarrow \begin{array}{c}
+\text{syllabic} \\
-\text{consonantal}
\end{array} \ ([-\text{syllabic}]) \]

\[(1.13) \ R[-\text{core}] \rightarrow ([-\text{consonantal}] \ R[+\text{core}]) \]

\[(1.14) \ R[-\text{core}] \rightarrow R[+\text{core}] /s/ \]

In order to build syllables we also need a syllable rule such as \( \sigma \rightarrow O \ R \), and rules for the composition of words in terms of syllable sequences, usually via intermediate levels of metrical structure such as metrical feet.

This is not formalisation for its own sake. Just like the English declarative statements in (1.7)–(1.10), the context-free rules are declarative in expression. The usual benefits claimed for formal statements apply, such as explicitness, transparency of expressive power, logical completeness and consistency (if they obtain), and so on. But beyond these technical considerations, a context-free grammar has many possible imperative interpretations. For example, it may be used as a generator of well-formed strings, as in Chomsky (1957: 26–7). Alternatively, there are a number of general parsing algorithms for context-free grammars that may be used to determine the structure of syllables defined by the rules on the basis of a well-formed string of segments, or alternatively to determine the ill-formedness of an ill-formed string of segments in the event that the parsing algorithm, which is provably correct, is unable to compute a structure for some input. These, it seems to me, are the minimal requirements of a generative grammar, and they are more obviously satisfied in declarative grammars than in informally stated grammars, especially imperative grammars which invariably involve many intermediate transformational levels.

This is not to imply that declarative grammars are all context-free: context-free grammars are used here as perhaps the most familiar example of declarative grammar. A more general imperative interpretation of a set of declarative constraints derives from techniques of theorem proving and automatic deduction studied in logic and some branches of computer science. In this approach, a task such as ‘determine the syllable structure of /bwej/’ is cast as the question ‘Can it be proved that /bwej/ is a well-formed
syllable? The following chain of inference can be drawn out by an automatic theorem-proving algorithm:

1. **Question:** Can it be proved that /bwej/ is a well-formed syllable of Spanish?
2. **Rule:** A well-formed syllable consists of a sequence of an onset and a rime.
3. **Question:** Does /bwej/ consist of a sequence of an onset and a rime?
4. **Rule:** An onset consists of a non-sonorant consonant and an optional sonorant.
5. **Question:** Does /bwej/ consist of a sequence of a non-sonorant consonant and an optional sonorant and a rime?
6. **Rule:** A rime consists of an optional non-consonantal segment, a syllabic segment and an optional non-syllabic segment or /s/.
7. **Question:** Does /bwej/ consist of a sequence of a non-sonorant consonant and an optional sonorant, and an optional non-consonantal segment, a syllabic segment and an optional non-syllabic segment or /s/.
8. **Rules:** /b/ is a non-sonorant consonant, /w/ is non-consonantal, /e/ is syllabic and /j/ is non-syllabic.
9. Therefore (7, 8), /bwej/ consists of a sequence of a non-sonorant consonant and an optional sonorant, and an optional non-consonantal segment, a syllabic segment and an optional non-syllabic segment or /s/.
10. Therefore (5, 6), /bwej/ consists of a sequence of a non-sonorant consonant /b/ and a rime /wej/.
11. Therefore (3, 4), /bwej/ consists of a sequence of an onset /b/ and a rime /wej/.
12. Therefore (1, 2), /bwej/ is a well-formed syllable of Spanish.

The chain of reasoning that such a proof employs in testing the initial question (1) against the rules of the grammar follows the linguistic structure of the string in determining the well-formedness of the string. By keeping a note of the inferences regarding the constituency of each segment, a complete structural analysis of the string is arrived at in the course of addressing the initial question. The formal similarities between proofs in logic and derivation in generative grammar run very deeply: it is probable that the notion of derivation is inherited from the idea of a proof, since Chomsky was familiar with formal logic and the concepts of proof and algorithm.
1.4 Linguistic representations and cognitive representations

I shall consider here what phonological representations in generative phonology denote. A large number of views concerning this question have been advanced over the years, and it seems unlikely that agreement will be reached in the near term. Some phonologists (e.g. Goldsmith 1976: 16, Dogil 1984, Browman and Goldstein 1986) believe they denote articulatory scripts. Some (e.g. Halle 1985) believe they denote mental scripts. Some believe they denote abstract, purely phonological, objects (Foley 1977). Some believe they are just convenient fictions, ephemeral constructions to aid in the development of theoretical ideas.

Despite the cognitive overtones of the word ‘representations’, I shall not attempt to construct a theory of mental or cognitive representations. Such a goal seems premature to me at present, though recent research such as Pierrehumbert and Nair (1995) is promising and suggestive that the surface orientation of Declarative Phonology is psychologically attractive. However, to the extent that cognitive representations of linguistic knowledge are ‘computed’ by the human brain, it seems reasonable to demand that the representations proposed in linguistic theories are at least capable of supporting computation of some kind (e.g. generation, acceptance, transduction or parsing). Contemporary methods of artificial computation may bear only a remote relationship to natural cognitive computation, but at least the theory of computation is sufficiently advanced to recognise the limits of its computing devices, and is capable of considering abstract limits and properties of computational problems, such as computational complexity. Using such techniques, it becomes possible to argue that, linguistic object X, represented according to theory Y, is at least in principle capable of solution in finite, linear or real time, using finite computational resources. To the extent that the same constraints appear to be found in cognitive computation, we can say that theory Y is cognitively plausible, in a weak sense.

If, however, nothing is known about the formal or computational properties of a theory, then cognitive plausibility cannot be claimed. If a linguistic theory is so powerful that it allows the successive application of context-controlled deletion rules, then there is not normally a computationally reliable method to accept or parse the linguistic representations of such a theory, an observation which is greatly at odds with elementary observations of human linguistic computation. This particular issue is pursued further in chapter 3.
1.5 Notation

Scholars in many disciplines employ specially constructed notations to assist them in their work. The special notations of mathematics, chemistry, astrology, meteorology, proofreading, electronics, music and even ordinary writing, are all purpose-built, artificial constructs. Various branches of linguistics too, including syntax, phonetics and phonology, have their own special notations.

In some areas, the study of such notations has been advanced to a level that the notations seem indispensable to the subject. Without their special notations, the discourse of mathematicians and logicians would be very different. Notation is so important to these subjects that they gave rise to the formalist hypothesis that reasoning and proofs could be reduced to purely notational operations on a set of axiomatic sentences. The formal approach pursued here does not, however, indicate commitment to that hypothesis.

1.5.1 Formalism

Linguists have long been concerned with theory-internal debates and considerations, both practical and philosophical, quite independently of mathematical considerations. Why, then, does this work, like much of contemporary linguistics, attempt to express linguistic principles with mathematical formality?

For me, there are two main answers to this question. Firstly, historically speaking, other areas of linguistics (notably syntax and semantics) have progressed in leaps and bounds since adopting a formal approach. Whether there is a causal connection between these two observations is impossible to tell, but since syntax and phonology are not altogether different in kind (both being concerned with combinatory and structural issues, for example), the hope of many linguists is that phonology can be put on a similar footing. This first answer is subject to the criticism of Harris (1987: 514) that 'one of the . . . lessons that may have been learnt from the great debates which have racked linguistics theory during the past quarter century was just that: almost anything linguistic . . . can be formalized if need be, and in any number of ways'. (I thank James Higginbotham for bringing this remark to my notice.)

The second answer will perhaps be more important: since formality demands (or at least promotes) a certain kind of explicitness and (in theory, if not always in practice) clarity of presentation, the formalist programme offers phonologists the possibility of pursuing certain debates in a rigorous
and impersonal fashion, thus allowing the prospect of avoiding chasing non-issues. This goal is not always met: Batóg (1967) and Brainerd (1971) do not attain it, because they were not tuned in to the concerns of phonologists of their time. Bird (1995) will be a more successful example of this aim, I hope. A suitable formal framework could offer phonologists of different schools a lingua franca for constructive debate. Formality is also required in order to construct computational implementations of linguistic theories, so as to be able to rigorously test their theoretical claims. The partial grammar of English phonological structure in chapter 7 has been computationally implemented and tested in this way. (See also Coleman (1990a), for a computational implementation of Kaye, Lowenstamm and Vergnaud’s (1985) ‘Charm theory’ of phonology, and see also Dresher and Kaye’s (1990) computational implementation of parameter-setting in Metrical Phonology.)

The formalism developed here draws on several areas of finitary mathematics (set theory, logic, graph theory, algebra and formal language theory) that have already been found to be of considerable utility in other areas of linguistics. Set theory and logic are widely employed in semantic theory, to express concepts such as entities, properties, relationships, modalities, time, and reference, which I shall argue are found in phonological theory too. Elementary set theory is at the root of all the other mathematical disciplines adopted in this thesis. Graph theory is extensively used in contemporary approaches to phonology (see Coleman and Local 1991 and chapter 4) as well as in Unification Grammar (Shieber 1986: 20, Knight 1989). Algebra is extensively used in syntactic theory. The algebra of concatenation used in formal language theory is the basis of the theory of phrase structure and classical transformational grammar. The algebra of unification is widely used in non-transformational theories of syntactic structure, notably GPSG (Gazdar et al. 1985: 27) and Lexical–Functional Grammar (LFG: Kaplan and Bresnan 1982). Formal language theory studies artificial, auxiliary languages as well as natural languages with a mathematical rigour. A familiarity with formal language theory is assumed at the level that is presented in Syntactic Structures, Aspects of the Theory of Syntax and Wall (1972). I assume some knowledge of the Unification Grammar formalism (Shieber 1986), in particular its development in Gazdar et al. (1985), and competence in basic logic and set theory (e.g. Allwood, Andersson and Dahl 1977, Dowty, Wall and Peters 1981, McCawley 1981).

Formalisation can sometimes actually be an obstacle to understanding,
however. It can be harder to see the problems in a confused argument when that argument is immersed in a private formal language, the meaning of which is known only to the author. Even extremely sound formal studies, such as Brainerd (1971) and Batóg (1967), may fail to achieve the wider recognition which they deserve, since in their rigour they are not sufficiently accommodating to the relatively non-formal background of most linguists. Since I regard formalism as a tool, rather than an end in itself, I have attempted to strike the right balance between formalism and informality. I recognise, however, that this strategy risks falling between both stools: too inaccessible for linguists, and too informal for logicians.

1.5.2 Phonological Notation

Notation and linguistic representation are not just products of the mid 1970s phonological theories: they have been a central interest of language theorists for centuries. The motivation for this interest has changed with the times, of course. While in an earlier age the development of a written form of representation for language was regarded with religious status, succeeding centuries saw similar concerns directed towards improvement or rationalisation of spelling, development of a ‘Real Character’ (an artificial language in which falsehoods would be inexpressible), or the establishment of writing systems for previously unwritten languages (Westermann 1927, Robins 1967: 114–15).

Before the rise of generative phonology in the 1960s, most phonologists did not recognise rules in their theories, so there could be no distinction between rules and representations. Phonology was almost exclusively concerned with discovering and motivating the phonological units and representations of each language. In recent decades, notation and representation have been prominent in, if not central to, phonetics and phonology. Although phoneticians have shaken themselves free of this focus on purely notational issues, phonologists remain firmly embroiled in a seemingly endless discussion of representation.

Drawing on techniques from formal language theory, graph theory and other branches of finitary mathematics, in chapters 5–7 I shall rigorously examine several phonological notations. Of necessity, my selection is only a minority of the many systems that have been proposed, so I have picked a representative sample that exemplifies a variety of representational styles, a set of historically related theories, and a set of theories that have had most influence on contemporary phonological theory. Each system will be considered from four angles: (i) its historical context; (ii) the representation of
phonological objects, categories, or primes; (iii) the representation of phonological structure; (iv) the representation of phonological phenomena such as assimilation, epenthesis, etc.

I shall attempt to determine which theories or (parts of theories) of phonological representation are equivalent to each other, and which parts of theories are genuine original developments. In common with other linguists, I shall assume that parts of theories which are equivalent, albeit perhaps notational variants, do not merit extensive discussion. Determining formal equivalences between theories serves to highlight the idiosyncrasies, which are important, for it is these proposals which may prove ground-breaking.

1.6 Overview

Taking the historical development of ideas in a field as a framework for discussion is a standard monographic technique. For expository convenience, I pursue a ‘Whig history’ view of the development of thinking about phonological representations in which the line of history is followed without deviation from the IPA through phoneme theory, SPE, and non-linear phonology until the ‘ideal theory’, declarative phonology, is developed in the final chapters but prefigured by aspects of all its predecessors. The good historian of the field, of course, will recognise the selectivity and partiality of this approach to the presentation of ideas. Naturally, a number of important by-ways have been omitted, especially Firthian Prosodic Phonology and Stratificational Phonology. In a more encompassing study these and other approaches would deserve detailed treatment, but their historical side-lining prevents me from giving them space here. Some more recent phonological work receives little treatment since it does not raise any significant new issues of representation. For that reason the reader will find no mention of Optimality Theory, for instance. Recent works broadly consistent with the views expressed here, but which due to their recent appearance I shall not discuss at length, include Bird (1995), Burzio (1994) and Broe (1993).

In chapter 2 I prepare the introductory groundwork for the rest of the book by looking at a segmental phonetic notation, the International Phonetic Alphabet. The purpose of this example is to illustrate the terminology and a number of key concepts in the study of linguistic representation, in preparation for the following chapters. The ‘critical’ chapters (3 and 4) examine segmental and non-segmental theories of phonological representation in a quasi-historical fashion, an arrangement which reflects the
development of ideas and practices, and which concludes at the present dis-parate state of development of the various theories. In chapters 5, 6 and 7 I attempt to take the discipline a step further by constructing and exemplifying a new, declarative theory of phonology.

Chapter 2 sets out some fundamental notational resources used in many theories in portraying phonological representations. I shall argue that a phonological notation, in this sense, is an artificial, auxiliary language for portraying phonological representations. Using insights gained from the study of other artificial, auxiliary languages (the notations of set theory, symbolic logic and generative grammar), I illustrate many fundamental concepts used throughout the book, including notation, denotation, letter, alphabet, string, transcription, concatenation, graph. To illustrate this discussion, I examine one of the many progenitors of contemporary phonological notations, the alphabet of the International Phonetic Association (IPA 1949). Although the IPA is not a phonological notation proper in contemporary phonological theory, its historical status as a precursor to segmental phonological notations makes it an ideal vehicle through which to explain and illustrate some of the formal machinery used in the following chapters. I show how the distinction between phonetic form (notation) and interpretation (denotation), and the notation of relaxed segmental interpretation, allows epenthesis, metathesis, elision, assimilation and coarticulation to be modelled phonetically, rather than phonologically.

Chapter 3 further explores componential, segmental theories of phonological representation. I show how componential notations are suitable for the expression of phonological relations (oppositions and transformations). A consideration of phonetic variability and the distinction between redundant and predictable features leads to an investigation of the power of different classes of rules. I present a proof that phonological formalisms containing rules of the form A \(\rightarrow\) B / C \_ \_ D, where B is possibly the empty string, are essentially completely unconstrained, because they are equivalent to unrestricted rewriting systems. I argue that there is a trade-off between the power of grammars and the complexity of phonological representations in the analysis of phonological phenomena, and that the goal of constructing a maximally constrained theory of phonology can be achieved by relaxing the class of representations. I present a technique for constructing phrase-markers for all of the types of grammar in the Chomsky hierarchy, a method which provides a formal basis for declarative grammars of various kinds, as well as a means of comparing the strong generative capacity of grammars.
In chapter 4 I discuss the two most influential non-segmental theories of representation, Autosegmental Phonology and Metrical Phonology, with a brief consideration of Dependency Phonology. These theories date from a period in the mid 1970s which saw an explosion of new ideas about phonological representations. I argue that the phonological rule formalisms which each of these theories employs are excessively unconstrained. I show that the parameter-setting account of Metrical structure codified by Dresher and Kaye (1990) benefits from formalisation using the resources of unification-based grammar. I conclude that current approaches to phonological notation are all theoretically unsatisfactory, though each makes some promising contributions.

Having examined the properties of phonological notations, in chapter 5 I argue that declarative formalisms such as Unification Grammar offer phonologists a notation and formalism which meets their theoretical requirements, and which does not suffer from the problems identified in the preceding chapters. A phonology based on Unification Grammar differs from other contemporary phonological frameworks in that it does not use rewriting rules or other derivational operations. In order to do this, it employs non-segmental representations and a phonetic interpretation relation. Thus, rejection of rewriting rules and segmental representations are complementary.

In chapters 6 and 7 I illustrate and motivate the use of Unification-based Grammar in phonology by presenting case study analyses of extensive fragments of Japanese and English phonology. In each of these, I present reasonably complete descriptions of the phonotactic structure of syllables, and the distribution of features at different places in structure. I also discuss the structure of polysyllabic words in those languages. The intention of these studies is to demonstrate the flexibility of Declarative Phonology in the description of two unrelated and rather different natural languages.
2 Segmental representations and their phonetic interpretation

Fundamental to almost all research in phonetics is the concept of speech sound. It is the basic unit around which every textbook of phonetics has been organized from A. M. Bell’s ‘Visible Speech’ of 1867 to Peter Ladefoged’s ‘Course in Phonetics’ of 1975. In fact if the notion of speech sound were to be excluded from phonetic discourse – by decree of an all-powerful dictator – work in the field would come to a virtual standstill.

(Halle and Stevens 1979)

Most linguists would not know a systematic phonetic description if they saw one.

(Ladefoged 1977)

2.1 Introduction

Since there is no general notation for phonological representations, theories of phonological representation are proposed and defended using theory-internal notations. Classical phonemic phonology used strings of lower-case romanic symbols for its notation. Firthian Prosodic Phonology used Greek and upper-case romanic symbols, with subscripts and superscripts. Transformational-generative phonology used two-dimensional matrices of distinctive features, and non-linear phonological theories make use of two- and three-dimensional graphical notations to portray phonological relations such as constituency, dependency and association.

Faced with two apparently different phonological representations from two different theories, it is not possible to determine by inspection alone whether they are notational variants which denote the same phonological object in different notations, or whether they denote different phonological objects. In order to understand how theories of phonological representation work, it is helpful to understand first how phonological notations work. In this chapter I shall examine phonological notation from a cross-theoretical perspective. I shall explore those properties of phonological notation which occur in several, if not all, of the theories examined in the following chapters.
Exploring the distinction between phonological representation and notation is assisted by the circumstance that most of the key concepts of phonological notation can also be revealed in segmental phonetic notations such as the IPA alphabet. Since the two phonological theories I will consider in the next chapter, phonemic phonology and transformational-generative phonology, employ notations which are historical descendants of segmental phonetic notation, the IPA is a convenient starting-point. However, in order for my analysis of phonological representations to be sufficiently general to cover feature-matrix notation and the graphical notations of the non-linear phonological theories, I shall go on to develop a general framework for our study of phonological notations, based on a number of temporal and spatial relations.

The structure of this chapter is as follows. After considering phonetic notation in a historical perspective in section 2.2, I shall discuss the different kinds of segmental phonetic notation in section 2.3. The central theme of this chapter, introduced in section 2.4, is the relationship between notation and denotation. I develop this theme further in the following two sections, by considering the syntax (notation) of IPA transcriptions in section 2.5, and the semantics (denotation) of IPA transcriptions in section 2.6. A comparison with two componential segmental notations, the Organic Alphabet and feature-matrix representation, is briefly made in sections 2.7 and 2.8. In section 2.9, I show how the structure and temporal interpretation of strings of atomic and componential symbols can be elucidated in a uniform fashion, by using string-graphs. (In chapter 4, these graph-theoretic concepts will be applied to an examination of Autosegmental Phonology.) Finally, in section 2.10, I argue that the distinction between notation and denotation sheds light on phenomena at the interface of phonetics and phonology, including assimilation, epenthesis, metathesis, elision, 'hard' coarticulation and suprasegmentals.

2.2 Historical context of phonetic notations

A phonetic notation is an artificial, auxiliary writing system in which speech may be represented to a greater or lesser degree of detail. Of the many phonetic notations which have been developed over the centuries, most have been alphabetic, modelled on one or more of the romanic writing systems used in various European languages.

The development of the alphabetic approach to phonetic notation is closely tied up with the alphabetic model of speech, which developed over
a considerable period of time, since at least the time of Plato. Nearly all the
elements of the alphabetic model of speech are represented in the mediaeval
tory of the 'letter', which derives via Priscian from the Greeks
(Abercrombie 1949, Robins 1967). In this theory, written letters and spoken
sounds were not distinguished from each other as they are today. Instead,
a letter (littera) was viewed as a triune abstract object, whose various aspects
were 'figure' (figura) or written form, 'power' (potestas) or spoken form and
'name' (nomen) or conventional name. For instance, in Modern English we
give the conventional name 'aitch' to the figure 'h', with spoken form [h].

Alphabetic writing systems invariably involve some conventional depart-
tures from a simple one-to-one relationship between 'figures' and their
'powers', and it has long been recognised that in the English writing system,
for example, the orthographic sequence 'sh' represents 'phonetically
simple' (i.e. non-dynamic) [ʃ] (a two-to-one mapping), and that
orthographically simple 'x' represents the phonetic sequence [ks] (a one-to-
two mapping). Furthermore, careful attention to the sounds corresponding
to written letters reveals a degree of regular, predictable variability; for
example, the sound represented by 'p' in *pit* is different from its powers in
*tip* and *spit*. From such observations arose various schemes for 'improving'
the orthographies of many languages (Baugh 1951: 388, Kelly 1981), often
attempting to restore 'logical harmony' by adjusting spelling conventions
or introducing a number of new letters in order to achieve a one-to-one
mapping between letters and their powers. In the nineteenth century, the
requirements of a new 'scientific' discipline (philology), coupled with
characteristic Victorian enthusiasm for efficiency and modernity, saw a bur-
geoning of interest in spelling reform, the development of shorthand, and
phonetic alphabets. Not by accident, these developments went hand in
hand (Kelly 1979). Two of the many progenitors of the IPA, Henry Sweet
and Isaac Pitman, developed revised spelling systems, alphabets for more
detailed work in philology, dialectology and phonetics, shorthand systems
and auxiliary alphabets such as the Initial Teaching Alphabet, as well as
making important contributions to phonetic theory.

Among his various achievements, Sweet developed two alphabetic nota-
tions, Broad Romic and Narrow Romic (Henderson 1971: 230), and an
iconic phonetic notation, the Organic Alphabet (Henderson 1971: 255).
The Organic Alphabet was Sweet's revision of A. M. Bell's Visible Speech
(Bell 1867, 1882). Sweet credits one of many earlier attempts at an organic
alphabet for consonants to Brücke (see Henderson 1971: 254). Broad and
Narrow Romic were succeeded by the IPA, and are still reflected in its two
primary modes of use, broad and narrow transcription (a distinction which Sweet modelled on Ellis's Glossic and Universal Glossic; see Henderson 1971: 233), and in the Principles of the IPA (IPA 1949). Although the Organic Alphabet has long since ceased to be used as a phonetic notation, it provides an elegant and interesting example of a non-roman, articulatory-parametric notation, predating its most well-known successor, Chomsky and Halle’s (1968) articulatory distinctive-feature notation, by a century.

2.3 Segmental phonetic notation

Segmental phonetic notations are the most common type of phonetic notation, which includes alphabetic notations (e.g. IPA 1949), componential alphabetic notations (e.g. Bell’s Visible Speech, Sweet’s Organic Alphabet), and componential, analphabetic notation (e.g. Jespersen 1889, Pike 1943, Jakobson, Fant and Halle 1952: 43–5). The model of speech organisation behind all of these is the same: speech is held to consist of finite sequences of sounds. ‘Sounds’, in this restricted sense, can thus be called ‘sound-segments’, or more usually, just ‘segments’. In this chapter I will examine exemplars of three kinds of segmental phonetic notation, namely the IPA, the Organic Alphabet and feature-matrix notation. Since the IPA is the best-known segmental phonetic notation, I will examine it in far more detail, although several aspects of my account of the IPA apply to the other two kinds of segmental phonetic notation as well.

The use of a segmental notation-system does not commit a phonetician to a segmental view of speech. It is equally possible to adopt a non-segmental view of speech while nevertheless employing a segmental notation-system, or to subscribe to the segmental view of speech while rejecting segmental notations and employing non-segmental methods of representation, such as a continuous-time representation of the speech wave. The first of these two possibilities is exemplified by the work of Kelly and Local (1989); the second by many speech scientists (e.g. Rabiner and Juang 1993: ch. 8).

Segmental notation-systems employ a taxonomy to categorise segments (and by extension the letters used to represent segments) in order to facilitate the selection of appropriate symbols in the process of transcription. For instance, the segment-type represented by ‘p’ in the IPA is the type of token which involves simultaneous voicelessness, bilabiality, complete oral closure etc. Thus, observation of simultaneous voicelessness, bilabiality,
complete oral closure and all the other taxonomic properties of [p] is a necessary and sufficient criterion for selection of ‘p’ as the appropriate letter of representation. The taxonomic categories are not binding, however, and can be altered without affecting the nature of the transcription system. The principles of use of the IPA allow symbols to be used with non-standard denotations if such a departure permits unusual phonetic categories to be transcribed using ‘more romanic’ symbols. Roman symbols without diacritics are regarded as ‘more romanic’ than roman symbols with diacritics, which in turn are regarded as ‘more romanic’ than invented symbols.

In componential notation-systems such as the Organic Alphabet or column-vectors of distinctive features, on the other hand, such phonetic properties as voicelessness, closure and bilabiality are encoded in the form of the notation, not just in categories of the denotation. The basic objects of the notation-system represent not segments but subsegmental components, a term proposed by Harris (1944). These components are combined into complex letters (either alphabetic or column-vectors), which may then be concatenated into sequences to represent sequences of segments. The prototype of this kind of notation is perhaps the Pictish Ogham alphabet (Diringer 1962: 164, Halle 1969).

The IPA is a hybrid of the componential and non-componential methods. It would require a vast number of new symbols to be invented in order to have a separate, unique letter for each segment-type in the IPA taxonomy, but since new symbols are regarded as less preferable to modifications of roman symbols, a degree of componentiality is used, that is diacritics. But this componentiality is not thoroughgoing or systematic, and so I categorise the IPA as essentially a non-componential system with componential additions.

2.4 Notation and denotation

Since it is clear from the preceding discussion that the notation and denotation of segmental phonetic representations are, to an extent, separate issues, let us consider the relationship between them in a little more detail. The distinction between notation and denotation is a theme to which I shall return many times. Simply put, ‘notation’ refers to the symbols and diagrams employed in the expression of phonological arguments in print; ‘denotation’ refers to the linguistic objects (e.g. phonemes, features, syllables etc.) that an item or expression in the notation-system represents. I shall also employ the term denotation to refer to the mapping between
2.4 Notation and denotation

notation and linguistic objects. As I stated in chapter 1, there are various views as to the nature of denotations; but in every case notation is just print. If a clear distinction between linguistic notations and the objects they denote is not maintained, it becomes easier to accidentally confuse properties of the notation with properties of the denotation. A brief mention of two examples will be instructive. First, in phonological theories such as Dependency Phonology, it is sometimes claimed that phonological rules may not refer to the absence of a unary feature, and that therefore such a notation is more restrictive than the use of binary features, with which rules may refer to either value. Yet this claim is only true by stipulation: it is not an inherent property of notation-systems which only employ unary features. It would be equally possible for advocates of binary feature-systems to stipulate that, say, the unmarked value of a feature may not occur in a rule or representation, and it would be equally possible for advocates of unary features to employ rules that refer to the absence of a feature (e.g. the ~ notation of Dependency Phonology; see Anderson and Ewen 1987: 127). The argument rests on a confusion between the notation used in a particular phonological theory and its denotation, the objects of the theory. Second, in chapter 4 it is argued that the so-called 'No Crossing Constraint' of Autosegmental Phonology does not constrain the class of well-formed phonological representations (i.e. denotations) in that theory, as proponents of that theory claim. It only constrains the class of well-formed diagrams, that is illustrations in journals etc. It is thus not a linguistic or phonological constraint at all, but a stylistic or notational constraint. The No Crossing Condition is employed in Autosegmental Phonology as if it were a denotational constraint on phonological representations. However, if the argument of chapter 4 is correct, it is merely a notational constraint, a condition on pictures.

Confusion between notation and denotation of phonological representations is endemic in the literature. Consider a third example of this confusion. After describing at some length their model of articulatory dynamics, Browman and Goldstein (1986) propose 'to develop a formalism for articulatory phonological representation' (Browman and Goldstein 1986: 241). At first they are careful to maintain the distinction between partial phonetic descriptions (which are symbolic objects) and the physical events which they denote:

The descriptions in this section differ from those in the previous section only in terms of the degree of detail in the description . . . the descriptions are coarser-grained and more qualitative. Thus we are referring to the
same set of dynamically-specified gestures, but this time using symbols which serve as indices to entire dynamical systems. These symbolic descriptions highlight those aspects of the gestural structures that are relevant for contrast among lexical items. (Browman and Goldstein 1986: 241)

So far, their model maintains a distinction between symbolic phonological representations and the physical, phonetic, articulatory gestures which they denote. Then, without justification, Browman and Goldstein baldly state that 'the minimal units in the lexical representations are dynamically-defined articulatory gestures' (Browman and Goldstein 1986: 241). This seems quite contradictory to the former statements, since gestures, according to Browman and Goldstein, are not symbolic, notational objects, but dynamic, physical, denotational objects.

Confusion between phonological notations and their phonetic denotations is also found in generative phonology and some of its successors, such as Government-based Phonology (Kaye, Lowenstamm and Vergnaud 1985). Kaye, Lowenstamm and Vergnaud's theory, like almost every other variety of generative phonology, is intended to define not only just the right set of segmental phonological categories, but also just the right set of segmental phonetic categories. They state that 'the primary unit of segment constitution is the ELEMENT, which is a fully specified matrix, phonetically interpretable as in SPE theory or some equivalent formulation'.

But the number of categories defined by Charm theory is much less than the number of phonetic categories needed to describe minor but linguistically significant inter- and intra-speaker variation. Therefore, the set of phonetic categories made available in Charm theory is greatly over-restricted (Coleman 1990a). Confusing phonological for phonetic categories results in a theory with either too many phonological categories or too few phonetic categories.

2.4.1 Phonetic transcription and phonetic transcriptions

Having examined the general distinction between phonetic notation and denotation, let us now consider how they relate to each other.

The mapping between a segmental phonetic notation and the physical events which are its denotation is made explicit in the activity of phonetic transcription. Observing some speech produced by an informant, a phonetician writes down some symbols from a notation system, which are intended to denote that speech. The complementary activity, production, is rarer, observable only in phonetics classes, perhaps.
In transcribing an utterance, the phonetician maps the sound-tokens of the utterance to the letter-tokens of their written record, via the categories or types of sounds and letters in their notation-system. Segmental phonetic transcription, therefore, is a composed, one-to-one sequence-preserving mapping from sound-tokens to sound-types, from sound-types to letter-types, and from letter-types to letter-tokens. The product of the phonetic transcription activity/mapping is also customarily called a phonetic transcription, though Kelly and Local (1989) propose the term 'record' to avoid confusion between the activity and its result.

Segmental phonetic transcription thus involves four kinds of objects: sound-tokens (phones), sound-types (phonic categories), letter-types (symbols) and letter-tokens (inscriptions); two ordering relations: one temporal 'precedes' \( <_T \) (over phones), the other spatial 'left (or right) of' \( <_L \) (over inscriptions); two many-to-one mappings 'instance of'; a one-to-one mapping 'represents'; and a one-to-one mapping 'transcribes' (fig. 2.1). The labels on this diagram \( (<_T <_L, D \text{ and } E_b) \) are discussed later in this chapter. The objects and relations on the right of figure 2.1 constitute the notation or syntax of IPA transcriptions, whereas the objects and relations on the left, and the relationships between the right and the left, constitute the denotation or semantics of the IPA. The terms 'syntax' and 'semantics' in this and later sections are not used with the meanings normally found in linguistics, but those of the logicians, notably Tarski (1933), Morris (1938) and Montague (1972). Essentially, 'syntax' refers to the construction of representations; 'semantics' to the relationships between representations and the objects and events which they denote. Firth (1935a) was unusual in regarding 'semantics' as a
concept relevant at all levels of linguistic analysis, including phonology and phonetics.

2.5 The syntax of IPA transcriptions

I shall consider two aspects of the syntax of IPA transcriptions. First, I shall examine the syntax of IPA symbols, which can be quite complex objects with their own internal syntax of diacritics and special devices such as ligatures. This discussion will also lay the foundations for the examination of componential notation systems later in the chapter. I shall then proceed to consider the way in which IPA symbols are put together to form complete transcriptions.

2.5.1 The syntax of IPA symbols

In a segmental phonetic notation-system such as the IPA there is a finite alphabet of symbols (extensible if desired), including symbols marked by diacritics. The elements of this alphabet (i.e. all of the symbols) shall be termed ‘basic expressions’, elements of the set \( E_b \). Informally, in segmental phonetic theory, \( E_b \) will include at least as many basic expressions as there may be distinct sounds. Consequently, it is clearly a tall order to specify \( E_b \) in its entirety in advance. Courses lie open that are more realistic: one such is to derive basic expressions by a mechanism of symbol-generation that is assured of generating at least enough symbols to represent all possible distinct sounds. (Kelly and Local (1989: 26) call such a set of categories ‘indefinitely extensible’.) If we take this course, we can define \( E_b \) to be the set of expressions that represent phonetic categories. Phonetic categories are classes of possible utterances of a speech-sound (phone); to each and every speech-sound there relates a phonetic category (irrespective of speaker or occasion of utterance), and to each phonetic category corresponds a basic expression.

The existence of a set of diacritics, and appropriate operations for recursively applying them to alphabetic symbols, is sufficient to ensure a large enough stock of symbols for phones. These operations are considered in subsection 2.5.3 below.

It is not necessary to develop the syntax of IPA symbol-generation in full in order to observe its basic characteristics, but it is apparent that the result of pursuing this investigation further would be quite a complicated two-dimensional grammar of individual symbols, with many subcategories of diacritics, and special conditions to ensure their correct combination. For instance, since ‘+’ and ‘−’ are both recursive diacritics, it is possible to form
both ‘t−+’ and ‘t+-’ from ‘t’. Are either of these likely to be accepted as legitimate symbols? Clearly the answer to this question depends very much on the interpretation given to them. In the case of the IPA such an expression is semantically anomalous, as the ‘+’ and ‘−’ diacritics have opposing interpretations, but it is not the purpose of syntactic descriptions to prevent semantically anomalous interpretations. Consequently, it might be reasonable in such cases to generate more symbols than are necessary (i.e. over-generate) and use semantic constraints to filter out such symbols, rather than attempting to encode semantic constraints in the syntax. In a componential system it would, however, be justifiable to enforce such constraints in the syntax.

Recursive diacritics are necessary in order to generate a sufficient number of phonetic categories, but they present a problem to the use of the standard theory of formal languages, in which it is usually assumed that the set of categories is finite. It is necessary, therefore, to treat diacritics and letters as separate basic objects, and their composition as complex objects derived by syntactic rules of the notation-system.

2.5.2 The syntax of segmental phonetic expressions
Having briefly considered how basic expressions (i.e. possibly complex symbols) of the IPA are constructed, let us now consider how such basic expressions are put together into complete expressions.

The syntax of IPA expressions is very simple: if E_b is the set of basic expressions, then all strings formed by concatenating elements of E_b are well-formed expressions. Concatenation is the main operation used to form IPA expressions, although a special ‘vertical’ form of concatenation, cocation, can be regarded as a separate operation. Cocatenation is not frequently used in published IPA representations. While cocatenation is therefore of secondary importance to a consideration of IPA syntax, it is of equal importance to concatenation in the Organic Alphabet and other componential notation systems. I consider cocatenation in some detail in the next subsection.

2.5.3 Cocatenation and commutative cocatenation
Suppose that, in addition to concatenation of symbols, symbols may be written one on top of the other, an operation for which I have coined the term ‘cocatenation’, symbolised ©. Cocatenation is exactly like concatenation, in that it is associative, non-commutative and has the same identity element 0. Cocatenation is just ‘vertical’ concatenation. Since © is non-commutative, we can adopt the convention that x © y means
Segmental representations

\[
x
y
\]

i.e. \( x \) is on top of \( y \).

There are several examples of this operation in the IPA. For instance, ligature symbols such as \( \text{t} \text{s} \) can be regarded as consisting of the concatenation of \( t \) and \( s \), to which is cocatenated the \( - \) diacritic. Thus,

\[
(2.1) \quad \text{t} \text{s} \text{ is equivalent to } - \Theta (\text{t}s)
\]

In the Organic Alphabet cocatenation is used to combine components into ‘letters’. For instance, the vowel symbols are composed of a ‘stem’, to which is added either the ‘dot definer’ or the ‘hook definer’ (Henderson 1971: 50) to show respectively the distinction between ‘narrow’ and ‘wide’ vowels (in contemporary terms, roughly ‘normal’ vs. ‘expanded pharynx’ cf. Henderson 1971: 267). The position of the ‘dot’ or ‘hook’ definer with respect to the ‘stem’ is crucial: above denotes high vowels, and below denotes low vowels. Ignoring here the fact that the orientation of the definer is also different in the two cases, we can characterise the form of high-vowel symbols as Definer \( \Theta \) Stem, and low vowel symbols as Stem \( \Theta \) Definer.

Kelly and Local (1989: 75-7) use vertical alignment of symbols to indicate phonetic variability. Thus, for them, a transcription does not denote an utterance, but a set of (similar) utterances. In the places where such utterances differ from each other, the transcription may include a stack of symbols for the different variants observed at those places. This is not quite the same kind of vertical catenation as cocatenation, however. In the two cases of cocatenation discussed above, it matters that the ligature is placed over the conjoined symbols, or whether the ‘definer’ is placed above or below the ‘stem’; hence cocatenation is non-commutative. For Kelly and Local’s notation of variants, however, the vertical order of symbols in the column does not matter. This kind of cocatenation is commutative. To distinguish it from non-commutative cocatenation, I shall use a different symbol, \( \oplus \). (\( \oplus \) is intended to resemble the non-commutative arithmetic operator \( - \), and \( \odot \) the commutative arithmetic operator \( + \).)

Commutative cocatenation is used in the construction of feature-matrix representations, in which

\[
\begin{bmatrix}
    a \\
    b
\end{bmatrix}
\stackrel{\Theta}{=}
\begin{bmatrix}
    b \\
    a
\end{bmatrix}
\]

The Organic Alphabet has a symbol \( \odot \), called the ‘link’ (Henderson 1971: 269), which represents cocatenation – probably commutative cocatenation.
The set of transcriptions that may be built from $E_b$, concatenation, commutative concatenation and non-commutative concatenation can be treated as a formal language of phonetic transcriptions. In the next section, we will consider the interpretation or meaning of this language.

2.6 The semantics of IPA transcriptions

Having defined a set of transcriptions, what do they mean? The meaning of a transcription cannot be just the (ordinary) meaning of the words of the utterance it represents. I propose that the meaning of a transcription is the utterance which it denotes. This claim, at least, squares with the idea in philosophical semantics that the meaning of an expression (sentence, phrase etc.) is a function of the meaning of its parts and the way in which they are put together, the so-called (Fregean) Principle of Compositionality (Dowty 1979: 2, Dowty, Wall and Peters 1981: 8, 42-3). It is hard to see how the ordinary meaning of the transcription of the Japanese sentence Uma ni notta ('I rode a horse') in my notebook is in part determined by the fact that the first vowel was preglottalised and slightly centralised. But it is fairly straightforward to see that the phonetic denotation of the transcription is a product of the denotations and the arrangement of the letters and diacritics in the transcription. We shall call the relationship which links IPA transcriptions to their denotations the denotational semantics or phonetic interpretation of the IPA.

Having argued that the phonetic interpretation of a notation-system such as the IPA is a denotational relationship akin to the semantic interpretation of syntactic theory, it becomes natural to adopt ideas from semantic theory in order to characterise the nature of phonetic denotation. There are two common approaches to semantic interpretation, model-theoretic semantics and axiomatic semantics (see e.g. Dowty, Wall and Peters 1981: 50–3). Of these, only the model-theoretic approach maintains a distinction between expressions and their interpretation in the real world that parallels our discussion of phonetic notation and denotation. Although model theory is the more recent approach to natural-language semantics, it is the most perspicuous approach to employ in the present context.

2.6.1 Model theory

Model theory is an approach to semantic theory which constructs simple but abstract formal models of parts of the world in order to provide an
account of the meaning of expressions in mathematics, logic or linguistics. In semantic theory, objects such as names, numerals or letters of the alphabet are used to represent objects such as people, imaginary beings, locations and periods of time, and set-theoretic relations over such names etc. are used to model real-world relations, such as 'brother of'. In this way, the rigorous methods applied by logicians, mathematicians and computer scientists to the description of the meaning of the 'languages' of mathematics, logic or computer programming have been carried over with much success into the domain of natural-language semantics. Especially notable in this domain is the work of Richard Montague (Montague 1974, Dowty, Wall and Peters 1981) and his followers (Partee, 1976, Dowty 1979, Barwise and Perry 1983).

Model theory is particularly suitable to the development of accounts of the semantics of more elaborate logics than simple first-order predicate calculus. It has been fruitfully applied to the integrated development of intensional, modal and temporal logics (Montague 1974, Dowty, Wall and Peters 1981, van Benthem 1988). Given the importance of problems of the notation of time in phonetic and phonological representations, consideration of some details of temporal logic will be fruitful. In anticipation of these developments, a discussion of model theory and phonetic notation, with illustrations, will be helpful.

2.6.2 Model theory and segmental phonetic notation

A model for a phonetic theory is a mathematical relation between phonetic events (such as utterances) and phonetic expressions in some particular notation. Although this relation is intrinsically non-procedural, it may be viewed in directed fashion as a relation of transcription (that defines phonetic expressions given utterances by a speaker or speakers on an occurrence or occurrences), or as a formal characterisation of what Jones (1918: 8–9) terms 'catenation', that is, having read a phonetic expression, the act of producing one or more utterances from the class of utterances defined by a particular model with respect to that particular phonetic expression. A model, then, relates real-world events to phonetic expressions by defining the class of possible events that a phonetic expression denotes (of which actual events are a subset), and conversely relates phonetic expressions, which are symbolic representations that are held to be constant (irrespective of speaker or occasion of utterance) to real-world phonetic events that are necessarily respective to particular speakers on particular occasions.
In the discussion of IPA semantics which follows, I shall refer to a model \( M \) that will determine for all IPA phonetic expressions a class of possible sound-type sequences (abstracted away from speakers and occasions of utterance), and for a given phonetic expression, speaker and occasion of utterance, an actual sequence of sound-tokens. \( M \) relates the set of expressions \( E \) (constructed out of \( E_b, \neg, \oplus \) and \( \ominus \) – basic expressions, concatenation, commutative concatenation and non-commutative concatenation) to a set of phonetic categories \( C \), a set of times \( T \), and temporal relations including temporal precedence \( < \) and temporal co-occurrence or overlap. I shall use the notation \([x]\) to refer to the phonetic interpretation of expression \( x \) according to model \( M \).

As with the account of the syntax of the IPA given in section 2.5, the discussion of its semantics will proceed in two steps. First I shall consider the semantics of individual IPA symbols, and then strings. Model-theoretic semantics is hard going if it is presented too formally, and it is not my intention to give a tutorial here. Consequently, my treatment of details will be informal. The semantics of individual symbols will be developed in stages, with subsections devoted to the categories of phonetic events employed in the model, the temporal arrangement of those events, and the treatment of componential notations, which will provide a useful introduction to the theme of phonetic interpretation of phonological representations pursued in later chapters.

2.6.3 The semantics of IPA symbols

The semantics of segmental phonetic symbols can be considered in several stages. First, there is the question of what phonetic categories a (possibly complex) symbol denotes (categorial interpretation). This is considered in subsection 2.6.4. Secondly, there is the question of how such phonetic categories relate to each other in time (segmentation theory). This is discussed in subsection 2.6.5. Thirdly, there is the question of how the categorial interpretation and segmentation theory interact. This will be discussed later, in subsection 2.10.

2.6.4 Categorial interpretation

The categorial interpretation of (possibly complex) IPA symbols can be determined by reference to IPA (1949). The denotation of each symbol is expressed in two ways. Firstly, by reference to ‘sounds’ of particular languages. This can be characterised as a ‘prototype’ approach to phonetic interpretation, and I shall ignore it, since it simply places the problem of
defining the denotation of symbols at a further remove. The second method used in IPA (1949) is reference to the charts of place and manner of articulation for consonants and vowels. These charts define the value of each symbol in terms of the categories \{Consonant\} and \{Vowel\}; for consonants, in terms of \{Bilabial, Labiodental, Dental and Alveolar, Retroflex, Palato-alveolar, Alveolo-palatal, Palatal, Velar, Uvular, Pharyngeal, Glottal\} and \{Plosive, Nasal, Lateral, Lateral Fricative, Rolled, Flapped, Rolled Fricative, Fricative, Frictionless Continuants and Semivowels\}; and for vowels in terms of \{Front, Central, Back\}, \{Rounded\} and \{Close, Half-close, Half-open, Open\}. Consonants are also defined in terms of \{Voiced\} and \{Voiceless\}. Though these labels are not written on the IPA chart in IPA (1949), they are used in the textual notes and are widely accepted categories, used in, for example, the *Journal of the International Phonetic Association*, and are employed in the 1993 edition of the IPA chart (IPA 1993).

The denotation of each IPA symbol is thus a combination of phonetic categories. Each of these categories can be equated with a set of cardinal sounds, or with the characteristic function of such a set. For instance:

\[
\begin{align*}
\text{Bilabial} & = \{[p], [b], [m], [f], [w] \ldots \} \\
\text{Plosive} & = \{[p], [b], [t], [d], [g], [k], [g], [q], [G], \ldots \} \\
\text{Voiced} & = \{[b], [d], [z], [j], [g], \ldots \}
\end{align*}
\]

Combinations of categories can be formed by the intersection of such sets of properties. (In the simplest case, the intersection operation takes two sets as its arguments. It is possible, and in the present context necessary, to form intersections of several sets. Intersection of several sets will be represented by placing the intersection symbol in front of an expression denoting the set of intersecting sets.) For instance:

\[
\cap \{\text{Bilabial, Plosive}\} = \{[p], [b]\}
\]

No importance attaches to the order of elements in such an expression: \(\cap \{\text{Plosive, Bilabial}\}\) gives exactly the same result.

When the intersection has a single member (other than the empty set), we have a minimally complete characterisation of that object. For example:

\[
\cap \{\text{Voiced, Bilabial, Plosive}\} = \{[b]\}
\]

The denotation of an IPA symbol, therefore, is the intersection of the sets of its properties. It is not just the set of its categories, since these categories also contain many other cardinal sounds which are not part of the denotation of that symbol.
Diacritics contribute a single property to an expression. Thus:

\[(2.5) \quad \hat{\cdot} = \{\text{Nasalised}\}\]
\[\hat{\cdot} = \{\text{Breath}\}\]

Practically all of the IPA diacritics can be interpreted as accretion of properties in the denotation. Some diacritics, however, alter the denotation of a letter by substituting one property for another. For instance, the diacritic meaning 'less rounded' denotes that a phonetic category is less rounded than the denotation of the unmodified letter. The semantics of this diacritic might be regarded as a function mapping the property \{Rounded\} onto the property \{Less Rounded\}, that is a substitution of \{Rounded\} by \{Less Rounded\}. In the face of examples such as this, the compositional interpretation of combinations of letters and diacritics becomes a little thornier. I will sidestep this issue here, as it is peripheral to the thrust of this study.

2.6.5 *Segmentation theory: internal temporal interpretation*

The model developed so far is insensitive to the internal temporal structure of phones. In the preceding section, the denotation of a symbol was given exclusively in terms of the categories (sets) by which it is defined. Since the denotation of basic expressions are physical events, which occur in time, there is a sense in which the categories by which a basic expression is defined are themselves organised in time, even if it is only as simple an organisation as 'they all occur for exactly the same stretch of time’. If this were the case, we could say that the categorial operation of intersection is coupled to a temporal operation, such as temporal overlap. Logically, however, the model is agnostic as to temporal interpretation. The arrangement in time of the defining physical properties of the categories \{Pulmonic, Egressive, Voiceless, Labial\} etc. could be any arrangement: for example, there is no requirement for the physical properties of the category \{Labial\} to coincide with or even overlap the physical properties of the category \{Voiceless\} at all. This means that the expression *pan* could denote an utterance such as [dagm] in the present model, since the physical properties of \{Voiceless\} in the denotation of 'p' are found simultaneously with the denotation of 'a', the physical properties of \{Labial\} in the denotation of 'p' are found simultaneously with the denotation of 'm', and the physical properties of \{Alveolar\} in the denotation of 'n' are found simultaneously with the denotation of 'p'!

This laxity of temporal organisation will be worthy of further consideration in later chapters, especially in connection with the phonetic
interpretation of phonological representations. But it is out of keeping with
the normal use of the IPA, which requires the physical properties of the cat-
egories of a basic expression to co-occur in time. There may be some slight
degree of temporal freedom, such as a little variability in the temporal
alignment of categories, but there must be some time where all of the phys-
ical properties of a set of categories co-occur.

Our account of phonetic denotation must include reference to a set of
temporal objects (e.g. moments in time or intervals of time) T. If the objects
in T are moments of time, we can say that the categorial combinator ∩ is
paralleled by the temporal combinator ∧, which equates moments of time.
If the objects in T are intervals of time, we can say that the categorial com-
binator ∩ is paralleled by the temporal relation Overlaps ⊓ which denotes
transitive overlap (overlap which satisfies the constraint 'if a Overlaps b and
b Overlaps c then a Overlaps c').

We can build both moments and intervals into the model in at least two
very simple ways. Firstly, we can define T to be a dense set of moments of
time t_n totally ordered by a reflexive, transitive, asymmetric relation ≼, and
define an interval I to be a sequence [t_n, t_m] (cf. Dowty 1979: 139). We can
then define the relation Includes ⊑ and (transitively) Overlaps ⊓ as follows:

\[(2.6) \quad [t_a, t_b] \text{ Includes } [t_c, t_d] \text{ if and only if } t_a \leq t_c \text{ and } t_b \geq t_d\]
\[\text{[t_a, t_b] Overlaps [t_c, t_d] if and only if } [t_a, t_b] \cap [t_c, t_d] \neq \emptyset\]

Alternatively, we can define T to be a set of intervals, and add either inclu-
sion or overlap to the model as a primitive (see van Benthem 1988: 59–60).
With intervals introduced in the second way, a moment can be defined to
be a minimal subinterval (i.e. a least element in the inclusion relation). The
difference between a model with intervals and transitive overlap and a
model with moments and equation of moments is of very little consequence
in the present context, however, and I shall henceforward assume that T and
< model the co-ordination of either moments or intervals (≥ can be
derived from <), and either exact momentary co-occurrence or transitive
overlap, without further distinction. I will thus use the symbol ≥ to denote
transitive overlap also, unless a specific distinction needs to be drawn
between ≥ and ⊓.

In general phonetic theory, another important aspect of the temporal-
categorial interpretation of a category is the continuous motion in time of
the acoustic or articulatory phonetic parameter(s) of the categorial inter-
pretation, such as the slope, direction, speed and target of formant motions
which cue place of articulation in consonants. Since IPA (1949) and the
phonological theories examined in later chapters ignore them, such category-internal dynamics will not be considered here.

This section began by assuming, in line with the preceding development of the model, that categories and intervals were separable kinds of objects, and now concludes that every category in the denotation of an expression is indexed with an interval. The combination of a category or property $p$ and an interval $i$ is called an event $(e)$ by Bird and Klein (1990), making explicit the accounts of Russell (1914), van Benthem (1983: 113) and others, for whom events are primitives. Generally speaking, event logics are homomorphic to interval logics, since each event $e=(i,p)$ maps to its interval component $i$, and event precedence, overlap and inclusion have the same properties as interval precedence overlap and inclusion. Because of this homomorphism, it is not crucial most of the time to distinguish the logic of intervals from the logic of events, and as with the point/interval distinction, I shall employ the operators $<$ and $@$ for precedence and transitive overlap of points, intervals or events, making finer distinctions between them only where it is necessary. I refer the reader to the excellent discussion in Bird and Klein (1990) for a fuller account of an event logic for phonetic and phonological description.

2.6.6 The semantics of IPA strings
Having considered the interpretation of basic expressions (single possibly complex symbols), I shall now consider the interpretation of strings of basic expressions. As with the semantics of basic expressions, we can consider the categorial interpretation of strings of basic expressions independently of their temporal interpretation, to begin with. We shall consider the categorial interpretation only briefly. In IPA (1949) there are a few symbols which could be regarded as 'string diacritics' that affect the categorial interpretation, such as the breve ~, which 'placed over a letter indicates the weaker element of a diphthong e.g. au. It is not as a rule necessary to insert this mark' (IPA 1949: 17).

I shall assume that ~, like :, can be treated as a diacritic modifier of basic expressions, as above. There are also a few symbols which must be regarded as 'string diacritics', namely the stress marks ' and ,, which indicate that the following syllable has respectively strong and medium stress. Whatever these two categories of 'stress' may be, it is clear that the intended effect of ' and , (a) is categorial and (b) usually affects more than one segment. Since these cases of categorial modification of strings are exceptional, though, we shall ignore them, and suppose that the interpretation of strings is purely temporal.
Temporal interpretation of IPA strings

The only means of representing the temporal arrangement of reference categories in the IPA are concatenation of diacritics and basic expressions, which was considered above, and concatenation of basic expressions. (The IPA ligature symbol ~ indicates simultaneous production of the denotation of two categories. It is thus, like the Organic Alphabet’s ‘link’ symbol, just a notational device for treating any basic expression as if it were a diacritic, in order to modify another basic expression.) To take account of the denotation of concatenation we must include in the model an ordering on the set of times. This ordering is represented with the symbol <. We need not say much about <, except that it is a transitive, asymmetric relation on times. Other details which refine the submodel of time or temporal sequence are considered further below.

So far, I have considered temporal precedence and temporal overlap to be completely separable relations. It is an artificial distinction, however, and ideally precedence and overlap should be treated in an integrated fashion. For a moment-based temporal structure, this is easy: for instance, we could define @ and < in terms of a total ordering ≤ in the usual way:

\[(2.7) \quad a @ b \text{ if and only if } a \leq b \text{ and } b \leq a \]
\[a < b \text{ if and only if } a \leq b \text{ and } \neg a @ b\]

With an interval-based temporal structure, the interaction of precedence and overlap is more complicated, though still tractable. We can distinguish at least two possible alternative accounts, which I shall now consider (a third is taken up later in this chapter in connection with the interpretation of componential notations).

If ‘precedes’ is a total ordering (i.e. for every a and b, either a precedes b or b precedes a), we shall call the segmentation theory strict. If ‘precedes’ is a partial ordering (i.e. there may be some a and b for which neither a precedes b nor b precedes a, in which case we say that a and b overlap), we shall call the segmentation theory relaxed. In general, a segmental phonetic transcription must be either relaxed or detailed. To see why this is so, consider a phonetic transcription of a voiceless, aspirated stop releasing into a vowel, for instance [pʰaː]. Despite the fact that [ʰ] precedes [a], in such transitions the quality of the vowel is also audible during the aspiration phase of the stop. In a ‘relaxed’ theory of segmentation, we simply say that the denotation of [aː], and thus the detail of the vocalic quality during the aspiration is denoted by the symbol for the vowel following the symbol for the aspiration. The total ordering of the letters in the transcription does not
2.6 Semantics of IPA transcriptions

Figure 2.2 ‘Relaxed’ segmentation

Figure 2.3 ‘Strict’ segmentation

imply that the denotations of each symbol are also totally ordered. In this case, for example, the aspiration interval is included in the vowel interval. The ordering of the denotations of each symbol in time can be illustrated in a segmentation diagram such as figure 2.2.

In such segmentation diagrams, the horizontal dimension denotes the relative duration and the sequence or overlap of intervals of time, although the horizontal dimension of the diagram is not drawn to a precise timescale. The vertical dimension is not significant in general, although specific theories of phonological notation could give a theory-internal significance to the vertical dimension.

If the possibility of denotational overlap is not entertained in the transcription theory (segmentation is strict), however, the phonetic transcription needs to include more detail about the quality of the aspiration. For instance, the transcription could be written [pʰɑː], denoting a strict sequence of segments, illustrated in figure 2.3. In these two simple examples, the transcriptions are isomorphic sequences of letters: it is in their denotations that they differ in whether they permit objects to overlap or not, and this distinction makes it necessary to employ different symbols in the two cases.

The semantics of the IPA can thus be summarised as follows:

1. the denotation of each member of the set of IPA symbols is the intersection of a set of physical properties;
the denotation of a diacritic is a function from denotations of basic expressions to denotations of basic expressions;
the denotation of concatenation is temporal adjacency (i.e. sequence and/or overlap) in which the denotation of the second symbol ends no earlier than the denotation of the first symbol;
the denotation of cocosmentation is temporal overlap.

2.7 Another example: Sweet’s Organic Alphabet

The account of the syntax and semantics of the IPA developed above introduced all of the fundamental concepts required for accounts of the syntax and semantics of many other notation systems too. In this section I shall briefly consider Sweet’s Organic Alphabet, which has been mentioned several times already as a good example of a componential notation system.

Sweet’s Organic Alphabet is a phonetic notation for recording speech in terms of articulatory reference categories alone. The vocal tract is divided into parts, and each part of the vocal apparatus is assigned a symbol. If any part of the vocal tract functions in the production of an utterance, then the symbol for that part is included in the representation of that utterance. Of course, this simple idea suffers from the fundamental problem of observation: every part of the vocal apparatus ‘functions’ in some way, and thus the set of reference categories is necessarily selective. The Organic Alphabet is segmental, in that utterances must be cross-parametrically segmented by the observer before each segment can be classified and transcribed according to the reference categories.

The aim of the Organic Alphabet, like its predecessor, Bell’s Visible Speech, was to make the mechanisms of speech more ‘visible’, for the sake of academic phoneticians, learners of foreign languages, and as an aid to help deaf-mutes to form speech. It attempts this in two ways:

1. Each reference category has a unique symbol (unlike IPA categories). This makes the system simpler and more systematic and perhaps therefore easier to learn.
2. Every reference category is physiological, as far as possible, and each symbol is a simple depiction of the relevant articulator or organ. For instance, the IPA category ‘voicelessness’ corresponds to the Organic Alphabet category ‘open glottis’ (Henderson 1971: 257) or ‘breath’ (the names of these reference categories are somewhat variable) symbolised O, whereas the IPA category ‘glottal
2.8 Feature-matrix representation

stop' corresponds to the Organic Alphabet category 'closed glottis', symbolised \( \_\). The \( \_\) and \( \_\) are intended to be simple pictures of the appearance of the vocal cords (open and closed, respectively), as seen from above. (\( \_\) lives on in the IPA as the \([_\_]\) (breath) diacritic.)

Both place and manner of articulation are related to articulators and their states, each articulator or state being provided with a unique symbol.

The syntax of the Organic Alphabet is quite simple. Basic symbols, representing reference categories, can be combined together into compound symbols representing complete (or more complete) descriptions of vocal-tract state, and compound symbols are then concatenated to denote sequences of vocal-tract states.

Examples of the Organic Alphabet in use come nowhere near the phonetic ideals intended for it, however. The compound symbols are never complete in their representation of vocal-tract state. In its use the Organic Alphabet is strongly shaped by phonological considerations, and it is only in the transcription of non-linguistic sounds that Bell or Sweet recorded vocal-tract configurations in the maximal detail possible within the capability of the notation-system. In general use, the Organic Alphabet is just like any other alphabet, with the added disadvantages of being exotic for the learner and printer alike. In this respect, it is hardly any different from exotic experimental alphabets, such as the Shaw Alphabet (Read 1962).

Being segmental, the Organic Alphabet is also open to the criticism that 'pretheoretic' segmentation violates the spirit of phonetic iconicity intended for the Organic Alphabet. This criticism is strengthened when it is acknowledged that the segmentation required by the Organic Alphabet is not actually 'pretheoretical', but is highly theoretical – the theory being the (segmental) phonological theory for which phonetic records are intended to be the preliminary data.

2.8 Feature-matrix representation

Whereas the denotation relation for the IPA is a one-to-many mapping from symbols to sets of reference categories, the Organic Alphabet has the explicit goal of a one-to-one denotation relation between symbol(-parts) and reference categories. In other words, the denotation relation of the Organic Alphabet is simpler and more systematic than in the IPA.
Feature-matrix representation shares this property with the Organic Alphabet, but goes further, in making the names of phonetic reference categories themselves the formal elements from which phonological categories are constructed. Thus the denotation relation for features is simply the identity function, feature concatenation being systematically interpreted as temporal co-occurrence (or overlap – more on this in section 2.10). Feature-matrices are often (Chomsky and Halle 1968, Brainerd 1971) described as \( n \times m \) rectangular matrices of features, but this is not generally the case, as the columns of the ‘matrix’ are not necessarily of equal length. A more accurate definition is that a feature-matrix is a string of column-vectors, where a column-vector is a set of feature-value pairs. An additional but rarely stated condition on the syntax of column-vectors is that each feature occurs only once and may have only one value. In other words, column-vectors are functions (Goldsmith 1976: 22) from the set of features to the set of values (partial functions if the value of any feature is undefined for some column-vector; see Gazdar et al. 1985: 24–5). Because of the condition that each feature may occur only once in a column-vector and may have only one value, the operation by which feature-value pairs are combined to form column-vectors cannot be concatenation, or even commutative concatenation, however. We can extend commutative concatenation to include the single-value constraint to form a new operator \( \cup \) (unification), with the following elementary properties:

**Unification of two categories**

1. \( C_1 \) unifies with \( C_2 \) if and only if for all feature-value pairs \([f: v]\) in \( C_1 \), \([f: v]\) unifies with \( C_2 \).

**Unification of a single feature-value pair \([f: v]\) with a category \( C\)**

2. If \( f \) does not occur in \( C \), the unification of \([f: v]\) and \( C \) is \([f: v] \oplus C \) (i.e. the concatenation of \([f: v]\) and \( C \)).

3. If \([f: u]\) occurs in \( C \), \([f: v]\) unifies with \( C \) if and only if \( v \) unifies with \( u \).

**Unification of atomic feature-values \( u \) and \( v \)**

4. \( u \) unifies with \( v \) if and only if \( u = v \) (i.e. \( u \) and \( v \) are identical).

For example, the category \([\text{coronal}: +, \text{voice}: -, \text{nasal}: +]\) unifies with \([\text{nasal}: +, \text{back}: -]\) by the following line of reasoning:

**Clause 1**: \([\text{coronal}: +, \text{voice}: -, \text{nasal}: +]\) unifies with \([\text{nasal}: +, \text{back}: -]\) if and only if \([\text{coronal}: +], [\text{voice}: -] \) and \([\text{nasal}: +] \) unify with \([\text{nasal}: +, \text{back}: -]\).
Clause 2: [coronal] does not occur in [nasal: +, back: -], so \([\text{coronal: +}] \cup [\text{nasal: +, back: -}] = [\text{coronal: +}] \oplus [\text{nasal: +, back: -}] = [\text{coronal: +, nasal: +, back: -}]\).


Clause 3: [nasal] occurs in [voice: -, coronal: +, nasal: +, back: -] (i.e. as [nasal: +]), so we must check that the values of [nasal] unify.

Clause 4: + \cup + if and only if + = +. The values of [nasal] unify, so the unification of [coronal: +, voice: -, nasal: +] and [nasal: +, back: -] is [voice: -, coronal: +, nasal: +, back: -].

This brief and simplified definition of unification will be supplemented by more detailed discussions in chapters 3 and 5.

The way in which feature-matrix representation works is generally like the Organic Alphabet, then, and the syntactic resources of concatenation and cocatenation and the semantic resources of event logic and model theory discussed earlier in this chapter are sufficient to construct and interpret IPA, Organic Alphabet and feature-matrix representations. Thus, despite the superficial differences between representations in these three different notation-systems, they all exploit similar formal resources.

By employing componential symbols, a judicious selection of phonetic or phonological features can help in the description of regularities that cut across otherwise dissimilar phonetic or phonological classes. For example, the fact that \([h]\) and \([k]\) are both 'back', in that they involve raising of the tongue dorsum towards the velum, is reflected in the fact that their segmental descriptions according to Chomsky and Halle (1968) include the feature \([+ \text{back}]\) in both cases.

2.9 Segmentation and temporal overlap

The definition of \@ given in section 2.6.7 holds good as an account of simultaneous (or coterminous) intervals of time, and even non-coterminous, overlapping intervals of time, provided that they form an equivalence class, or in other words, provided that < is a partition. I shall call the first case an example of 'strict' segmentation, and the second 'relaxed' segmentation. The first case can be illustrated by the segmentation diagram of figure 2.4, and the second by figure 2.5, both of which portray the relative temporal arrangement of the intervals in (2.8).
Strict segmental interpretation is frequently employed in spectrographic studies (e.g. Fant 1969, Roach et al. 1990), where it is called 'acoustic segmentation'. Relaxed segmental interpretation is frequently encountered in articulatory studies, and is the standard theory of speech organisation.
(Ladefoged 1975: 52, Fowler 1983, Mohanan 1986, Browman and Goldstein 1989). It is attractive because, while remaining segmental, and thus offering a simple interface to segmental phonology, it admits of a degree of overlapping between adjacent segments as a consequence of slight differences in the intrinsic durations of subsegmental components.

### 2.10 Superweak segmental interpretation

A number of phoneticians have proposed that the speech phenomena involving articulatory overlap discussed above – assimilation, epenthesis, metathesis and elision – are phonetic phenomena, attributable to the inherent physical properties (e.g. inertia) of the human speech organs (see Ohala 1990: 266–71). The fact that these phenomena are not universal, and may vary in their details across languages which do exemplify them, shows that a purely physical phonetic account is not tenable, however. Mohanan (1986: 163) observes: ‘The fact that [epenthesis] appears in American English, but not in South African English, shows that it is a physiologically motivated linguistic phenomenon, not a purely physiological phenomenon, and therefore must be represented in the outputs of grammars of languages.’ A similar debate concerning consonant–vowel coarticulation has been engaging phoneticians for many years. Generally speaking, it is possible to distinguish two kinds (or ‘senses’) of coarticulation, termed ‘hard’ and ‘soft’ by Fujimura (1981: 109). The term ‘hard’ coarticulation refers to the involuntary, non-linguistic, language-universal, phonetic aspects of coarticulation which are believed by articulatory phoneticians to be consequences of inalterable physical properties of the vocal organs, such as muscular inertia. ‘Soft’ coarticulation refers to language-specific, low-level phonological aspects of coarticulation. There has been an extensive debate over the relative contribution of ‘hard’ and ‘soft’ coarticulation to observed coartulatory phenomena, such as vowel–(consonant–)vowel transitions. Whalen (1990) summarises this debate, and presents some experimental results that suggest that ‘soft’ coarticulation predominates. Klatt in Allen, Hunnicutt and Klatt (1987: 109–12) shows that there are two aspects to coarticulation in American English: the effect of coarticulation on formant transitions from consonant to vowel can be modelled by a ‘locus theory’ linear equation, provided that the vowels are divided into three sets, ‘front’, ‘back unrounded’ and ‘back rounded’. In this account, the locus theory models the ‘hard’ aspect of coarticulation, whereas the division of the vowels into three phonological classes models the ‘soft’ aspect of coarticulation.
Segmental representations

Hard coarticulation can be modelled by regarding consonants as being 'coproduced' with (i.e. concatenated to) vowels, rather than concatenated to them (Öhman 1966, Perkell 1969, Gay 1977, Mattingly 1981, Fowler 1983) (fig. 2.6). In order for 'hard' coarticulation to be modelled in this way, temporal overlap cannot be modelled by the @ relation. In the temporal relations underlying figure 2.6, @ does not form equivalence classes, and < is not a partition. The problem is that @ is only suitable for modelling overlap of intervals when the sequence of segments is not destroyed by the overlapping of neighbouring segments. This is only the case if there is a non-empty subinterval of each interval in @, a circumstance which does not obtain with respect to figure 2.6. In order to include examples such as figure 2.6 in the theory of segmental phonetic interpretation, the notion of segmentation must be relaxed even further, to allow for the intransitive overlap of components, because although C1 and V overlap, and V and C2 overlap, it is not the case that C1 and C2 overlap. Transitive overlap (i.e. @) ensures that every segment is in general more narrowly bounded in time than segments whose components may overlap intransitively. The duration of segments constructed by transitive overlap is bounded above by a multiple of the duration of the longest component, whereas the duration of segments constructed by intransitive overlap is bounded above by the sum of the durations of its components. I shall call segmental interpretation with intransitive overlap 'superweak'.

If transitive overlap is replaced by intransitive overlap, there is no need for 'hard' coarticulation to be implemented in the phonological component (as Griffen 1985 and Bird and Klein 1990 suggest), since it can be made a consequence of the phonetic interpretation of segmental phonetic representations, according to the coproduction model.

It should be apparent from figure 2.6 that this 'superweak' segmental interpretation is also adequate to model suprasegmental events, just as vowels may be regarded as 'suprasegmental' to consonants in the above account of coarticulation. Thus IPA suprasegmental symbols such as ' and , while segmental in their syntax (in that they are concatenated to other

![Figure 2.6 Coarticulation modelled by temporal overlap (coproduction)](image)
symbols in strings no differently from the symbols for segmental categories), can nevertheless be given a suprasegmental denotation.

2.11 Conclusion

In this chapter, I have examined the form and content of a segmental phonetic notation, the IPA. I employed this example to illustrate many fundamental concepts of phonetic notations and their denotation, such as 'sound' and 'letter', 'type' and 'token', 'relaxed' and 'strict' segmentation, componential representation and the key characteristics of model theory.

In the discussion of the relaxed and superweak interpretations of segmental phonetic representations, I briefly showed that a number of phonetic phenomena, including epenthesis and hard coarticulation, could be regarded as deriving from the overlap of phonetic components. In chapter 5, I shall extend this list to include some instances of assimilation, elision and metathesis, and in the following chapters I shall argue that the development of purely phonological accounts of these phenomena in transformational-generative, prosodic and 'non-linear' theories of phonology are unnecessary, or at least overcomplicated, since these phenomena can be regarded in part as matters of phonetic interpretation.
3 Segmental and transformational phonology

The very existence of phonological phenomena such as alternation, systematic regularity, and diachronic and synchronic sound changes require, ipso facto, that some type of segmental level be postulated in order to capture significant linguistic generalizations. In describing the sound structure of a given language, then, a level of segmental representation is required in order to account for the idiosyncratic and predictable regularities in the sound pattern of that language.

(Pisoni and Luce 1987: 34)

3.1 Introduction

In this chapter I examine the two principal segmentally based theories of phonological representation, phonemic and transformational phonology. These were the first ‘class’ of phonological theories which historically arose from segmental phonetic theory. After a brief historical resumé, I shall discuss the theory of phonemic phonology, the representation of distinctive segmental phonological units, and phonological representations made up of those units. In section 3.4, I describe transformational-generative phonology, a derivational theory of the mapping from lexical (phonological) to surface (phonetic) segmental representations. I illustrate some formal problems of transformational grammars with a transformational phonological analysis of three kinds of allophony in Japanese. Despite the wide use of and apparent empirical support for the rewrite-rule notation and the primitive operations of insertion, deletion and replacement in generative phonology, I argue in section 3.5 that the grammar formalism employed in transformational-generative phonology is excessively unconstrained, being equivalent to unrestricted rewriting systems. I also argue that the constraints on transformational phonology which have been proposed cannot be relied upon to overcome the problems of excessive power. In section 3.6 I present two enrichments of the theory of phrase-markers – multiple dominance and complex symbols – which permit generalisations
about linguistic representations to be expressed directly (without the inter-
vening notion of 'derivation'), without relaxing the restrictiveness of the
classical phrase-structure grammars of the Chomsky hierarchy. I suggest
that the declarative theory of linguistic representations which results is
preferable to the procedural SPE-type approach, a topic which leads into
the subject of the next chapter, 'non-linear phonological representation'.

3.2 Historical context

While the development of phonetic notations such as the IPA provided a
tool by which the (phonetic) irregularities and ambiguities of conventional
orthographies might be overcome, it permitted a far greater degree of dis-
crimination and detail to be recorded than is necessary for many practical
purposes, such as the indication of pronunciation in dictionaries or foreign-
language teaching material. Phonetic notation is, in many respects, too
distant from language and too close to speech for such purposes. Phonetic
notation actually confounded the early modern phonologists' desire for one
symbol per distinctive sound-type in the language. It was realised from the
early years of this century that superfluous detail is unwelcome: Sweet
(1911, quoted in Henderson 1971: 252), for example, wrote: 'the pronuncia-
tion of these diphthongs [in high and how] varies so much in different parts
of the English-speaking territory, and the distinctions are so minute that it
would be inconvenient to express them in writing; and as these distinctions
are non-significant, it would be useless to do so'. The problem of predict-
able variation in the pronunciation of a letter in different contexts does not
usually present any practical difficulty to students of foreign languages,
however, and a one-to-one mapping between the putative minimal, func-
tional, concatenative units of speech (segments) and the minimal concatena-
tive units of writing (letters) was established as a sufficient and efficient
basis for linguistic notation. Thus was born the modern avatar of the medi-
aeval 'letter', the phoneme.

The history of segmental phonology is almost commensurate with the
history of modern phonology. Consequently, it is not possible or appropri-
ate to give anything more than the briefest sketch here. Throughout the
heyday of structuralist phonology, the phoneme theory was the corner-
stone of phonological theory. In almost every school of linguistics, despite
an extensive variety of interpretations, the phoneme was generally regarded
as the minimal unit of language, the point at which sound and meaning met.
(The London school of Prosodic Phonology is a prominent exception from
this period.) With the development of transformational-generative phonology in the 1960s, the phoneme was finally usurped by the distinctive feature as the prime unit in the analysis of speech, but the phonological segment still has a place in the scheme of things to the present day, even in non-linear theories, such as Autosegmental Phonology, which reject segmental phonetic representations, but retain segments in 'underlying' phonological representations. Since Schane (1971) there has been a gradual reinstatement of the phoneme as the basic concatenative element of lexical phonological representations.

The theory of generative grammar which Chomsky, Halle and others applied to phonology derived originally from research on the syntax of sentences. Phonology was regarded as an interpretive component of grammar: that is, it was dependent on syntactic structure. Consequently, although generative grammar did develop one widely used non-linear form of representation, syntactic phrase-markers, it retained a unilinear, concatenative view of phonological representation. This may have been in part because transformational grammarians did not develop graphical representations of phrase-structure for grammars more powerful than context-free, such as the type of rules used by generative phonologists. By employing powerful systems of quite simple rules to describe analytically challenging phenomena such as English stress or Finnish vowel harmony, phonologists had little reason to consider using any other form of phonological representation than strings of segmental symbols.

The derivational theory of grammar employed in generative phonology is historically associated with the computational model of sequential computing machines (automata). There is a close relationship between the Chomsky hierarchy of generative grammars, and a hierarchy of automata of various kinds (see section 3.5 below). In retrospect, it seems that the appeal of the sequential machine metaphor in generative grammar arises from the fact that at the time that generative grammar originated there were no other discrete computing models in existence (Bird 1995: 27 makes a similar point). Parallel and distributed computing, and object-oriented and declarative programming had not been developed, so there were no other linguistic-processing metaphors available at that time.

### 3.3 Phonemic phonology

Let us begin by examining the historical progenitor of transformational-generative phonology, phonemic phonology. It was perceived from quite early on by such phonologists as J. Baudouin de Courtenay (see Jakobson
3.3 Phonemic phonology

and Waugh 1979: 14), Saussure (1916), Sapir (1925) and Firth (1935a) that the meaning of a word or phrase is not signalled by exactly how it is pronounced – if it was, physiological differences between people and the acoustic consequences of those differences would make speech communication almost impossible – but how it differs from the other words or phrases which might have occurred instead. For example, slight variations in the precise degree of aspiration occurring towards the beginning of a token of the word *pit* are irrelevant within certain limits, so long as it is sufficiently distinguished from *bit*. It is for this reason that the [p] in English is interpreted as /b/ (less aspirated than voiceless [pʰ] = /p/), whereas French [p] is /p/, being less voiced than [b] = /b/. Sapir's (1925) presentation of the importance of relations in phonology is particularly lucid:

Mechanical and other detached methods of studying the phonetic elements of speech are, of course, of considerable value, but they have sometimes the undesirable effect of obscuring the essential facts of speech-sound psychology. Too often an undue importance is attached to minute sound discriminations . . . It is true that no two individuals have precisely the same pronunciation of a language, but it is equally true that they aim to make the same sound discriminations, so that if the qualitative differences of the sounds that make up A's pattern from those that make up B's are perceptible to a minute analysis, the relations that obtain between the elements in the two patterns are the same. In other words, the patterns are the same pattern. A's s, for instance, may differ markedly from B's s, but if each individual keeps his s equally distinct from such points in the pattern as *th* (of *think*) and *sh* and if there is a one-to-one correspondence between the distribution of A's s and that of B's, then the difference of pronunciation is of little or no interest for the phonetic psychology of the language. (Sapir 1925: 20–1)

From this 'relational' view of the phoneme, which abstracts away from the phonetic content of its manifestations in various contexts and in the mouths of different speakers, the idea that phonemes are the linguistic invariants of speech originated, an idea which remains popular today, according to the work of many speech researchers (e.g. Halle and Stevens 1979, Stevens and Blumstein 1981, Pisoni and Luce 1987), speech technologists (e.g. Newell et al. 1973, Ainsworth 1976, Shoup 1980, Allen, Hunnicutt and Klatt 1987, Bladon et al. 1987, Holmes 1988) and psychologists (e.g. Eimas 1985).

3.3.1 Variability in phonemic phonology

An issue which all phonological theories must therefore face is the question of reconciling invariance of phonological representations with variability
of their phonetic denotations (assuming the physical view of phonetic denotations discussed in chapter 2). Distinctive-feature notations offer an intuitively satisfactory mechanism for the description of many aspects of such variability, as I shall illustrate in section 3.4.2 below. However, let us first examine the nature of such variability and its description in terms of classical phonemic phonology. In general, such variance is one of several kinds: (i) free (random, unexplained, apparently non-causal) variation; (ii) external variation (variation which is apparently systematic, but not conditioned by grammatical factors, such as social or stylistic variation); (iii) linguistically conditioned variation; (iv) automatic (mechanical, phonetic, involuntary) variation.

The boundaries between these categories are fuzzy. Linguistically conditioned variability is generally regarded as a phonological matter; automatic variation a phonetic one. I shall not consider free and external variation further. Automatic variation will be considered only slightly here, for example in section 3.4.1, in which automatic coarticulation is examined as far as is necessary to dismiss it as an account for a palatal assimilation phenomenon in Japanese. Since phonetic studies which examine purely mechanical aspects of speech typically examine non-linguistic utterances (e.g. reiterant speech) and do not usually control for cross-language variability, it is prudent to assume at first that the distinction between automatic and linguistically conditioned variation is unknown, and that every instance of variation which is systematic might be linguistically conditioned, and thus a suitable issue in the evaluation of phonological representations (see e.g. Whalen 1990). In the remainder of this section, I shall illustrate non-distinctive phonological variability with some examples from Japanese. Specifically, I shall discuss the well-known phenomenon of 'voiceless vowels' in Japanese and its relationship to 'vowel elision', 'syllabic fricatives', 'affrication', 'palatalisation' and their phonological determinants. I shall return to this set of phenomena in this and the following two chapters, and use it to illustrate aspects of phonological representation in various phonological theories.

### 3.3.2 Variability of /i/ and /u/ in Japanese

Japanese orthography contains, in addition to Chinese characters, two native Japanese syllabaries, in which the units represent CV combinations and syllable-final consonants. These units, called 'moras', are the accent-bearing units of Japanese. Here I shall argue that 'mora' is also a well-defined category of phonological constituents in Japanese. A full discussion of the place of moras in Japanese phonology is given in chapter 6.
3.3 Phonemic phonology

Table 3.1 ‘CV’ and ‘CyV’ moras

<table>
<thead>
<tr>
<th></th>
<th>/a/</th>
<th>/i/</th>
<th>/u/</th>
<th>/e/</th>
<th>/o/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k/</td>
<td>ka</td>
<td>ki</td>
<td>ku</td>
<td>ke</td>
<td>ko</td>
</tr>
<tr>
<td>/s/</td>
<td>sæ</td>
<td>fi</td>
<td>su</td>
<td>se</td>
<td>so</td>
</tr>
<tr>
<td>/t/</td>
<td>tæ</td>
<td>tfi</td>
<td>tsu</td>
<td>te</td>
<td>to</td>
</tr>
<tr>
<td>/n/</td>
<td>nae</td>
<td>ni</td>
<td>nu</td>
<td>ne</td>
<td>no</td>
</tr>
<tr>
<td>/h/</td>
<td>ha</td>
<td>çi</td>
<td>φu</td>
<td>he</td>
<td>ho</td>
</tr>
<tr>
<td>/m/</td>
<td>ma</td>
<td>mi</td>
<td>mu</td>
<td>me</td>
<td>mo</td>
</tr>
<tr>
<td>/y/</td>
<td>yæ</td>
<td>yu</td>
<td></td>
<td></td>
<td>yo</td>
</tr>
<tr>
<td>/r/</td>
<td>ræ</td>
<td>ri</td>
<td>ru</td>
<td>re</td>
<td>ro</td>
</tr>
<tr>
<td>/w/</td>
<td>wa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/ky/</td>
<td>kyæ</td>
<td></td>
<td>kyu</td>
<td></td>
<td>kyo</td>
</tr>
<tr>
<td>/sy/</td>
<td>sæ</td>
<td></td>
<td>fu</td>
<td></td>
<td>fo</td>
</tr>
<tr>
<td>/ty/</td>
<td>tæ</td>
<td></td>
<td>tfu</td>
<td></td>
<td>tfo</td>
</tr>
<tr>
<td>/ny/</td>
<td>nyæ</td>
<td></td>
<td>nyu</td>
<td></td>
<td>nyo</td>
</tr>
<tr>
<td>/hy/</td>
<td>çæ</td>
<td></td>
<td>çyu</td>
<td></td>
<td>cyo</td>
</tr>
<tr>
<td>/my/</td>
<td>myæ</td>
<td></td>
<td>myu</td>
<td></td>
<td>myo</td>
</tr>
<tr>
<td>/ry/</td>
<td>ræ</td>
<td></td>
<td>ryu</td>
<td></td>
<td>ryo</td>
</tr>
</tbody>
</table>

Note: [u] represents an unrounded close back vowel, [ʃ] represents voiceless alveo-palatal friction (whether as a fricative by itself, or as part of an affricate), [ɸ] represents voiceless bilabial friction, [n] represents palatal or alveo-palatal closure with nasality; [k], [t] represent aspirated stops, [ɾ] represents a tap auditorily similar to a lax [d] or short, clear [l]. [y] is used instead of IPA [j], after the practice of standard phonemic analyses of Japanese, such as Bloch (1946, 1950).

Table 3.1 shows the combinations of consonants and vowels that may form a mora in Japanese, with the exception of voiced obstruents, and /p/, a morphophonological alternant of /h/. The rows and columns are phonemic, and the entries in this table are systematic phonetic (i.e. allophonic) representations. To illustrate the phonetic variability of the entries in table 3.1, a selection of normalized extracts from my own impressionistic phonetic records of three native speakers of Japanese are presented in table 3.2. The examples in table 3.2 are representative of the moras whose systematic phonetic representations are given in lines 3 and 4 of table 3.1. Each of these is attested in a number of variants, some of which can be shown to be context-specific variants, that is, fricative nucleus, whispered nucleus and close vocalic nucleus. There are also many context-specific variants of the moras with half-open and open vocalic nuclei, but that variance is not
shown here, as it is not relevant to the present discussion. The superscripts [\textsuperscript{p}], [\textsuperscript{u}], and [\textsuperscript{o}] indicate the oral cavity resonance, that is, the secondary articulation or ‘vocalic colouring’ of consonantal articulations. Superscript [\textsuperscript{p}] indicates clear, palatal, front resonance, with no lip-rounding; [\textsuperscript{u}] indicates dark, velar, back resonance with protruded, spread lips; and [\textsuperscript{o}] indicates a more central quality, with notable lip-rounding.

The greatest variability is seen in the phonetic nucleus in each case. I have distinguished fricative nuclei from whispered and voiced vocalic nuclei. By ‘phonetic nucleus’, I mean the period of continuant articulations (e.g. vowels, fricatives etc.) that occupies most of a mora’s duration. Nucleus whispering in Japanese is a very commonly described phenomenon: in segmental terms, it is often said that close vowels are whispered or ‘devoiced’ when they occur either between voiceless consonants or utterance-finally

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>a) /tV/ moras:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative nucleus (syllabic continuant)</td>
<td>i\textsuperscript{f}</td>
<td>u\textsuperscript{t}</td>
</tr>
<tr>
<td>Whispered nucleus</td>
<td></td>
<td>t\textsuperscript{si}</td>
</tr>
<tr>
<td>Close vocalic nucleus</td>
<td>i\textsuperscript{i}</td>
<td>su</td>
</tr>
<tr>
<td>Half-open vocalic nucleus</td>
<td>te</td>
<td>o</td>
</tr>
<tr>
<td>Open vocalic nucleus</td>
<td>i\textsuperscript{a}</td>
<td></td>
</tr>
<tr>
<td>b) /sV/ moras:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative nucleus (syllabic continuant)</td>
<td>i\textsuperscript{f}</td>
<td>u\textsuperscript{s}</td>
</tr>
<tr>
<td>Whispered nucleus</td>
<td></td>
<td>s\textsuperscript{y}</td>
</tr>
<tr>
<td>Close vocalic nucleus</td>
<td>i\textsuperscript{i}</td>
<td>su</td>
</tr>
<tr>
<td>Half-open vocalic nucleus</td>
<td>se</td>
<td>o</td>
</tr>
<tr>
<td>Open vocalic nucleus</td>
<td>i\textsuperscript{a}</td>
<td>sa</td>
</tr>
</tbody>
</table>
3.3 Phonemic phonology

after voiceless consonants, subject to certain accentual restrictions (Ueda 1976, Hasegawa 1979a, b, and Haraguchi 1984).

3.3.3 Variability of /t/ and /ty/ in Japanese

Table 3.2(a) also showed that as well as /l/ and /u/ varying in the environment /t-/, /t/ varies according to the vowel phoneme which follows it. In other words, high vowels can vary depending on the consonant which precedes them, and some of those consonants vary depending on the vowel which follows them. As a result of these mutual conditions on variability, classical phonemic accounts of Japanese phonology invariably encounter difficulties in analysing the patterns of palatalisation and affrication of voiceless coronal consonants before close vowels (illustrated in table 3.2). If each mora is analysed as the concatenation of consonantal onset and vocalic nucleus, the phonetic realisations of line 4 moras as (roughly) [ta], [tʃi], [tsu], [te], [to], and line 3 moras as [sa], [ʃi], [su], [se], [so] suggest context-dependent variation of the form:

(3.1) 
/t/ is realised as:
  [tʃ] before /i/,
  [ts] before /u/,
  [t] before any of the other vowels.

(3.2) 
/s/ is realized as:
  [ʃ] before /i/,
  [ʃ] before any of the other vowels.

For instance, Daniels writes: ‘In all cases, [ts] may be regarded as a variant of [t] . . . under the influence of [u] or [u] . . . When prefixed to [i] or [t] . . ., [tʃ] may be regarded as a variant of [t], . . . and [ʃ] as a variant of [s], under the influence of these vowels’ (Daniels 1958: 58–9). Problems arise in placing the moras [tʃa], [tʃu], [tʃo] and [ʃa], [ʃu], [ʃo] into this scheme, for there is now, apparently, a set of mora-initial consonantal contrasts with [ta], [tsu], [to], [sa], [su] and [so]. Numerous analyses (e.g. Bloch 1946, 1950, McCawley 1968) have proposed to represent the distinctively palatalised series of onsets as /ty/, /sy/ and so on, giving a repertoire of moras that includes the following:

CV:  /tal/, /tul/, /tul/, /tel/, /tol/       
     /sal/, /si/, /su/, /sel/, /sol/  
CyV:  /tya/, /tyu/, /tyo/       
      /syal/, /syu/, /syo/  

(The CyV moras do not occur in the native Japanese vocabulary, but were introduced into Japanese phonology via words borrowed from Chinese.)
This analysis has two problems. The first is asymmetry. Lass (1984: 25) describes 'system symmetry' (a term I take from Kelly and Local 1989: 100) as a 'quasi-“aesthetic”' constraint on phonological descriptions. It seems to me that the desire for system symmetry derives from the usual scientific goal of finding maximally general, exceptionless descriptions of the data, because asymmetries are potential irregularities or exceptions in need of further explanation. Symmetry derives from the foundation of phonological analyses on combinatorial systems of oppositions or regular correspondences between classes of phonological units, and not just relations between isolated pairs of units, and is therefore to be expected, whereas asymmetries stand in need of explanation. Secondly, the treatment of palatalisation has not been carried through to include /ti/ and the other /Ci/ moras. In /ti/, the initial is interpreted as the automatically palatalised variant of /t/ that occurs before /i/, whereas for /tya/, /tyu/ and /tyo/, palatality is attributed to the presence of a 'glide' segment /y/. The classical example of such an analysis is Bloch (1950). In an attempt to overcome both of these problems, Bloch (1946) entertains the following interesting hypothesis: 'A syllable contains a syllabic, alone or with one nonsyllabic preceding . . . Vocalic syllabics are a, o, u, e, and ya, yo, yu, ye; the second group palatalise a preceding non-syllabic . . . But ye is [i], and will be written i for simplicity.' In this way he quite ingeniously turned an asymmetrical (5 + 3) phonemic division into a symmetrical (4 + 4) and more regular morphophonemic division, at the same time offering an explanation for why the phonetic exponents of /ti/ (morphophonemic tye) begin with a palatalised affricate: [tʃi]. But Bloch did not sustain this imaginative advance into a mode of argument that is now common, and in Bloch (1950) he retracted the analysis. The analysis illustrates the need for some precise criteria in establishing phonemic analyses, but also the multiplicity of possible phonemic solutions, depending on supplementary assumptions about phonetics and morphology.

3.3.4 ‘Minimal pairs’, distinctiveness and types of oppositions
The test *par excellence* for the establishment of phonemes in a phonological analysis of a language is the minimal-pair relation, which can be found in the works of Patañjali [second century BC] (Allen 1953: 81), Sweet (see Henderson 1971: 230), Trubetzkoy (1969: 31–45) and Jakobson, Fant and Halle (1952). In the minimal-pair relation, two phonetic segments are deemed to represent or exemplify two different phonemes if substitution of one for the other alters an utterance’s meaning. Thus, /p/, /b/ and /m/ are
3.3 Phonemic phonology

regarded as different phonemes in English, because substitution of /p/ for /b/ in *bit*, /b/ for /m/ in *mitt* or /m/ for /p/ in *pit* changes the meaning of the word in question.

The minimal-pairs test defines a partition of the set of sound-types \( T \) into a number of disjoint subsets \( P_n \) (see Brainerd 1971). The disjointness of phonemes is demonstrated by the following argument: suppose sound-types \( a \) and \( b \) were in phoneme \( P_1 \), and \( b \) and \( c \) were in phoneme \( P_2 \), so that there is some \( b \) in both \( P_1 \) and \( P_2 \), which are therefore non-disjoint. Then by the definition of minimal pairs, replacing \( a \) by \( c \) in some utterance changes the meaning of that utterance. Since \( b \) is not significantly distinct from \( c \), as they are members of the same phoneme, replacing \( a \) by \( b \) should also change some utterance's meaning. But \( a \) and \( b \) are both in phoneme \( P_1 \) and are *not* semantically distinctive. To avoid such a paradox, phonemes must be disjoint sets. This reasoning presumably lies at the heart of structuralists' prohibition of overlapping phonemes.

The number and composition of \( P_n \) will differ from language to language. Each \( P_n \) is a Jonesian phoneme whose members (sound-types) are called phones. According to some phonologists, notably Trubetzkoy, although a phoneme corresponds to a class of allophones, each phoneme was associated in particular with the principle or prototypical unmarked member of the class, or with the set of distinctive features that every member of the class had in common. The phones in each phoneme may be in free variation with other phones (i.e. their occurrence is not predictable), or their occurrence may be distributionally determinate, in which case they are termed allophones). As observed in chapter 2, in general it is phones, not sound-tokens, that are denoted by alphabetic letters in square brackets. Phonemes are written using alphabetic letters in slant brackets. The use of alphabetic letters to denote both sound-types and sets of sound-types can be misleading. Phones and phonemes are often treated as if they are the same sort of object. The only sense in which they are similar, according to the conception of classical phonemic theory, is that they are both kinds of sets (sets of sound-tokens and sets of sound-types respectively).

The minimal-pairs test rests on the notion of phonological opposition. Trubetzkoy (1969) developed a rich taxonomy for different types of opposition:

1 *Binary* (bilateral) vs. *n-ary* (multilateral) oppositions (Trubetzkoy 1969: 67–74). In a binary (bipolar or bilateral) opposition, sets of phonemes may be partitioned into two. For example, the opposition *voiced* (unaspirated)
Segmental and transformational phonology

vs. voiceless (aspirated) in English, exemplified by pairs such as *pit/bit*, *tip/dip*, *choke/joke*, and *cot/got*, exhausts the range of distinctive phonatory types in this structural position. (Nasal stops and voiceless unaspirated stops are not directly comparable with these four paired examples, as they do not occur with the same distribution in English. Syllable-initial nasal stops do not show the same four-way distinction in place of articulation, and voiceless unaspirated stops only occur after */s/.) Other languages exemplify three- or four-way phonatory oppositions. For example, Classical Greek is believed to have had a three-way phonatory opposition of voiceless aspirated */pʰ/, */tʰ/, */kʰ/* vs. voiceless unaspirated */p/, */t/, */k/* vs. voiced (unaspirated) */b/, */d/, */g/*. The four-way phonatory opposition in Sanskrit—voiceless aspirated vs. voiceless unaspirated vs. voiced unaspirated vs. voiced aspirated—is classically analysed as a product of two binary oppositions, although the three-, four- and five-term place of articulation distinctions were regarded by Trubetzkoy (1969: 68) as multilateral (e.g. n-ary, or scalar) oppositions.

2 Privative, gradual and equipollent oppositions (Trubetzkoy 1969: 75). An opposition in which the difference between one phoneme and another lies in the absence vs. presence of a mark (e.g. a distinctive feature) was termed privative. Examples of binary privative oppositions might be */m/* vs. */b/* and */n/* vs. */d/* in English, in which one phoneme is distinguished from its partner by the absence vs. presence of the nasality mark. A consequence of privative oppositions is that the distribution of the phoneme which lacks the mark will be a subset of the distribution of its marked correlate, because every distributional statement regarding the unmarked phoneme will also apply to its marked congener, but not vice versa. For example, every statement about the distribution of bilabial closure will also apply to bilabial closure with nasality, but statements specifically about the distribution of bilabial closure with nasality do not necessarily refer to bilabial closure in the absence of the nasality mark. In the English ‘place of articulation’ classification labial vs. alveolar vs. alveo-palatal vs. velar, the terms do not have the same distribution: alveo-palatal place of articulation is not found syllable-initially after */s/* or before */l/, */r/* or */w/*; the first consonant of three-consonant syllable-initial clusters must be alveolar */s/*; and in two-consonant clusters following long vowels or diphthongs, one of the consonants must be alveolar. We might consequently regard place of articulation as a multilateral privative opposition, in which the several terms may be placed on a scale in which each has one more mark than the one next to it in the
3.3 Phonemic phonology

scale, so that the phonemes may be ranked from least marked to most marked. The least marked phonemes have the most general distribution. In the English place of articulation opposition, 'alveolar' is the least marked term since it may occur in all consonantal positions in the syllable, and 'alveo-palatal' the most marked term, since in onsets it does not occur in consonant clusters. Using ‘F’ and ‘G’ as privative features, this pattern could be analysed by categorising 'alveo-palatal' as doubly marked (F and G), velars (F) and labials (G) as singly marked and alveolars as unmarked.

Alternatively, the pattern of English place of articulation distinctions might be analysed not as the accumulation of several marks, but as various degrees or gradations of the same property (e.g. location of constriction), in which case the opposition is said to be gradual. Another English example of a gradual opposition might be vowel height, which is not usually regarded as a question of absence vs. presence of features, but of the scale of values open – mid – close.

Oppositions which are neither privative nor gradual are described by Trubetzkoy as equipollent. Such an opposition is one in which all phonemes in the opposition are 'equally possible'. For example, in English onsets, the terms of the binary opposition voiceless (aspirated) vs. voiced (unaspirated) are regarded as 'equally' possible because both co-occur with each place of articulation (see above), and when the opposition is neutralised (i.e. after /s/), neither one nor the other may occur, but a third phonatory category 'voiceless unaspirated' is manifested, which has a combination of properties of both poles of the opposition.

Trubetzkoy’s discussion of these different types of opposition does not seem to recognise that the categorisation of any opposition rests on the theory of phonemic composition which is adopted. For instance, believing that the voiced/voiceless distinction is (universally?) characterised by the presence vs. absence of a ‘voicing’ mark, Trubetzkoy regarded it as a privative opposition. In an analysis in which ‘voicing’ and ‘voicelessness’ were regarded as features of equal status, such as the [± voice] of Chomsky and Halle (1968), the voicing opposition would be equipollent.

3 Functional classes of oppositions (see also Jakobson and Halle 1956: 20–2). Oppositions may also be classified in terms of their linguistic function. For example, the oppositions discussed above are all examples of (lexically) distinctive oppositions (Trubetzkoy 1969: 90–241). Non-distinctive oppositions have equally important linguistic functions, as various as the demarcation of linguistic units (e.g. words or morphemes; Trubetzkoy
Segmental and transformational phonology

1969: 273–97), cueing grammatical distinctions (as with mutation in Welsh or tone in many African languages, which are predictable according to morphosyntactic criteria; see Trubetzkoy 1969: 182), marking dialectal, social, ethnic or sexual affinity (see Trubetzkoy 1969: 46–8), or managing interaction (e.g. features of turn-finality vs. non-finality such as pitch and speech rate). Though these and other functional classes of oppositions are by and large ignored in contemporary theories of phonological representation, which concentrate almost exclusively on lexical distinctiveness, they are as important.

The names for the terms in the above discussion of oppositions were all traditional descriptive phonetic categories, such as ‘voiced’, ‘voiceless’, ‘aspirated’, ‘unaspirated’, ‘labial’, ‘alveolar’, ‘alveo-palatal’ and ‘velar’. Out of Trubetzkoy’s classification of phonological oppositions arose the hypothesis that the various properties of an utterance are not just names of classes of phonemes, but might be regarded as more fundamental linguistic units than the phonemes that they classify, like the subatomic units from which atoms are made. For example, since the ‘phonatory’ opposition in English appears to be binary and equipollent (privative, according to Trubetzkoy), whereas the ‘place of articulation’ opposition is quaternary and privative or gradual, it might seem that the terms of these two oppositions are not merely classificatory labels, but distinct types of subphonemic unit. This view forms the basis of the hypothesis that phonemes are not indivisible atoms, but are themselves constructed from a small set of primitive phonological features. It is a short step from regarding phonetic properties such as ‘voiced’ as the dimensions of phonological oppositions, to regarding them as semi-independent objects from which phonological segments are constructed.

3.3.5 Distinctive and predictable features

As a historical development of the hypothesis that phonemes are not indivisible units which are merely classified by phonetic properties, but complex units composed of phonetic features, some phonologists’ view of the relationship between phonetic and phonological representations was significantly altered. In the view of Jakobson (Jakobson and Halle 1956: 19, Jakobson and Waugh 1979: 21, 29), phonetic and phonological units were not distinguished by the arbitrary type–token relation; instead, the phonological units of Jakobsonian phonemics were partially unspecified phonetic representations, composed of features from a supposedly universal set of possible phonological oppositions. Each opposition and each phonological
feature required representation by itself as functionally important parts of segmental units. This meant that alphabetic segmental representations were replaced by analphabetic, componential, segmental representations.

In chapter 2 I described a partially componential system (the use of diacritics in the IPA), and a more thoroughgoing componential system, the Bell/Sweet organic alphabets. Analphabetic notations were emulated in phonological analysis by structuralist phonological analyses such as Hockett (1947a) (see fig. 3.1), Martin (1951) and Bloch (1950). In the latter paper, Bloch categorises the allophones of Japanese in terms of six places of articulation ‘qualities’: L (labial), D (dental or alveolar), P (prepalatal), F (front, prevelar), B (back, mediovelar), G (glottal); seven degrees of aperture: 1 (closure), 2 (affrication) . . . 7 (wide aperture); and three privative ‘dichotomies’: N (nasality), V (voicing) and Q (quantity). For example, the long, palatalised, voiceless dental stop allophone [ti T] is categorised ‘(D1 Q F5)’. Bloch refers in footnotes to Harris (1944), Hockett (1947a, b) and Jakobson and Lotz (1949). Harris (1944) likewise refers to conversations with Bloch and Jakobson, and says that ‘a phonetic system of this kind
without the phonemic limitations is Otto Jespersen's analphabetic system'. Analphabetic notations were also employed in some Firthian prosodic analyses, such as Robins (1957), Sprigg (1957: 108) and Kelly and Local (1989: 206–12), but the mainstream of development of feature-based notations arose from the work of Jakobson and Lotz (1949), Jakobson, Fant and Halle (1952), Jakobson and Halle (1956), Chomsky, Halle and Lukoff (1956), Halle (1959) and Chomsky and Halle (1968).

In Jakobson, Fant and Halle (1952), an example of an ‘analytic transcription’ using their distinctive features is presented (figure 3.2). This representation is a two-dimensional matrix, the columns of which denote the temporal sequence of segments, and the rows of which record the values of each distinctive feature for each segment in turn. At the head of each column a conventional alphabetic symbol provides a more readable supplementary definition of each segment. The names of the features are given at the left-hand side, though this is supplementary information, since the fact that the vertical sequence of the features is the same from segment to segment means that the feature which each cell records is precisely defined by the position of the cell in the table without this supplementary naming. According to this version of a distinctive-feature theory, all oppositions are binary, so that the possible values of each feature can all be drawn from a set of two distinct symbols, such as \{ +, - \}. Some features bear both ± values, but this is merely a technique to record two features on one line. Not every segment bears a value for every feature in the analytic transcription. The values of redundant features, for example, being predictable, are omitted, and a blank space is left. Representations of segments are formed by concatenating values of features. Since the vertical sequence of feature-values is meaningful, this concatenation is non-commutative (see chapter 2).

Following Cherry, Halle and Jakobson (1953), Halle (1959) employs ana-
lytic transcriptions of a related, but slightly different kind, an example of which is reproduced in figure 3.3. (Halle (1959: 43) used the 'matrix' form as well.) Instead of being lined up in columns, the values of each feature are listed in a fixed sequence, with blank space to delimit one segment from the next. Because only the feature-values and not the names of the features are represented, in order for each value to be correctly attributed to the feature which bears it, features which lack a value are written with a special symbol $0$, not just left blank. The number $n$ of features, both distinctive and redundant, being constant, each segment is represented as a sequence of $n$ symbols from the set \{+, --, 0\}. This form of notation is retained in the formal definitions of SPE phonology in Chomsky and Halle (1968: 390).

In this version of distinctive-feature notation, concatenation is not used at all.

The notation which predominates in Chomsky and Halle (1968), and which has been popular since, records the name of each feature and its value for each segment in turn. Though superficially more verbose than the earlier method, this convention has several advantages, some of which derive from the fact that the same notation can be employed in rules as well as in representations:

1 Features without a value, or which are irrelevant in the definition of a rule, can simply be omitted. There is no need to use a space-filling place-holder symbol such as 0 to indicate unspecified features.
The order of features in a segmental specification is irrelevant. The features of each segment are joined together by commutative cocatenation.

No special horizontal alignment is needed to pair values with features. Consequently, two different features can be written on the same line of type, for instance [+nasal] → [+voice].

Chomsky and Halle (1968) retain Jakobson’s preference for equipollent, binary oppositions, but include a notable example of a scalar feature (the integer-valued feature [stress]), and envisage all features in surface phonetic representations as scalar. A binary (resp. ternary . . . ) equipollent phonological opposition is regarded as a binary (resp. ternary . . . ) partition of the set of values of a scalar phonetic feature. This is illustrated in the following diagram:

<table>
<thead>
<tr>
<th>Phonological feature-values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>+F</td>
</tr>
<tr>
<td>-F</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Phonetic feature-values:</th>
</tr>
</thead>
<tbody>
<tr>
<td>0F</td>
</tr>
<tr>
<td>1F</td>
</tr>
<tr>
<td>2F</td>
</tr>
<tr>
<td>. . .</td>
</tr>
<tr>
<td>(n−1)F</td>
</tr>
<tr>
<td>nF</td>
</tr>
</tbody>
</table>

Thus, whether a feature is binary or scalar is not a property of the feature itself. One and the same feature may have binary values in phonological representations, and scalar values in phonetic representations. Chomsky and Halle (1968: 65, 169) refer to the phonological (discrete, categorial) and phonetic (continuous, articulatory, parametric) functions of a single set of distinctive features, and on page 400 to the ‘intrinsic content of features’. They consider the possibility of two separate sets of features (phonetic and phonological) with a consequent strict separation of phonetic and phonological representation, but reject it by the following argument:

It might be proposed, in the light of the distinction between phonological and phonetic distinctive features, that the two sets be absolutely unrelated – that . . . the rows be labelled entirely differently in the phonological and phonetic matrices . . . Only the phonetic features would now be ‘substantive’; the phonological features would be without physical content and would provide an arbitrary categorization.

. . . since all phonological rules would operate on the ‘empty’ categories A, B, etc., gradually filling them in and revising their entries, the grammar would now have to be supplemented with a set of rules operating at the point at which all matrices are fully specified and providing that phonetic features be associated with the categories; for example we would have rules providing that [α A] → [α vocalic], [α B] → [α consonantal] . . . But every grammar will have to have exactly these rules; hence they do not contribute in any way to the choice among grammars and can just as well be eliminated from all grammars. (Chomsky and Halle 1968: 169)
This argument is quite unsound, however, for two reasons. Firstly, it neglects a different and perhaps rather more plausible possibility, which is that discrete categorial features map directly onto ranges of parametric phonetic features without an intermediate level of scalar phonological features to mirror the parametric phonetic features. This intermediate level, on which the argument quoted above depends, is a straw man. And secondly, the specific claim that 'every grammar will have to have exactly these rules' implies that all phonetic differences between languages are phonological. More recent research, such as Lindau and Ladefoged (1986), Mohanan (1986: 154–81) and Pierrehumbert and Beckman (1988), suggests that the phonetic interpretation of phonological features may differ from language to language — or rather, that it is theoretically sensible to partition phonetic and phonological rules in such a way as to allow for language-specific differences between phonetic rules.

The four distinct types of opposition defined by the categorisation binary vs. scalar and privative vs. equipollent can be represented using Chomsky and Halle's notation: binary oppositions are indicated using the values + and −, for instance [+voice] and [−voice]; scalar oppositions are indicated using integer values; equipollent oppositions are those where all of a feature's possible values appear in some rules or lexical representations; and privative oppositions are those where some of a feature's possible values never appear in lexical representations, but must be supplied by rule. For example, although [+nasal] appears in lexical representations, [−nasal] is always predictable in the absence of [+nasal], and, being redundant, can be omitted from lexical representations.

It is possible to make a further kind of distinction between oppositions on the basis of the number of possible values which they may have: fixed vs. denumerable. Consider, for example, the SPE [stress] feature. The values of this feature are integers, and it is thus non-binary and equipollent. The number of possible values this feature may have depends on the number of times stress rules may apply, which depends in turn on the complexity of the syntactic structure. In particular, the largest stress value in a representation is a function of the depth of the syntactic tree, or in other words, the length of its derivation. For any particular representation, this number is finite, and can be computed by a recursive definition. Thus the number of possible values of the stress feature for any representation is denumerable.

It is not possible, however, to state once and for all what that number is without computing such a denumeration. And because syntactic structure is recursive, any such upper limit on the possible values of the [stress]
feature will be superseded in a more complex structure (Chomsky and Halle 1968: 23). In every case, however, the largest value of [stress] is denumerable. Thus, the [stress] feature in Chomsky and Halle (1968) can be said to be denumerable, although the cardinality of the set of values it may take cannot be fixed at a finite number.

Now consider an analysis in which the integers 1, 2 and 3 are used to denote a three-way distinction. For example, [1high], [2high] and [3high] might denote the height distinctions of open, mid and close vowels respectively. By the arguments given above, [high] is clearly a denumerable feature, but unlike [stress], it is also fixed in advance: any three distinct symbols would do equally well. The fact that integers are used is inconsequential, unlike the case of [stress], where the established relations between integers, such as 'precedes' or 'is greater than', and functions such as addition, are all made use of in the analysis and representation of stress.

This distinction between fixed and denumerable values of a feature is necessary in addressing an occasional criticism of integer-valued features, which is that employing a fixed set of integers as the values of, say, vowel height, fails to account for the infinite number of numerals which are not employed to represent vowel height in such analyses. Such a view mistakes the nature of the objects which are employed in such an analysis, though: they are not an arbitrary subset of the complete set of integers, but simply a finite number of distinct symbols, a miniature number-system. Non-numeric symbols might be used for fixed-valued features instead, with no formal difference to the analysis.

The relationship between phonetic and phonological segments and features in distinctive-feature theory is illustrated by the following diagram:

<table>
<thead>
<tr>
<th>Phonological segments</th>
<th>are included in</th>
<th>Phone</th>
<th>Minimal syntagmatic units</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>consist of</td>
<td>Phones</td>
<td></td>
</tr>
<tr>
<td>Distinctive and</td>
<td>are included in</td>
<td></td>
<td></td>
</tr>
<tr>
<td>redundant features</td>
<td></td>
<td>Phonetic features</td>
<td></td>
</tr>
<tr>
<td>Syntax</td>
<td></td>
<td></td>
<td>Paradigmatic 'atoms'</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Semantics)</td>
</tr>
</tbody>
</table>

The column labelled '(Semantics)' contains those objects which in chapter 2 were defined as categories in the interpretation of phonological units, not formal phonological objects. In IPA segmental theory, for example, segments are not made up of the categories voiced, bilabial, plosive etc. Those categories are the names of phonetic properties which realise phonological units. They are aspects of a unit's meaning, not its form.
What Chomsky, Halle and Jakobson achieved in developing feature-based representations of this kind was to embody the representation of ‘semantic’ categories (names of phonetic distinctions) in the ‘syntax’ of feature-based representations, so that the phonetic interpretation of their phonological representations is transparent. In their distinctive-feature theory, the traditional interpretive categories such as voiced, bilabial, plosive etc. are made component objects of the representation. They are no longer just ‘semantic’ categories, but phonetic and phonological units alike are ‘syntactic’. That is why I put ‘(Semantics)’ in parenthesis. It could be said that the interpretation of each segment and feature has been made explicit in its form to such an extent that form and content are no longer easily distinguished. According to generative phonology, every phonological unit and representation has an intrinsic phonetic interpretation (an inherent pronunciation; see Chomsky and Halle 1968: 169–70, 400). This claim is only tenable (let alone credible) because the distinctive features which make the alphabet of phonemes non-arbitrary are given the names of traditional phonetic categories, such as ‘voiced’, and where new categories are introduced, such as [delayed release], they are defined in conventional descriptive phonetic terms. If all the feature names were replaced by more abstract or even completely arbitrary labels, such as *vagrant, purple* or *Bangladesh*, the properties of the grammar formalism, and even the analyses encoded by particular grammars, would be unchanged. Yet generative phonologists have always regarded the use of non-arbitrary names as central to their enterprise, despite the fact that even *voiced, bilabial* and so on are, outside the English language, arbitrary and uninterpretable labels, devoid of any inherent phonetic reference.

Phonemic phonology, then, is a theory with two levels of representation, phonemic and allophonic. Whether ‘allophonic’ should be interpreted as ‘phonetic’ or not is open to question, but it is clear that the range of phonetic observations that a phonemic analysis seeks to account for are those observations represented in the allophonic representations. So we will regard allophonic representations as phonetic representations, especially as in transformational-generative phonology allophonic representations are called ‘systematic phonetic’ representations.

Chomsky and Halle’s distinctive-feature notation was part of an elaborate theory of the relationship between phonological and phonetic representations that provided an explicit framework for the formal examination of underlying phonological (lexical) invariance and superficial phonological (systematic phonetic) variability. In the next section, I shall examine the mechanisms of that theory in detail.
3.4 Transformational-generative phonology

Transformational-generative phonology is a theory which proposes a method for deriving strings of surface phonetic segments (allophones) from lexical strings of phonemes using mappings from minimally redundant lexical representations to fully specified allophonic representations, via intermediate representations of varying degrees of specificity and redundancy. These mappings are usually defined using an ordered sequence of rules of the form \( A \rightarrow B / C \_\_ D \), usually called ‘context-sensitive’, but which I shall argue below define unrestricted rewriting systems, since \( B \) may be the empty string. Rules of the form

\[
A \rightarrow B / C \_\_ D
\]

can be equivalently represented

\[
\begin{array}{c}
A \\
E
\end{array} \rightarrow \begin{array}{c}
B \\
E
\end{array} / C \_\_ D
\]

The appeal of transformational grammars to phonologists and speech researchers is perhaps due to the fact that the basic operations of a transformational grammar – deletion, insertion, permutation and copying – are apparently empirically instantiated by such well-established phonological phenomena as elision, epenthesis, metathesis, assimilation and coarticulation. The following examples illustrate these phenomena and their description using rewriting rules:

(3.3) **Copying (i): within-syllable assimilation**

\[
\begin{align*}
N & \rightarrow m / _\_ \{p, b, w\} & \text{rhombus [\text{romb}\,\!\text{əs}]}
N & \rightarrow m / _\_ \{f, v\} & \text{triumph [\text{traɪ\,\!\text{əm}f}]}
N & \rightarrow n / _\_ \{\theta, \delta\} & \text{tenth [\text{te\,\!\text{nθ}]}
N & \rightarrow n / _\_ \{t, d, s, z\} & \text{sand [\text{sænd]}
N & \rightarrow n / _\_ \{f, t, \theta, \phi\} & \text{hinge [\text{hɪ\,\!\text{ŋ}dʒ}]}
N & \rightarrow n / _\_ \{k, g\} & \text{hang [\text{hæŋ(g)]}
\end{align*}
\]

(3.4) **Copying (ii): coarticulation**

\[
\begin{align*}
k & \rightarrow k / _\_ \{i, e, æ\} & \text{keep [k\,\!\text{i:p}]}
k & \rightarrow k^w / _\_ \{o, u\} & \text{cool [k\,\!\text{u:l}]}
k & \rightarrow k / _\_ \{α, ά\} & \text{cart [k\,\!\text{ɔ:t}]}
\end{align*}
\]

(3.5) **Insertion: epenthesis**

\[
\begin{align*}
ns & \rightarrow nts & \text{mince [\text{mɪnts], pence [pents]}
l s & \rightarrow l ts & \text{pulse [\text{pʌlts]}
\end{align*}
\]

(3.6) **Deletion: elision**

\[
\begin{align*}
nd & \rightarrow n & \text{sandwich [\text{san\,\!\text{wɪtʃ}]}
\end{align*}
\]
3.4 Transformational-generative phonology

(3.7) Substitution (i): between-word assimilation
n → ɡ / _ {k, ɡ} ran quickly [raŋkwikli]

(3.8) Substitution (ii): within-word assimilation
n → m / _ {p, b, w, f, v} sandwich [samwɪtʃ]

(3.9) Permutation: metathesis (cf. Lass 1984: 188)
ær → ra burnt [brænt]
sk → ks ask [æks]
sp → ps wasp [waps]

A tacit assumption of this approach is that a systematic phonetic representation (a string of allophones) contains all the linguistically predictable information of the representation of an utterance, so that it can be mapped directly into sequences of values of a set of universal phonetic parameters (see Chomsky and Halle 1968: 65, 169, Ladefoged 1977: 229–32, Halle 1983, Mohanan 1986: 160–5 for partial accounts of phonetic interpretation in transformational-generative phonology). The complexity of the surface phonetic patterns is supposed to be entirely accounted for by the phonological rules. The integer values of the phonetic features in the systematic phonetic representation are regarded as the values of control-parameters representing the 'target' postures of the articulators (see Halle 1983).

3.4.1 Coarticulation and the phonology-phonetics interface

The division of labour between phonological rules and physical phonetic processes is addressed most sharply in the analysis and representation of coarticulation. In chapter 21 argued that coarticulation had both phonetic ('hard') and phonological ('soft') aspects. 'Hard' coarticulation was briefly discussed in that chapter, which dealt with phonetic representations. In this section, therefore, I shall consider a transformational treatment of only 'soft' coarticulation.

As a first and weakest hypothesis, we might regard the coarticulation of consonants to adjacent vowels as arising from the copying of the vowel features to the consonant, for example,

\[
C \rightarrow \begin{bmatrix}
\alpha \text{ back} \\
\beta \text{ high} \\
\gamma \text{ low} \\
\delta \text{ round}
\end{bmatrix} \% _{X} \begin{bmatrix}
\alpha \text{ back} \\
\beta \text{ high} \\
\gamma \text{ low} \\
\delta \text{ round}
\end{bmatrix} V
\]

where % _{X} X means 'in environment _{X} X and its mirror-image X _{X}'. (The 'mirror-image' notation for rule contexts was not employed in
Chomsky and Halle (1968), but was proposed somewhat later, in Langacker (1969). An example of such a coarticulation rule occurs in Saib (1978: 100).

This preliminary formulation represents the case when a consonant 'takes on' the features [back], [high], [low] and [round] of a neighbouring vowel, represented by the duplication of the variables α, β, γ in the two segments mentioned in the rule. These assimilations in backness, highness, lowness and roundedness are individually attested as possible coarticulatory processes in some languages, as the following paragraphs illustrate.

Back
In many varieties of English, where the phonetic distinction between 'advanced' [−back] and 'retracted' velar [+back] articulations is not lexically distinctive, this distinction is conditioned by the frontness/backness of the neighbouring vowel (the vowel following a syllable onset or preceding a syllable coda). Advanced velar articulations are found in the environment of front vowels, and retracted velar articulations are found in the environment of back vowels. That this phenomenon is 'soft', arbitrary, phonologically determined coarticulation and not just an automatic mechanical side-effect of articulation is demonstrated by the fact that in some languages and dialects precisely the opposite correlation of the value of [back] in neighbouring vowels and consonants is found, so that they are dissimilated. According to J. Local (personal communication), forms such as [bʊɡ] bag and [bæk] or [baeq] back are found in some North-Eastern English varieties. In my notebooks I have face-to-face impressionistic phonetic transcriptions which record retracted velar closure before front vowels in Japanese e.g. [ge], and retracted velar – nearly uvular – closure after front vowels in Sinhalese e.g. [ekʷ]. According to Stevens, Keyser and Kawasaki (1986), Steriade (1987) and Keating (1988a, b), if a feature is non-distinctive in some segment, it often receives its value by phonetic coarticulation with a nearby segment in which the feature is specified. According to this view, such dissimilation of a feature is only to be expected where the feature is phonologically distinctive. Though this is often the case, it is not universal. Arbitrary dissimilation of this kind is occasionally found.

High
In analyses of Russian such as Halle (1959), palatality is treated as a distinctive property of some consonants. When it is distinctive, it is specified.
When it is not, it is copied from the vowel. The references mentioned in the previous paragraph attribute this contextual filling-in of redundant features to an inherently assimilatory view of underspecification.

Low

In Arabic, a language with distinctive pharyngeal place of articulation in consonants (and also pharyngeal secondary articulation in velarised consonants), and a very simple three-vowel system (/i/, /u/, /a/), lower vowel qualities are observable for tokens of these three categories in pharyngeal or pharyngealised consonant environments than in non-pharyngeal(ised) environments. This example of vowels assimilating to consonants can be complemented by Klatt's (see Allen, Hunnicutt and Klatt 1987: 110-12) observations of soft coarticulation in English, in which one of three coarticulatory variants of alveolar and velar consonants found by statistical techniques was found in the environment of [−back, −round] vowels. This class of vowels in English happens also to be [+low], so the description [−back, +low] would be equally good, which would implicate [low] as a potential coarticulatory feature. Saib (1978: 101) discusses this kind of coarticulation in Tamazight Berber.

Round

In English there is no productive distinction between rounding and non-rounding in consonants independently from vowels. In other words, in analyses which distinguish rounded from unrounded vowels, there is no need to establish an opposition of rounded vs. unrounded consonants. The rounding or non-rounding of vowels is predictable from the phonological backness, height and structure of the nucleus. Since [round] is not a distinctive feature of vowels, there is no particular reason to restrict it to being a binary-valued feature, and a complete scale of degrees of rounding may be observed, from /i/ (most spread) through /e/, /a/, /æ/, /ɔ/, to /u/ (most rounded), and similarly for the long vowels. Except in those cases where rounding is a redundant feature, as notably in English /w/, /r/, /ʃ/, /tʃ/, and /dʒ/ (Brown 1981), consonants manifest the rounding of the vowel or part of the nucleus with which they are coproduced. In these cases, then, [round] too is coarticulatory.

Trubetzkoy (1969: 285, 289 n. 11) observes that in some Turkic languages with vowel harmony, the intervening consonants vary with the vowel-harmony stretches in a coarticulatory fashion (see also Comrie 1981: 63). In terms of distinctive features, the harmonising features of vowels, [back],
[high], [low] and [rnd], occur as the secondary articulation of intervening consonants.

3.4.2 Transformations in Japanese: variants of /i/ and /l/ul

As an example of the way in which transformational rules work, I shall illustrate the way in which a transformational phonology would analyse the examples of phonetic variability in Japanese discussed in section 3.3 above.

In section 3.3.2, two interacting phonological phenomena in Japanese were described: whispered high vowels and fricative nuclei. Nuclear friction is a descriptive term for the phenomenon which is usually presented as the accent-dependent deletion of close vowels following fricatives and affricates (Ueda 1976, example 4). Note, however, that in fricative nuclei the so-called 'underlying' vocalic quality is actually present in the secondary articulation of the fricative nucleus (Schane 1971: 510). If this observation is of phonological relevance (and I shall argue that it is), in a segmental generative account it might be proposed that the vocalic quality of the vowel is copied to the consonant as palatal secondary articulation (velar in the case of 'back-nucleus' moras), before the vowel is deleted.

If the analysis of nuclear friction includes the proposal that a segment is deleted, however, a concatenative theory of phonetic interpretation would predict that the material which remains after such deletion ought to be of shorter duration than a CV mora. This prediction is refuted by phonetic observations: notice in table 3.2 the greater length of the friction in fricative-nucleus moras relative to their whispered and close vocalic counterparts. In short, if a 'vowel-deletion' rule is proposed, there must also be a rule that reassigns syllabicity to the friction units, or a compensatory lengthening rule (De Chene and Anderson 1979, Ingria 1980, Prince 1984, Fukui 1986, Poser 1986, Wetzels and Sezer 1986) to account for its increased duration. The impressionistic observation of greater length in these cases is well supported by instrumental observations; see Port, Dalby and O'Dell (1987), Nishinuma (1983).

A simple transformational description of these two phenomena ('whispered high vowels' and 'fricative nuclei') employs two rewrite-rules to describe them:

1 High Vowel Devoicing (see Ueda 1976: 315 example 1, Hasegawa 1979b: 388, Haraguchi 1984: 147): high vowels ([i], [u]) are voiceless (whispered) when they occur between two voiceless consonants (3.11).
3.4 Transformational-generative phonology

(3.11) \[ V \rightarrow [-\text{voice}] / [-\text{voice}] \]
\[ [+\text{high}] / [-\text{voice}] [-\text{voice}] \]

For example [kįkai], [ʦįkai], [ʃίta], [ɕiṭo], [ʦųkue], [sūtte].

2. Voiceless Vowel Elision (see Ohso 1973: 13, Ueda 1976: 315 example 2, Haraguchi 1984: 146): unstressed voiceless vowels are deleted after voiceless fricatives (3.12). (The increased length of the remaining consonant is disregarded here.)

(3.12) \[
\begin{array}{c}
V
\rightarrow \emptyset
\end{array} / \]
\[ [-\text{voice}] [-\text{sonorant}]
\[ +\text{continuant}] \]

For example, /ʃita/ \(\Rightarrow\) [ʃita], /ɕiṭo/ \(\Rightarrow\) [ɕiṭo], /ʦųkue/ \(\Rightarrow\) [ʦkue].

Where a deletion rule operates, information can be lost from the representation. The only way of counteracting this loss is to ensure that the information which is removed is redundant, that is, either predictable from the presence of some other feature, segment or phonotactic configuration, or available elsewhere in the representation. This is the so-called condition on recoverability of deletion (Chomsky 1965: 144). Recoverability can be difficult to guarantee, however, since it requires that whenever information is deleted, the means for its restoration must never subsequently be lost. For example, the feature, segment or phonotactic configuration which makes the deletion recoverable must itself be made in some way undeletable.

Problems are almost inevitably incurred when a grammar contains rules that both extend and reduce strings, and where the interaction of the rules is not transparent. This is discussed at length in later sections. (For lucid, less technical discussions of this issue see Hofstadter 1979: 73–4 and Penrose 1990: 130–2).

I shall therefore assume that (3.12) is fed exclusively by (3.11): that is, there is a strict feeding order of application in which (3.11) precedes (3.12) and (3.12) can only apply if (3.11) has applied in order to derive the \(V[-\text{voice}]\) context of the left-hand side of (3.12). By collapsing these two rules, it can be seen that the left-hand side of (3.12), V, must also be predictably [+ high], and the right-hand environment of (3.12) is redundantly the same as (3.11): that is,

(3.13) \[ C[-\text{voice}] \]
The deletion in (3.12) is thus theoretically recoverable, since its environment can be inferred to be

\[(3.14) \begin{array}{c}
C \\
\text{[-voice]} \\
\text{[-sonorant]} \\
\text{[+continuant]} \\
\end{array} \rightarrow \begin{array}{c}
C \\
\text{[-voice]} \\
\end{array}\]

and its result

\[(3.15) \begin{array}{c}
C \\
\text{[-voice]} \\
\text{[-sonorant]} \\
\text{[+continuant]} \\
\end{array} \rightarrow \begin{array}{c}
C \\
\text{[-voice]} \\
\end{array}\]

cannot arise otherwise. This information could be used to define an inverse rule (3.16) that recovers deleted vowels under the relevant circumstances:

\[(3.16) \begin{array}{c}
C \\
\text{[-voice]} \\
\text{[-sonorant]} \\
\text{[+continuant]} \\
\end{array} \rightarrow C \begin{array}{c}
\text{[-voice]} \\
\text{[-sonorant]} \\
\text{[+continuant]} \\
\text{[+stress]} \\
\text{[+high]} \\
\end{array}\]

This analysis is complicated by another phenomenon in Japanese, however – Affricate Sequence Assimilation:

\[(3.17) t J \rightarrow s / s s / \]

for example /itʃi/+/sai/ \(\Rightarrow (3.11) /itʃi\ + sai/ \Rightarrow (3.12) /itʃai/ \Rightarrow (3.17) lissai/ ‘one year old’ in which /tʃ/ + /s/ \(\Rightarrow /ss/.

The presence of rule (3.17) in a transformational analysis of Japanese results in some unfortunate indeterminacies. Faced with the string /ss/, how does a hearer know whether the geminate fricative results from affricate sequence assimilation (the correct deduction, in this case), or from voiceless vowel elision – that is, from a putative underlying form /isu/+/sai/? Clearly, in this circumstance external information, whether morphological, syntactic or pragmatic, has to be employed in order to make the correct analysis. Although the effects of deletion by rule may be recoverable, the presence of other rules can introduce parsing indeterminacies. An important question which arises from this discussion, then, is: is guaranteeing recoverability of deletion a real problem for transformational phonological analysis, or is it just our abilities as analysts that are taxed? Put another way, can the recoverability constraint always be met, if we are sufficiently ingenious, or are deletion rules problematic in principle, despite the linguist’s analytical skills?
3.4 Transformational-generative phonology

3.4.3 Variants of /t/ and /ty/ in Japanese, and the representation of affrication

I shall now consider how a classical generative analysis (following Chomsky and Halle 1968) might model the phonological variability of /t/ and /ty/ in Japanese described in subsection 3.3.3 above. We would expect to be able to arrive at redundancy-free lexical entries and avoid 'phonemic overlapping' by employing a suitable feature-system, incrementally deriving the intended surface forms through the successive application of rules of the required generality and parsimony. For example, rule (3.18) applies irrespective of voicing and backness, compressing in a single statement the specification of affrication of coronal consonants (/t/, /d/), before high vowels (/i/, /u/).

\[(3.18) \quad C \quad \rightarrow [+\text{delayed release}] / [-\text{sonorant}] \quad V \quad [+\text{high}]\]

The feature [delayed release] has attracted criticism (see Ewen 1980, 1982; van der Hulst and Smith 1982b: 5 for a summary and further references) as it is the only feature in the SPE system with an implicitly dynamic interpretation. Its utility is dependent on considerations such as the following. Suppose that instead of rule (3.18) we employed a rule such as (3.19), which explicitly inserts a coronal continuant between the initial stop and the vowel.

\[(3.19) \quad \emptyset \rightarrow s / t_0 \quad \{i\} \quad \{u\}\]

A fuller, more general expansion of this, without the disjunction, is:

\[(3.20) \quad \emptyset \rightarrow C \quad / \quad C \quad \rightarrow \begin{cases} \begin{array}{c} \text{sonorant} \\ \text{continuant} \\ \alpha \text{ voice} \end{array} \end{cases} \quad \begin{cases} \begin{array}{c} \text{sonorant} \\ \text{continuant} \\ \alpha \text{ voice} \end{array} \end{cases} \quad V \quad [+\text{high}]\]

The occurrence of [α voice] in the adjacent consonant specifications in this rule ensures that the fricative part of the coronal affricate inserted after the stop part of the affricate will be voiceless if the stop is voiceless and voiced if the stop is voiced.

Rule (3.18) is usually judged to be preferable to (3.20) since it is more parsimonious. Furthermore, using [delayed release], the functional, distributional and structural unity of affricates is captured and a simple
concatenative two-segment CV mora structure is maintained. However, (3.20) also has certain advantages:

1. It states the unity of place of articulation of affricates explicitly, rather than implicitly as with [delayed release]. One advantage of this explicitness for a phonetically based phonological theory is that phenomena such as nasal or lateral release, such as [tn] and [t1], can be represented using feature-matrix structures that directly parallel that which specifies [ts].

2. A tempting reason for representing the affricate [ts] as a sequence of two feature-matrices, that is, spelling out the [s], is that in Japanese the phoneme /s/ is also palatalised before /i/, just like the affricate [ts]. In other words, rule (3.19) feeds rule (3.21), which is independently motivated, so that greater parsimony and integration than the rigidly phonemic analysis is achieved.

(3.21)  s → /—i

The situation is not so simple, however, since in the case of the partially palatalised affricate [t-ʃ], it must also be specified that the initial [t], in its turn, is palatalised:

(3.22)  t → /—ʃ

This rule would not be required in an analysis which employed [delayed release]. Using [delayed release] to express affrication, the ‘spreading’ or ‘regression’ of palatalisation to the stop portion is guaranteed, and we may express (3.21) more generally as (3.23):

(3.23)  C → [+high] /—[+high]  
        [—back]  [—back]

Since (3.23) is self-feeding, it can apply over any string of consonants that precedes V [+high, —back]. This may actually be desirable, since consonants immediately preceding /Ci/ moras (i.e. obstruent or nasal mora) are palatalised e.g. in [mattʃi] ‘matches’ or [geʃki] ‘health’. Since palatalisation rule (3.23) is self-feeding, representing affricates as sequences of segments is an attractive possibility.

There is no obvious way to choose between these two possible analyses of affrication in the classical transformational model. The advantages of the feature [delayed release] are its abbreviatory value and the fact that it reflects the functional unity of affricates; but these benefits are more-or-less negated by the cost of lost generalisations.
3.4.4 *Japanese palatalisation as 'soft' coarticulation*

In chapter 2 I argued that the phonetic aspects of consonant-vowel coarticulation ('hard' coarticulation) may be modelled by the temporal arrangement (coproduction) of the phonetic exponents of phonological units, and in section 3.4.1 I argued that some aspects of coarticulation ('soft' coarticulation) may be defined using a copying rule. In this section I shall examine whether Japanese palatalisation is purely phonetic, or whether both phonetic and phonological aspects need to be recognised.

Phonologists and phoneticians have frequently argued that the regressive spreading of palatality before high front vowels is not phonologically relevant (e.g. Campbell 1974, n. 11, Ladefoged 1975: 49), and have treated the phenomenon as an instance of universal coarticulation principles (Gay 1978; but see Hattori 1967: 542 for an appealing counter-argument). In other words, palatalisation or fronting of consonants is held to be completely predictable before /i/. This does not hold before non-high vowels, however, and so where palatalisation occurs in such an environment, it must be specified explicitly. Palatalisation is not, then, simply an automatic coarticulatory effect. The proposal of a supposed /y/-glide in some phonemic analyses can be defended in these terms. In this section I shall examine whether a universal account of coarticulation can be called upon to explain palatalisation in the environment of /i/ and /y/.

I shall argue, however, that in a transformational phonological analysis, certain clearly phonological, non-automatic processes must be ordered after the particular case of regressive palatalisation under consideration, and hence it cannot be interpreted as a necessary ('hard') coarticulatory process. I thus concur with Fujimura (1990: 218), who characterises Japanese palatalisation as a special, language-specific kind of soft coarticulation.

The self-feeding palatalisation rule, (3.23), is simpler than (3.24), which is more restricted in its application, and is not self-feeding.

\[(3.24) \quad C \rightarrow \left[ +\text{high} \right] / -\text{back} \quad V \left[ +\text{high} \right] / -\text{back} \]

Rule (3.23) is also observationally more satisfactory, since it will correctly spread palatality to preceding consonants. Rule (3.24) would necessitate the operation of a further, later rule in order that immediately preceding consonants also become palatalised. But if a later rule is to apply, then (3.24) cannot possibly represent a 'mechanical' coarticulation process, since such a process could only plausibly operate at the very end of a derivation.
Even if (3.23) is employed, it can still be shown that a later rule may operate. The description of the phenomenon given here is from an SPE-type fragment of Japanese phonology that appears in Ueda (1976: 315):

[i] and [u] normally disappear between a preceding voiceless continuant consonant and a following voiceless consonant. This may be formulated as:

\[
\begin{array}{c}
+\text{voc} \\
-\text{cons} \\
+\text{high}
\end{array}
\rightarrow \emptyset \\
\begin{array}{c}
-\text{voc} \\
+\text{cons} \\
-\text{voi} \\
-\text{nasal} \\
+\text{cont}
\end{array}
\]

Words like Misisippi 'Mississippi', tukusi 'horsetail' are thus pronounced [mi/Jppi], [tskufi].

Further examples given by Ueda include kisusi [kisʃ] 'kissing', which demonstrates that the respective ordering of palatalisation (3.23) and high-vowel deletion (3.25) must be:

\[ /\text{kisusi}/ \Rightarrow (3.23) \Rightarrow /\text{kisuʃi}/ \Rightarrow (3.25) [\text{kisʃ}] \]

and not:

\[ /\text{kisusi}/ \Rightarrow (3.25) \Rightarrow /\text{kiss}/ \]

for in the latter case there is no following high front vowel to palatalise the final [s]. Again, if a rule can be established which must follow palatalisation, as Ueda's analysis suggests, then palatalisation in Japanese must be a bona fide phonological process, not merely an automatic coarticulatory phonetic artefact.

The foregoing discussion of Japanese exemplifies two problems of transformational phonology: the invertibility problem and the rule-ordering problem. The invertibility problem, exemplified by the discussion of voiceless vowel elision in Japanese, can be expressed:

In a transformational grammar which employs deletion rules, is the underlying form of a string recoverable from the surface by inverse transformations?

The rule-ordering problem, exemplified by the preceding discussion of relative ordering of coarticulatory palatalisation and high-vowel deletion, in which it was shown that the language-specific (thus parochial phonological) deletion rule must follow the putatively universal coarticulatory palatalisation rule. The rule-ordering problem can be stated:
Is it possible to formulate a unique order of rule application in which all language-specific phonological rules precede all language-universal phonetic rules?

The rule-ordering problem includes, as a subcase, the ordering consistency problem, which can be stated as:

Is it possible to formulate an ordered set of rules which is consistent and free of rule-ordering paradoxes?

Examples and discussion of rule-ordering paradoxes may be found in Somerstein (1977: 175–6). Though some revisions of the SPE rule-ordering model can eliminate some of these paradoxes, Somerstein (1977: 187–8) notes some paradoxes which remain even if alternatives such as local ordering or partial ordering are adopted.

Although it is clear that these formal problems would have to be resolved for the transformational model to continue to be taken seriously, they have by and large not been well attended to, perhaps because they are rarely problematic. Furthermore, where such problems do arise, it has usually been found possible to resolve them by amendments to the original SPE framework. It might be asked, then, whether the existence of occasional persistent examples of these problems is due simply to a lack of ingenuity or attention on the part of phonologists, or whether they betray intrinsic problems in the SPE transformational model.

In the next section I shall show that the mechanism proposed in order to map from systematic phonemic to systematic phonetic representations is indeed too powerful, and that no amount of ingenuity can solve the invertibility problem caused by deletion rules. Grammars made with context-sensitive rules with deletion may sometimes happen to be contingently quite restricted. For instance, if the rules apply in a fixed order without cyclicity, they may be compiled into a finite-state transducer (Johnson 1972, Kaplan and Kay 1994). But I shall show that in general such conditions do not obtain, and that there is thus no guarantee that a recognition algorithm that implements the inversion of such a grammar will halt, which I shall show to be problematic for psychological realism in theories of performance and language acquisition.

### 3.5 The excessive power of transformational grammars

Before proceeding with my discussion of formal problems with SPE-type rule systems, it is necessary to define them precisely, and to consider their
relationship to various other kinds of grammar. In this section, therefore, I shall briefly review some of the mathematical and computational results and arguments concerning transformational grammars. This discussion will also help to motivate the declarative phonological theory developed in chapters 5 and 6.

3.5.1 Review of elementary concepts of formal language theory
According to the formal theory of generative grammars (e.g. Chomsky 1957: 13, Chomsky and Miller 1963: 283, Hockett 1966: 186, Brookshear 1989: 37), a 'language' is any set of strings. A string $x$ is said to be well formed with respect to a grammar $G$ if $G$ generates $x$. The string $x$ is said to be generated by a grammar $G$ if some ordered application of the rules of $G$ to a start symbol $S$ yields $x$. Likewise, the string $x$ is said to be accepted by $G$ if some ordered application of the inverses of the rules of $G$ to the string $x$ yields the start symbol. An alternative term for language in the above sense is 'string set', a usage which I shall adopt, as it is distinct from the ordinary, non-technical meaning of 'language'. If two grammars generate the same string set, they are said to be weakly equivalent and to have the same weak generative capacity. In general, this definition is without practical significance, since there is no general way to decide whether two grammars are weakly equivalent or not. But in particular cases, two grammars may be sufficiently closely related and well understood that weak equivalence can be decided (see Benson 1979: 107). Furthermore, if two grammars define the same or isomorphic sets of structural descriptions (i.e. derivation graphs) then they are said to be strongly equivalent; that is, they have the same strong generative capacity. Derivation graphs are discussed in section 3.6.

3.5.2 Rule Types
Rules of the form $A \rightarrow B / C \_ D$ (the type of rules used in transformational phonology) are only one of a number of rule types examined by Chomsky in the 1950s. The different rule types define a hierarchy of types of grammars (see table 3.3), each capable of describing a wider range of languages and linguistic phenomena than those below it in the hierarchy. $V_T$ is the set of terminal symbols, $V_N$ the set of non-terminal symbols, $X^*$ is a sequence of zero or more $X$s, $X^+$ is a sequence of one or more $X$s, $|X|$ is the length of $X$ (i.e. the number of symbols in $X$), and $\emptyset$ is the empty string. In transformational phonology, the distinction between terminal and non-terminal symbols is minimal. Since any segment may be rewritten, they must
Table 3.3 The Chomsky Hierarchy of generative grammars

<table>
<thead>
<tr>
<th>n</th>
<th>Rule type, with conditions on their symbols</th>
<th>Name</th>
<th>Automaton type</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>A → B, A ∈ (V_T ∪ V_N)<em>, B ∈ (V_T ∪ V_N)</em></td>
<td>Unrestricted</td>
<td>Turing Machine</td>
</tr>
<tr>
<td>1</td>
<td>A → B / C _ D, A ∈ V_N, B ∈ (V_T ∪ V_N)<em>, C, D ∈ (V_T ∪ V_N)</em></td>
<td>Context-sensitive</td>
<td>Linear bounded</td>
</tr>
<tr>
<td>1</td>
<td>A → B,</td>
<td>A</td>
<td>≤</td>
</tr>
<tr>
<td>2</td>
<td>A → B, A ∈ V_N, B ∈ (V_T ∪ V_N)*</td>
<td>Context-free</td>
<td>Push-down</td>
</tr>
<tr>
<td>3</td>
<td>A → aB, A ∈ V_N, B ∈ (V_N ∪ {∅}), a ∈ V_T</td>
<td>Right linear</td>
<td>Finite state</td>
</tr>
<tr>
<td>3</td>
<td>A → Ba, A ∈ V_N, B ∈ (V_N ∪ {∅}), a ∈ V_T</td>
<td>Left linear</td>
<td>Finite state</td>
</tr>
</tbody>
</table>

all be members of both V_N and V_T. Chomsky and Halle (1968: 5 n. 1) state that 'boundary symbols do not appear in phonetic representations'. Since phonetic representations are elements of V_T*, it can be inferred that boundary symbols such as #, ## and + are only in V_N, not V_T. Boundary symbols are supplied by morphosyntactic structure and may not be rewritten.

Grammars of type number n are called G_n, and the class of languages defined by grammars G_n is called L_n. Context-sensitive and monotonic grammars are weakly equivalent (the question of their strong equivalence is discussed in section 3.6 below), as are right linear and left linear grammars (which are trivially not strongly equivalent, since they have right-branching and left-branching derivation graphs respectively). It can be shown (see Wall 1972: 213) that L_1 is a proper subset of L_2, L_2 is a proper subset of L_3, and L_3 a proper subset of L_0; hence the term 'language (or grammar) hierarchy'. Grammars G_n defining a class of languages L_n (0 < n ≤ 3) are said to be more restrictive than grammars G_m if and only if n > m. Abstract computing devices called 'automata' capable of recognising (i.e. determining well-formedness) each class of languages are listed in the right-hand column. A Turing Machine is a hypothetical computing device with the ability to read and write symbols on an infinitely extensible, bidirectional memory tape according to a finite set of instructions (Brooksheer 1989: 138-42). A linear bounded automaton is an automaton similar to a Turing Machine, in which the distance the memory tape may move in each step is bounded by a finite number. A push-down automaton is a linear bounded automaton in which the tape can move in one direction only, but with an additional 'first-in-last-out' auxiliary memory tape (a 'push-down store') (cf. Chomsky 1963: 338, Brooksheer 1989: 74–7), and a
finite-state automaton is like a Turing Machine with a unidirectional tape and no auxiliary 'working memory' at all (Brookshear 1989: 28–32).

The power of grammars is a two-edged sword: although more powerful grammars can describe complex phenomena with ease, the kind of computational machine required to implement them must also be considered. Turing Machines, for instance, can be used to compute the value of any recursively enumerable function. But the power of a Turing Machine is its own downfall, since there is no guarantee that such a machine can determine in a finite number of steps (i.e. in a determinate period of time) whether or not a solution can be found. It can thus carry blindly on forever, pursuing false leads. More modest computing machines may be unable to generate, recognise or parse such complex languages, but at least they will always either find a solution or terminate the search in a finite amount of time. Functions which can be computed by a procedure which will always either find a solution or terminate (i.e. problems which can be solved in practice) are called recursive functions, and are said to have effective or algorithmic solutions. (It is important not to confuse ‘effective’ with ‘efficient’. The human brain might, for all we know, embody an effective procedure for parsing, say, a certain non-context-free subset of the context-sensitive languages, which at the same time is nevertheless a highly inefficient algorithm involving, perhaps, multiply redundant parallel processing. There is no reason to expect the natural ‘algorithms’ of the brain to be efficient. Inefficiency might, for example, be an advantage in evolutionary or genetic terms, for example, to overcome limitations imposed by the fact that millions of brain cells die every day.) Since in everyday life humans appear to be able to generate, recognise and parse well-formed utterances, and to recognise ill-formed strings as ill-formed in a finite time (and initiate ‘repair’ strategies, or work out that the string is just not in their language), it is usually assumed by generative grammarians that natural languages are recursive, and that natural-language processing is algorithmic. (This assumption is occasionally called into question; see e.g. Chomsky 1981: 11, Langendoen and Postal 1984, Manaster-Ramer 1983, Perrault 1984: 173–4. This issue is pursued in section 3.5.4 below.) The goal, then, in developing a formal theory of natural-language syntax or phonology, is to use a type of grammar which is as powerful as necessary, but as restrictive as possible.

Deletion rules and excessive power
In the preceding discussion of Japanese, it was shown that deletion rules can give rise to the invertibility problem in a transformational grammar. In this
section I shall demonstrate that this problem is intrinsic to the rewrite-rule formalism, and not just to the particular set of rules which I discussed above.

Deletion rules (i.e. rules of the form \( A \rightarrow \ldots \emptyset \ldots \)) present special problems for algorithmic implementations of grammatical competence. Thus it is of some consequence that phonologists make frequent use of deletion rules in modelling phenomena such as elision. While the addition of deletion rules to a finite-state grammar or a context-free grammar has no deleterious effects (Chomsky 1963: 367), the addition of deletion rules to a context-sensitive or monotonic grammar (note the restriction to \((V_T \cup V_N)^+\), not \((V_T \cup V_N)^*\)) turns it into an unrestricted grammar with Turing Machine power. The excessive power of phonological rewriting rules is established by the following theorem (see also Levelt 1976: 242–3):

**Theorem 1** (Révész 1983: 53).

For every type 0 grammar (i.e. unrestricted rewriting system) there is an equivalent grammar \( G' \) in which every rule has any of the following forms:

1. \( S \rightarrow \emptyset \)
2. \( A \rightarrow a \)
3. \( A \rightarrow B \)
4. \( A \rightarrow BC \)
5. \( AB \rightarrow AC \)
6. \( AB \rightarrow CB \)
7. \( AB \rightarrow B \)

where the initial symbol \( S \) may occur only on the left-hand sides of the rules.

In this theorem, \( A, B \) and \( C \) are single non-terminal symbols, \( a \) is a terminal symbol and \( \emptyset \) is the empty string. All of these kinds of rule are endemic in generative phonological analyses, with the exception of the first, which is only needed in grammars of languages in which the empty string is a well-formed sentence, and in natural-language grammars can be ignored. To wit:

1. \( A \rightarrow a \) is exemplified by rules in which a phonological unit \( A \) is rewritten as a systematic phonetic segment \( a \), which is not subsequently altered by the phonological component.
2. \( A \rightarrow B \) is exemplified by phonological rules in which a segment \( A \) is modified in some way by being rewritten as a segment \( B \), which is in turn altered by other rules.
3. \( A \rightarrow BC \) is exemplified by 'fission' rules creating a contour or diphthongal unit from a non-contour or monophthongal unit, and
phrase-structure rules such as metrical structure rules. Although such rules are sometimes prohibited in SPE-type phonology (e.g. Pulleyblank 1986: 161), and require the X symbol in rules of the form $A \rightarrow X / U _{–} V$ to be two symbols long, they can always be replaced by rules in which X is restricted to a length of one symbol, as follows: every rule of the form $A \rightarrow BC$ can be replaced by three rules $A \rightarrow D, \emptyset \rightarrow C / D _{–}, D \rightarrow B$, where D is an auxiliary symbol introduced to prevent a C from being inserted after Bs in other environments i.e. not deriving from As.

4 $AB \rightarrow AC$ is exemplified by right-context phonological rules of the form $B \rightarrow C / A _{–}$.

5 $AB \rightarrow CB$ is exemplified by left-context phonological rules of the form $A \rightarrow C / _{–} B$.

6 $AB \rightarrow B$ is exemplified by right-context phonological deletion rules of the form $A \rightarrow \emptyset / _{–} B$.

No special hope should be placed in the absence of rules with two-sided environments from this list, since it is now known (Penttonen 1974, Révész 1983: 47–52) that every two-sided context-sensitive grammar has an equivalent left- and right-sided correlate.

Two restrictions on the context-sensitive rules favoured by phonologists may at first sight appear to redeem them. The first, due to Chomsky and Halle (1968: 332), is that in phonological rules of the form $A \rightarrow B / C _{–} D$, $|A|, |B| = 0$ or 1. While this is certainly a ‘restriction’, it does not reduce the power of the context-sensitive rule format, since the rules in theorem 1 all observe that restriction. A second restriction, due to Johnson (1972: 45–56), is that if context-sensitive rules are prohibited from applying to their own output, the resulting grammar is finite-state. However, such cyclic application of context-sensitive rules is quite customary in segmental generative phonology, and would seem to be unavoidable in the segmental analysis of suprasegmental phenomena such as the account of English stress in Chomsky and Halle (1968), or aspects of Chamorro phonology according to the analysis by Chung (1983). Support for the possibility of cyclic rule application has waned recently in light of non-cyclic analyses of a number of phenomena long taken as exemplars of cyclicity. Szpyra’s (1992) non-cyclic analysis of Polish yers is a case in point. Lauri Karttunen (personal communication) has pointed out to me that Johnson’s result includes the iterative application of rules in which the focus of the rule ‘migrates’ monotonically through the string to the left or right with each
3.5 Excessive power of transformational grammars

application. The SPE formalism is not explicitly subject to such a restriction. It is not clear whether it is implicitly subject to this restriction (see Kaplan and Kay 1994: 365). Johnson (1972: 110) also shows that certain extant phonological analyses employing integer-valued features are not finite-state. Johnson's restriction is thus not applicable to the rule formalism used in transformational phonology, and we conclude that the 'context-sensitive' rewrite rules employed in generative phonology are, in general, not strictly context-sensitive, but transformational. Chomsky and Halle themselves categorise the SPE rules as 'local transformations' (see Chomsky 1965: 99). Joshi and Levy (1977) show that all local transformations have no more than context-sensitive power. Theorem 1 demonstrates that Chomsky and Halle's claim that SPE rules are local transformations underestimates the power of SPE rule-systems.

Although generative phonology has moved on from the SPE model, many of its original practices, including the use of transformational rules, have been retained in subsequent work. For example, McCarthy (1979/1982: 200) says: 'we cannot dispense with nonsyntactic transformational rules entirely . . . Since it is impossible to express a metathesis rule by anything except transformational formalism, we must conclude that phonological rules do have this richer formalism available to them.' This claim is completely erroneous: the permutation of adjacent elements requires only context-sensitive generative power. Hudson (1980: 108–10) and Kaye (1985: 289) present non-transformational phonological descriptions employing partially ordered strings that are quite adequate for describing metathesis. See also the discussions of metathesis in chapters 2 above and 5 below.

Arguments such as Anderson's (1982) that the extension of Autosegmental and Metrical theories into the segmental domain considerably weakens phonological theory miss the point: it hardly matters that a grammar which is already too unrestrictive is weakened even further. It appears, then, that a variety of formal arguments all point to the following conclusion:

*Excessive power of rewrite rules*

The simple context-sensitive rules with deletion employed by generative phonologists define completely unrestricted rewriting systems – the most powerful, least constrained kind of grammar in the Chomsky hierarchy.

Similar formal problems with the excessive power of transformational grammars led syntacticians of various viewpoints to abandon the use of
such rules in the early 1980s and attempt to find more restrictive grammatical formalisms. In the subsections which follow, I shall consider two questions: (i) why classical transformational grammars are no longer used in syntactic analyses; and (ii) what lessons can be learned from syntactic theory which may help phonologists to overcome the excessive power of transformational phonology. I shall examine the causes and nature of this excessive power, in an attempt to devise ways of overcoming it.

3.5.3 Unrestricted transformational grammars and their string-sets
Quite independently of the formal problems with unrestricted transformational grammars, linguists now accept that transformational grammars of the Aspects type (Chomsky 1965) are too unrestricted for the realistic analysis of natural languages. Chomsky and others have consequently attempted to establish a number of universal principles and constraints, such as subjecacy, by which the range of possible languages that contemporary transformational theory defines is further restricted (see Chomsky 1981, van Riemsdijk and Williams 1986). Advocates of GPSG (see Gazdar 1982, Gazdar et al. 1985) and several other frameworks, such as LFG (Bresnan 1982), and Head-driven Phrase-Structure Grammar (Pollard and Sag 1987), claim that transformations are not needed at all, and the belief that the grammars of all natural languages are not more powerful than context-sensitive is widely expressed. (The possibility of applying the techniques developed in non-transformational syntactic theories to phonological analysis is taken up in detail in chapters 5 and 6.)

By contrast, the continued preponderance of all manner of transformational rules in generative phonology suggests that phonologists, on the whole, are unaware of the formal problems that led syntacticians to give up transformational rules, or that they regard phonology to be governed by different considerations from syntax (see Bromberger and Halle 1989). Consequently, I shall present the issues involved in the 'generative capacity debate' informally at first, with references where applicable to more technical presentations. Because the weak generative capacity of a grammar is not affected by whether its terminal symbols are words, morphemes, phones, feature-matrices or segments of whatever extent, this discussion is just as relevant to assessments of the restrictiveness of phonological analyses as to syntactic analyses.

As background to this debate, it should be appreciated that Chomsky's original claim that natural languages cannot be adequately characterised by even context-sensitive phrase-structure grammars (Chomsky 1957,
3.5 Excessive power of transformational grammars

1963), far from being the great 'advance' which later theorists tenaciously attempted to defend, was a retraction from the stronger proposition that grammars of natural languages are all highly restricted in weak generative capacity.

It used to be believed (Chomsky 1965: 208 n. 37, Wall 1972: 295, Levelt 1976: 243) that the class of languages generated by transformational grammars was a proper subset of the recursive sets, and that they could thus be recognised effectively. Peters and Ritchie (1973) put paid to this belief by showing that Standard Theory (Aspects) transformational grammars defined in Chomsky (1965) generate the type 0 languages, those with recursively enumerable string sets (Wall 1972: 294–5, Peters and Ritchie 1973, Levelt 1976, Newmeyer 1980: 175–6). Since generative grammarians by and large believe that natural languages constitute a proper subset of the recursive sets, which are properly included in the recursively enumerable sets, Standard Theory transformational grammars are, according to Peters and Ritchie's demonstration, capable of defining string-sets which are not possible human languages.

This result should be treated with care. Lapointe (1977) shows that it depends crucially on transformational rules of a kind that has never been encountered in actual grammars of natural languages. If such rules are prohibited, grammars which define recursive sets and yet still retain transformational rules are produced. The argument then becomes one of expressiveness, parsimony etc., not weak generative capacity per se. In a similar vein, Berwick and Weinberg (1984) point out that weak generative capacity should not be afforded undue attention, as more restricted grammar classes can nevertheless define exotic string-sets with properties that are quite unlike those of attested natural languages (for example, languages consisting entirely of palindromes are all context-free), and point out that the languages defined by highly restricted grammars may be exactly 'covered' by less restrictive, but more elegant grammars. For example, the syllables of a language form a finite-state language (see Carson 1988, Carson-Berndsen 1990), but are more elegantly described using a context-free phrase-structure grammar (see Randolph 1989: 59), since they have a hierarchical, branching structure. The use of transformational grammars might also be defended in this way.

Such excessive power is only problematic for linguistic theory because of the belief that natural languages are recursive sets. (In this discussion, the set of strings generated by a generative phonology from the lexicon is regarded as a language.) Alternatively, it could be that the belief that
natural languages are not recursive sets noted earlier is correct, even though it is a less popular view. In the next subsection, let us examine the issue of recursiveness in more detail, in order to determine why most generative grammarians believe natural languages to be recursive.

3.5.4 Recursiveness

The argument that natural languages are recursive can be summarised informally as follows. Generative grammarians generally believe that speakers of a natural language can judge with equal ease the grammaticality or ungrammaticality of any arbitrary finite string in that language. This is what is meant by the statement that natural languages are recursive sets, since a recursive set is a set whose complement is also recursive. If speakers can judge both that well-formed strings are well-formed and that ill-formed strings are ill-formed, then the set of well-formed strings and the set of ill-formed strings have each other as a recursive complement, and are hence recursive sets. Bresnan and Kaplan (1982:xl) describe the recursiveness of natural-language grammars as a ‘reliability constraint’, and express it ‘a speaker of a language is regarded as a reliable arbiter of the sentences of his or her language’.

If natural languages were recursively enumerable, but not recursive, then a speaker could not be guaranteed to determined that an ill-formed string (e.g. a foreign string) was ill formed, though they might still be able to effectively recognise well-formed strings. Since speakers of natural languages seem able to distinguish well-formed from ill-formed utterances with equal ease, the class of natural languages is believed to lie within the more restricted class, recursive sets. But Standard Theory transformational grammars define a much wider, less restricted, range of languages than the recursive sets, and consequently Standard Theory transformational grammar is too unrestricted to be a realistic theory of natural language.

Psychological realism: learning and parsing

The excessive power of transformational grammars (including SPE phonology) is consequently also relevant to the specific claim of generative linguists that theories of grammar should be psychologically realistic. In essence, this boils down to the following requirements of a generative grammar:

1. It must be learnable;
2. its strings must, therefore, be recognisable; and consequently
3. it must be decidable (i.e. recursive).
Learnability theory is a complex and technical field of study, which rests on many idealisations and assumptions which might be questioned. It is not necessarily the case that more restrictive grammars are always more (easily) learnable than less restrictive grammars, so I shall concentrate on points 2 and 3 only. A grammar G is decidable if the language it defines L(G) and the complement of that language are recursive, that is, a procedure which instantiates G (e.g. an abstract automaton equivalent to G) can always decide in a finite number of steps (i.e. a finite period of time) whether any string (e.g. any string of segments) is well formed or ill formed.

A grammar G is parsable if there is some effective procedure which can assign one or more structural descriptions according to G to any well-formed string of L(G). Criticisms of transformational and context-sensitive grammars (Gazdar 1982, Sampson 1983) have centred on the fact that while effective procedures are known that in the worst case parse context-free languages in time proportional to the cube of the length of the string, and in ideal cases in linear time, no such psychologically interesting results are available for less restrictive classes of grammars. In itself, this does not constitute a formal objection to less restrictive grammars, since the possible psychological 'realism' of a grammar is not dependent on our ability to find effective algorithms. Berwick and Weinberg (1984) argue convincingly that the asymptotic performance of very general algorithms is not directly relevant to linguistic arguments, for very specific algorithms are known for parsing less restricted languages (e.g. context-sensitive languages) in nearly linear time. They also argue that space costs are just as relevant as time costs to the evaluation of parsing efficiency, and that whereas general parsing algorithms for context-free languages may be time-bounded, they are not exactly thrifty in their use of computational space. This problem has been made painfully apparent in parsers for GPSG, where the implementation of ID/LP rules and metarules causes an explosion in the computational time and space required (see Thompson 1982, Barton, Berwick and Ristad 1987). Experimental evidence that ill-formed strings are not especially troublesome to recognise is given in several psychological studies, such as Greenberg and Jenkins (1964). I take such results as indicative of recursiveness.

3.5.5 Summary of problems with transformational phonology

In this section I shall summarise the problems with transformational phonology that have been discussed above. I shall conclude this chapter by considering two strategies for overcoming the excessively unconstrained nature of transformational phonology. In section 3.5.6 I shall discuss a
number of constraints on transformational phonology which have been pro-
posed in the literature. I shall suggest in section 3.6 that the power of the
grammar formalism needed to account for natural phonological phenomena
might be restricted if segmental, concatenative representations were to be
given up in favour of other forms of representation. After a brief discussion
of the historical limit of segmental phonological representation, supraseg-
mental phonology, I shall proceed in chapter 4 to examine a clutch of alter-
native 'multilinear' theories of phonological representation.

The argument of this chapter has so far been developed as follows. There
are a number of 'in principle' computational arguments against the seg-
mental–transformational model of phonology, centring on the contention
that since it sanctions deletion rules, it may define non-recursive languages,
and thus be overpowerful, unlearnable or both. Since any computational
procedure can be computed by an unrestricted rewriting system, the use of
such systems constitutes the null hypothesis about human linguistic com-
petence that it is a computational procedure of some otherwise uncon-
strained kind: 'if all we can say about a grammar of a natural language is
that it is an unrestricted rewriting system, we have said nothing of any inter-
est' (Chomsky 1963: 360).

There are also a number of practical and empirical arguments against the
view that the transformational phonology model is appropriate:

1 Because of the great richness of the class of rules which may be
expressed in the transformational model, it may be difficult, if not
impossible, in practice to derive an optimal transformational
phonology. This problem is compounded by the many interactions
that may obtain between rules, and is highlighted by the existence
of rule-ordering paradoxes.

2 Structural domains such as syllable, onset, foot etc. are not explic-
itly represented in transformational phonology, so that structural
and non-adjacent contextually conditioned variation cannot be ex-
pressed felicitously. (This argument is pursued further in chapter 4.)

3 Deletion and insertion rules permit arbitrarily long portions to be
removed from or added to a string. However, real phonological
processes do not employ such powerful mechanisms. Most phono-
logical processes involve rearrangement of material, rather than
deletion or insertion, and many apparent cases of 'deletion' (e.g.
the Japanese data examined above) do not involve deletion at all.

I believe that these problems can be attributed in part to the following
cause: generative phonology attempts to cast phonetic and phonological
3.5 Excessive power of transformational grammars

representations in an identical formalism. ‘Phonetic’ representations are merely the end result of a long process of mappings from strings to strings. If phonetic representations are considered instead to be the denotations of phonological representations, it can be seen that generative phonology attempts to do for phonology what generative semantics did for syntax.

Although transformational rules are now widely acknowledged to be too powerful, the strategy which has been adopted for limiting this power, namely the preference for general constraints rather than specific rules, has not been shown to appropriately restrict the class of languages that can be defined. Furthermore, transformational rules (in particular, deletion and insertion rules) are resurrected whenever it is convenient. Such rules are only required because strings continue to be central to phonological representations.

3.5.6 Constraints on transformational phonology

In this subsection I shall examine a number of additional constraints which it has been proposed should be added to the classical transformational model of phonology examined above. I shall attempt to elicit the circumstances under which the proposed constraints would succeed in ameliorating the problems of excessive power discussed above.

Strict cyclicity
Rubach (1984, see also references therein) argues that strict cyclicity constrains the remoteness of underlying representations in the following way: underived forms will not be subject to any cyclic rules at all (only postcyclic rules), which requires the lexical representation of stems to be very similar to the systematic phonetic representation (the degree of difference depends only on the non-cyclic rules, which by definition can apply once only). The phonological remoteness of the lexical constituents of a derived form is in like manner constrained by the inability of a rule block to return to an earlier cycle. Strict cyclicity thus appears to enforce bounds on the length of derivations.

‘Naturalness’ constraints
Postal (1968) argues that since phonological transformations model morphophonological alternations, the systematic phonemic representation of non-alternating forms (or the parts of an alternating form not directly involved in an alternation) should not be different from the systematic phonetic representation. For example, although it is argued within transformational phonology that the systematic phonetic segment /ʃ/ in electrician
Segmental and transformational phonology
derives transformationally from an underlying /k/ in the systematic phonemic representation (cf. electric), surface /ʃ/ in non-alternating forms (e.g. dish) cannot be derived from an abstract underlying form non-distinct from the surface form.

For non-alternating forms, Postal's Naturalness Condition places a bound on the length of the derivation: if 'nondistinctness from the surface form' is taken to mean 'distinct only from the surface form by virtue of addition of redundant features', then the length of the derivation is bounded by the number of redundant features to be added – a small, finite quantity. In alternating forms, the Naturalness Condition serves to predict that the deep structure of the lexical entry is nondistinct from the underived surface form. For example, the deep structure of electriʃian contains a /k/ and not, say, /ʃ/, because the underived form electric surfaces with a /k/, and the lexical representation of that underived form must be non-distinct from the surface, according to the Naturalness Condition. However, it is questionable whether the Naturalness Condition places any limits at all on the number of steps or 'craziness' of the route that a derivation might take in transforming a 'natural' lexical representation into a 'natural' surface representation. Since transformations may add, delete or substitute segments, the derivation of electriʃian might proceed from an underlying electrikian via intermediate forms of unconstrained exoticness. For the Naturalness Condition to be most constraining, it is necessary to prohibit deletion and substitution rules, the kinds of rules which it was argued above were most problematic for the power of the transformational formalism.

A stronger version of Postal's Naturalness Condition, the True Generalisation Constraint (Hooper 1976: 13), holds that systematic phonemic and intermediate representations must also conform to the generalisations and constraints which hold in surface structure. For example, phonotactic constraints observed to hold of surface forms cannot be violated in intermediate or deep structures. 'Surface-trueness' places an important limit on the extent to which lexical phonological representations may differ from their surface forms.

A consequence of surface-trueness is that it rejects analyses in which a constraint holds for part of a derivation, but is then relaxed or ignored for another part of the derivation. An example of such an analysis is Wiese (1990), in which the Obligatory Contour Principle is regarded as holding over lexical representations, but is then suspended at later levels of representation. Another example is the use of the Obligatory Contour Principle or the No Crossing Constraint in Autosegmental Phonology as devices for
3.5 Excessive power of transformational grammars

'fixing-up' ill-formed intermediate representations so that they 'come out right in the end' (see section 4.3.3). Since these analyses employ an intermediate level of representation which does not adhere to constraints which hold on the surface, the analysis is not 'surface-true' and is thus potentially excessively unconstrained. A further example is the analysis presented in Borowsky (1989: 145), in which 'After Level 1, Structure Preservation is turned off, and as a result, syllable structure is less restrictive, allowing larger codas and making vowel shortening unnecessary.'

Structure preservation

The principle of structure preservation, first formulated by Kiparsky (1985), is expressed in Borowsky (1989: 148) as follows: 'Language-particular structural constraints holding for underlying representation hold also for derived representation, and vice versa . . . By this we mean that structural constraints persist, and no new constraints may be introduced during the derivation.' Like strict cyclicity and 'naturalness' constraints, structure preservation would be genuinely restrictive if destructive rules such as deletion were abandoned. But though structure preservation is apparently declarative in spirit, according to Borowsky's definition it is structural constraints, not the phonological representations themselves, which are 'preserved'. For instance, a deletion rule like Halle and Mohanan's n-Deletion (Halle and Mohanan 1985: 96, Mohanan 1991), which deletes the final /n/ of hymn to form the isolation form /him/ (cf. hymnal) would not violate Borowsky's definition of 'structure preservation', since it is not introduced or suspended at any stage in the derivation, but is always available to apply (in its appropriate stratum). 'Structure preservation' does not preserve structure at all, but preserves 'structural constraints', a different matter entirely. In this respect a generative phonology incorporating 'structure preservation' is no different from any other normal kind of rewriting system, in which all rules are available throughout the derivation.

Another example of 'structure preservation' can be seen in the account of metrical structure assignment in Kiparsky (1979). In contradistinction to the analysis of Liberman and Prince (1977), in which all metrical structure is erased ('deforestation') at the end of each cycle prior to reconstruction in subsequent cycles, in Kiparsky's analysis 'the assignment of metrical structure is cyclic in an essential way . . . there can be no Deforestation . . . The effect is that metrical structure assigned in earlier cycles is kept, insofar as it is not redrawn by the reapplication of (1b) [assignment of feet]' (Kiparsky 1979: 422; emphasis added). The caveat which I have emphasised
in this quotation rather weakens the structure preservation constraint. But it is weakened still further by being violated by some steps in the analysis which Kiparsky develops. For example: 'This patterning is obliterated by the effects of the Rhythm Rule in those word types which are subject to it. For reasons which will become clear below, I will assume that the Rhythm Rule is an operation on metrical trees' (Kiparsky 1979: 424). If the structure preservation constraint were observed, the resulting grammar would be monotonic, which would indeed be a significant restriction on an otherwise excessively powerful formalism. But since it is not observed, I conclude that it does not constrain orthodox generative phonology at all.

It is notable that in the Chomsky hierarchy, the critical determinant of the power of a rule-system is rule type, not the type of symbols that are used. Therefore, given the choice, it may be better to encode problematic phenomena by using richer forms of representation rather than more powerful rules. This is the strategy that has been adopted by successful alternatives to transformational grammar, such as GPSG (Gazdar et al. 1985) and LFG (Kaplan and Bresnan 1982). In chapters 5 and 6 below I shall pursue the consequences of this hypothesis in much more detail. In the following two sections, however, I shall examine two aspects of linguistic representation – syntagmatic and paradigmatic structure – and show that by employing richer forms of syntagmatic and paradigmatic representations, the notion of 'rewriting rules' can be completely dispensed with.

3.6 Phrase-markers and representations of derivations

3.6.1 Order in derivations

In the theory of generative grammar, the concept of (representations of) phrase structure was originally parasitic on the notion of derivation (see Chomsky 1957: 28). Only more recently, in works such as Stowell (1981) and Gazdar et al. (1985: 44), has the study of syntagmatic structure in linguistic representations been liberated from the notion of 'derivation'. In generative phonology, change has been slower, and there are still two popular ways of thinking about derivations. To illustrate these two views, consider a derivation 1 \Rightarrow i_1 \Rightarrow i_2 \Rightarrow \ldots i_n \Rightarrow s from a lexical representation 1 to a systematic phonetic representation s via intermediate representations i_m. In the first view, a rule is a means of defining a single link i_m \Rightarrow i_{m+1} in the relation 1 \Rightarrow * s that relates surface and lexical representations. In contemporary terminology, it could be said that a rule licenses a step in a derivation. This relational view is agnostic with respect to possible implementations of rules. A transformational phonology of this kind simply encodes grammatical competence, and could be employed in a performance
3.6 Representations of derivations

theory to derive lexical representations from surface representations (in speech perception, for example) or vice versa (in speech production). It should be noted that in formal descriptions of generative grammar (e.g. Chomsky and Miller 1963: 292), rules and the derivations they license are indeed relations of this kind: rules \( A \rightarrow B \) are elements in the set of productions \( P \), and could all be written equivalently as \((A, B) \in P\), without an arrow in sight to suggest any sort of primacy to the left-hand side of the rule or the derivation it licenses. For example, Chomsky (1965: 9) emphasises that

To avoid what has been a continuing misunderstanding, it is perhaps worth while to reiterate that a generative grammar is not a model for a speaker or a hearer. It attempts to characterize in the most neutral possible terms the knowledge of the language that provides the basis for actual use of language by a speaker–hearer. When we speak of a grammar as generating a sentence with a certain structural description, we mean simply that the grammar assigns this structural description to the sentence. When we say that a sentence has a certain derivation with respect to a particular generative grammar, we say nothing about how the speaker or hearer might proceed, in some practical or efficient way, to construct such a derivation.

In the second view, a rule is a description of a phonological ‘process’, a mapping from string \( s_1 \) to \( s_2 \), involving modifications made to \( s_1 \) in order to generate \( s_2 \). This view embodies (and partly derives from) a wholly procedural view of (human or abstract) computation, the sequential machine metaphor (S. Bird, personal communication) of linguistic competence. There are two important senses in which this second view is a simpler and more faithful view of how generative phonologists conceive of their analyses. Firstly, despite Chomsky and Halle’s occasional ex cathedra decrees about the incorrectness of this view, such as that quoted above, they themselves present their analyses in these terms most of the time. Bromberger and Halle (1989: 51, n.1) play down the distinction between the declarative and procedural metaphors: ‘We deliberately eschew in this discussion the use of “declarative rules” and “procedural rules” in characterizing the differences between syntax/semantics on the one hand and phonology on the other hand. That terminology, which carries a number of associations from the domain of computational linguistics, strikes us as unhelpful.’ Chomsky (1981: 90), too, states:

[Theory] Ia, assumes as above that Move-\( \alpha \) forms these S-structures from base-generated D-structures . . . [Theory] Ib assumes that the base generates S-structures directly . . . It is not easy to find any empirical evidence to distinguish between Theories Ia and Ib. It may be that they are to be understood as two realizations of the same more abstract theory.
It is notable that Halle's and Chomsky's descriptions are nevertheless presented in a procedural manner, and these caveats are relegated to footnotes. To illustrate the bias towards generation of generative phonology in actual practice, consider the following examples. First, Chomsky (1965: 82) contains this discussion of a phonological rule:

\[(19) \ [+\text{continuant}] \rightarrow [+\text{voiced}] / \_ [+\text{voiced}] \]

This will convert [sm] into [zm], [fd] into [vd], [sg] into [zg], etc.

[Emphasis added.]

In a footnote he goes on to state: ‘Rule (18) [should be (19) – J.C.] applies to any segment specified as [+continuant] (hence to [s]) in a context which is specified as \_ [+voiced] (hence to the context \_ m), \textit{converting} the segment to which it applies to a voiced segment’ (Chomsky 1965: 213 n. 14; emphasis added).

Second, Bear (1990) demonstrates that a typical transformational phonological analysis actually only works in the ‘forwards’ direction (i.e. from lexical to surface representations). In order to make such analyses invertible or bidirectional (and thus made more obviously descriptions of phonological competence rather than one aspect of performance viz. production), the rules of the analysis have to be substantially altered, for example by adding features which would be redundant in generation, but which are necessary to restrict the search space in analysis.

More recently, Chomsky (1994: 6) explicitly defends the derivational view and the opacity of underlying representations:

My own judgement is that a derivational approach is nonetheless correct . . . There are certain properties of language, which appear to be fundamental, that suggest this conclusion. Under a derivational approach, computation typically involves simple steps expressible in terms of natural relations and properties, with the context that makes them natural ‘wiped out’ by later operations and not visible in the representations to which the derivation converges.

In contradistinction to the context-sensitive, rewriting model of grammar, a context-free phrase-structure grammar could be said to be non-derivational because any order of application of a particular set of rules to a given string assigns the same phrase-marker to that string (excepting structural ambiguity). For example, the contemporary view of stress assignment and syllable structure assignment (the details of which will be discussed further in section 4.4.3) is essentially a context-free theory which can be implemented by simultaneous application of independent foot- and syllable-structure context-free rules. The precise details of the rules in question are
subject to some degree of debate, but the following analysis will illustrate the feasibility of a non-derivational account. Firstly, according to the version of Metrical Phonology in Hayes (1982), Myers (1987) and Borowsky (1989), there are metrical (phrase-)structure rules for two kinds of syllable, those that have a branching rime (heavy syllables), and those whose rime does not branch (light syllables). According to this version of Metrical Phonology, -VC syllables are light in English, even though the rime apparently branches, because intervocally the C is parsed with the following syllable, and word-final C is regarded as extrametrical. -VCC and -VVC syllables are also possible heavy syllables. In these two cases, the final C is regarded as extrametrical, and the remaining -VC or -VV is a binary branching rime, which is thus heavy. Following Hayes (1982), extrametrical constituents will be marked with the diacritic feature [+em]. We can write a simple feature-based context-free grammar to describe these principles as follows:

(3.26) \[ \sigma \rightarrow O \ R \]  
\([\alpha \text{ heavy}] \quad [\alpha \text{ heavy}]\)

(3.27) \[ R \rightarrow X \ X \]  
\([+\text{ heavy}]\)

(3.28) \[ R \rightarrow X \]  
\([-\text{ heavy}]\)

We can also write metrical phrase-structure rules for various kinds of feet. Three disjoint cases can be identified:

(3.29) \[ \Sigma \rightarrow \sigma \]  
\([+\text{ heavy}]\)

(3.30) \[ \Sigma \rightarrow \sigma \ \sigma \]  
\([+\text{ heavy}] \quad [-\text{ heavy}]\)

(3.31) \[ \Sigma \rightarrow \sigma \ \sigma \ \sigma \]  
\([-\text{ heavy}] \quad [-\text{ heavy}] \quad [+\text{ em} ]\)

For greater generality, the first two rules can be collapsed to

(3.32) \[ \Sigma \rightarrow \sigma \left( \sigma \right) \]  
\([+\text{ heavy}] \quad [-\text{ heavy}]\)

The third case, the ternary foot rule, is given a degree of internal structure by some phonologists (e.g. Liberman and Prince 1977, Kiparsky 1979), which can be expressed:

(3.33) \[ \Sigma_{\text{ternary}} \rightarrow \Sigma_{\text{binary}} \ \sigma \]  
\([-\text{ heavy}]\)
Segmental and transformational phonology

where \( \Sigma_{\text{binary}} \) is expanded as in rule (3.30). Alternatively, if a final light syllable is analysed as a daughter of the word rather than of the final foot, only two foot rules are required:

\[
\begin{align*}
(3.34) \quad \Sigma & \rightarrow \sigma \\
& \quad [+\text{heavy}]
\end{align*}
\]

\[
\begin{align*}
(3.35) \quad \Sigma & \rightarrow \sigma \quad \sigma \\
& \quad [-\text{heavy}]
\end{align*}
\]

To encode the fact that the foot which is the main-stress domain comes at the right-hand edge of the word, the following word-level metrical phrase-structure rule is sufficient:

\[
(3.36) \quad \omega \rightarrow (P_w) \Sigma_s
\]

More precisely, this rule relates to the constituents called \( \alpha \)-constituents in Selkirk (1984a), level-1 constituents in Church (1985) and stratum-1 constituents in Mohanan (1986) and Halle and Mohanan (1985). The structure of the optional \( P \) (pretonic) constituent is given by the rule \( P \rightarrow (\sigma[-\text{heavy}] (\sigma[-\text{heavy}])) \Sigma^* \). (I have allowed for up to two initial light syllables in the pretonic because some words, such as \( \text{stereoscopist} \), may be spoken with two unfooted initial unstressed syllables rather than footed into a stressed-unstressed pattern, yielding, for instance, \( \text{stèreoscopist} \), with secondary stress on the initial syllable, though that metrical pattern is also well formed.) In words of only one or two light syllables (which will not satisfy the above word rule for want of the appropriate syllables to form even a normal binary foot), the first is stressed:

\[
(3.37) \quad \omega \rightarrow \sigma_s \\
& \quad [-\text{heavy}] \left( \sigma_w [-\text{heavy}] \right)
\]

The way in which these rules are applied need not be regarded derivationally. This is because the same selection of phrase-structure rules can be applied top-down, bottom-up, left-to-right, right-to-left, outside-in, inside-out, breadth-first, depth-first or even head-first to yield the same set of hierarchical phrase-structure relations. For examples see figures 3.4 and 3.5. The particular metrical structure defined by both of these examples is that of words such as \( \text{stereoscopist} \): Bresnan and Kaplan (1982:xlv) refer to this property of context-free grammars as 'order-free composition'. Another term for this same property is 'declarativeness' (Shieber 1986: 11, 24). The equivalence between context-free phrase-structure rules and partial linguistic structures on
which the order-free composition of context-free phrase-markers is based is a significant weakening of the established dichotomy between rules and representations. In declarative grammar, the weakening of the distinction between rules and representations is exploited to such an extent that it is possible to regard such grammars as non-derivational, representation-based, rewrite rule-free formalisms.

It is not immediately obvious that a similar non-derivational view could be established for more powerful kinds of phrase-structure grammar, not least of all because more powerful kinds of phrase-structure grammars lack an equivalent graphical construct to context-free phrase-markers by which rules may be regarded as partial representations. I shall now show that such a general definition of phrase-marker is available, which will allow us to treat phrase-structure grammars of all kinds in the same declarative fashion as the context-free case discussed above. Since the general theory of phrase-markers which I propose below defines comparable sets of linguistic structures for all kinds of phrase-structure grammar in the Chomsky hierarchy, this development will afford us a way of comparing the strong generative capacity of phrase-structure grammars. This is an important (albeit hitherto rather neglected) development, because, as Chomsky (1957: 61) says: 'discussion of weak generative capacity marks only a very early and primitive stage of the study of generative grammar. Questions of real linguistic interest arise only when strong generative capacity (descriptive adequacy) and, more important,
explanatory adequacy become the focus of discussion." By providing a precise account of the strong generative capacity of a grammar, the following discussion counters any charge that the previous discussion focussed too narrowly on the weak generative capacity of phonological rule-systems.

Révész (1983) gives a general method of constructing derivation graphs for all kinds of phrase-structure grammar, which in the particular case of context-free phrase-structure grammars defines phrase-markers which are trees (though they are not identical to Chomsky's phrase-structure trees). Révész's method is as follows. First of all, each general rewriting rule $x_1x_2$
3.6 Representations of derivations

\[ x_1 \ldots x_p \rightarrow y_1y_2y_3 \ldots y_q, \] where \(|y| = q\) may be greater than or equal to \(|x| = p\) (i.e. in monotonic grammars), or less than \(p\) (i.e. in grammars with deletion rules), is assigned a rule number \(R\). Each application of rule \(R\) is represented in the derivation-graph by a subgraph consisting of:

1. the sequence \(x_1x_2x_3 \ldots x_p\);
2. arcs drawn from each \(x_n\) to a single node which is labelled with the rule number or label \(R\);
3. the sequence \(y_1y_2y_3 \ldots y_q\);
4. arcs drawn from the node labelled \(R\) to each \(y_m\).

Some examples of the contribution that individual rules of different kinds make to general derivation-graphs are the following:

1. The monotonic (context-sensitive) rule \(ABC \rightarrow Dab\)

\[
\begin{array}{ccc}
A & B & C \\
 & \downarrow & \\
R_1 & \\
 & \downarrow & \\
D & a & b
\end{array}
\]

2. The type-0 deletion rule \(ABC \rightarrow Bd\)

\[
\begin{array}{ccc}
A & B & C \\
 & \downarrow & \\
R_2 & \\
 & \downarrow & \\
B & d
\end{array}
\]

3. The context-free rule \(A \rightarrow Dab\)

\[
\begin{array}{ccc}
A \\
 & \downarrow & \\
R_3 & \\
 & \downarrow & \\
D & a & b
\end{array}
\]
Although representations of context-free (CF) derivations employing Révész's general phrase-markers are not identical to Chomsky's representation of context-free derivations, by collapsing the left-hand non-terminal nodes with the rule number node that they dominate, Révész's CF phrase-markers can be converted into Chomsky's CF phrase-markers. Furthermore, this correspondence can be inverted, so that Révész's CF phrase-markers can be produced from Chomsky's CF phrase-markers.

To illustrate the generality and the utility of Révész's method, I shall now present three example derivations in three different kinds of phrase-structure grammar, together with their derivation-graphs.

The monotonic language $a^n b^n c^n$

Révész (1983: 5) gives the following well-known grammar for the language whose strings are all of the form $a^n b^n c^n$

$$
R_1: S \rightarrow abc \\
R_2: S \rightarrow aXbc \\
R_3: Xb \rightarrow bX \\
R_4: Xc \rightarrow Ybcc \\
R_5: bY \rightarrow Yb \\
R_6: aY \rightarrow aaX \\
R_7: aY \rightarrow aa
$$

The derivation of $a^3 b^3 c^3$ is:

$$
S \Rightarrow aXbc \quad (R_1) \\
\Rightarrow abXc \quad (R_2) \\
\Rightarrow abYbcc \quad (R_4) \\
\Rightarrow aYbbcc \quad (R_4) \\
\Rightarrow aaXbbcc \quad (R_4) \\
\Rightarrow aabXbbcc \quad (R_4) \\
\Rightarrow aabbXcc \quad (R_4) \\
\Rightarrow aabbYbccc \quad (R_4) \\
\Rightarrow aabbYbbccc \quad (R_4) \\
\Rightarrow aaYbbcc \quad (R_4) \\
\Rightarrow aaabbbccc \quad (R_7)
$$

The way in which this grammar works is not immediately obvious from inspection of either the rules or the derivation. Some of the rules introduce an additional a ($R_6$, $R_7$), b ($R_4$) or c ($R_4$), but it is not obvious why new a's are introduced by two different rules, whereas new b's and c's can only be introduced by the same rule. X and Y can be moved about among the b's, according to rules $R_3$ and $R_5$: specifically $R_3$ allows an X to migrate right-
Figure 3.6 Derivation-graph of $a^3b^3c^3$

wards through a string of b’s, and $R_5$ allows a Y to migrate leftwards through a string of b’s.

Figure 3.6 makes the interactions between rules in this derivation more explicit. (Symbols not further rewritten – the terminal string of the derivation graph – are italicised for clarity.) From the graph it is easy to see how this grammar works. Firstly, $R_2$ introduces the first a, b and c, and a special control symbol X. Inspection of the a’s, b’s and c’s shows that they come through the derivation unchanged from beginning to end. The migrations of X and Y through the b’s, which arise from the interaction between rules, can be clearly seen as the diagonal lines which criss-cross the graph. The derivation-graph makes explicit the non-local composition of the purely local interactions between rules that encode the cross-serial dependencies
of this string. Furthermore, the function of these migrations and their localisation to strings of b's can be clearly seen: the migrations 'ferry' information between the (non-adjacent) a's and c's. This is done as follows. The only way to get rid of the non-terminal X introduced by R₂ is to make it migrate rightwards using R₃ until R₄ can apply. When R₄ applies, it introduces a new b and c at the interface between the b's and c's, together with a new non-terminal, Y. The purpose of the Y will be to introduce a new a to match the b and c that R₄ introduces. The only way to get rid of the non-terminal Y from the derivation is to migrate it leftwards through the b's, until R₆ or R₇ applies. Both R₆ and R₇ introduce the new a which is needed to balance the b and c introduced at the beginning of Y's migration. R₆ differs from R₇ only in that it also introduces a new X to start the whole process again to commence another recursive cycle in order to introduce yet more a's, b's and c's in equal numbers.

3.6.2 A phonological example: Japanese High Vowel Devoicing
Consider the application of rules of the form A → B / C __ D, an example of which is the Japanese rule that devoices vowels which occur between two voiceless consonants. This rule was formulated above as:

(3.11) \[ V \rightarrow [-\text{voice}] / C C \]
\[ [+\text{high}] [-\text{voice}] [-\text{voice}] \]

The application of this rule to a string such as /kikai/ is graphed using Révész's method as follows:

```
  k i k a i
 / \ / \
 R(3.11) / \
  k i k a i
```

Alternative order of application of two rules
The two rules \( R₁ = ABC \rightarrow DBC \) (i.e. \( A \rightarrow D / _BC \)) and \( R₂ = BCE \rightarrow BCF \) (i.e. \( E \rightarrow F / BC _\)) with complementary environments _ BC and BC _ yield the same result when applied to ABCE, whatever their order of application. This (vacuous) structural ambiguity is illustrated in the following partial derivation-graphs:
It is apparent from this diagram that the ambiguity arises from the fact that the material that is changed by rule \( R_1 \), A, is not mentioned in rule \( R_2 \), and the material that is changed by rule \( R_2 \), E, is not mentioned in rule 1. The material that is mentioned in both rules, BC, is not altered by either rule.

Vacuous ambiguity is undesirable, for in an analysis which employed such a pair of rules in this way, the representation should reflect the intuition that, since the different order of application of rules yields the same result, the representations of the two derivations ought to be non-distinct. As with phrase-markers for context-free phrase-structure grammars, if the order of application is immaterial to the yield of the grammar, there should be a single phrase-marker.

If the rules of the grammar are translated into 'phonology' format, that is, as rewriting a single symbol in a particular environment, then a different kind of derivation-graph can be given which does not suffer from this vacuous ambiguity. (Manaster-Ramer and Kac (1990: 351) call context-
sensitive (CS) grammars in phonology format and CF grammars unisinistral.) The above derivation can be graphed:

\[(3.38) \quad A \leftarrow / B C \rightarrow E \]

\[D \quad F\]

where a downward arc denotes 'derives', \(\leftarrow\) denotes 'in the context of' and \(\rightarrow\) (resp. \(\leftarrow\)) denotes 'before' (resp. 'after') \(X\).

This representation is notationally equivalent to the canonical context-sensitive derivation-graphs discussed by Hart (1980) and by Hart and Révész in Révész (1983 ch. 10). These graphs, like the phrase-markers of context-free languages, are trees: since context-sensitive phrase-structure grammars in 'phonology' format are unisinistral, the effect of the 'rewrite' part of each rule is represented with a single mother node representing the single category of the left-hand side and one or more daughters representing the one or more categories of the right-hand side. A second kind of arc labelled \(<_c>\) for context-dependency (like the \(\leftarrow\) and \(\rightarrow\) arcs used above) show the dependency between the mother node and the left and right environments. Finally, a third kind of arc labelled \(<_f>\) for free is drawn between the leftmost daughter and the left-hand environment and between the rightmost daughter and the right-hand environment to show that the particular rule application is completed, and that the daughters and the environment are free to form part of the input to another rule.

3.6.3 *General derivation graphs as a foundation for declarative grammars*

There are at least two adequate methods for constructing derivation-graphs of non-context-free grammars, then: a general method, and a method specific to unisinistral context-sensitive phrase-structure grammars. Both methods are similar to the tree-graph phrase-markers used to show derivations in context-free grammars, and the usual context-free case is a subcase of both of the methods discussed above, which gives them considerable intuitive appeal.

Just as context-free derivation-graphs provide a foundation for the declarative, non-derivational interpretation of context-free phrase-structure grammars, the following important result is now transparent:

*Theorem 2*

Any rewriting system, even unrestricted, monotonic and context-sensitive grammars in 'phonology' format, can be given a declarative interpretation.
Sketch of proof: The declarative interpretation for the set of rewrite rules employed in an analysis is the join of the representations of each rule into a single well-formed derivation-graph. How the rules are actually put together – how they feed or bleed each other, and how they 'top out' at the root(s) and 'bottom out' at the frontier – does not need to be specified in cases of unambiguous yields.

3.6.4 Segments, features and rule application

The discussion above assumed without comment that the generative grammars under consideration employ simple atomic categories such as a, b, c, [b], [A], /η/, /al/. In SPE phonological rules, however, the categories are binary-encoded complex symbols such as:

\[
\begin{array}{c}
+\text{consonantal} \\
-\text{vocalic} \\
-\text{nasal} \\
+\text{voice} \\
-\text{sonorant} \\
-\text{continuant} \\
+\text{grave} \\
+\text{compact}
\end{array}
\]

If the simple atomic symbols of standard phrase-structure grammars (such as those of the sort defined in table 3.3) were simply replaced by their equivalent encodings in a binary feature-system, the grammars will not be materially changed, and will still define exactly the same classes of languages. Halle (1969) and Gazdar et al. (1985: 20) point out that such grammars are standard phrase-structure grammars which happen to employ calligraphically ornate symbols.

Grammars which employ such symbols do so in order that their rules may apply to a class of symbols with some features in common, and not just to a single specific symbol. This makes each rule of a grammar which uses complex symbols correspond to a set of rules using atomic symbols. For example, a rule which uses feature-based symbols to express the fact that all voiceless stops are aspirated in a language with four places of articulation in the stop system will require a single feature-based rule such as

\[(3.39) \quad \begin{array}{c}
-\text{sonorant} \\
-\text{continuant} \\
-\text{voice}
\end{array} \rightarrow [+\text{aspirated}]
\]

as opposed to four rules employing atomic symbols, e.g.
Segmental and transformational phonology

\[
\begin{align*}
p & \rightarrow p^h \\
t & \rightarrow t^h \\
c & \rightarrow c^h \\
k & \rightarrow k^h
\end{align*}
\]

The feature-based rule expresses what is common to the four rules of the atomic-symbol based grammar. To the extent that these four rules are disjoint, the feature-based grammar eliminates a disjunction from the theoretical machinery of the grammar, a move which makes the grammar more explanatory, since disjunction is an 'anti-generalisation' mechanism (see Manaster-Ramer and Zadrozny 1990: 196). In addition, the feature-based rule would apply equally well in a grammar with any number of distinctive places of articulation, or with four different places of articulation (e.g. p, t, k, q), and therefore captures the independence of the voicelessness/aspiration generalisation from other language-specific statements.

'Calligraphic ornateness' in the set of symbols would be pointless unless it were put to some use. In order to realise the potential descriptive utility of complex symbols to the full, it is necessary to modify the criteria for rule applicability. In a grammar which uses only atomic symbols, a rule of the form \( X \rightarrow Y \) can only be applied to a string of the form \((p)Z(q)\) if \(X = Z\); in other words, a rule can be applied to a string if the rule's left-hand side is a substring of the string to which the rule is applied. The symbols used in representations are drawn from the same stock as the symbols used in rules. With feature-based grammars, however, the situation is somewhat different: the symbols used in rules are typically less specific than the symbols in the strings to which they are applied. Thus the criterion which must be satisfied for a rule to be applicable must be altered: employing the above notation, it is not necessary to require that \(X = Y\) – that would not produce the desired outcome. For example, the left-hand side of rule (3.39) might not be equal to the representation of the segments to which it applies, which contain additional features, and are thus more specific. All that is required in a feature-based grammar is that the left-hand side of the rule matches the string to which it is applied, without conflict in the value of any feature.

Since the set of pairs of strings in the 'equals' relation is a proper subset of the set of pairs of strings in the 'matches without conflict' relation, it might initially be supposed that feature-based grammars are less restrictive than the corresponding classes of grammars using only atomic symbols. This is not the case, however. It is trivial to show that every grammar employing atomic symbols is strongly equivalent to a grammar using complex symbols. Furthermore, by mapping the set of all possible complex
3.6 Representations of derivations

symbols one-to-one onto a (possibly quite large) set of atomic symbols, it
is possible to construct a strongly equivalent atomic grammar for every
feature-based grammar. Though each rule in the feature-based grammar
may correspond to several rules in the atomic grammar, the set of distinct
categories and the set of structural descriptions defined by each grammar
are isomorphic, and thus the grammars are strongly equivalent. Conse-sequently, relaxing the 'equals' criteria for rule applicability to
'matches without conflict' does not formally weaken the grammar formal-
ism at all, but serves only to allow the set of rules to be abbreviated and
additional cross-categorial generalisations to be captured.

However, I have not yet precisely defined 'matches without conflict'. This
is an important step since there are at least two natural relations between
categories which would work, but with different consequences for rule
applicability. If complex symbols are considered as sets of features, the
subset relation is one possibility. In other words, the complex symbol F
could be said to match the complex symbol G without conflict if every
feature f in F was also in G, which is the same as saying F ⊆ G. The left-
hand side of the rule discussed above matches the segments to which it
applies by this definition. But suppose the left-hand side of the rule con-
tained an additional feature, say [+round], which does not occur with any
value at all in any lexical specification (i.e. it is a redundant feature). It is
clear that the revised left-hand side still does not conflict with the segments
to which it applies, since none of them contains [-round], yet nevertheless
[+ round] is not a member of the complex symbols to which the rule applies,
and thus the rule would now fail to match those symbols if the 'subset'
definition of 'matches without conflict' is employed. The weaker version of
'matches without conflict' is the 'unifies with' relation mentioned in chapter
2. The unification operation might be preferred to the subset relation on the
grounds that it is independently usable as the operation by which features
are combined with each other to form complex symbols. I shall conclude
this section by revising the standard definition of (atomic) phrase-structure
grammars to allow for grammars with complex symbols. In formal lan-
guage theory (see Brookshear 1989: 51), a phrase-structure grammar is a
quadruple \((V_N, V_T, S, R)\), where \(V_N\) is the set of (atomic) non-terminal
symbols, \(V_T\) is the set of (atomic) terminal symbols, \(S\) is the (atomic) dis-
tinguished (root) symbol and \(R\) (the rules) a subset of \(V^*_N \times (V_N \cup V_T)^m\),
where the values of \(n\) and \(m\) determine the type of grammar, as summar-
ised in table 3.3 above. (Note that this definition does not require \(V_N\) and
\(V_T\) to be disjoint.) A grammar with complex symbols has four additional
ingredients: (i) a set of feature-names \( F \) and (ii) a set of feature-values \( V \), from which the complex symbols will be constructed by joining feature-value pairs together using (iii) a category combinator \( \otimes \). Finally, as discussed above, a grammar which employs complex symbols must have a well-defined relation \( M \) which defines which strings match which. A feature-based phrase-structure grammar is thus an octuple \((F, V, \otimes, V_N, V_T, S, R, M)\), where \( V_N \) and \( V_T \) are drawn from the set of complex symbols \( C \) which is defined recursively as follows:

1. Every pair \((F, V)\) (conventionally written \([VF]\)) is in \( C \).
2. If \( A \) and \( B \) are in \( C \), then \( A \otimes B \) is in \( C \).
3. Nothing else is in \( C \).

(I shall ignore the question of finite closure of the set of categories which arises if members of \( C \) may also be members of \( V \), so that category-valued features are introduced. Refer to, for example, Gazdar et al. (1985: 35–40) for fuller consideration of this matter.) A system in which \( \otimes \) is cocatenation of features (cf. chapter 2) will allow categories to be generated with more than one value for a given feature. Such a possibility has occasionally been considered, either favourably or unfavourably (see Wang 1968: 704, Goldsmith 1976: 21, Hoard 1978: 70), usually concerning the representation of contour segments. More usually, however, \( \otimes \) is unification, which requires that features with opposite or more generally different values may not be combined. A system in which \( M \) also is taken to be unification, rather than equality or the subset relation, is thus a rather simpler structure \((F, V, \sqcup, V_N, V_T, S, R)\). Grammars of this form may be called unification-based phrase structure grammars, or UPSG's for short, with the usual finite-state, context-free, context-sensitive and unrestricted subclasses defined by restrictions on the form of the rules.

### 3.7 Conclusion

In this chapter I described the logical development of segmental transformational grammar from the theory of distinctive features in phonemic phonology and the desire to provide a formally explicit account of the lexical representation of words and their relation to the observable 'surface' form of words. I demonstrated that some of the technical problems which beset the transformational theory are deeply rooted consequences of the excessive power of the \( SPE \) rule formalism. I argued that the constraints which some phonologists have proposed in an attempt to rectify this exces-
3.7 Conclusion

Sive power do not constrain the formalism as desired. I presented a method for the direct, declarative characterisation of linguistic structures which facilitates the development of a declarative approach to phrase-structure grammar. Finally, I argued that the employment of 'complex symbols' in SPE phonology produces a revised hierarchy of phrase-structure grammars called unification-based phrase-structure grammars. While these are no more powerful in weak generative capacity than classical phrase-structure grammars, they permit cross-categorial generalisations and much more parsimonious rule schemata to be defined. In the next chapters I shall show that the greater ability of UPSGs (compared with classical PSGs) to express linguistic generalisations, both paradigmatic and syntagmatic, makes them a more restricted alternative to the transformational SPE formalism.
Non-linear phonological representations in contemporary generative phonology

In a little paper at the end of 1973 called ‘Tonemic Structure’, I suggested a notation for falling toned vowels:

V
  \___/ H L

The observation was made then – by Paul Kiparsky, I think – that this would account for what, in this thesis, is called ‘stability’ of tone melodies. As a notation, this was simple enough; as a theory of anything, it was quite inadequate.

(Goldsmith 1976: 1)

4.1 Introduction

Two recent phonological theories may, in some combination, offer the sort of firm proposals regarding the form of phonological representations that earlier work on suprasegmental and subsegmental phonology lacked. Autosegmental Phonology can be viewed as a generative theory of the distribution of prosodic units with respect to segmental units. Metrical Phonology offers a formal account of the phonological constituent structures and hierarchy of domains in which prosodic units may be located. I begin by examining Autosegmental Phonology, showing in sections 4.3.4–4.3.10 a foundational problem in the Autosegmental formalism. After briefly examining the tentative proposals which have been made regarding the interaction of Metrical and Autosegmental structures in phonological representations, I shall show how Unification-based Phrase-Structure Grammars allow Metrical structures and Autosegmental units to be closely integrated in a declarative, generative version of Prosodic Phonology.
4.2 Origins of non-linear generative phonology

Despite certain similarities to Prosodic Phonology discussed in Broe (1988), Coleman (1991), Goldsmith (1992) and Ogden and Local (1994), Autosegmental Phonology developed out of segmental generative phonology without extensive historical links with Firthian Phonology. If a historical basis is sought, it should be in two earlier American linguistic theories, 'long-component' analysis (Harris 1944) and phrase-structure grammar. (See, for example, phrase-markers such as example (59) in Chomsky (1965: 108-9), in which both one-to-many constituency relations and many-to-one 'are properties of' relations are represented by 'association' lines, in the sense of Hockett (1966: 164).) It is interesting to wonder whether Autosegmental Phonology's historical relationship to Hockett's phonological work on non-linear phonological representation (a term first proposed in Hockett 1947a; cf. Goldsmith 1990: 3-4, 8) ever brought either of the following two quotations to Goldsmith's attention:

An association from A to B assigns, to each element of A, at least one element of B. Here are two associations:

\[
A = \{a_1, a_2, a_3, a_4\} \quad A = \{a_1, a_2, a_3, a_4\}
\]

\[
B = \{b_1, b_2, b_3, b_4\} \quad B = \{b_1, b_2, b_3, b_4, b_5\}
\]

D1

\[
D_3
\]

(Hockett 1966: 164)

(This is not a phonological example: Hockett is talking about mappings between any two sets.)

Abstractly, a phonon stepmatrix \( s = |K_1 | \ m = |K_1 K_2 \ldots K_m | \) has nothing to do with time. But we introduce, first, quantized pseudo time \( \tau \) by specifying that each character \( K_i \) 'terminates' at time \( \tau_i \), as suggested by Figure 4.

\[
| K_1 K_2 K_3 \ldots K_m |
\]

\[
\tau_0 \quad \tau_1 \quad \tau_2 \quad \tau_3 \quad \tau_m
\]

Figure 4

(Hockett 1966:234)
Goldsmith (personal communication) thinks that he might have come across Hockett's terminology while working on the development of Autosegmental Phonology.

The use of phrase-markers to show the assignment of units at one level of phonological description to units at another level of description had been used in some phonological analyses prior to the inception of Autosegmental Phonology (see Cheng 1966: 146, Fudge 1969). In its emphasis on phonological representations, with much less emphasis on the rule-systems used to define those representations than in SPE phonology, Autosegmental Phonology fulfils Chomsky's repeated injunctions that the proper aim of linguistic theory is to elucidate the nature of the structural descriptions defined by a grammar (i.e. its strong generative capacity) rather than simply sets of strings (i.e. weak generative capacity).

In a similar way, the arboreal Metrical theory (one of two approaches to phonological representation proposed in Metrical Phonology) developed out of the representation of syllable structure using context-free phrase-markers. (The term 'arboreal Metrical theory' for tree-based, as opposed to grid-based, Metrical Phonology, is due to Poser 1986: 181.) Metrical Phonology is greatly indebted to Anderson and Jones's (1974) paper on phonological representations in Dependency Phonology, since the class of graphical techniques which Anderson and Jones proposed for the representation of within-syllable structure and between-syllable prominence relations (stress) are notationally equivalent to those employed in the arboreal Metrical theory. Anderson and Jones (1974) is cited by Kahn (1976: 19–20).

Autosegmental Phonology and Metrical Phonology share a common ancestry in Leben's (1973) thesis Suprasegmental Phonology. Goldsmith (1976: 1) records that Leben's argument that 'in some languages, even short vowels could bear two successive tones' could be expressed using the notation quoted in the epigraph above. Goldsmith goes on to say:

In a paper on tone in Sanskrit by Robert May and myself, I suggested that the syllable might well be considered an autosegmental level. In this sense, the string of familiar C and V segments could be broken up into an autosegmental representation where the second tier was composed of syllables:

\[ C \rightarrow V \rightarrow C \rightarrow V \rightarrow C \rightarrow V \rightarrow C \]

\[ \Sigma_1 \rightarrow \Sigma_2 \rightarrow \Sigma_3 \rightarrow \Sigma_4 \]
... this approach would remove the glaring restriction in the theory of autosegmental tonology that says that only syllabic segments can associate with tonemes. Given syllabic structures like (i), CVCV would become (iii).

\[
\begin{align*}
\text{(a)} & \quad \text{C} \atop \text{V} \atop \text{C} \atop \text{V} \\
\text{(iii)} & \quad \Sigma_1 \atop \Sigma_2 \\
& \quad H \quad L
\end{align*}
\]

While Goldsmith went on to concentrate on the application of the Autosegmental theory to tonal phenomena in particular, others pursued the Autosegmental theory in other directions. Clements (1976) examined its applicability to the representation of vowel harmony, and McCarthy (1979/1982) investigated its applicability to non-concatenative morphology. The Autosegmental treatment of syllable structure proposed by May and Goldsmith was pursued further by Kahn (1976), which laid an important foundation for Liberman and Prince's (1977) exposition of Metrical Phonology. The Autosegmental theory of syllable structure was later consolidated by Clements and Keyser (1983).

4.3 Non-linear phonological representation

Autosegmental Phonology, Dependency Phonology and Metrical Phonology employ graphical and other diagrammatic forms to represent various aspects of phonological structure. Such diagrams are presented as having explanatory import, rather than being purely illustrative. For instance, Goldsmith (1976: 16) states that 'Autosegmental phonology is a particular claim . . . about the geometry of phonetic representations.'

The incorporation of a graphical component in phonology may be likened to the adoption of a phrase-structure component in syntactic theory, since the graphs encode information about phonological constituency and other relations. This kind of development in phonological theory raises new questions about the nature, variety and complexity of possible graphs in phonological theory, a task which has been pursued by phonologists over the last decade and a half, resulting in a number of important proposals concerning phonological representations in non-linear theory. In the following sections I shall examine some of the claims that have been made about phonological graphs, and I shall attempt to
explicate some of the properties that characterise Autosegmental, Metrical and Dependency representations.

In the previous chapter I showed that the generalisation from strings of atomic symbols to strings of complex symbols does not affect the weak generative capacity of a grammar. This is not, however, the case with the generalisation of string grammars to graph grammars of the type exemplified by multilinear Autosegmental Phonology, for in addition to rules that change elements in some way, as in string grammars, there may also be rules that change the relations holding between elements. For example, one of the operations permitted in Autosegmental Phonology, 'delinking', is the deletion of an association line. If one of the units incident to such a line is linked to the rest of the representation by that line alone, delinking has the effect of separating that unit from the rest of the representation. Units which are separated from the representation are said to be 'floating'. A unit which remains floating in systematic phonetic representations is said to be 'stray'. Stray units are not regarded as part of the representation for purposes of phonetic interpretation. Consequently, delinking of a unit can have an effect on the result of phonological derivations equivalent to deletion of that unit. But delinking is not the same as deletion, which is another operation. 'Delinking' is a kind of 'temporary deletion' operation without any correlate in segmental-transformational phonology, since delinked units are not removed from the phonological representation, and may be relinked with the rest of the representation at a later point in the derivation.

4.3.1 An autosegmental analysis
To illustrate the kind of analyses which Autosegmental Phonology makes possible, I shall show how the proposals made by Autosegmental Phonologists can be applied to the case of palatalisation and affrication of Japanese coronal consonants which were shown in the previous chapter to be a little problematic for segmental-transformational phonological theory. If the problems of that SPE-type analysis can be attributed to the strict segmentality of its surface phonological representations, an Autosegmental analysis might be more satisfactory than a transformational analysis. For instance, the canonical Autosegmental treatment of affrication (Clements and Keyser 1983) unites the sequence of features [-continuant][+continuant] under a single C node, which itself bears a single matrix of features that encapsulates the homorganic articulation of
an affricate, [ts], for example, can be represented in multilinear fashion along several simultaneous tiers (Prince 1984) as:

\[
\begin{array}{c}
\text{–voice} \\
\text{–sonorant} \\
\text{+coronal} \\
\hline
\text{C} \\
\text{[–continuant]} \\
\text{[+continuant]}
\end{array}
\]

The values of the feature [continuant] are written on the tier below the C node, as they express subsegmental information. Thus the SPE feature [delayed release] is obviated, and affricates are treated as paradigmatically unitary, but syntagmatically binary. The phonetic parallels between affricates and stops with nasal or lateral release is reflected in representations such as (4.2a–c), though the functional, phonological relevance of these parallels must still be established on phonological grounds.

\[
\begin{array}{c}
\text{–sonorant} \\
\text{[–cont]} \\
\text{[+cont]}
\end{array}
\]

\[
\begin{array}{c}
\text{–sonorant} \\
\text{[–cont]} \\
\text{[+cont]}
\end{array}
\]

\[
\begin{array}{c}
\text{–sonorant} \\
\text{[–cont]} \\
\text{[+cont]}
\end{array}
\]

Affrication Nasal plosion Lateral release

The transformational affrication rule (3.19), translated into Autosegmental terms, is:

\[
\begin{array}{c}
\text{–sonorant} \\
\text{[–cont]} \\
\text{[+cont]} \\
\text{[–son]} \\
\text{[+son]} \\
\text{[+nasal]}
\end{array}
\]

\[
\begin{array}{c}
\text{–sonorant} \\
\text{[–cont]} \\
\text{[+cont]} \\
\text{[–son]} \\
\text{[+son]} \\
\text{[+lateral]}
\end{array}
\]

Example (4.4) expresses the regressive association of palatality with consonants:
Rule (4.4) may apply over a sequence of consonants, just like the segmental rule (3.23). This example shows that Autosegmental Phonology allows more solutions than transformational phonology in cases of assimilation (but see Anderson 1982). Because it is constructed on segmental and transformational foundations, it may use copying rules (e.g. 3.23), or, alternatively, its own device, association rules (e.g. 4.4). Rule (4.4) differs from (3.23) in employing autosegmental association, rather than copying. In uniplanar representation, reassociation is more restrictive than copying, since it falls under the strictures of the Well-formedness Condition (see below for further discussion), and is consequently always locally bounded, whereas copying may be completely unbounded.

The spreading mechanism may sometimes be problematic because of its unbounded nature. For example, in the case of the Japanese 'vowel elision' phenomenon reported by Ueda that results in such forms as [kisj], the leftward spreading of palatalisation must be blocked so that [kisj] does not derive [kiJJ]. In order to do this, the following structural restriction must be formally instantiated:

Palatality may extend as far back as the final consonant of a preceding syllable, but not so far as that syllable's initial consonant.

(In order to appreciate this restriction, it should be noted that [s] is syllable-initial in [kisj], which is phonologically /kisusi/.) An Autosegmental analysis might propose that the constituents of a / Ci/ mora can be represented by two syllable terminals, labelled C and V. In [tsi], the C terminal is lexically associated with the segmental matrix for /t/, and the V terminal with the matrix [+high, -back]. (As in chapter 3, I assume that the [-low] specification of close vowels may be derived by default from [+high].)

This does not settle whether C and V units or their features are to be represented on one tier, as in the quotation from Goldsmith (1976: 1) above, or on two independent tiers. I shall consider each of these possibilities in turn. With just two tiers – the CV tier and a segmental tier – the derivation could proceed as in (4.5a–d).
At this stage, in order to represent compensatory lengthening of the [s] which accompanies devocalisation (see chapter 3), [s] must become associated to the floating V ('syllabic') node. To achieve this, the association line linking [+high, -back] with the C node must be 'swung out of the way' to ensure that it does not cross the association line between [s] and V, which would be a violation of the No Crossing Constraint. The position of the V features relative to the C features in the representation is unimportant, as long as they are associated with each other via the CV tier.

Consider now how the derivation proceeds if the vocalic and consonantal features are set apart, on independent tiers. In (4.7a–d) I have written the V features above the CV tier, and the C features below the CV tier. Since 'consonant' and 'vowel' features are not restricted to C and V slots respectively, and since they are not necessarily intercalated, but may occur with both C and V units, I term 'consonant' and 'vowel' features stricture and
resonance features, accordingly. The form of (4.3) will consequently be amended to (4.6) to accord with this decision.

(4.6) Affrication 5

$$\begin{align*}
\text{Resonance tier} & \quad [+\text{high}]  \\
\text{CV tier} & \quad C \rightarrow \frac{C}{V}  \\
\text{Stricture tier} & \quad [-\text{cont}] [-\text{cont}] [-\text{cont}] \quad [-\text{sonorant}]  \\
& \quad [+\text{coronal}] 
\end{align*}$$

(4.7) a. \quad \begin{array}{c}
[+\text{high}]  \\
[-\text{back}]  \\
C \quad V  \\
\downarrow  \\
t
\end{array}
\quad \text{Palatalisation \downarrow}

b. \quad \begin{array}{c}
[+\text{high}]  \\
[-\text{back}]  \\
C \quad V  \\
\downarrow  \\
t
\end{array}
\quad \text{Affrication \downarrow}

c. \quad \begin{array}{c}
[+\text{high}]  \\
[-\text{back}]  \\
C \quad V  \\
\downarrow  \\
\quad \text{Devocalisation and compensatory lengthening \downarrow}
\end{array}

D. \quad \begin{array}{c}
[+\text{high}]  \\
[-\text{back}]  \\
C \quad V  \\
\downarrow  \\
\quad \text{t s}
\end{array}

Analysis (4.7) is more satisfactory than analysis (4.5). In the multiplanar mode of representation, stricture (consonantal) and resonance (vocalic) features are represented on separate tiers: this allows such phenomena as 'vocalic colouring' of consonants, consonantal syllabics, and vowel-frica-
4.3 Non-linear representation

tive alternation to be characterised with ease. The motivation for this decision was, however, to avoid violating the restriction that association lines may not cross. But if such a restriction is to carry any force, it should not be possible to subvert it by transferring any subgraphs whose association lines are likely to cross onto separate tiers or planes.

Other problems can be avoided if statements of sequence are kept quite separate from statements of association of items in separate tiers. We have seen, for instance, that Japanese has CV, but not VC moras. In lexical representations, constant duplication of the information that C precedes V in CV moras is highly costly, since if a C and V are parts of one mora, it is completely predictable that their relative surface order is 'C first, V second', as there are no VC moras. Where vowel-consonant sequences do occur, there is always a mora boundary between the vowel and consonant, as that consonant either begins the next mora or is a complete mora itself. The relative order of consonants and vowels within a mora is thus totally predictable, as is the location of mora boundaries. Consequently, the most parsimonious analysis of Japanese is one in which consonants and vowels are grouped into moras, syllables etc., but not explicitly ordered.

The informational content of linear precedence in phonological representations has to my knowledge never been discussed in Autosegmental Phonology. It is just assumed that linear precedence in phonological or phonetic structure is represented by the order of printed items at no cost (see Cheng 1971).

4.3.2 Well-formedness in Autosegmental Phonology

Autosegmental Phonology, then, is a phonological theory which employs multitiered representations rather than strings as its central data structure (Goldsmith 1976, 1990; van der Hulst and Smith 1982a). Phonological processes such as assimilation, harmony etc. are given derivational accounts similar to those employed in string-based approaches to phonology. But in line with the trend towards declarative formalisms in linguistic theory, Autosegmental Phonological representations and derivations are sanctioned not by extrinsically ordered grammar rules, but by general 'principles' tempered by 'constraints', together with some language-specific rules. Well-formedness of an Autosegmental Phonological representation and its derivation are assessed by their adherence to and satisfaction of these 'principles', 'constraints' and rules. In the remainder of this section I shall consider the nature of principles, constraints and rules in Autosegmental Phonology in detail.
4.3.3 Well-formedness constraints

The set of well-formed phonological representations in Autosegmental representation is defined by

1 the initial associations of a representation, which are specified in the lexicon and are thus language-specific;
2 the universal association conventions, which define how associations not defined in the lexicon should be filled in;
3 language-specific rules, which may alter the pattern of associations defined in the lexicon or by the association conventions (though the association conventions may reapply after a rule has applied);
4 universal well-formedness constraints on the form of representations; in particular the No Crossing Constraint and the Obligatory Contour Principle.

The initial associations are only constrained by the universal well-formedness constraints. Since these are examined in detail below, I shall not discuss initial associations further.

Rules

Despite the 'universal principles'-based orientation of Autosegmental Phonology, a reading of multi-language studies such as Goldsmith (1976, 1990) and Clements and Keyser (1983) shows Autosegmental Phonology to depend more on language-specific rewriting rules than on general principles. After more than ten years of Autosegmental Phonology, a set of parameters and their settings such as those proposed by Babić (1988: 9) (reproduced below) for the varieties of accent-systems in Croatian dialects shows little progress towards a universal account of tone parameters, when compared with Zubizaretta (1982: 198) or Clements and Sezer (1982) (mentioned in chapter 1).

Croatian accent systems appear to vary in terms of the following major parameters:

(A) number of tonal classes:
1 one (H-tone)
2 two (H-tone/L-tone)

(B) distribution of H-tones
1 per mora
2 per syllable
4.3 Non-linear representation

(C) contrast between H-tone and L-tone words
1 on monomoraic and bimoraic syllables
2 on bimoraic syllables only
3 on monomoraic syllables only
4 on neither (= A1)

(D) restrictions on distribution of tones on syllables
1 L only on monomoraic syllables
2 L only on bimoraic syllables
3 L unrestricted

(E) distribution of accent
1 not final syllable
2 free
3 double

(F) distribution of long syllables
1 free
2 not preceding accent
3 not following accent
4 only accented syllable

Like the statements of Zubizaretta or Clements and Sezer quoted in section 1.2, when this statement is compared with the highly explanatory set of metrical parameters to be discussed in the next section, it is clear that such lists of putative parameters do not constitute a highly restrictive, language-universal set of autosegmental parameters so much as an ad hoc descriptive tabulation of the data. Although such tabulations are relevant to the elucidation of underlying principles, they should not be mistaken for such principles. The form of language-specific autosegmental rules has not been defined by Autosegmental Phonologists other than by exemplification. A survey of the examples in the literature reveals that autosegmental rules are a graphical extension of the SPE-type context-sensitive string-rewriting rules, modified so that as well as inserting, changing or deleting units (segments, features and autosegments), autosegmental rules may also insert or delete association lines. I established in chapter 3 that the deletion ability of SPE rules renders them excessively powerful. A corresponding result for grammars which map autosegmental graphs onto autosegmental graphs has not yet been proved. Consequently, although it cannot be definitively said that (node and line) deletion rules in autosegmental graph grammars make them equivalent to arbitrary graph rewriting systems, it should be noted that in the ‘Chomsky-like hierarchy’ of graph grammars presented in Nagl (1979), no family of non-monotonic graph grammars (i.e. those that
permit node deletion rules) less powerful than unrestricted graph grammars (which define all recursively enumerable sets of graphs) has yet been identified. Unless and until a type of graph grammar intermediate between unrestricted and monotonic (context-sensitive) is identified, caution dictates that Autosegmental deletion rules too should be regarded with suspicion.

Association conventions
Besides putative autosegmental parameters and language-specific rules, which I have argued are currently essentially unconstrained, Autosegmental theory proposes a small number of universal well-formedness constraints. The first example of these, in Goldsmith (1976: 24) was the following:

Well-formedness Condition (initial statement)
1 All vowels are associated with at least one tone; all tones are associated with at least one vowel.
2 Association lines do not cross.

The two clauses of this condition are now also known as the Association Conventions and the No Crossing Constraint respectively. In this section I shall examine the force of the Association Conventions. I shall defer discussion of the No Crossing Constraint until after I have considered the other constraints.

In their first formulation (and in many subsequent formulations) the Association Conventions are specific to the associations between tones and tone-bearing units. They are an attempt to encode the default spreading or ‘filling-in’ of lexically unspecified tonal associations. Goldsmith (1976: 130) noted that the fact that ‘the mapping is left-to-right, for one thing, follows from nothing we have seen so far. Yet all the examples we have seen, or will see, of non-accentual systems bear this property.’ The Association Conventions were subsequently modified to include this ‘fact’. For example, Clements (1985: 74) formulates them as follows (an equivalent definition is given in Clements and Keyser 1983: 63–4):

(24) Association Conventions
(i) Given a continuous string S consisting of one or more free P-segments and an open string T occurring in its domain, associate (free) P-segments in S to (free) P-bearing units in T in a one-to-one manner from left to right (such association is nonoverlapping, maximal, and exhaustive in the sense of Clements and Ford (1979));
(ii) Given an open string T remaining after the operation of (24i), associ-
ate each (free) P-bearing unit in T with the P-segment in whose domain it falls (giving precedence to the P-segment associated with a P-bearing unit occurring to the left of T).

This procedural mechanism encodes the observation that once lexically specific autosegmental associations are taken into account, the mapping from autosegments to their foci (to borrow a term from Firthian Prosodic Phonology) is one-to-one, except at the right-hand edge of the string, where the final autosegment associates with all remaining foci. It is not usually clear what linguistic unit the string must be (whether words, morpheme or whatever). According to Clements (1985: 73), the set of foci or P-bearing units for each autosegment is defined on a language-specific basis, by parameter-setting. Goldsmith's emphasis on left-to-right association follows the tradition employed for simplifying tone transcriptions in many dictionaries and grammars of African languages, following the orthographic recommendations of Westermann and Ward (1933: 145). Autosegmental Phonology thus rests on the hypothesis that apart from lexically determined associations, rule-determined associations or associations at the right edge of words (or strings of some unspecified size), autosegmental-to-segmental association is one-to-one. Except for a certain number of quite specific cases, therefore, Autosegmental Phonology is actually no different from the segmental treatment of suprasegmental units. I shall examine the specific cases of 'autosegmental' representation that are not one-to-one in the next subsection.

In Goldsmith (1990: 14) the Association Convention is changed yet again:

Association Convention
When unassociated vowels and tones appear on the same side of an association line, they will be automatically associated in a one-to-one fashion, radiating outward from the association line.

Revision of the principles of a theory in the light of subsequent research is appropriate in any theory, within reasonable limits. If the principles of the theory have to be changed so often that the only constancy is in the names given to the principles, it becomes reasonable to doubt the credibility of those aspects of the theory. Whether the revisions of the Association Conventions are granted or not, the latest formulation, quoted above, is more inadequate than the old left-to-right version. Consider, for instance, the effect of bidirectional radiation of association on the following initial representation:
When the Association Convention applies, there will be conflict between \( P_1 \) and \( P_2 \) over \( X_3 \) and \( X_4 \):}

\[
\begin{array}{c}
X_1 \ X_2 \ X_3 \ X_4 \ X_5 \ X_6 \\
\end{array}
\]

The Association Convention does not specify whether the spread of association lines from \( P_1 \) and \( P_2 \) proceeds ‘equitably’, so that \( X_3 \) is allotted to \( P_1 \) and \( X_4 \) is allotted to \( P_2 \), or ‘competitively’, so that both \( X_3 \) and \( X_4 \) are allotted to either \( P_1 \) or \( P_2 \). The left-to-right version of the conventions would give a different result, for example, in which both \( X_3 \) and \( X_4 \) are allotted to \( P_1 \). The most charitable critical position that can be adopted is that Autosegmental Phonology is simply unspecific about non-lexical or non-rule-governed association.

Supra- and subsegmental associations
The fact that the association relation is undirected means that there is effectively no formal distinction between many-to-one and one-to-many associations. In this section I shall argue that there is a linguistically significant difference, though, which can be summarised as follows. If we take ‘\( X \)-to-Y’ to mean ‘\( X \) autosegments to \( Y \) segments’, then ‘one-to-many’ mappings are those where a single suprasegmental autosegment is associated to two or more segments, and ‘many-to-one’ mappings are those where two or more autosegments are associated with one segment. I shall argue, however, that, since ‘many-to-one’ is never more than ‘two-to-one’, there is a formal distinction between ‘many-to-one’ and ‘one-to-many’ associations that is uncaptured by Autosegmental Phonology.

There are two cases where two-to-one associations are employed in Autosegmental analyses. The first case is in describing subsegmental structure, as in prenasalisation, nasal or lateral release, affrication, pre- or post-aspiration, short diphthongs etc. It is conceivable that subsegmental structure might involve a greater than two-to-one association: for example, stops might be analysed into a sequence of closing, hold and release subsegments. But in any such cases, a finite bound can still be placed on subsegmental association, unlike suprasegmental associations, whose strength is that they can spread to fill a domain. The second use is when two
suprasegmental autosegments 'share' a segment at their juncture. Both of these cases are described equally well by a unification-based phrase-structural description, which has the two added advantages of capturing the boundedness of subsegmental structure, and integrating with phrase-structure of other kinds, such as syllabic and metrical structure. Subsegmental structure can be described using context-free phrase structure rules of the form \( s \rightarrow a_1 \, a_2 \ldots \, a_n \), where \( s \) is a segmental category, and \( a_m \) are subsegmental categories. Two-to-one association at prosodic junctures, by contrast, cannot be analysed by context-free rules alone, but unification must be employed. (A fuller account of this analysis of prosodic juncture is given in section 5.5.8.)

The configuration in question is portrayed in (4.10). The associations of interest are

\[
\begin{align*}
\text{(4.10)} & \\
& a_1 \quad \text{and} \quad a_2 \\
& s_3 \quad s_3
\end{align*}
\]

Because unification is idempotent (i.e. \( x \cup x = x \)), we can 'split' any unit into two copies of itself, providing we can express the fact that the two copies are token-identical. For example, if we split the segment \( s_3 \) into two copies, \( s_{3a} \) and \( s_{3b} \), (4.10) can be reexpressed as (4.11):

\[
\begin{align*}
\text{(4.11)} & \\
& \ldots \quad a_1 \quad \ldots \\
& s_1 \quad s_2 \quad s_3 = s_{3b} \quad s_4 \quad s_5 \ldots
\end{align*}
\]

The structure in (4.11) can be described by the two phrase-structure rules \( a_1 \rightarrow \ldots \, s_1 \, s_2 \, s_{3a} \) and \( a_2 \rightarrow s_{3b} \, s_4 \, s_5 \ldots \), together with a mechanism to ensure that \( s_{3a} \) and \( s_{3b} \) are equated by unification. The nature of such a mechanism is described in section 5.5.8.

Even apparently unbounded one-to-many prosodic associations can be defined using the resources of phrase-structure grammar, rather than an independent 'association component' specific to autosegmental phenomena. (I take it as uncontroversial that bounded one-to-X prosodic associations where \( X \) is a finite and invariant bound can be defined using phrase-structure rules.) Firstly, consider the one-directional left-to-right spreading pattern in (4.12):
In a phrase-structure based theory, this pattern of spreading can be expressed using the right-branching linear structure in (4.13):

\[ (4.13) \]

Unidirectional right-to-left spreading can be represented in a similar fashion using a left-branching linear structure. In (4.13), the topmost a is a '(prosodic) autosegment', \( s_1 \) is its 'focus' and \( s_2 \) to \( s_n \) is the sequence of a-bearing units to which a 'spreads'. The unboundedness of a-spreading follows from recursion through a.

Bidirectional spreading, which might be regarded as a strength of Autosegmental Phonology over Metrical Phonology, also presents no problem to an account based on phrase structure. The bidirectional pattern of spreading illustrated in (4.14) can be represented in phrase-structure terms as the context-free structure shown in (4.15).

\[ (4.14) \]

In this structure, the fact that a is 'focussed' on \( s_2 \) and 'spreads' out in both directions is represented by the topmost level of structure in (4.15), which is licensed by the rule \( a \rightarrow a s_2 a \).

Common to the phrase-structure representations of bidirectional and unidirectional spreading is the recursive immediate dominance (ID) relation: \( a \rightarrow a, f \). In the unidirectional pattern, the direction of spreading will be left-to-right if the linear precedence relation (LP) \( f < a \) is observed, and right-to-left if the linear precedence relation \( a < f \) holds. Bidirectional
spreading, in addition, employs the ordered phrase-structure relation \( a \rightarrow a a \) at its topmost level of structure. Thus the direction and type of spreading (left-to-right, right-to-left or bidirectional) follows from the interaction of immediate dominance and linear precedence statements in a unification-based context-free phrase-structure grammar. (For a formal description of immediate dominance and linear precedence statements, and their relationship to phrase structure, see Gazdar et al. 1985: 44–50 and 5.6.2 below.) Since phrase-structure mechanisms (whether unification-based or otherwise) are required for the definition of metrical structure, as I shall show in section 4.4.3, I conclude that there is no need for an independent autosegmental spreading mechanism.

Well-formedness constraints as a ‘repair’ mechanism
The final weakness of the ‘universal well-formedness constraints’ in Autosegmental Phonology is that they do not hold of all representations, but only surface representations. In other words, an intermediate representation may actually violate the well-formedness constraints (or other constraints, such as the obligatory contour principle), in which case the ‘constraints’ serve to repair the errant representation in order that the surface representation will be well formed. Bird (1990: 3–4) argues that this ‘repairing’ interpretation of autosegmental constraints is highly problematic:

The well-formedness condition . . . divides the set of autosegmental diagrams \( \mathcal{A} \) into the well-formed diagrams \( \mathcal{A}_w \) and the ill-formed diagrams \( \mathcal{A}_i \).

There is a radically different notion of well-formedness which comes from logic and which has been adopted in some quarters of linguistics, which may be summarized as follows:

The Well-Formedness Constraint: Each syntactic rule operates on well-formed expressions of specified categories to produce a well-formed expression of a specified category. (Partee 1979: 276)

If Goldsmith's language were recast in such a system then it would be necessary to pick which of \( \mathcal{A} \) and \( \mathcal{A}_w \) constitutes the well-formed expressions. If we employed his algebraic structures (1976: 28–30) as the formal semantics for this language the resulting system would either be unsound or incomplete relative to the class of intended models. If the set of well-formed expressions is \( \mathcal{A} \) then most well-formed expressions do not have an interpretation (unsoundness). If the set of well-formed expressions is \( \mathcal{A}_w \) then some expressions which have an interpretation are ill-formed (incompleteness . . .). If a logic is not both sound and complete, the connection between its expressions and the objects those expressions
describe is seriously impaired. Either phonological representations can describe non-existent utterances or some utterances cannot be represented.

Not only is the autosegmental spreading mechanism unnecessary, then, but it is also incoherent. In the remainder of this section I shall examine the remaining proposed constraint of Autosegmental Phonology, the No Crossing Constraint (Hammond 1988), and show that it too is fundamentally inadequate as a means by which to constrain the class of well-formed Autosegmental Phonological representations.

4.3.4 *The No Crossing Constraint*

In this and later subsections I shall examine a disquieting problem concerning the ‘constraining power’ of the ‘No Crossing Constraint’ (NCC) with respect to multiplanar Autosegmental Phonological representations. (The No Crossing Constraint is the statement that in a well-formed Autosegmental diagram, lines of association may not cross.) I shall argue that the 'No Crossing Constraint' is not a constraint at all, since it does not reduce or restrict the class of well-formed Autosegmental Phonological representations. Since the details of this argument have been published elsewhere (Coleman and Local 1991), I shall restrict myself here to an abbreviated version of the argument.

Goldsmith's original (1976) formal description of Autosegmental Phonology, which underpins all subsequent work in Autosegmental Phonology, defines Autosegmental Phonological representations to be *graphs*. Given this basis, the argument which we shall set out concerning the failure of the No Crossing 'Constraint' to constrain Autosegmental Phonological representations can be summarised as follows: graphs which conform to the No Crossing Constraint are *planar graphs*. Some 'multiplanar' Autosegmental Phonological representations are planar graphs, but some are (*necessarily*) *non-planar graphs*. The No Crossing Constraint does not restrict the class of non-planar graphs, so by Occam's razor we are faced with two possibilities. Either we continue to recognise non-planar Autosegmental Phonological Representations as well formed and drop or modify the No Crossing Constraint, or we retain the No Crossing Constraint, and thus cease to recognise non-planar Autosegmental Phonological representations as well formed. We consider two sets of examples of Autosegmental Phonological representations, one from the literature and one of our own, which are necessarily non-planar, and thus conclude that the first alternative must be selected. I conclude that the No
4.3 Non-linear representation

Crossing Constraint is not a constraint, since it does not apply to necessarily non-planar Autosegmental Phonological Representations.

Work in Autosegmental Phonology has been largely of an empirical nature, discovering and exploring the descriptive potential of tiered phonological representations. The development of a formal and rigorous presentation of Autosegmental Phonology is a project which now deserves attention, and work is in progress on several fronts. Building on the work of Sagey (1988), Bird and Klein (1990) have studied the temporal interpretation of Autosegmental Phonological representations. Kay (1987) has examined Autosegmental Phonological representations from a computational standpoint, and has elegantly demonstrated that n-tier Autosegmental Phonological representations can be generated or parsed by n+1-tape finite-state transducers. Developments of this idea have been pursued by Kornai (1991), Wiebe (1992) and Bird and Ellison (1994). Our work in this subsection explores the consequences of Goldsmith’s (1976) original topological presentation of Autosegmental Phonology.

The core of my argument can be sketched as follows: a distinction must be drawn between Autosegmental Phonological representations, and diagrams of those Autosegmental Phonological representations. Diagrams are not linguistic objects, but pictures of linguistic objects, and may have properties such as perspective, colour etc. which are of no relevance to linguistic theory. The NCC is a constraint on diagrams, not on Autosegmental Phonological representations. When the conditions by which the NCC restricts the class of diagrams are examined and linguistically irrelevant factors such as width or straightness of lines are removed, it is apparent that the intent of the NCC is to enforce the following planarity constraint: Autosegmental Phonological representations are planar graphs. Thus, the planarity constraint is the defining distinction between the two varieties of Autosegmental Phonology, planar and non-planar (i.e. multiplanar).

In planar Autosegmental Phonology the No Crossing Constraint has no place in linguistic theory, since it is universally the defining characteristic of planar graphs. In non-planar Autosegmental Phonology the NCC is unrestricted, because all graphs can be portrayed as three-dimensional diagrams with no lines crossing.

4.3.5 Autosegmental Phonology and graph theory

We proceed by considering the syntax of Autosegmental Phonological Representations, introducing definitions of the terminology which we employ in the subsequent argument. To begin, we must be careful to
distinguish Autosegmental Phonological representations (APRs), which are linguistic objects, from both diagrams (which are pictorial objects) and graphs (which are mathematical objects). Later, we shall conclude that APRs are graphs.

In Autosegmental Phonology, a phonological representation consists of a number of phonological objects (segments, autosegments and timing slots) and a two-place relation, called association, A, over those objects. In addition, the phonological objects in an Autosegmental Phonological representation are partitioned into a number of well-ordered sets, called tiers. For example, in figure 4.1, the objects d, b and m are segments, [+H] and [+R] are autosegments and the Xs are timing slots.

Goldsmith (1976: 28) defines an Autosegmental Phonological representation as a set of sequences $L^i$ of objects $a_j$ (each of which is a tier, which Goldsmith calls ‘levels’), together with an ordered sequence $A$ of pairs of objects whose first and second members are taken from disjoint tiers. In his own words, he states: ‘Each autosegmental level is a totally ordered
sequence of elements, \( a^i_j \); this is the \( j^{th} \) element on the \( i^{th} \) level. Call the set of segments on the \( i^{th} \) level level \( L^i \). Since each \( L^i \) is a sequence, each tier is a function from the set of natural numbers to a set of objects \( a^i_j \). Let us give each set of objects \( a^i_j \) in \( L^i \) the label \( O^i \). Since the objects in each tier form a set, each \( a^i_j \) is unique and the function is one-to-one. The natural numbers are totally ordered by \( \leq \), which maps one-to-one to a total order \( \prec \) for each \( L^i \). Each \( L^i \), therefore, is a set of objects \( O^i \) for which there exists a total order \( \leq^i \). From \( \leq^i \) it is possible to define \( \prec \) in the usual way: viz. \( a \prec^i b \iff a \leq^i b \) and \( \neg (b \leq^i a) \). \( \prec^i \) provides the usual definition of adjacency, namely \( a \) is adjacent to \( b \) if and only if \( a \prec^i b \) or \( b \prec^i a \) and there is no \( c \) such that \( a \prec^i c \prec^i b \) or \( b \prec^i c \prec^i a \).

Following an illustration, Goldsmith continues with his formal presentation: 'In addition to these . . . sequences of segments, there is a totally ordered sequence of pairs – essentially the association lines . . . Call the set of these pairs “A”. I shall use the symbol \( \preceq \) to denote the total ordering of A. Since there may be more than two tiers in an Autosegmental Phonological representation, A can be partitioned into subsets \( A^{i,j} \), each of which contains the set of pairs (association lines) linking \( L^i \) with \( L^j \). The partitions \( A^{i,j} \) are now called charts by Goldsmith (1990). The relationship between \( \preceq \) and \( \leq \) is as follows: let \( \preceq \) be the total ordering on A, let \( \{(a,b),(c,d)\} \subseteq A \) and let \( \{a,c\} \) and \( \{b,d\} \) be in disjoint tiers. Goldsmith states (1976: 28) 'A in a sense organizes the other levels' (i.e. the endpoints of A). Formally

\[
4.16 \quad (a^i, b^i) \preceq (c^i, d^i) \iff a^i \leq^i c^i \text{ and } b^i \leq^i d^i
\]

When Goldsmith states that A is a set of pairs forming a totally ordered sequence, it must be understood to mean that each chart \( A^{i,j} \) is totally ordered. Consequently, \( \preceq \) is not total, but can be partitioned into total orderings \( \preceq^{i,j} \).

According to Goldsmith's definitions, then, an Autosegmental Phonological representation is a structure \((O,A)\), where \( O \) is the set of phonological objects and \( A \) is a subset of \( O^i \times O^j, i \neq j \). Such a structure is a graph. This graph may be augmented with maps \( \pi_1 \) and \( \pi_2 \) from A to O which pick out the endpoints of each association line. These are defined by Goldsmith as follows:

Define projection \( \pi_1 \) from \( 2^A \) (the set of subsets of A) to \( 2^{L^1} \) (the set of subsets of \( L^1 \)) in the natural way:

\[
\pi_1 \left( \{(a^1_1,a^2_1),(a^1_1,a^2_2),(a^1_2,a^2_1),\ldots\} \right) = \{a^1_1,a^1_2,a^2_1,\ldots\}
\]
Subsequent to Goldsmith's original account of Autosegmental Phonology (Goldsmith 1976), a number of extensions and refinements to the above account have been explored by various phonologists. These extensions are considered in detail in Coleman and Local (1991). However, Goldsmith's original definitions are applicable to all subsequent versions of Autosegmental Phonology.

In Autosegmental diagrams (such as figs. 4.1 or 4.5), phonological objects are represented by alphabetic symbols, features or vectors of features, and the association relation by straight lines connecting each pair of objects that is in the association relation. Tiers are portrayed in Autosegmental diagrams by sequences of objects separated by spaces.

A bipartite graph is one in which the set of vertices can be partitioned into two disjoint subsets $L^1$ and $L^2$, such that the set of edges is a subset of $\{(a,b) \mid a \in L^1, b \in L^2\}$. Since in Autosegmental Phonology the association relation holds only over objects on separate tiers $L^1$ and $L^2$ (i.e. an object may not be associated with any other object on the same tier), the existence of tiers in Autosegmental Phonology ensures that its graphs are (at least) bipartite. APRs with a single tier are just segmental strings. To merit the name, an APR must have at least two tiers. For instance, $L^1$ might be the set of timing slots, and $L^2$ the set of melody segments.

Graphs in which the set of nodes can likewise be divided into $n$ disjoint subsets are called $n$-partite graphs. An Autosegmental Phonological representation with an anchor tier and $m$ melody tiers is thus maximally an $(m+1)$-partite graph, though it might, depending on what conditions are imposed on the association relation, be minimally merely a bipartite graph. This is the case if melody units can only be associated to anchor units, and not to each other. Pulleyblank (1986: 14) advances this condition. It is also the case in the theory of tier organisation proposed by Clements (1985), in which each segment is a tree of melody units, which makes Autosegmental Phonological representations at most bipartite (since all trees are bipartite graphs). In multitier two-dimensional Autosegmental diagrams, association lines do not usually 'meet' tiers except at a node. The graphs denoted by these diagrams are also (at most) bipartite. Since $m$-partite graphs where $m > 2$ are therefore prohibited in one way or another from occurring in the several varieties of Autosegmental Phonology, Autosegmental diagrams portray (amongst other things) bipartite graphs.
Although Autosegmental diagrams portray bipartite graphs, we cannot make a simple equation between Autosegmental Phonological representations and bipartite graphs. There are some bipartite graphs which are not APRs: in other words, APRs are a proper subset of the bipartite graphs. An example of a bipartite graph which is not an APR by any current definitions is the cubic graph shown in figure 4.2.

Having examined the structural properties of APRs and other graphs in a little detail, we shall now focus our attention on one of the defining properties of well-formed APRs, the No Crossing Constraint. I shall first show that the NCC follows directly from the total ordering of A.

Two association lines \((a^i,b^i)\) and \((c^i,d^i)\) are said to cross if and only if \(a^i <^i c^i\) and \(d^i <^i b^i\). The NCC is the statement that in an APR, there is no pair of association lines which cross.

*Theorem 3*

Within a chart, the NCC follows from the total ordering of \(A^i\).

*Proof 1*

If 'lines cross' (i.e. if \(a^i <^i c^i\) but \(d^i <^i b^i\)) then by (4.16) neither \((a^i,b^i) \lesssim (c^i,d^i)\) nor \((c^i,d^i) \lesssim (a^i,b^i)\) in which case \(\lesssim\) is not total and \(A^i\) is not well defined. \(A^i\) is only well defined if lines do not cross, or in other words, if the NCC holds. 

Since the NCC follows from mathematical properties of Autosegmental Phonological representations, it is not a specifically *linguistic* constraint. Additionally, since the NCC follows from the total ordering of \(A^i\), it is not necessary for Autosegmental Phonology to contain both the NCC and the total ordering of \(A^i\). Either one or the other statement could be dropped, and so, by Occam's razor, should be.
Since all Autosegmental Phonological representations are graphs on which some further restrictions have been placed (such as total ordering of tiers), all of the universal properties of graphs hold of Autosegmental Phonological representations, together with some special properties. Autosegmental Phonological representations are a special kind of graph, but they are also subject to all the universal properties of graphs.

We shall now consider the relationship between the two kinds of Autosegmental diagrams (two- and three-dimensional), and two kinds of graphs, planar graphs and Euclidean (non-planar) graphs.

### 4.3.6 Planarity

A Jordan curve in the plane is a continuous curve which does not intersect itself. A graph $G$ can be embedded in the plane if it is isomorphic to a graph which can be portrayed in the plane with points representing the vertices of $G$ and Jordan curves representing edges in such a way that there are no crossings. A crossing is said to occur if either (i) the Jordan curves corresponding to two edges intersect at a point which corresponds to no vertex, or (ii) the Jordan curve corresponding to an edge passes through a point which corresponds to a vertex which is not one of the two vertices which form that edge.

A planar graph is a graph which can be embedded in a plane surface. A Euclidean graph is a graph which can be embedded in Euclidean 3-space, that is, normal, three-dimensional space. All planar graphs are Euclidean, but not all Euclidean graphs are planar. That is, there are some Euclidean graphs which cannot be embedded in the plane. Such graphs are called (necessarily) non-planar graphs.

The two kinds of graphs and diagrams we are considering are related as follows:

**Graphs:** planar graphs $\subseteq$ Euclidean graphs

**Diagrams:** two-dimensional diagrams $\subseteq$ three-dimensional diagrams

I have singled out the planar/Euclidean distinction for particular consideration, since it might be thought that there is a simple one-to-one relation between planar graphs and two-dimensional diagrams, and Euclidean graphs and three-dimensional diagrams. I shall demonstrate that this is not the case, and that this mistaken view underlies a number of problems with the NCC.

By definition, every planar graph can be portrayed in the plane of the paper as a flat or perspectiveless network of points and non-crossing lines
(a two-dimensional diagram); and every flat network of points and non-crossing lines represents a planar graph. By definition, every three-dimensional network of points and non-crossing lines represents a Euclidean graph. I now show that the reverse case also holds.

**Theorem 4**

Every graph can be embedded in Euclidean 3-space.

**Proof** (Wilson 1985: 22)

We shall give an explicit construction for the embedding. Firstly, place the vertices of the graph at distinct points along an axis. Secondly, choose distinct planes (or 'paddles') through this axis, one for each edge in the graph. (This can always be done since there are only finitely many edges.) Finally, embed the edges in the space as follows: for each edge joining two distinct vertices, draw a Jordan curve connecting those two vertices on its own 'paddle'. (We assume there are no edges joining a vertex to itself.) Since the planes or 'paddles' intersect only along the common axis along which all the vertices lie, none of the Jordan curves corresponding to the edges of the graph cross.

I shall illustrate the construction used in this proof by showing how the APR portrayed in figure 4.1 can be embedded in Euclidean 3-space. Not all the nodes of the APR portrayed in figure 4.1 are uniquely labelled. To rectify this, we shall subscript the X slots from left to right X₁, X₂, . . . X₅. We shall label the roots of the syllable-structure trees from left to right with the labels a_r and a_r. We shall label the branching daughter node of o_r with the label R_p and for the sake of thoroughness we shall allow for the possibility that o_r has a non-branching daughter node between o_r and X₅ which we shall label R₂. The two [+H] nodes must be disambiguated, so we shall subscript them from left to right [+H]₁ and [+H]₂. (We shall use these unambiguous labels to refer to the nodes in the embedding of the graph.) The embedded graph has an axis of fifteen nodes in any order. Because the order of nodes in the axis is irrelevant to the construction, we can always make the order of nodes in the axis conform to their order in the tiers of the APR. The embedding operation never requires the order of nodes within tiers to be altered. Suppose that the order of nodes in the axis is σ₁, σ₂, R₁, R₂, X₁, . . . , X₅, d, b, m, [+H]₁, [+H]₂, [+R]. There are fourteen association lines in the APR, so the construction will have fourteen 'paddles' intersecting at the axis. On each 'paddle' draw a Jordan curve between the two vertices of the axis which correspond to the pair of nodes linked by the association line corresponding to that 'paddle'. Viewed one by one, the paddles are as follows:
Given an embedding of a graph G in Euclidean 3–space, we can draw the embedding in a three-dimensional diagram by projecting it into the plane using projective geometry (Lord and Wilson 1984: 32), for instance, using one-point perspective projection. It is consequently simple to prove that:

**Theorem 5**
Every graph G can be portrayed in a three-dimensional diagram as a network of points and (in perspective) non-crossing lines.

**Proof 3**
Embed G in Euclidean 3–space. Project the embedding into the plane. ■

The fact that in Autosegmental Phonological representations, the set of vertices is partitioned into tiers, each of which is totally ordered, does not affect the validity of these theorems. The single axis of vertices required for the construction used in the proof of theorem 5 may be partitioned into totally ordered subsets of objects without affecting the result.
4.3 Non-linear representation

It is harder to show that a graph is necessarily non-planar than that a diagram is three-dimensional. For a diagram to be three-dimensional it merely has to appear to be three-dimensional. A necessary and sufficient criterion for the non-planarity of a graph $G$ is:

*Theorem 6* (Kuratowski 1930)

$G$ is non-planar if and only if it contains a subgraph which is homeomorphic to either of the two graphs $K_5$ (the fully connected graph over five vertices) or $K_{3,3}$ (the fully connected bipartite graph over two sets of three vertices), shown in fig. 4.3.

A reasonably approachable presentation of the proof is Gibbons (1985: 77–80). (Two graphs are *homeomorphic* if they can both be obtained from the same graph by inserting new vertices of degree 2 into its edges.) In particular, note that Kuratowski's theorem requires only that a subgraph of a graph is shown to be non-planar in order to show that the whole graph is non-planar.

The need for association lines to be *straight* for the NCC to work can be demonstrated as follows. Graphs have the general form $(V,E)$. Consider the graph $\{(t_1,t_2,x_1,x_2),\{(t_1,x_2),(t_2,x_1)\})$. The set of vertices $V = \{t_1,t_2,x_1,x_2\}$, and the set of edges $E = \{(t_1,x_2),(t_2,x_1)\}$. Define a partition into tiers $L_1 = \{t_1,t_2\}$, $L_2 = \{x_1,x_2\}$ and the order $t_1 < t_2$, $x_1 < x_2$ (fig. 4.4a). If the NCC requires association lines to be straight, then this Autosegmental Phonological representation cannot be portrayed without crossing lines (fig. 4.4a), and it would thus be excluded by the NCC. But if there is no such restriction on the straightness of association lines, this Autosegmental Phonological representation *can* be portrayed without crossing lines (fig. 4.4b), and thus the NCC does not prohibit this APR.

This demonstrates that the No Crossing Constraint is a condition on pic-
tures, not phonological representations, since straightness of lines is a property of pictures, not linguistic representations. The straightness of association lines is conventional rather than formal: it has never been explicitly defended in Autosegmental Phonology; it does not follow from other principles of the theory; and it is sometimes abandoned when it is convenient to do so (see e.g. McCarthy 1979/1982: 140, Archangeli 1985: 345, Prince 1987: 501, Pulleyblank 1988: 256, 259, Hayes 1989: 300, McCarthy and Prince 1990: 247). If the lines denoting the association relation need not be straight, then the NCC will sometimes necessarily hold and at other times only contingently hold. The cases in which the NCC contingently holds are those like figure 4.4, in which if the lines need not be straight, the NCC can be circumvented. In such cases, the NCC is non-restrictive, and therefore cannot be linguistically relevant. However, in the cases in which the NCC necessarily holds, it is indeed restrictive, for it limits the class of Autosegmental Phonological representations to planar graphs. In these cases adoption of the NCC is equivalent to support for the hypothesis that Autosegmental Phonological representations are planar graphs.

We have shown that every planar graph can be portrayed as a three-dimensional diagram. However, not every three-dimensional diagram represents a planar graph. Some three-dimensional diagrams represent necessarily non-planar graphs. In order to understand the NCC, we are especially interested in the class of properly (i.e. necessarily) non-planar graphs, which cannot be portrayed in two-dimensional diagrams without crossing lines.

4.3.7 Planarity and the No Crossing Constraint

In the early days of Autosegmental Phonology (Goldsmith 1976), all Autosegmental diagrams were drawn as if to lie entirely in the plane of the paper (e.g. figure 4.5). As I showed in the preceding subsection, however, if the No Crossing Constraint applies to Autosegmental Phonological
4.3 Non-linear representation

Figure 4.5 A planar graph (Goldsmith 1976)

Representations, not diagrams, it defines a general, topological sense of planarity:

**Planarity Condition**
A graph is planar if and only if it can be embedded in a plane surface with (by definition of 'embedding') no edges crossing.

Not all graphs can fulfil this requirement, however they are portrayed, and it is therefore necessary to determine whether all Autosegmental Phonological representations can, if only in principle, be portrayed in the plane. If some cannot, then APRs are in general non-planar (whether they are portrayed as such or not), and the No Crossing Constraint is not restrictive.

If the NCC applies to diagrams, it is a drawing convention, not a part of linguistic theory. But if it applies to Autosegmental Phonological representations, the Planarity Condition and the No Crossing Constraint are equivalent: the No Crossing Constraint has no specifically linguistic status, in that it is the defining characteristic of planarity. We foresee that proponents of the NCC might wish to argue that if the NCC is retained, it is the Planarity Condition which is vacuous. But this argument is inadequate in two respects. Firstly, in addition to prohibiting Autosegmental Phonological representations which are necessarily non-planar, the NCC also excludes Autosegmental Phonological representations which may be portrayed in diagrams whose lines only contingently cross, even if there is some other way of drawing them in which no lines cross. Thus the NCC prohibits some Autosegmental Phonological representations merely on the
basis of the way in which they might sometimes be portrayed. The planarity condition, on the other hand, is a condition on Autosegmental Phonological representations, not diagrams of Autosegmental Phonological representations. It thus constrains linguistic (phonological) representations, not diagrams of linguistic representations. Secondly, to the extent that 'no crossing' is a universal property of planar graphs and not just those planar graphs employed in phonological theory, there is no reason for the mathematical definition of planarity to be imported into linguistics as a 'principle of Universal Grammar'. The relevant principle is 'Autosegmental Phonological representations are planar graphs' (if such is the case), not 'association lines must not cross'. For these reasons, given the equivalence of the NCC and the Planarity Condition, those Autosegmental Phonologists seeking to defend the planarity hypothesis would do better to adopt the Planarity Condition directly (since it is a constraint on linguistic representations) than the NCC, which is a constraint only on drawings of linguistic representations.

Various phonologists (e.g. Archangeli 1985, Goldsmith 1985, Pulleyblank 1986) have found it convenient to generalise the planar representations of early Autosegmental Phonology to three-dimensional, 'paddle-wheel', representations, in which several independent autosegmental planes intersect along a distinguished tier of timing units. There is, however, an important distinction between convenience and necessity. It is widely believed and commonly assumed by Autosegmental Phonologists that the necessity of three-dimensional representations has already been uncontentiously demonstrated. Archangeli (1985: 337), for instance, writes 'McCarthy's (1979; 1981) analysis of Semitic forced a truly three-dimensional phonological representation' (emphasis added). Yet McCarthy (1979; 1981) gives not a single diagram which even appears to be non-planar, let alone a necessarily non-planar representation.

Because the belief in the necessity for non-planar Autosegmental Phonological representations is widespread, it has rarely been defended in the literature. I have found only one article containing three diagrams of Autosegmental Phonological representations which are demonstrably non-planar (see figure 4.11 below). I shall discuss these examples in more detail below.

I shall prove the non-planarity of Autosegmental Phonological Representations by establishing a necessary and sufficient criterion for a graph to be (necessarily) non-planar. I shall then use this criterion to test the logic of the argument and examples adduced in support of the claim
that phonological representations have already been shown to be necessarily non-planar. I shall argue that the falsity of claims in the literature about three-dimensionality arise from a failure to distinguish diagram conventions from genuine and uncontroversial universal properties of graphs.

Since the association relation is a bipartite graph, according to Kuratowski’s theorem, homeomorphism to $K_{3,3}$ is a necessary and sufficient condition for an Autosegmental Phonological representation to be necessarily non-planar. The only instances of homeomorphism to $K_{3,3}$ to be found in the Autosegmental literature and in the remainder of this subsection are cases of identity to $K_{3,3}$.

4.3.8 Three-dimensional diagrams in the Autosegmental literature

Figure 4.1, taken from Archangeli (1985), is typical of those diagrams of the association relation that are purported to be necessarily non-planar. Archangeli’s logic is inexplicit, but seems to be as follows (see Archangeli 1985: 337): suppose there is a tier above the anchor tier (for instance, a consonant melody), and another tier below the anchor tier (for instance, a vowel melody). Then there are at least two ‘paddles’, one in the plane of the paper above the anchor tier, the other in the plane of the paper below the anchor tier. Now if yet another independent tier is called for (syllable structure, perhaps), yet another paddle, separate from the two in the plane of the paper, is required. Thus phonological representations with more than two melody tiers on separate paddles are necessarily non-planar. This argument is erroneous. Figure 4.1 has three independent paddles, but it is not homeomorphic to either $K_5$ or $K_{3,3}$, and can be portrayed in the plane with no lines crossing. Figure 4.6 is one possible plane embedding of figure 4.1. All the other examples of three-paddle Autosegmental diagrams that I am aware of from the literature (with the exception of those in figure 4.7 which I discuss below) also have plane embeddings. The graph of which figure 4.6 is a plane embedding is in no way affected by the manner in which it is portrayed. Since it is unchanged, it retains all the structure of figure 4.1, still supporting reference to all the relevant notions of locality (i.e. adjacency or association; see Hoberman 1988) and accessibility as in figure 4.1. Such an embedding is nothing other than a different way of looking at the same formal object.

The argument which Archangeli offers is one of only a few published attempts to establish the non-planarity of Autosegmental Phonological representations. However, Archangeli’s hypothesis has been generally accepted by Autosegmental Phonologists, presumably because it is undeni-
ably convenient to use three-dimensional diagrams in Autosegmental Phonology. A number of other three-dimensional examples putatively establishing the non-planarity of Autosegmental Phonological representations are discussed in Coleman and Local (1991), and are shown, like Archangeli's example, not to portray necessarily non-planar APRs.

A detailed examination of the Autosegmental Phonology literature in Coleman and Local (1991) showed that the common belief that a necessarily non-planar Autosegmental Phonological representation already exists is mistaken, even though a demonstration that just one Autosegmental Phonological representation is necessarily non-planar would be sufficient for rejection of planar Autosegmental Phonology and two-dimensional diagrams as inherently too restrictive.

One such APR has been portrayed in the Autosegmental literature, in
4.3 Non-linear representation

4.3.9 A necessarily three-dimensional diagram

Consider a phonological representation with three anchor units on one tier, three autosegments on one or more other tiers and a line of association between each anchor and each autosegment. Such a graph cannot be portrayed without crossings on a plane surface, since it is homeomorphic to $K_{3,3}$. There is no linguistic reason why such a representation might not be motivated in certain cases. Wetzels's example (figure 4.7) is one such case. Two more are illustrated in figure 4.9, which shows the distribution of backness, rounding and nasality over three timing units in the pronunciation of the words *room* and *loom* by a Guyanese English speaker, data that will be
discussed in more detail below. Both of the graphs portrayed in figure 4.9 are homeomorphic to $K_{3,3}$, and thus they are necessarily non-planar. In order to demonstrate that the graphs portrayed in figure 4.9 are the correct Autosegmental Phonological representations of the two words, Coleman and Local (1991) show that the features [back], [round] and [nasal] are independent autosegmental features that are lexically associated with independent segments, and therefore spread independently. For this to be the case, they must lie on independent tiers. Liquid onsets are lexically associ-
ated with [±back], the nucleus with [±round] and the coda with [±nasal].
These three autosegments then spread to each of the other syllable termi-
nals, to produce the Autosegmental Phonological representations por-
trayed in figure 4.9.

We thus showed that planar graphs are not adequate for Autosegmental
Phonological representations of Guyanese English, because the
Autosegmental Phonological representations of room and loom cannot be
planar. Given that the principles which interact to produce this result are
not particularly special and are individually attested elsewhere, I have no
reason to believe that Guyanese English is either unnatural or special in this
respect, and thus planarity (i.e. the No Crossing Constraint) is too severe
to be a universal constraint on Autosegmental graphs.

4.3.10 The NCC does not constrain three-dimensional Autosegmental
Phonology
Since Autosegmental Phonology is necessarily non-planar, the No Crossing
Constraint has no force, because all graphs, however complex, can be por-
trayed in three-dimensional diagrams without edges crossing. The fact that
some versions of Autosegmental Phonology employ 'paddle wheel' graphs,
rather than unrestricted (i.e. Euclidean) graphs, does not affect this result.
I conclude that the No Crossing Constraint is not a constraint at all, since
it either incorrectly restricts the class of phonological graphs to planar
graphs, or else it carries no force.

I have demonstrated that APRs are graphs, and concluded that the NCC
does not restrict the class of such graphs. The conclusion follows from the
definition of APRs advanced by Autosegmental Phonologists from the
inception of Autosegmental Phonology. To avoid this conclusion, those
fundamental definitions would have to be changed.

The NCC is a semi-formal way of stating that the mapping between two
tiers induced by the association relation is sequence-preserving. This goes
back to Goldsmith's original formal definition, in which it is which
ensures sequence preservation. Autosegmental Phonology therefore makes
a twofold claim: firstly, that phonological representations are divided into
tiers; secondly, that the tiers are sequenced.

The NCC could be revised as follows: 'Within a chart, association lines
may not cross' (see Kitagawa 1987). This alteration to Autosegmental
Phonology is not sufficient by itself to rescue the NCC, however, since Plane
Reduction (Halle and Vergnaud 1987: 55) may still derive a chart which vi-o-
lates the revised NCC from two charts which independently adhere to it.
Contemporary generative phonology

This operation, originally called *tier conflation* by McCarthy (1986), is *chart conflation*, according to Goldsmith's terminology.*

Our result concerning the non-restrictiveness of the NCC does not condemn the Autosegmental approach to phonological representation. However, it does mean that the NCC cannot be invoked as a constraining 'principle' of Autosegmental theories. The conclusion I draw from this is that while the Autosegmental model provides a useful notation for the study of non-linear phonological representations, it is *not* an adequate grammatical framework for the definition of those representations.

I shall now turn to the arboreal Metrical theory, a non-linear phonological theory which employs phrase-structure-type phonological representations, and show how its 'Principles and Parameters' approach can be formally defined using a Unification-based Phrase-Structure Grammar. I shall preface this exegesis of Metrical Phonology with a short discussion of headed prosodic structures in Dependency Phonology, which probably influenced the basic ideas of Metrical Phonology.

4.4 Constituency, heads and modifiers in phonological theory

In *SPE* phonology (Chomsky and Halle 1968), the paradigm phonological theory of the 1970s, there were just three 'sizes' of phonological objects: features, segments and strings. Segments were sets of features, and strings were sequences of segments.

In the intervening period, this extremely parsimonious position has been relaxed, as it has been found that admitting a richer variety of different-sized objects, far from complicating phonological theory, enables phenomena which hold over various phonological domains to be modelled more elegantly and with a more 'principle-based' formalism. Consequently, in most contemporary phonological models, a variety of 'sizes' of phonological objects are recognised, such as metrical feet (Liberman and Prince 1977, Selkirk 1984a), syllables (Fudge 1969, 1987; Selkirk 1984a, b) and the parts of syllables, onset, rime, nucleus and coda (Clements and Keyser 1983). These objects are joined together in phonological representations by the relation of constituency. Nucleus and coda are usually taken to be constituents of the rime (but cf. Clements and Keyser 1983), onset and rime are constituents of the syllable, metrical feet are formed from syllables, and utterances are made up of metrical feet. Along with the widespread adoption of phonological constituent structure, there has also developed a less widespread but nevertheless popular subtheory, the 'dependency' theory of
constituency (Anderson and Jones 1974). This theory maintains that not all constituents are alike: in every phonological (or morphological or syntactic) structure, some constituents are 'heads' and some are 'non-heads'. In the following subsections, I will briefly review the manner in which headed constituent structures are represented in several contemporary phonological theories, and what interpretations have been attributed to headedness. (This discussion is also a necessary preliminary to the theory of the phonetic interpretation of headed phonological structures which I will present in chapter 5.) I shall then conclude the present chapter by showing how the highly explanatory 'parameter-setting' approach to the definition of metrical structure benefits from being expressed in the UPSG formalism.

4.4.1 Dependency Phonology
Dependency Phonology was one of the first theories to adopt phonological constituent structure and the head/non-head distinction (Anderson and Jones 1974). In Dependency Phonology, the head/non-head distinction has both a phonological and a phonetic function. One of the phonological (specifically, distributional) characteristics is that heads are obligatory constituents, whereas non-heads ('dependents') are optional constituents. One of the phonetic characteristics is that heads are more prominent than non-heads. 'Prominence' means different things in different contexts. In the account of stress and accentuation proposed in Dependency Phonology, 'prominence' is stress or accent. In the account of syllable structure, 'prominence' is sonority or 'vowellness'.

Figures 4.10 and 4.11 are examples of phonological representations in Dependency Phonology. Figure 4.10 represents a syllable consisting of just a vowel, such as the word 'I', perhaps, and figure 4.11 represents a syllable consisting of an initial consonant, a vowel and a final consonant, the word *pit*, perhaps. Dependency graphs such as these encode information about constituent structure, linear order and headship. The head of a structure is the object written directly below the mother of that structure in the phonological representation. The rime is the obligatory part of the syllable, and the nucleus is the obligatory part of the rime, so the nucleus node is written below the rime node and the rime node is written below the syllable node. Onset and coda are optional constituents. The head/non-head distinction offers an explanation for why vowels may form syllables by themselves, whereas consonants (generally) cannot. The SPE theory, which lacked the head/non-head distinction, could not really do this, except by explicitly
Dependency Phonology permits both dependent + head constructions and head + dependent constructions. However, Ewen (1982) argues that the relative order of heads and non-head sisters changes from level to level of phonological structure. Thus, rimes govern onsets (right-to-left) to form syllables, stressed syllables govern unstressed syllables (left-to-right) to form metrical feet, and tonic feet govern non-tonic feet (right-to-left) to form tone groups. If this pattern of alternation is extended to the fine structure of onsets, the theory would predict that the onset-initial /s/ position governs stop + glide clusters (left-to-right) in onsets such as /spr/, and the onset-final glide position governs core obstruents (right-to-left), assigning a right-branching structure to onsets. I shall call this the Alternating
Government Hypothesis. Thus, as well as accounting for phonetic prominence and certain facts about the distribution of phonological objects, the notion of dependency (or, conversely, government) has also been invoked to account for the ordering of phonological units according to a general principle, the Alternating Government Hypothesis.

4.4.2 Metrical Phonology

Metrical Phonology is similar to Dependency Phonology as far as they address the same domain of data, that is, stress and syllable structure. (Dependency Phonology also addresses the internal structure of segments, on which subject Metrical Phonology is silent.) There are some differences, certainly (see Ewen 1982). For instance, in Metrical Phonology, constituent structures are almost always binary-branching. (Pierrehumbert and Beckman (1988: 21) propose n-ary branching metrical structure.) One sister of each set of nodes directly dominated by the same mother is labelled ‘s(tong)’ and the other(s) ‘w(ak)’, corresponding to Dependency Phonology’s ‘head’ and ‘non-head’ respectively. The notational practices of Metrical Phonology are somewhat different from Dependency Phonology, partly as a consequence of having labelled nodes, but these differences are not germane to the present discussion.

A consequence of headedness in phonological representations is illustrated by the following example. In both Dependency Phonology and Metrical Phonology, the difference in the location of the main and secondary stresses — and hence the vowel qualities — in a morphologically related pair of words such as photograph and photographer is ultimately attributed to the fact that the number of syllables is different, so the constituent structures of the metrical feet are different, and thus the (metrical) head is a different syllable in the two cases.

As an illustration of Metrical Phonology, I will briefly discuss some aspects of the representation of stress placement in English words. (This discussion will presage a fuller treatment in chapter 7.) In English, there is a close relationship between morphological structure and lexical metrical structure (i.e. lexical stress). A compound word has two constituents, of which in this case the first is stressed (e.g. bläckbird). The lexical constituent structure and the lexical metrical structure are congruent (binary branching); see figure 4.12. Similarly, words of Germanic origin can be divided morphologically into prefixes and stems. It is the stem which bears the stress, relative to the prefix. Again, the lexical constituent structure and the lexical metrical structure are congruent (figure 4.13).
In words of Latinate extraction or form, however, morphological structure and metrical structure are not congruent. For instance, the morphological structure of the word *photographical* is *photo*+*graph*+*ic*+*al* (figure 4.14), whereas the metrical structure places a foot-division between *photo* and *graphical*. As a rule, Latinate stress falls on the rightmost heavy syllable of the last three syllables of the word (Chomsky and Halle 1968: 33, Lass 1988: 114). As we saw in section 3.6.1 earlier, a metrical representation of Latinate words is a right-dominant word-tree (see rules (3.36) and (4.17) below) consisting of a 'beginning part' of zero or more feet which are of no relevance to main stress placement, and a 'final foot' from the rightmost heavy syllable of the last three syllables of the word up to the end (fig. 4.15). By doing this, a declarative, structural characterisation of Latinate stress...
4.4 Constituency, heads and modifiers

Latinate stress

Pretonic word
Main stress domain

* e.g. *photo-*

Head syllable
zero to two light syllables

-graph
-ical

Figure 4.15 ‘Latinate’ metrical structure

place...t stress falls on the first syllable of the last foot.’ The validity of the division of Latinate words into a pretonic part of indeterminate length, and a final foot main stress domain of finite length, is demonstrated by the fact that the final foot, as a whole, is more strongly stressed than the pretonic. (The internal structure of the pretonic will not be considered here.) When the parts of this analysis are drawn together, it will be seen that Metrical Phonology can be employed to assign relative prominence to every syllable in the string. The representation-based approach to stress assignment proposed in Metrical Phonology is attractive because it does not require complex stress-manipulating transformations as in SPE phonology, and yet it provides a conceptually simple and descriptively adequate account of stress placement.

4.4.3 Declarative Phonology and parameters of metrical structure

I stated in chapter 1 that the originators of Metrical Phonology, like those of Autosegmental Phonology, attempted to dispense with SPE-type context-sensitive phrase-structure rules. Metrical Phonology took its lead instead from Chomsky’s ‘Principles and Parameters’ approach to syntactic description, which at the time of Metrical Phonology’s origins was a relatively new model. Stowell’s (1981) thesis was particularly influential in arguing that syntactic phrase-structures do not need to be described by language-specific phrase structure rules, but can be characterised by a small set of quite different but language-universal principles. The differences between languages arise from the language-specific settings of a number of ‘parameters’ or options for variability which are circumscribed by the ‘principles’.
The set of parameters proposed by Zubizaretta (1982) for Japanese metrical structure is indicative of early proposals made by Metrical Phonologists in the 'Principles and Parameters' manner. In the years since then, a more definitive set of metrical parameters has been developed. In this subsection, I shall examine the parameters of metrical structure listed by Dresher and Kaye (1990) in their discussion of a stress-learning algorithm based on parameter-setting. To the extent that this small, simple and precisely formulated set of parameters is capable of both describing the stress patterns of a large variety of languages, and also accounting for acquisition of those stress patterns, the parametric metrical theory can fairly be described as one of the most successful and insightful developments in contemporary linguistic theory, and especially so for adherents of the 'Principles and Parameters' approach. I shall show that the unformalised parameters which Dresher and Kaye present are insightfully formalised using the resources of Unification-based Phrase-Structure Grammar.

The eleven parameters which Dresher and Kaye enumerate are as follows:

P1: The word-tree is strong on the [Left/Right].
P2: Feet are [Binary/Unbounded].
P3: Feet are built from the [Left/Right].
P4: Feet are strong on the [Left/Right].
P5: Feet are quantity-sensitive (QS) [Yes/No].
P6: Feet are QS to the [Rime/Nucleus].
P7: A strong branch of a foot must itself branch [No/Yes].
P8A: There is an extrametrical syllable [No/Yes].
P8: It is extrametrical on the [Left/Right].
P9: A weak foot is defooted in clash [No/Yes].
P10: Feet are non-iterative [No/Yes].

It is not necessary in this discussion to exemplify the evidence for this particular selection of parameters, which will be discussed individually below. The empirical support for the coverage of this account is so extensive that it suffices to state it here, and refer the reader to Dresher and Kaye (1990) and the references they cite for further support. Nevertheless, it would be a mistake to regard this success as demonstrating the preferability of the parametric approach to the phrase-structure grammar approach. Phrase-structure grammar has changed greatly under the developments in unification-based theories of grammar such as GPSG, so that Stowell's original criticisms of classical phrase-structure grammar do not extend to
4.4 Constituency, heads and modifiers

Unification-based Phrase-Structure Grammar. In fact, as I shall now show, ten of the eleven parameters which Dresher and Kaye set out can be directly expressed in terms of UPSG-type ID/LP rules. I shall argue that regarding the parameters of metrical structure as constraints in a UPSG complements the ‘Principles and Parameters’ approach by casting its English-language definitions of the parameters and their settings into the formal language of Unification-based Grammar.

PI: The word-tree is strong on the [Left/Right]

An unformalised ‘principle’ lies behind this parameter, which is that words consist of feet, one of which is stronger than the other(s). This can be formally expressed as the immediate dominance rule-schema (4.17):

\[(4.17) \omega \rightarrow \Sigma_s, \Sigma_w\]

where \(\omega\) is the category of words, \(\Sigma\) is the category of feet, \(s\) denotes ‘strong’, \(w\) denotes ‘weak’ and \(x^p\) denotes \(p\) occurrences of category \(x\). (This latter usage makes (4.17) a rule-schema, not a rule.) The two values of PI correspond to the only two possible linear orderings of \(\Sigma_s\) with respect to \(\Sigma_w\):

\[(4.18) \text{If PI} = \text{Left then} \Sigma_s < \Sigma_w\]
\[(4.19) \text{If PI} = \text{Right then} \Sigma_w < \Sigma_s\]

Note that because Dresher and Kaye’s parameters are not defined in a precise formalism such as UPSG, there are many other logically possible but unattested forms which PI could take. For example, no independent requirement of the ‘Principles and Parameters’ approach would have prevented PI from being, for example, ‘the word-tree is strong on the [Left/Right/3rd foot from left/Center/ . . . ]’. The two values which PI may take are just the only two possibilities which the ID/LP format permits. Unification-based Phrase-Structure Grammar, therefore, far from being a needlessly complex alternative to the ‘parameter-setting’ model, in fact makes precise the formulation of that model, with interesting and explanatory consequences. Such consequences will be seen in the UPSG formalisation of the other metrical parameters too.

P2: Feet are [Binary/Unbounded]

The unstated principle behind this parameter is ‘feet consist of syllables’. In a similar fashion to the word-rule, this principle can be formalised using the ID rule-schema (4.20):
(4.20)  $\Sigma \rightarrow \sigma_\alpha, \sigma^+_w$

(This rule-schema could be more concisely formulated $\Sigma \rightarrow \sigma^+$, but would then have to be revised to take account of P4.) The two possible settings of P2 correspond to two constraints on the value of n in (4.20). These are:

(4.21)  If P2 = Binary then $n \leq 1$

(4.22)  If P2 = Unbounded then $n \geq 0$

Although there are various other possible ways n might be parameterised than these two (in either Dresher and Kaye's statement or in the UPSG formulation), (4.21) is better than Dresher and Kaye's formulation, because as well as setting an upper bound on the size of feet, the combination of (4.20) and (4.21) also allows feet of just one syllable in languages where the 'Binary' classification is appropriate. Although this minor failing of Dresher and Kaye's statement could be rectified simply by restating it as, for instance, 'Feet are at least one syllable long and at most [Binary/Unbounded]', expressing this statement using the formal resources of phrase-structure grammar demands that the minimum and maximum limits must be determined in order to actually formulate the rule-schema and its limits correctly.

P3: Feet are built from the [Left/Right]

Although the wording of this parameter statement is formulated in 'procedural' language, it is susceptible of a simple declarative formulation based on ID and LP rules that is no different in kind from the formulation of P1, for instance. Dresher and Kaye's verbal statement of this parameter is quite different from that of P1, however, so it can be seen that the UPSG formulation reveals a formal similarity between P1 and P3 that is uncaptured in the verbal statement.

The unformulated principle which lies behind this parameter is that when a word is divided into feet according to the other parameters and principles, a weak syllable (or possibly more, depending on the word and the settings of the other parameters) may remain which is not incorporated into any foot. For example, in Dresher and Kaye's example (8b), the Warao word *yiwaranae* is parsed into two sw feet (P4 = Left) from the right (P3 = Right) – *yi.wara.nae* – which leaves the initial syllable *yi* unassigned to a foot. (Dresher and Kaye assign it to a 'defective' foot of one syllable with no syllable-level s/w relations, of course. We shall propose a slightly different analysis here, though the difference is formally inconsequential to this argu-
4.4 Constituency, heads and modifiers

Hayes (1982) proposes 'under orthodox metrical assumptions' that such 'stray' syllables are dominated by the word node directly in the initial syllables of Latin examples such as *refectus*, and English examples such as *agēnda*, *paréntal* and *homonymous*, though his analysis differs from ours by then adjoining such stray syllables to adjacent feet. Pierrehumbert and Beckman (1988: 148) also propose that these stray syllables are directly dominated by the word node, a choice which I shall follow here. To accommodate this possibility, the word rule (4.17) must be revised to (4.23):

\[(4.23) \quad \omega \rightarrow \sigma^n_w, \Sigma_s, \Sigma^e_w\]

The two possible values of P3 can now be seen to encode whether the stray syllables \(\sigma^n_w\) precede or follow the feet \(\Sigma\) and any syllables that may be contained in those feet. If the LP statements that encode P3 are insensitive to the distinction between strong and weak feet, then there are only two categories mentioned in (4.23), \(\Sigma\) and \(\sigma\). There are, consequently, only two possible ways of ordering these two categories, which correspond to the two possible settings of P3:

\[(4.24) \quad \text{If } P3 = \text{Right then } \sigma < \Sigma \]
\[(4.25) \quad \text{If } P3 = \text{Left then } \Sigma < \sigma \]

Again, alternative possibilities, such as 'Feet are built from both ends at once (or alternately)', would require a rather more complex formal definition, such as sensitivity to the difference between strong and weak feet, which the present analysis does not include. By formalising the parameters in the UPSG formalism, we can expropriate the cross-linguistic generalisations of the 'Principles and Parameters' model, and its account of language acquisition, by making its formal expression explicit.

**P4: Feet are strong on the [Left/Right]**

This parameter provides the two possible LP statements that could co-occur in a UPSG along with the foot rule (4.20) stated above:

\[(4.26) \quad \text{If } P4 = \text{Left then } \sigma_s < \sigma_w \]
\[(4.27) \quad \text{If } P4 = \text{Right then } \sigma_w < \sigma_s \]

**P5: Feet are quantity-sensitive (QS) [Yes/No]**

Dresher and Kaye (1990: 145) describe this parameter as follows: 'Quantity sensitivity, in the theory being adopted here, means that a branching rime (or a branching nucleus, depending on the setting of P6) may not occupy a
weak position in a foot.' Quantity sensitivity, therefore, is a relationship between the s/w property of syllable nodes, and the branching/non-branching nature of rime nodes. An unstated assumption of Dresher and Kaye's statement is that rimes are constituents of syllables. This is appropriately formalised by the conventional phrase-structure rule for syllables:

\[(4.28) \quad \sigma \rightarrow O R\]

In order to formalise the relationship between different subcategories of \(\sigma\) and different R structures, we need to be able to distinguish branching rime or nucleus structure dominated by R from non-branching rime or nucleus structure dominated by R. This can only be achieved in a context-free account by subcategorising R by some feature, such as \([\pm\text{branching}], [\pm\text{heavy}]\) or [weight: heavy] vs. [weight: light]. The distinction between these is entirely arbitrary, so I shall employ \([\pm\text{heavy}]\) without further discussion. A rime node will be subcategorised as \([+\text{heavy}]\) if and only if it dominates a branching structure; otherwise it will be subcategorised as \([-\text{heavy}]\) (or it could be left unspecified for weight, perhaps). The formulation of quantity sensitivity will be simplified if it is assumed that syllables are also subcategorised with the same weight specification as their rime:

\[(4.29) \quad \sigma:\text{heavy} = R:\text{heavy}\]

This is so that quantity sensitivity can be treated as a relation between s/w and \([\pm\text{heavy}]\) that is entirely local to the syllable node. The account developed so far makes no link between the s/w and \([\pm\text{heavy}]\) markings of the syllable node, so no more rules are required to capture the 'P5 = No' parameter-setting. The 'P5 = Yes' setting corresponds to the inclusion of the following additional constraint in the grammar:

\[(4.30) \quad \text{If } P5 = \text{Yes then } [+\text{heavy}] \supseteq \sigma,\]

Note that this constraint only prohibits heavy, weak syllables. It does not prevent light syllables from being strong, which is the case, for example, in a quantity-sensitive language such as English, although as Dresher and Kaye (1990: 145) require, it does ensure that 'in quantity-sensitive stress systems all heavy syllables are stressed'. Whether or not light syllables may be strong depends on the setting of the independent parameter P7.

P6: Feet are QS to the [Rime/Nucleus]

The discussion of P5 did not formalise the conditions by which R nodes are subcategorised for the feature \([\pm\text{heavy}]\). That depends on the setting of P6, which assumes in either case the phrase-structure rule-schema:
4.4 Constituency, heads and modifiers

The setting 'P6 = Rime' corresponds to the unbracketed extension of this rule-schema:

(4.32) If P6 = Rime then \( R \rightarrow NC \) [+heavy]

The alternative setting 'P6 = Nucleus' corresponds to the equivalent formulation of the nucleus structure:

(4.33) If P6 = Nucleus then \( N \rightarrow V_1V_2 \) [+heavy]

For the value of \([± heavy]\) to percolate from the nucleus node up to the rime node and thence to the syllable node (4.33) must be accompanied by the constraint \( R:heavy = N:heavy \). (This could in fact be assumed to hold, albeit uselessly, even in non-QS systems – in other words, it could be specified irrespectively of the settings of P5 and P6.)

This formalisation of P6 also allows the possibility, not apparent in Dresher and Kaye's statement, that a language could be quantity-sensitive to both the rime and the nucleus, so that closed syllables and syllables with long vowels would be heavy. Thus the formal difference between the UPSG account of P6 and, for example P1 and P4 (LP statements), or P2 (value of n) is consequential. Cases such as P1, P4 and P2 represent mutually exclusive possibilities, which is reflected in their formal descriptions, whereas the possible settings of P6 are merely different, and could logically co-occur, a distinction which is only apparent in the UPSG statements, and not in Dresher and Kaye's informal statements.

P7: A strong branch of a foot must itself branch [No/Yes]

Dresher and Kaye (1990: 146) describe this parameter as follows: 'The value of P7 in these languages is No, which means that the dominant (i.e. s) branch of a foot does not itself have to branch (i.e., contain a heavy syllable). If we set the value of P7 to Yes, then only heavy syllables can anchor a foot.' Note that in the statement of this parameter, weight is described in terms of branching, whereas in Dresher and Kaye's formulations of P5 and P6, the term 'quantity', rather than 'branching' was used. The formal phrase-structure-based account which I give here makes explicit the unity of the phenomenon referred to by this parameter, by employing the \([± heavy]\) feature again, which was explicitly related to the branching of the rime and/or nucleus node in the definitions above.
The parameter-setting ‘P7 = Yes’, then, simply corresponds to the inclusion of (4.34) in the set of phrase-structure constraints. The setting ‘P7 = No’ corresponds to the omission of (4.34).

(4.34) \( \sigma_5 \leftrightarrow [+\text{heavy}] \)

English is an example of a language in which P7 is set to ‘No’: for example, the suffix -icity bears stress on the antepenultimate syllable even though it is light, simply because the ultimate is extrametrical and the penultimate is light, and therefore weak. Dresher and Kaye cite Aguacatec Mayan as a language in which P7 must be set to ‘Yes’.

P8A: There is an extrametrical syllable [No/Yes]

In the immediately preceding paragraph, I stated (following the highly elegant analysis of English stress proposed in Hayes 1982) that the final syllable of -icity is extrametrical. The support for this claim is simply that if the final syllable in words with antepenultimate stress is analysed as extrametrical, the remaining syllables in the word (with the exception of initial stray syllables) can be exhaustively parsed into binary-branching feet (‘P2 = Binary’).

As with the analysis of stray syllables, the extrametrical syllable, if permitted in the language, will be analysed as a daughter of the word node, rather than as a constituent of a foot. The word rule (4.23) must consequently be revised yet again, to (4.35):

(4.35) \( \omega \rightarrow \sigma^n_w, \Sigma_s, \Sigma^p_w, \sigma^m \)

[+em]

(If it were not for the fact that initial stray syllables are also daughters of the word node, the [+em] feature would be unnecessary. Perhaps this could be taken in support of the claim that stray syllables are not daughters of the word node, but are remiss constituents of a foot with an empty strong ictus syllable.) With the principle behind P8A now made explicit in the phrase-structure rule (4.35), the two possible settings of P8A can be formalized:

(4.36) If P8A = No then m = 0

(4.37) If P8A = Yes then m \leq 1

P8: It is extrametrical on the [Left/Right]

This parameter refers to the relative linear order of the [+em] syllable if it is permitted and occurs, with respect to the other syllables. These other syllables can be referred to as [-em], though they do not need to be explicitly
specified with this feature. The unification mechanism will ensure that whichever of the following two LP parameters is set will require a [+em] syllable to precede or follow all the other syllables in the word, as appropriate.

\[(4.38) \quad \text{If } P8 = \text{Left then } [+em] < [-em]\]

\[(4.39) \quad \text{If } P8 = \text{Right then } [-em] < [+em]\]

The independent facts that \([\pm em]\) is defined to be a binary-valued feature, and that LP is irreflexive, combine to allow only two possible ways in which an extrametrical syllable might be ordered with respect to the other syllables in the word. Once again, the UPSG formulation seems rather more explanatory than the English-language expression of the rule.

P9: A weak foot is defooted in clash [No/Yes]
The procedural language in which this parameter is expressed does not render it transparent to a declarative reformulation. The dissimilatory prohibition of two adjacent strong syllables which it describes does not translate into UPSG terms as easily as the other parameters. That does not mean that it cannot be expressed in UPSG terms, however. The setting 'P9 = Yes' means that the single strong syllable in a weak (i.e. non-branching) foot is recategorised as unspecified for weight if it occurs immediately before a sw foot, or after a ws foot. I am not sufficiently familiar with actual cases of languages in which this parameter would be set to 'Yes' to state a reformulation in UPSG terms. Dissimilation of this kind is perfectly expressible in a UPSG, however: see example (7.6) below.

P10: Feet are noniterative [No/Yes]
This parameter, which determines whether or not secondary stresses (i.e. weak feet) occur in a language, sets a limit on the value of p in the word rule (4.35).

\[(4.40) \quad \text{If } P10 = \text{No then } p = 0\]

\[(4.41) \quad \text{If } P10 = \text{Yes then } p > 0\]

As Dresher and Kaye (1990: 147) explain: 'P10 [No] means that only the strongest foot in a word tree will count, while other feet are suppressed.'

4.4.4 Recent history
In order to make precise the claims of the Principles and Parameters approach to metrical theory, I have shown the explanatory power of the kind of unification-based extension to context-free phrase-structure
grammars proposed by Gazdar et al. (1985). In doing this, I hope to have shown the importance of the contributions of both phrase-structure theory, and the 'parameter-setting' theory. These two models were quite unnecessarily set against each other by their respective advocates during the 1980s: phrase-structure grammar was given up by many linguists in favour of unformalised 'Principles and Parameters' descriptions, in view of arguments promoted in Chomsky (1972, 1981) and Stowell (1981) that phrase structure was an epiphenomenon of independent principles of the categorial component, such as subcategorisation and X-bar theory. The parameter-setting model also offered a simple yet linguistically sophisticated account of the great - but not unlimited - variety to be found in languages, and the way in which any language can nevertheless be acquired apparently straightforwardly and equally well by practically all children.

These were attractive selling-points for the new framework, and its remarkable success in metrical theory brought about its popular ascendancy. Yet I have shown that by allowing categories to be factorised into their component feature structures, the formal innovations of unification grammar permit phrase-structure grammar to formally express constraints on phrase-structure trees such as the X-bar schema, projection of subcategorisation from the lexicon, independent dominance and precedence constraints, and feature-climbing or percolation. I have demonstrated that these resources provide the formal explicitness of statements which is the 'parameters' model's greatest want. In light of the above discussion, it seems warranted for advocates of phrase-structure grammar to answer the charge that phrase-structure grammars are *ad hoc* collections of rules with the reply that the principles and parameters of UPSG (the metrical part, at least) are a universal set of constraints, no different in kind from those envisaged by advocates of the 'Principles and Parameters' model, but with the notable advantage of being explicitly formalised, with consequent well-understood mathematical-linguistic properties.

### 4.4.5 Combining Autosegmental and Metrical Phonology

Having shown the benefits of employing UPSG to define the phonological representations proposed by Autosegmental and Metrical Phonology, it remains to consider how the complementary contributions of the two theories might be integrated with each other under the UPSG 'umbrella'. Before examining how the UPSG accounts of Autosegmental and Metrical Phonology might be combined, however, I shall briefly examine two earlier views regarding the integration of these two theories.
In the ‘paddle-wheel’ view of Autosegmental Phonology, metrical structure was simply regarded as being represented on another autosegmental plane (see figure 4.1). The only way in which the non-metrical and metrical parts of the representation were integrated was that they shared the same skeletal tier of CV (or later, X) slots. Even the ‘lines’ in the metrical and non-metrical parts of such representations should not be regarded as the same sorts of objects. As Scobbie (1991) argues, the lines in the metrical plane are the conventional asymmetric (directed) ‘dominance’ lines, whereas the lines used in non-metrical planes are symmetric (undirected) ‘association’ lines. Clements and Keyser (1983) represent syllable structure in autosegmental terms, but their proposal is a notable departure from orthodox metrical theory.

In the revised theory of autosegmental structure proposed in Clements (1985), the undirected association lines of Autosegmental Phonology are replaced by the directed ‘dominance’ lines of Metrical Phonology and phrase-structure theory. The interface between the autosegmental and metrical components of the representations is still the segmental X-slots (now called ‘root nodes’, since they are the roots of segmental phrase structure), as a consequence of which we can draw the following correspondences:

\[
\begin{align*}
\text{Metrical structure} & \quad = \quad \text{suprasegmental (dominating the root nodes)} \\
\text{Autosegmental structure} & \quad = \quad \text{subsegmental (dominated by the root nodes)}
\end{align*}
\]

These correspondences have the curious consequence of requiring a special explanation for suprasegmental autosegments: ‘downward’ many-to-one segment-to-autosegment associations. This is formally disastrous for the restrictiveness of Autosegmental Phonology, however, because, as Karlgren (1976: 14) shows, grammars with rules of the form \( A \rightarrow B \rightarrow \ldots \rightarrow C \rightarrow D \) (e.g. ‘downward’ autosegmental rules) as well as \( A \rightarrow B \rightarrow C \rightarrow \ldots \) (e.g. metrical rules) are not context-free, but equivalent to completely unrestricted rewriting systems. This still obtains if the ‘metrical’ constituency rules are all constrained to be binary-branching, since Karlgren’s proof concerns the addition of unbounded rules of the form \( A \rightarrow B \rightarrow \ldots \rightarrow C \rightarrow D \) to context-free grammars, irrespectively of whether or not those grammars are in Chomsky Normal Form.

Pierrehumbert and Beckman (1988: 152–60) propose a theory which integrates metrical and autosegmental representations in a fairly weak way. In their view, autosegmental tiers are separate from metrical trees, but the
autosegmental units on those tiers may be associated with any node in the metrical structures, whether terminal or non-terminal. In addition, however, they propose the following constraint on autosegmental-to-metrical associations (Pierrehumbert and Beckman 1988: 156): ‘The substantive [i.e. autosegmental] elements linked to a [metrical] node having any given height either (1) are also linked to the node’s head or (2) are peripheral to elements linked only in the node’s interior.’ (Since autosegmental and metrical nodes are separate types of object in Pierrehumbert and Beckman’s theory, association is a directed relation, for them.)

This constraint, and the theory within which it is proposed, translates straightforwardly into the UPSG formalism. The fact that autosegmental elements may be associated to terminal or non-terminal nodes squares with the fact that both terminal and non-terminal categories in UPSGs are complex symbols. Thus, instead of drawing a line from autosegmental units to metrical nodes, Pierrehumbert and Beckman might have equivalently simply have suffixed the autosegmental categories to the metrical nodes. In line with Pierrehumbert and Beckman’s constraint, we can formally distinguish peripheral from head features, and in order to represent the placement of peripheral features, we shall formally distinguish left from right edge features. These distinctions can be encoded using three category-valued features [left], [right] and [h]. Using these features, Pierrehumbert and Beckman’s figure 1.7 (reproduced here as figure 4.16) can be repre-
4.4 Constituency, heads and modifiers

Figure 4.16 translated into a UPSG representation

...presented in UPSG terms as figure 4.17. (In this diagram, circled numbers denote token-identical sharing of feature structures.)

...Pierrehumbert and Beckman's constraint on the distribution of features in prosodic trees can be implemented as the following disjunctive constraint on such representations:

**Prosodic Feature Licensing Constraint**

A feature at a non-terminal node must trickle down to the frontier: [h] features trickle down the head path, [left] features trickle down the leftmost path, and [right] features trickle down the rightmost path.

...The discussion of Autosegmental and Metrical Phonology presented above can be drawn to a conclusion by considering their complementary contributions to non-linear phonology. I argued that the principal concern of Autosegmental Phonology is the focus and spreading (or domain) of prosodic features with respect to the CV structural tier, whereas Metrical Phonology proposes a highly restrictive theory of a number of possible prosodic domains. I suggested that the respective contributions of each of these theories can be beneficially formalised using context-free phrase structures. Unidirectional and bidirectional spreading was formalised using recursive linear phrase structures, parameterised for direction of spreading by alternative LP statements in the case of unidirectional spreading. The ID rules that defined these structures also singled out one daughter as the prosodic...
‘focus’. Metrical structures were defined using ID rule-schemata, parameterised by bounds on iteration, alternative LP statements, and suppression or retention of certain constraints. Since Autosegmental and Metrical Phonology can both be recast in a unitary but still restrictive formalism, I conclude that the UPSG account of prosodic structure is preferable to the relatively *ad hoc* and independent Autosegmental and Metrical formalisms.

### 4.5 Conclusion

In this chapter, I showed in detail that Autosegmental and Metrical Phonology contribute complementary formal proposals concerning the distribution (focus and spreading) of non-segmental features in prosodic structures, thus making explicit the mechanisms of a declarative, non-segmental view of prosodic phonological representations. I argued that the ‘Principles and Parameters’ approach to Autosegmental Phonology is singularly unspecific, and that the only substantive constraint which it proposes, the ‘No Crossing Constraint’, fails to constrain the class of well-formed Autosegmental representations. I showed how focus and spreading of autosegments (both unidirectional and bidirectional) and the ‘parameters and principles’ of metrical theory can be expressed using the formal techniques of Unification-based Phrase-Structure Grammar.

These applications of UPSG in phonology exemplify and make explicit some of the benefits to be gained by using UPSGs in phonology, which were raised in the last chapter as a possible solution to the excessive power of the *SPE* formalism and some aspects of its successors. In the next chapter I shall consider the phonological applications of UPSGs in more detail, before proceeding in chapters 6 and 7 to exemplify the formalism with grammars of substantial fragments of Japanese and English phonological structure.
5 Phonological representations in Declarative Phonology

5.1 Introduction

In the previous two chapters I examined the development and content of various theories of phonological representation, segmental and non-segmental, procedural and declarative. I argued that segmental theories of representation require the introduction of excessively powerful, procedural, rewriting rules, whereas a non-segmental approach to representation can avoid the need for rules of such a powerful kind, by expressing phonological relations in the representations, rather than in the rules. I also made some more specific proposals regarding a non-segmental, non-procedural approach to the definition of phonological representations. In chapter 3 I presented a general method for representing 'rewrite rules' in phrase-markers, with special cases for (strictly) context-sensitive rules and the well-known case of phrase-markers for context-free rules. Each derivation (or rather, each class of equivalent derivations) can then be represented as the join (unification) of a set of local graphs, each of which represents a rule. Because unification is associative, the notion of 'order of rule application' becomes rather meaningless. In this way, I showed that any grammar can be interpreted declaratively, rather than procedurally.

In chapter 4 I demonstrated the great value of Unification-based Phrase-Structure Grammars as a common formalism for integrating the key concepts of Autosegmental and Metrical Phonology. I showed that the focus and spreading of autosegments and the prosodic hierarchies of Metrical Phonology can both be formalised using only context-free immediate dominance and linear precedence constraints. Moreover, I showed that UPSGs are an appropriate and enlightening formalism for the representation of the 'principles and parameters' of metrical theory.

In this chapter and the next I shall propose a declarative theory of phonological representation, constructed from the theoretical devices discussed in previous chapters in connection with existing theories of phonological representation (such as structures of distinctive features,
Declarative Phonology

hierarchical constituent structure and unification). Here, I set out the principles of the theory, and in the following chapter I shall illustrate the formalism with fairly extensive and exhaustive analyses of the phonological structure of words in Japanese and English. Since the declarative theory of phonological representation set out in this chapter is largely a synthesis of proposals examined in other contexts in the previous chapters, the initial motivation for each of the theoretical devices has already been discussed. This chapter will consequently define, rather than discuss, those proposals and devices.

5.2 Historical context

In chapter 3, the factors leading to the adoption of unification-based phrase structure theories of grammar by some syntacticians were discussed, and in chapters 1 and 4 the 'Principles and Parameters' approach developed by syntacticians and phonologists associated with Chomsky was examined. At present, Unification-based syntax and 'Principles and Parameters' grammar are theories whose practitioners are in apparent conflict. I showed in my discussion of metrical theory that this need not be the case as far as the theories are concerned, however. It is perfectly possible – and arguably desirable – to employ a highly restrictive theory such as ID/LP grammar in order to formalise the central concepts of the 'Principles and Parameters' approach. In reconciling these two approaches, there is no place for rhetorical obstacles. Both approaches gain from this exercise: the 'Principles and Parameters' gains formal expression, which, as I showed in chapter 4, can reveal new generalisations and principles which were concealed by defining the parameters in ordinary English prose. The UPSG formalism can gain by overcoming the biggest weakness of phrase-structure grammar, the arbitrariness and language-specific nature of its rules. By implementing putatively universal parameters and principles, Declarative Phonology and the 'Principles and Parameters' approaches stand to arrive at a precisely defined set of language-universal constraints, such as those discussed in relation to metrical theory. For the present, I can make only a partial contribution to this programme by attempting to give a precise definition of a UPSG-based formalism for phonological representation, and detailed examples of analyses employing that formalism.

The use of structured phonological representations, prosodic features, feature propagation principles etc. does not rest solely on empirical considerations. Equally important, and in some respects more important,
5.3 A non-derivational approach

is the issue of the generative power of a formalism, which was discussed at length in chapter 3. There are (at least) two approaches to the development of a theory which is empirically adequate and yet not excessively unconstrained. One approach is to construct a formalism which appears to be empirically adequate, and then check that it is not excessively unconstrained. This is the route explored by most phonologists, but work in this approach has in most cases not yet got beyond the empirical stage. It could turn out that any or all of these formalisms are excessively unconstrained. This concern has prompted myself and other researchers to try a second approach: to settle upon a formalism which is known to be formally constrained, see what adaptations need to be made in order to accord with empirical phonological and phonetic considerations, and determine the formal consequences of such adaptations. These two approaches, the 'exploratory' and the 'cautious', are complementary, since both aim for empirical and formal adequacy, and phonologists in both traditions must have both kinds of consideration in mind.

5.3 A non-derivational approach to phonology

In order to formally represent phonological structures, I and a growing number of other researchers are exploring the use of a formalism which was originally developed for the representation of syntactic information, the Unification Grammar formalism (Shieber 1986). Unification Grammar extends formal language theory to include complex categories, such as sets of (possibly nested) feature-value pairs, and incorporates insights from the type-theories employed in logic and computer science. The details of specific varieties of Unification Grammar, such as GPSG, LFG and HPSG are extensively described elsewhere (see Shieber 1986 and section 3.5.3 above for references).

Unification Grammar regards linguistic rules (such as phrase-structure rules) as partial descriptions of linguistic objects (such as trees or strings). (Contrast this with Chomsky's (1957: 27) presentation of trees, which was simply a declarative notation for the sequential application of a set of phrase-structure rules.) In this way, the formal distinction between rules and representations is eliminated. Rules are just partial representations. According to this view, the yield of the grammar is a set of representations which are formed by 'combining' lexical representations (encoding non-predictable information) with 'rules' (partial structural representations encoding predictable information) using the technique of constraint
satisfaction. This 'combining' operation involves two operations, concatenation and unification. Concatenation acts as syntagmatic 'glue', and unification acts as paradigmatic 'glue'. A Unification Grammar, then, yields a set of linguistic structures (e.g. trees), the frontiers of which define a set of well-formed strings.

Under this 'direct' interpretation of Unification Grammar, rules are not devices which map strings to strings in an ordered sequence of steps. Since unification and concatenation are both associative operations, the constraint-satisfaction mechanism is intrinsically unordered. This has two consequences, one of linguistic importance, the other of computational importance. First, a unification grammar defines a single level of representation, and different orders of 'rule application' do not have different results. Analyses which require rule-ordering of any kind are thus inexpressible in this approach. Second, the absence of rule-ordering and the presence of a single level of representation makes the formalism computationally simpler than theories with rule-ordering and transformational derivations. In particular, declarative formalisms such as Unification-based Phrase-Structure Grammar are independent of any particular procedural implementation, and admit of a variety of parsing or generation interpretations. In this respect, unification-based phrase-structure grammars are closer to being a 'competence' model than generative grammars based on the ordered application of rewriting rules.

5.3.1 Phonetic interpretation in Declarative Phonology
Although Declarative Phonology is in some respects radically different from most other current frameworks, it is in many respects a rather conservative reincarnation of many of the principles of European structuralist approaches to phonology. In this eclectic framework, phonology is only concerned with the encoding of information in speech patterns. It is thus primarily a combinatorial discipline, whose descriptions are in terms of structure, alternation, opposition, distribution and composition – essentially algebraic concepts. Many linguists view phonology and phonetics as forming a continuum, with phonological descriptions presented in the same formal language as phonetic descriptions, for instance in the assemblage of supposed universally distinctive phonetic properties. Though this is the most orthodox position, reflected in phonology textbooks, it is not a view which is shared by everyone. Foley (1977) calls this hypothesis 'Transformational Phonetics', rather than 'generative phonology', a terminological distinction which I shall adopt, since it is possible to support
the generative approach to phonology without extending that approach to phonetic interpretation. The non-derivational approach to phonology demands a rigorously maintained distinction between phonology, on the one hand, and phonetics, on the other. This commitment follows Trubetzkoy’s dictum that

The data for the study of the articulatory as well as the acoustic aspect of speech sounds can only be gathered from concrete speech events. In contrast, the linguistic values of sounds to be examined by phonology are abstract in nature. They are above all relations, oppositions, etc., quite intangible things, which can be neither perceived nor studied with the aid of the sense of hearing or touch. (Trubetzkoy 1969: 13)

Phonological descriptions in Unification Grammar, like syntactic descriptions, are formal, contentless – that is, not inherently meaningful – algebraic, formulated in the domain of set theory, and agnostic with respect to acoustic, perceptual or articulatory manifestation. Phonetic descriptions are necessarily domain-dependent (i.e. acoustic, perceptual or articulatory), contentful, formulated in physical or psychological domains. Declarative Phonology is thus a model with two qualitatively dissimilar levels of description. This model is similar to some earlier, segmental, speech synthesis models, in that phonological representations and parametric phonetic representations are qualitatively different. Such theories have been criticised at length by Fowler (1980) and Fowler et al. (1980), since they require an explicit ‘translation’ component to mediate between the two types of representation. Recently, Browman and Goldstein (1986, 1992) have attempted to develop a model in which phonological representations are simply partially specified parametric representations. The goal of such models is the avoidance of a ‘translation’ component (Browman and Goldstein 1986: 222). However, such a component (the phonetic interpretation or exponency function) is required even in models such as Browman and Goldstein’s. The apparent lack of a phonetic interpretation function in Browman and Goldstein’s model is due only to the systematic confusion between notations and other forms of representation, on the one hand, and the physical denotations of those representations, on the other.

I shall argue below that phonetically arbitrary phonological phenomena are conveniently ignored by the purely phonetically based view of phonology such as Browman and Goldstein’s, and in later sections I shall present an explicit and testable theory of phonetic interpretation as a demonstration that the ‘abstract’ position, including extrinsic timing, is tenable.
5.3.2 Transformational Phonetics and Generative Semantics

An analogy can profitably be made between Transformational Phonetics and Generative Semantics (Newmeyer 1980). Despite the name, this latter was an entirely syntactic theory, which sought to reduce semantic relations to syntactic relations, just as Transformational Phoneticians such as Halle (1983) and Browman and Goldstein (1986, 1989) are attempting to reduce all phonological relations to phonetic relations. For instance, the semantic relation of logical consequence in (5.1) was given a syntactic formulation in which *kill* was taken to be a form of the phrase 'cause to become not alive' and *die* a form of 'become not alive'.

(5.1) a. John killed Bill.
    b. Bill died.
    c. Bill is not alive.

Logical consequences such as (5.1b) and (5.1c) were then derived by purely syntactic operations (deletions etc.) from (5.2), the 'underlying' form of (5.1a):

(5.2) John caused Bill to become not alive.

While not without certain merits, Generative Semantics eventually proved itself too baroque to be taken seriously as a unified theory of syntax and semantics. Eventually, Interpretive Semantics, classically exemplified by Chomsky (1972), Jackendoff (1972) and Thomason (1976), gained more widespread acceptance. This is a more conservative approach, modelled on the specification of the syntax and semantics of formal languages, such as logic or computer programming languages. In this approach, syntactic and semantic descriptions are formulated in qualitatively dissimilar terms, and the meaning of an expression is determined by interpretation of syntactic descriptions (see Gazdar et al. 1985 chs. 9 and 10 for an extended example). Interpretation is a bidirectional relation holding between two different kinds of representation, rather than a function from syntactic representations to qualitatively similar semantic representations. The Interpretive theory does not necessarily mean that syntax and semantics are entirely unrelated, however, nor that the mapping between them is unprincipled. Though arbitrary, in the Saussurean sense, it is a systematic mapping, since it is subject to the Principle of Compositionality, which says that the meaning of an expression should be a function of the meanings of its constituent parts and the syntactic operation by which they are combined (Dowty, Wall and Peters 1981: 8).
5.3 A non-derivational approach

5.3.3 Transformational vs. Interpretive Phonetics

Generative phonology and much current phonetic theory seem to be recapitulating some of the mistakes of Generative Semantics. By constructing phonological categories from the names of common phonetic categories, such as 'voiced, strident, continuant, consonant', the appearance of a seamless unity to phonology and phonetics has been created. This unity is completely illusory. Being merely symbols, phonological labels such as 'voice' have no intrinsic phonetic denotation. The phonetic distinctions between *bitter* and *bidder*, for instance, include distinctions in vocal-cord activity, tongue dynamics (giving rise to quality variations throughout the entire articulation of both these words), velic-port aperture and timing distinctions (Lisker 1986, Kelly and Local 1989: 158–61). And yet 'getting the right names' for phonological oppositions is often regarded as one of the phonetician's main tasks (Keating 1988b) – compare the [ATR]/[TENSE]/[EXPANDED PHARYNX] debate with the various cross-language differences in the phonetic interpretation of 'ATR harmony' described in Lindau and Ladefoged (1986: 470–1). The equation of phonological with phonetic categories in Transformational Phonetics can be likened to a syntactic theory which confuses grammatical gender with sex. Some grammatical categories such as gender and number clearly have a semantic basis or motivation, but they are nevertheless arbitrary and non-semantic, in that no simple equation of grammatical gender and sex is possible. A derivational theory of gender would encounter serious difficulty when faced with 'counter-semantic' grammatical gender, such as *ship* having feminine gender in English, or more clearly, a 'girl' (*das Mädchen*) being neuter in German.

Similarly, the phonological category [+voice] sometimes denotes phonetically voiceless articulations. Although the label [voice] was chosen to describe the phonological opposition which may sometimes denote a distinction of phonetic voicing, this does not mean that [voice] always denotes phonetic voicing, nor that in the absence of phonetic voicing that [–voice] obtains. We do not need a rule to change lexically 'voiced' items into phonetically 'voiceless' items in a derivational fashion. It is adequate and conceptually simpler just to claim that the phonological object which for convenience we categorise as 'voiced' is not intrinsically phonetically voiced, and does not always denote phonetically voiced articulations. This position is close to the Glossematic view, expressed by Anderson (1985: 164) as follows: 'In [Hjelmslev's] terms, that is, it is the fact that [German] /b/, /d/, /g/, etc. undergo syllable-final devoicing that establishes the link
among them, not their phonetic similarity.' A further example of why the vocabulary of phonetic and phonological descriptions should be kept separate is shown by the sometimes contradictory use of the categories 'long' and 'short' in phonetic and phonological descriptions. For example, the English phoneme symbol /u/ is generally used to denote a phonetically short vowel. Yet in English, this category has two historical sources: an original phonologically short vowel category, as in words like put, and an original phonologically long vowel category, as in words like took, and for some speakers, room and tooth. The phonological test par excellence that illuminates the former status of these phonological categories is the fact that the Great Vowel Shift affected long vowels, but left short vowels unaffected. Put was unaffected by the Great Vowel Shift, whereas took was shifted from a mid back vowel to a high back vowel. If the consequences of the Great Vowel Shift are regarded as legitimate synchronic phonological phenomena, as evidenced by regular vowel-quality alternations in contemporary English, the phonological category of the vowel in took should be regarded as distinctively 'long' in non-transformational phonology, even though its phonetic interpretation is relatively short. There is no contradiction here if phonetics and phonology are kept apart as two separate levels of description. If any confusion arises, it is because the same terms, 'long' and 'short', are used with different senses in different domains of description.

Another illuminating example is the Southern British English pronunciation of /æ/ in many environments as [əː]. That this is indeed phonologically a short vowel is supported by the fact that it corresponds to an uncontentiously short-vowel category in cognate dialects; it retains a phonetically short exponent in some words, such as mass; it shows variability in other words, like plastic; and it can occur in structural configurations in which long vowels are prohibited, for instance tautosyllabic with /sk/ codas. Despite being a phonologically short vowel, however, it is phonetically similar, if not identical, to the phonetic exponents of the phonologically long-vowel category /aː/ of words such as bar, and is therefore usually categorised as a phonetically long vowel. (The fact that bar is an open monosyllable is a reliable diagnostic that it is phonologically long.) In view of examples such as this, I conclude that two different phonological representations may have the same, or sometimes the same, phonetic interpretation.

Let us consider an example of phonological arbitrariness which, despite appearing to have a phonetic motivation, cannot be accounted for by artic-
5.5 A non-derivational approach

A non-derivational approach to this example concerns English rime phonotactics. A reliable generalisation about native English syllables is that the rimes /ait/, /aik/, /auk/, /aup/, /aub/ do not occur, although most other combinations of a diphthong and obstruent are well formed. A possible account of these distributional facts is that in the context of diphthongs with an 'open' vocalic peak, the 'palatal' closure of /t/, /d/, /s/ may not combine with the 'palatal' rising diphthong /ai/, and the 'labial' closure of /p/, /b/, /m/, /f/, and the 'velar' obstruents /k/, /ɡ/ may not combine with the 'labial-velar' rising diphthong /au/. This analysis claims that the distributional facts are accounted for by a phonetically motivated principle of dissimilation, which can be called Gravity Dissimilation. In Declarative Phonology, this principle can be formulated as a constraint on rime constituency which prohibits the nucleus and coda agreeing in the value of the feature [grave] if the nucleus is an open rising diphthong. Since this phonotactic constraint is missing in other languages (and, indeed, in some dialects of English), it is a parochial fact about English phonotactics.

It is possible to describe this phenomenon in articulatory terms, using, for example, the categories of Browman and Goldstein’s gestural model. In the nucleus /au/, there is a sequence of tongue-back (TB) gestures. The first gesture is a [wide, palatal] gesture, and the second is a [narrow, velar] gesture. The second tongue-back gesture is co-ordinated with a [narrow, protruded] LIPS gesture. In /ai/, there is a [wide, pharyngeal] TB gesture, followed by a [narrow, palatal] TB gesture without an associated LIPS gesture. Generalising over these two vowel categories, it can be seen that the dissimilation constraint concerns sequences of [wide] and [narrow] TB gestures (in other words, open rising diphthongs).

In /k/ and /ɡ/ there is a [clo, velar] TB gesture, in /p/, /b/ and /m/ a [clo, labial] LIPS gesture, and in /f/ and /v/ a [crit, labial] LIPS gesture. A description which appropriately generalises over these categories is the disjunction ‘[velar] TB or [labial] LIPS’. Disjunctions are not, however, a part of Browman and Goldstein’s formal phonological vocabulary. Nor are ‘cover features’ such as [grave].

In /t/ and /d/ there is a sequence of [clo, palatal] and [crit, palatal] TT (tongue tip), and in /ʃ/ a [crit, palatal] TT gesture. (Since place and manner of articulation are parts of the specification of a gesture, and the tiers of scores are phonetically fixed to be the set of active articulators, the unity of place of articulation of such phonological categories as affricates and nasal-stop clusters in English and other languages cannot be expressed by...
factoring the place-of-articulation descriptor onto an independent ‘tier’.

The generalisation concerning these gestures is that they are all [palatal] TT gestures. The dissimilation constraint can now be expressed as follows: after sequences of [wide] and [narrow] TB gestures, [velar] TB or [labial] LIPS gestures may not follow [velar] TB or [labial] LIPS gestures, and [palatal] TT gestures may not follow [palatal] TB gestures.

Although it is thus possible to express Gravity Dissimilation using Browman and Goldstein's formalism, it leaves much to be desired from a simply descriptive standpoint, let alone from an explanatory standpoint. The description does not offer an account of why this constraint includes [velar] TB, [labial] LIPS and [palatal] TT, and not, say, [pharyngeal] TB, [dental] LIPS and [alveolar] TT. The usual explanation for the frequent association between [velar] and [labial] articulations, and their opposition to [palatal] articulations in phonological systems, rests upon the acoustically motivated explanation that [velar] and [labial] articulations are made at around the ends of the oral-cavity tube, whereas [palatal] articulations are made at around the middle of the oral cavity, dividing it into two resonating chambers, the two configurations consequently having very different acoustic properties (Jakobson, Fant and Halle 1952: 30). Although Browman and Goldstein recognise the need to allow for the possibility of acoustically motivated explanations of phonological phenomena such as Gravity Dissimilation, their model does not provide a mechanism for the unified specification of [grave] gestures. Odden (1991) presents further evidence for the category [grave], which he calls [Back~Round].

Even though the phenomenon of Gravity Dissimilation can be provided with a phonetic motivation, it is not phonetically deterministic in the sense that natural phonetic processes such as coarticulation and stop-epenthesis in nasal–fricative clusters arise mechanically in Browman and Goldstein’s model under different phasing conditions of two or more gestures. It cannot be predicted on articulatory grounds that English rimes may not contain ‘equally grave’ close-diphthongal-offglide-plus-obstruent constellations after open vowels. The existence of this phenomenon is phonologically arbitrary, although perhaps phonetically motivated, a contingent fact about English, not speech. The historical factors which have given rise to Gravity Dissimilation and the presence of a phonologically long but phonetically short vowel category and a phonologically short but phonetically long vowel category in English are, of course, circumstantial. But the presence of unusual, surprising and phonetically counter-intuitive phono-
Logical phenomena in the phonology of any language is far from unusual. Although phonological phenomena often have what Stampe (1979) calls natural processes as their phonetic basis or origin, such natural processes are not, by themselves, a sufficient basis for phonological theory. It is precisely the arbitrariness of phonology which gives languages their expressive potential. Arbitrariness does not mean that these phenomena are unsystematic, however. While some phonological phenomena may indeed be unsystematic, and thus must be represented in the lexicon, even arbitrary phenomena such as those discussed immediately above are almost completely regular, and it is appropriate to state that regularity with descriptive generalisations of some kind. I therefore maintain that far from being phonetically deterministic, phonological representations are arbitrary, though phonetic interpretation is systematic.

Although the declarative approach to phonology presented here is non-segmental, features are used to represent phonological distinctions, drawing upon the (by now) extensive study of feature-value categories (Gazdar et al. 1985: 17–42, Shieber 1986, Gazdar et al. 1988), which is discussed more fully in section 5.5 below. Every phonological feature has to be assigned an interpretation, and that interpretation is sensitive to its structural context. The adoption of an interpretive rather than a transformational phonetics does not mean that phonetic interpretation of phonological representation is completely unsystematic, though. Just as in semantic interpretation, the Principle of Compositionality must be observed.

If the phonetics–phonology relationship is likened to the syntax-semantics relationship, the parallels shown in figure 5.1 between syntax, semantics, phonology and phonetics can be made. This analogy was originally
proposed in Wheeler (1981), a thesis which set out the core characteristics of Declarative Phonology:

1 Phonological representations are purely formal, constructed in an algebraic domain, whereas phonetic representations are descriptions of continuous events, constructed in the physical domain. The relationship between phonological and phonetic representations is essentially arbitrary (although nevertheless systematic). From this standpoint, for instance, the assimilated 'velar' coda closure observed in some tokens of in that case is considered not as an altered /t/, but as a particular, context-specific token. Phonological representations are not 'turned into' phonetic representations by rewrite rules. (Phonology is monostratal.)

2 In phonological representations, the assignment of prosodic nodes to model phonological units of various extents (as in the UPSG accounts of Autosegmental and Metrical Phonology in the previous chapter) eliminates both the need for 'spreading' rules and the hegemony of short-domain (segmental) units. (Phonological representations are structured.)

3 When phonological representations are structured, interpreted and strictly distinguished from phonetic representations, deletion and insertion operations are not needed. The class of phonological operations is no longer problematic, consisting solely of (various) gluing and phonetic interpretation operations. Being non-destructive, phonological operations are monotonic.

I take phonological relations, rather than symbols or operations, to be of primary importance in linguistic structure. Since phonological symbols are never required except as the terms in phonological relations, it seems sensible to define relations as primes, and units/symbols intensionally or derivatively. Features, then, are just a way of notating poles of relations (typically paradigmatic relations, though in principle this could be extended to syntagmatic relations such as the s/w relation in Metrical Phonology). Desiring a non-transformational, monostratal theory, I adopt a 'type–token' approach to most forms of phonological variability. Phonological representations and phonetic representations are not just two ends of a sequence of phonological operations, using essentially the same kinds of symbol. Indeed, phonetic representation is not a symbolic level at all. Symbols merely do duty for the actual denotata of phonological representations, much as, say, logic formulas may do duty for denotata of expressions in semantic
Phonological processes in Declarative Phonology

In chapter 3 it was shown that the primitive operations of a transformational grammar—insertion, deletion, copying and permutation—are apparently instantiated by phonological phenomena such as coarticulation, assimilation, elision, epenthesis and metathesis. I shall now briefly summarise how the variety of phonological phenomena discussed in section 3.4 from the transformational perspective are amenable to non-transformational descriptions that are arguably less arbitrary. In the examples I shall use alphabetic symbols to abbreviate phonological structures and to represent parts of parametric representations in the usual fashion. This is purely expository, however: no alphabetic symbols appear in the phonological or phonetic representations in our theory, and no intrinsic relationship between phonological and phonetic objects abbreviated by the same alphabetic symbol is intended. I shall first give a brief account of Declarative Phonology's view of the various phonological phenomena listed above. Some of these will be further exemplified later in this chapter and in the next chapter. Several other cases are addressed in chapter 3 of Bird (1995).

As Declarative Phonology draws a sharp distinction between phonology and phonetics, it will be seen from the following examples that in each case it is possible to talk about the phenomena in phonological or phonetic terms. Determining whether a given case is best treated in one way or the other is an empirical issue. In general, however, we use the categorial vs. gradient distinction as a diagnostic to determine the most appropriate choice of analysis. I regard it as a strength of Declarative Phonology that this choice has to be made in each case, as that appears to fit with the experience of working phonologists (see e.g. Hayes 1992: 282). SPE-based phonology, in contrast, does not make this distinction in a sharp way, and hence does not force the analyst to decide whether a phenomenon is phonological or phonetic.

5.4.1 Allophony

Phonological allophony in Declarative Phonology arises from the application of feature-filling rules applying to a category at one place in structure
but not at another. For example, the phrase-structure rule O[-back] → /l/ allows /l/ to occur as an unbranching onset position and assigns the feature [-back] to the onset node, expressing the fact that the clear allophone of /l/ occurs in this position. In all other positions, that is, in branching onsets and in codas, /l/ is [+back] in many English dialects, a fact which could be expressed using a default or another feature-filling rule. Consequently, just as in derivational phonology, a given symbol may acquire different features at different places in structure through the operation of phonological constraints.

Phonetic allophony is the different interpretation of the same element in different structural contexts, rather than as involving several slightly different phonological objects instantiating each phoneme (fig. 5.2). I have used this idea as a basis for speech synthesis (see Coleman 1992). A few additional details are given later in this chapter. For the application of this approach to German /t/ allophony, see Walther (1993).

5.4.2 Assimilation
Phonological assimilation can also be modelled non-destructively by unification (figure 5.3). (See Local (1992) for a more complete description.) Since phonological assimilation involves copying or sharing features between two parts of a phonological representation, it is not particularly problematic for Declarative Phonology, and has been studied in detail by Scobbie (1991). The Declarative approach to assimilation requires the undergoer to be underspecified for the features involved in the assimilation. In order to get the features of the unassimilated form to surface correctly,
5.4 Phonological 'processes'

Coda  Coda  Coda
   [Tongue-back: closure]  [Tongue-back: closure]
   C       C       C
[nasal: +]  [nasal: -]  [nasal: +]  [nasal: -]
N   k       η   k

Figure 5.3 A declarative view of assimilation (∪ denotes unification; = denotes non-distinctness)

- [ii]  [uu]  [uu]
  [k]  [p]  [k]  [l]  [k]

keep  cool  cart

Figure 5.4 Declarative (coproduction) model of coarticulation

where appropriate, it may be necessary to employ defaults, either grammatical (across-the-board) defaults, where the surface is easily predicted, or lexical defaults in other cases. However, an increasing number of cases that have been regarded in the past as cases of phonological assimilation have been shown to be the result of gradient, phonetic processes rather than arising from the application of phonological rules or constraints. For example, assimilation of word-final alveolar consonants to the place of articulation of a following word-initial consonant in English is argued by Nolan (1992) to be such a case.

5.4.3 Coarticulation

As I proposed at the end of chapter 2, the phonetic aspects of coarticulation are simple to model if parametric phonetic representations may be 'glued together' in parallel, rather than simply concatenated. Consonants may then be modelled as being coproduced with vowels, rather than simply concatenated to them (see Öhman 1966, Perkell 1969, Gay 1977, Mattingly 1981, Fowler 1983). This analysis is implemented in the temporal interpretation of phonological objects in the mapping from phonological to phonetic representations. The temporal arrangement of the coarticulated objects can be portrayed using representations of overlapping events (see Griffen 1985, Bird and Klein 1990), for instance figure 5.4. As I mentioned
in section 2.10 and shall discuss further below, not all aspects of coarticulation can be treated in this fashion: phonetic interpretation needs to be sensitive to a certain amount of phonological differentiation or allophony too.

5.4.4 Epenthesis
Phonologically, the epenthesis operation is feature-filling, and thus presents no challenge to Declarative Phonology. Where epenthesis takes place without exception in a certain structural configuration (such as an empty C or V), a feature-filling default rule can be employed. For example, rule (5.3) inserts the features of schwa into any empty V position:

(5.3) \[ \text{V} \rightarrow [+\text{back}, \text{-high}, \text{-low}, \text{-round}, \text{-ATR}] \]

V positions bearing feature-values that are different from any of the features in the rule will not undergo the rule, as it is inconsistent with such descriptions. It thus applies to empty Vs, as well as V slots that bear some or even all of the features of schwa.

In certain cases, a consonant or vowel which alternates with zero in morphophonologically underived forms is not easily analysed as epenthetic. For example, linking /t/ in British English is no more plausibly a default consonant than linking /j/ or /w/ in, for instance, /wej/ in/way in/ or /gaw on/go on/. In such cases it seems necessary to represent the alternation with zero in the lexical representation. For example, using the notation (X) as shorthand for ‘X or zero’, nuclei that take linking /t/ can be represented, for example, /ɜ(r)/ ear, /ɛz(r)/ air, /ɔz(r)/ are, /ɔ(ː)z(r)/ or, /ɔ(ː)z(r)/ err, and /uəz(r)/ (as in poor, for some speakers). In order to complete the analysis, it is necessary to provide constraints which determine which alternant (i.e. zero or filled) occurs in which environment. For instance, if /t/ is prohibited from occurring in syllable-final position but is permitted to occur syllable-initially, the zero disjunct is the only legal possibility in isolated words or before consonant-initial words and the /t/ disjunct will be the only legal possibility before vowel-initial words. (See also Bird 1995: 95–7.)

Some instances of epenthesis may be analysed not in the phonology as the insertion of a segment into a string, but as due to minor variations in the temporal co-ordination of independent parameters (Jespersen 1933: 54, Anderson 1976, Browman and Goldstein 1986, Mohanan 1986:162–4) (fig. 5.5). It has been demonstrated many times (e.g. Fourakis 1980, Kelly and Local 1989: 135–9, 150–3) that such epenthetic elements are not phonetically identical to similar non-epenthetic elements. The transformational
5.4 Phonological 'processes'

<table>
<thead>
<tr>
<th>Closure</th>
<th>Friction</th>
<th>=</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasality</td>
<td>Non-nasality</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Closure</th>
<th>Friction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasality</td>
<td>Non-nasality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>n</th>
<th>s</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>t</td>
</tr>
</tbody>
</table>

Figure 5.5 Declarative characterisation of epenthesis (\(\sim\) denotes non-distinctness)

analysis, however, holds that the phonetic implementation of a segment is dependent on its features, not its derivational history (‘a [t] is a [t] is a [t]’), and thus incorrectly predicts that an epenthetic [t] should be phonetically identical to any other [t]. Where epenthesis has been historically lexicalised, for example in the word *bram[b]le* (from OE *bremel*, *brembel*; cf. *braml* in various British English dialects, recorded in Wright 1905), it is not necessary to introduce an epenthetic phonological segment /b/ into the phonological representation, or to import the temporal co-ordination analysis into the phonology. It is sufficient to fix the phonetic interpretation function for just these words with lexicalised epenthesis, so that the usual under-determinacy of phonetic interpretation is, in these cases, more specific. For speakers who habitually have epenthetic [t] in *mince, pulse* etc. there is little reason to believe that their phonological representation contains a three-consonant cluster in the coda if, for example, *mince* and *mints* are nevertheless distinct. Over time, such epenthetic segments may come to be included in the phonological representation, at which stage phonetic variability would no longer be expected. This is the case in the pronunciation of *bramble* by standard English speakers, who have no sense that [b] is variable, or any different from the [b] in *rhombus*, in which it is not epenthetic.

5.4.5 Deletion

In order to ensure monotonicity (and hence recursiveness), the formalism does not permit deletion as a phonological operation. On the basis of careful first-hand observation of a number of languages, including the English and Japanese examples of 'elision' in figure 5.6, and sections 5.6.6 and 6.7 (further examples may be found in Kelly and Local 1989: 132, 139–40), I have found that in many apparent cases of deletion, a ‘trace’ of the ‘deletion’ often remains, in the form of non-local phonetic exponents of the ‘deleted’ element. This experience makes me sceptical of the acuity with which many reported cases of ‘deletion’ have been reported. (Many such
cases, of course, are reported in the literature on the basis of second-hand phonemic descriptions, which are not well suited to noting the non-local traces of deletion that I have observed. Cases of phonological zero-alternation are a separate matter that present no problem for the monotonicity of a context-free theory.) Under a type-token interpretation, what has been 'deleted' is not any part of the phonological representation, but simply one or two of the phonetic exponents of a phonological unit. Since phonetic exponents typically come in large bunches, distributed in time often over quite considerable stretches of utterance, the absence of a particular exponent is not problematic.

In many cases, phonetic deletion or "elision" is the phonetic counterpart of epenthesis (fig. 5.5), and is thus in a sense the same phenomenon. Taking the 'unelided' form as more primitive than the 'elided' form, as is done in transformational analyses, is both unnecessary and meaningless in the declarative, interpretive account. Phonologically, there are two ways of looking at deletion in Declarative Phonology. First, and of least interest, rules of the form X → 0, termed 'deletion' in the SPE tradition, do not violate monotonicity in context-free or regular grammars. (Recall from chapter 2 that such rules do not increase the weak generative capacity of context-free grammars. They are only damaging when added to context-sensitive rewriting.) Such rules license preterminal symbols dominating empty categories. In phonological representations based on context-free derivation trees, both the X and the 0 are available in the representation, and thus the term 'deletion' is inappropriate. In formal language theory, the term 'empty production' is more usual. Rules of this form can be used to express quite general grammatical possibilities such as 'the coda node may be empty'. In the more specific cases in which SPE-style deletion is used, however, such rules are rarely adequate, and a more specific form of constraint must be used.

This second possibility, like the phonological treatment of epenthesis above, represents alternation with zero directly in lexical representations.
5.4 Phonological ‘processes’

Examples from English are the alternating final /n/ of _hymn ~ hymnal, autumn ~ autumnal_ etc. In these cases, the suffix forming an adjective from a noun is -al, the regular form occurring in, for example, _tone ~ tonal_. The /n/ alternates with zero in a very specific set of lexical items. The /n/ fails to surface in the isolation form: if it did, the coda /mn/ would be a phonotactically idiosyncratic consonant cluster not attested anywhere else in the lexicon. We shall regard it as an ill-formed coda. Thus, as with the analysis of linking /l/ discussed above, /n/ manifests if and only if it can be parsed as an onset to a vowel-initial suffix. (It is necessary to formulate this constraint with care in order to ensure that it does not dock onto a following vowel-initial word in, for example, _autumn air_. This is a little tricky, but merely a technicality.) As with phonological epenthesis, then, such forms include alternation with zero in their lexical representation, for instance, /him(n)/ _hymn_. The difference between phonological ‘deletion’ and ‘epenthesis’ resides not in the representations, but in the form of the constraints that select whether zero manifests in derived or underived forms:

(5.4)  \((X)\) manifests as 0 in Traditional name
a. morphologically underived forms epenthesis
b. morphologically derived forms deletion

5.4.6 Metathesis

Again, both phonetic and phonological analyses are possible. Phonetically, metathesis is another instance of different temporal synchronisation of an invariant set of phonetic exponents (figure 5.7). A Declarative Phonological account of metathesis is possible if linear ordering is factored separately from constituent structure in phonological representations. I shall argue below that such a step is motivated by two considerations: (i) linear order is highly redundant in richly structured phonological representations. By the usual principles of generative grammar, therefore, either order or structure should be limited or even abolished. Since numerous
analyses in non-linear phonology point to the importance of constituent structure, I shall cut down on linear ordering. (ii) According to the 'Principles and Parameters' view of metrical structure discussed in section 4.4.3 above, parameters of linear order such as 'Feet are strong on the [Left/Right]' are best formalised by stating language-specific selections of linear order separately from cross-linguistic structural constraints such as 'Feet are made of syllables.'

The phonetic treatments of epenthesis, elision and metathesis proposed above may all be regarded as instances of a more general phenomenon of non-significant variability in *phasing* (the temporal co-ordination of parallel events), arising from the underdeterminacy of the phonetic interpretation mapping. From the above survey, it can be seen that the hypothesis that phonological representations are sequences of symbols (strings) generated by a rewriting system can be eschewed in favour of the hypothesis that phonological representations are immutable structures. Consequently graphs, not strings, can be employed as the central data structure for phonological representation. (The 'box diagrams' presented above simply illustrate the temporal arrangement of the exoneny of the phonological units in a structured representation. They do not constitute a formal phonological notation, which will be discussed in more detail in the next section.)

5.5 A formalism for phonological representation

5.5.1 Feature-logic and the representation of linguistic information

As phonological structures are formal representations of one type of linguistic information, I shall adapt a formalism for the representation of syntactic information to the phonological domain: the Unification Grammar formalism (Shieber 1986), with which I shall assume some familiarity. Unification Grammar is an extension of formal grammar theory to include the use of complex categories, such as sets of (possibly nested) feature-value pairs, and incorporates insights from the type-theories employed in logic and computer science.

Two dimensions of linguistic structure are represented in Unification Grammar. Firstly, syntagmatic structure, which shows how smaller representations are built up into larger representations, is represented using graphs. For instance, syntagmatic structure may be represented using trees or similar kinds of data structure. Secondly, paradigmatic structure, which shows how informative differences between representations are encoded, is
represented using feature-value structures. For instance, the syntactic distinction between singular and plural may be encoded using a feature [number] with two values 'singular' and 'plural'. The distinction between masculine, feminine and neuter gender in a language such as German or the third-person pronouns in English may be encoded using a three-valued feature [gender] with values 'masculine', 'feminine' and 'neuter'. Taken together, these two features define six categories:

\[(5.5) \quad \begin{align*} & \text{[number: singular, gender: masculine]} \\
& \text{[number: singular, gender: feminine]} \\
& \text{[number: singular, gender: neuter]} \\
& \text{[number: plural, gender: masculine]} \\
& \text{[number: plural, gender: feminine]} \\
& \text{[number: plural, gender: neuter]} \end{align*} \]

Categories are usually written vertically, with features concatenated into columns, as the sequence of features is not an informative part of the representation. I shall employ comma-separated lists of features occasionally, however, for brevity. Features and their values are in practice rarely given long names, and abbreviations are common. For reasons which will become clear, feature-value pairs are usually written in the order [feature: value]. The opposite order, such as [\(+\text{voice}\)] or [3 \text{ person}] may be used as a short form.

Within a category, each feature can only have one value. This requirement ensures that all categories can be given a sensible interpretation, and that inconsistent categories such as those in (5.6) do not occur.

\[(5.6) \quad \begin{align*} & \text{[number: singular]} \quad \text{or} \quad \text{[+\text{voice}]} \\
& \text{[number: plural]} \quad \text{or} \quad [-\text{voice}] \end{align*} \]

The restriction to single values is imposed by requiring that categories are \textit{functions} from the set of features to the set of values. Since in any category not all features have to be assigned a value, they are partial, not total, functions. This leads declarative grammarians to make a distinction between \textit{description} and \textit{object}. The partial phonological structures encountered as lexical items and as grammatical constraints alike are descriptions of classes of actual linguistic objects. For example, the lexical representation /him(n)/ discussed above is a description consistent with both the isolation form /him/ and the derived form /himnal/. (An alternative kind of bracketing could be used to clarify this distinction, but in general we shall be concerned only with descriptions, not with the objects that they license, the actual phonological representations.)
5.5.2 Logical descriptions

Employing the description/object distinction enables us to write expressions that do not correspond to any specific phonological form. For example, there is no single lexical item /him(n)/ in which the /n/ is variably present or absent. Another example: we could represent the stem of electric, electricity and electrician, in which the stem-final consonant is [k], [s] or [ʃ], by employing the disjunction "[k] or [s] or [ʃ]" as part of the lexical representation (informally, /elektrɪ{k,s,ʃ}/). Because Declarative Phonology encodes many alternations in the lexicon, it is more lexicalist than other generative approaches.

The requirement that each feature have only a single value does not prevent several values being combined in some way, and then employed as a 'complex value'. For example, if it is known that a word is either dual or plural in a singular/dual/plural system, the disjunction operator ∨ may be used to form the complex feature-value pair [number: dual ∨ plural], equivalent to [number: dual] ∨ [number: plural]. An equivalent method of describing the same set of values could employ negation. For instance: [number: ¬ singular], equivalent to ¬ [number: singular]. Though the syntax of these descriptions is different, the categories described, [number: dual], [number: plural] and all extensions of them, are the same.

These formal resources are all available in standard generative phonology. For example, disjunction is normally represented by curly brackets; conjunction is expressed 'silently', simply as co-occurrence of structural elements in a representation or rule; negation is represented in the expression of filters such as *[+nasal, ¬voice]. These logical resources alone are the basis of a powerful description language. Even the use of unification is not especially novel in generative phonology, as pattern-matching conformant with non-contradiction of feature-structures is employed in SPE to determine whether a rule applies to a representation. Partial descriptions have always been used in generative phonology, for generality. Other binary operators sometimes used in Unification Grammar include union, by which set-values may be constructed, dot, by which list-values are constructed, and concatenation, by which string-values are constructed. Apart from the list constructor, dot, these combinators are also employed more generally in generative grammar. Thus Declarative Phonology can be seen as employing the logical operations already proposed in generative grammar, in a more extended range of circumstances and with the purpose of logic-based description more explicitly stated than hitherto. Any phonological theory must have a certain ability for expressing logical propositions involving
negation and conjunction, which, as Quine (1981) shows, is a sufficient basis for the expression of all truth-functional modes of composition, including disjunction, implication and, as we shall see, existential and universal quantification. Since such simple logic is very powerful, we cannot avoid the ability to express a much greater range of grammatical constraints using feature-logic than has typically been done in generative phonology before. In earlier work the expressive capabilities were not recognised and put to use, with the effect that all kinds of additional, undesirable and unnecessary additions to the formalism were proposed, such as powerful transformations. A final use of disjunction in Declarative Phonology is disjunction of rules. (I refer here to exclusive disjunction, \( \boxtimes \), defined as \( A \boxtimes B \) if and only if \( (A \lor B) \land \neg (A \land B) \), i.e. A or B, but not both.) Disjunctive rule application was possible in SPE phonology in certain circumstances: in fact, it was favoured. For example, Chomsky and Halle (1968: 63) propose that 'abbreviatory notations must be selected in such a way as to maximize disjunctive ordering . . . This principle seems to us a natural one in that maximization of disjunctive ordering will, in general, minimize the length of derivations in the grammar.' Conjunction of constraints is commonplace: for example, every complex constraint can be broken down into a number of conjoined subclauses, and normal 'feeding' order of constraint interaction is simply logical conjunction. It is not surprising that disjunction of constraints appears to be necessary too. A clear case is the selection of two contradictory possibilities in a parameterisation scheme, such as 'Feet are maximally binary or unbounded', meaning 'Only one of the following constraints holds: (i) feet are maximally binary; (ii) feet are unbounded.' Of course, a foot could be binary even if the 'unbounded' option has been selected in a grammar expressed in this form, because such short feet are licensed either way. But the disjunction between the two constraints is absolute in a given grammar. The use of disjunction should be avoided where possible, since, as Manaster-Ramer and Zadrozy (1990) note, arbitrary statements may be disjoined, making disjunction the antithesis of generalisation.

Another example, discussed in Scobbie, Coleman and Bird (1996), is counter-feeding rule interaction. Consider a grammar containing rules expressing a chain shift:

\[
\begin{align*}
th & \rightarrow d \\
d & \rightarrow z
\end{align*}
\]

(5.7) The intention of this grammar is that stops weaken in the following way: voiceless stops become voiced and voiced stops become fricatives. A
number of synchronic examples of such a pattern could be given, such as mutation in North Welsh. In a transformational phonology, there is a pitfall: if both rules occur in a grammar, we need to prevent (5.7a) from feeding (5.7b), since it is not the case that /t/ alternates with /z/. In the derivational approach, it is necessary to explicitly order (5.7b) before (5.7a), though this is a purely technical move to prevent an undesirable side-effect of the fact that rule ordering is used to denote conjunctive application. In Declarative Phonology, disjunction permits a number of possibilities. For example, Scobbie, Coleman and Bird (1996) show that by encoding the alternations using disjunctions such as \{t, d\} and \{d, z\} in lexical representations, the question of rule ordering does not arise, as lexical representations do not 'interfere' with one another anyway. Alternatively, in the system of phonological constraints proper, one could simply explicitly state that (a) and (b) are disjoint: that is,

\[(5.8) \quad (t \rightarrow d) \otimes (d \rightarrow z)\]

This requires stipulation, as in the rule-ordering analysis, but does not actually employ extrinsic ordering, and is hence consistent with Declarative Phonology. A third alternative builds on the distinction between phonology and phonetics. If (5.7a) and (5.7b) are rewritten as phonetic interpretation rules, as in (5.9), the problem of undesired feeding does not arise, because the items on the right-hand side are terminal symbols.

\[(5.9) \quad /t/ \rightarrow [d] \quad /d/ \rightarrow [z]\]

5.5.3 Negation and feature-switching

A feature [f] whose value is undefined may be represented [f: T]. ('T', denoting 'true', contributes no additional information, and is consistent with all values of [f].) A category-valued feature whose value is undefined may be represented [f: []]. The categories in which [f] is undefined are denoted ~[f]. A feature [f] which is known or required to have some value, though perhaps 'unknown' until supplied by some other piece of information, can be represented [f: ANY], that is, [f: T]. The ways of referring to the single value that a feature may take may be augmented by unary operators. For instance, the negation prefix ~ (or *) can be used to prohibit a feature from having a stated value. Combining ~ with ANY provides a way of referring to features that do not bear any value: [~f: ANY]. A constraint C1 referring to [~f: ANY] differs from a similar formulation C2 referring to [f: T] or [f: []] in that if T or [] are filled in, C2 is not violated, but C1 is. (See example (7.96) below.) The unification of a negated feature or category with a non-negated
feature or category is given by equation (5.10), read "*A unifies with B if and only if A does not unify with B."

\[(5.10) \quad \star A \sqcup B \leftrightarrow \neg (A \sqcup B)\]

In a binary feature system, \(\neg\) is equivalent to the alpha-switching prefix \(-\), as in expressions such as \([- \alpha\) voice\]. Permutation operators or one-to-one mapping functions which map each value in the range of a feature into some other value in the same set may be especially useful in modelling apparently systematic 'feature-changing' operations, such as cyclic vowel shift, without the need for transformations. I shall give an example of this in section 7.1.4 below.

### 5.5.4 Category-valued features and feature equations

The value of a feature does not have to be an atomic symbol such as 'singular': it may be a category. For instance, the present tense of the irregular verb *to be* displays a variety of inflected forms:

\begin{equation}
\begin{array}{ll}
\text{number: sg} & \rightarrow \text{am} \\
\text{person: 1} & \\
\text{number: sg} & \rightarrow \text{are} \\
\text{person: 2} & \\
\text{number: sg} & \rightarrow \text{is} \\
\text{person: 3} & \\
\text{number: pl} & \rightarrow \text{are}
\end{array}
\end{equation}

In order to express the constraint that subject and verb must agree in these features, it is convenient to bundle them up as the value of a higher-order feature \([agr]\); thus:

\begin{equation}
[agr: [\text{number: sg, person: 1}]] \rightarrow \text{am}
\end{equation}

If \([agr]\) is also defined for nominal categories, then subject–verb agreement amounts to the condition that the value of the feature \([agr]\) is identical for the subject and the verb. Such an identity condition can be enforced using an equation such as (5.13):

\begin{equation}
\text{subject:agr} = \text{verb:agr}
\end{equation}

I shall show below how such sharing of feature structures can be used to model 'phonological agreement' such as soft coarticulation and assimilation. Dissimilation constraints can be expressed using inequality, rather than equality, as an operation on structures. For example, inequality (5.14)
Declarative Phonology requires that category-valued features \([u]\) and \([v]\) (which may derive from separate constituents) must not have the same value of feature \([f]\).

\[(5.14) \quad u:f \neq v:f\]

5.5.5 Implication

A system of logic containing conjunction and negation is powerful enough to express implications, that is, conditional statements 'if \(X\) then \(Y\)', conventionally written '\(X \supseteq Y\) or \(X \implies Y\)'. Such an expression is truth-conditionally equivalent to '\(\neg(X \land \neg Y)\)'. For example, the statement 'if it rains tomorrow, I shall stay inside' is equivalent to 'it is not the case that (a) it rains tomorrow and (b) I don't stay inside'. Implicational statements have been employed in generative phonology in two guises in the past. First, some rather complex rules in \(SPE\) phonology used arrangements of angle brackets and conditional constraints to collapse several partially overlapping cases of rules into a single statement. For example, Chomsky and Halle (1968: 181) give the following rule:

\[
\begin{align*}
V \rightarrow [+\text{tense}] & / \left\{ \begin{array}{l}
[\alpha \text{ low}] \\
[\beta \text{ stress}] \\
[\#, \text{ where } \beta = + \text{ if } \alpha = +] \\
[-\text{high}] \\
C_{\text{I}} \end{array} \right. \\
\end{align*}
\]

The second use of implication is in redundancy rules such as (5.15):

\[(5.15) \quad [+\text{nasal}] \rightarrow [+\text{voice}]\]

Although (5.15) is conventionally rewritten in the form of a rewriting rule, it does not mean that the feature [+nasal] is removed and the feature [+voice] inserted in its stead. If the [+nasal] segment bears the value [−voice], (5.15) will change that value to [+voice]. But redundancy rules differ from rewriting rules in that they fill in redundant features, which by common practice are omitted from lexical representations. At the time at which (5.15) applies, the [+nasal] segment typically bears no value for [voice]. Redundancy rules do not change any feature values, but simply fill them in. They thus have the character of implications. Rule (5.15) could be logically expressed as (5.16):

\[(5.16) \quad [+\text{nasal}] \supseteq [+\text{voice}]\]

As I showed above, such a statement does not need to be conceived as a mapping from strings at one level in a derivation to strings at a different
level. It is logically equivalent to the constraint $\neg(\text{[+nasal]} \& \neg\text{[+voice]})$, holding of a single level. Even in a derivational framework, redundancy rules have this character of logical conditionals true at each step in the derivation. In order to eliminate sequential rule application, Declarative Phonology builds on this idea by reanalysing feature-changing rules as feature-filling rules wherever possible. These rules are then seen to be conditionals of a single level of phonological representation.

5.5.6 Quantification

Logical conjunction and disjunction are also a basis for universal and existential quantification ('for all . . . ' and 'there exists . . . '). For instance, if $x$ is a variable ranging over the set of instances $\{x_1, x_2, \ldots, x_n\}$, the logical formula $\forall x \ P(x)$ ('all instances of $x$ have the property $P$') means the same as $P(x_1)$ and $P(x_2)$ and $\ldots$ $P(x_n)$. In similar fashion, the existential form $\exists x \ P(x)$ ('some instance of $x$ has the property $P$') is logically equivalent to $P(x_1)$ or $P(x_2)$ or $\ldots$ $P(x_n)$, where 'or' is inclusive (that is, more than one instance of $x$ might possibly have property $P$).

Quantified statements are commonplace in phonology. In general, grammatical generalisations involve universal quantification: for example, 'every foot contains one or more syllables'. Other constraints employ existential quantification. For example, Bird (1995: 64) characterises prosodic licensing constraints as having the general form 'for every node of type $x$, there is some node of type $y$ which dominates it'. (The link between licensing and domination is not crucial to this illustration: other kinds of licensing relation have the same general form.) Scobbie, Coleman and Bird (1996) propose that while grammatical constraints or rules are universally quantified, lexical entries (which are also constraints) are existential statements of the form 'there exists a word such that . . . (e.g. it has two syllables, it begins with a /p/, stress falls on the second syllable etc.)'.

5.5.7 Types, subsumption and the Elsewhere Principle

The remaining principal contribution of Declarative Grammar to phonological theory is the idea that feature-structures are typed or sorted and form a hierarchy of information from maximally general (least informative) to maximally specific (most informative). Typing is in part a matter of logical hygiene: as Scobbie, Coleman and Bird (1996) express it, 'we need to specify somewhere in our grammar that a vowel is the sort of entity which can be [round] or [unrounded] but not [nominative] or [animate]'. But typing gives us more than this: 'Once we have defined a vowel to be the sort of entity that expresses rounding, . . . [rounding] means [round] or [unrounded]. In other
words, a vowel is "unspecified" for rounding in a different way than a noun: rounding is an appropriate sort of information for a vowel, but is inappropriate for a noun' (Scobbie, Coleman and Bird 1996). A scheme of types can be used to establish a hierarchy of relationships between types (see Pollard and Sag 1987, ch. 8). For instance, following Bird (1995: 59), 'C' and 'V' are subtypes of the type 'segment', 'k' a subtype of 'C', 'syllable' and 'mora' are not in a subtype relation, though one-mora (light) and two-mora (heavy) subtypes of syllable could be set up in a system of types.

Like other aspects of phonological representation, types denote sets of representations. For example, the type 'syllable' refers to the set of representations denoting syllables, whereas the subtype 'light syllable' refers to a subset of the set of syllables. The designation 'light syllable' is more specific, less general and more informative than the designation 'syllable', because the set of light syllables is a proper subset of the set of syllables. The same kind of information hierarchy can be seen in feature-structures, even without a typing scheme. For example, the description [-sonorant] denotes one class of phonological objects, of which the description [-sonorant, +coronal] picks out a subset. The informativeness of the latter description can be compared to the former without even having to count the number of representations that each description licenses because of a special relationship which obtains between them: the feature-structure [-sonorant] is contained in the feature-structure [-sonorant, +coronal]. We say that the more general description [-sonorant] subsumes the more specific description [-sonorant, +coronal], because the set of representations that the more general description denotes includes the representations denoted by the more specific form (see Shieber 1986: 14–16; Gazdar et al. 1985 say that B is an extension of A, for A subsumes B). Because of this deep-rooted relationship between the syntactic property of containment of one feature-structure within another and the corresponding semantic notions of informativeness, generality and specificity, subsumption is an important and powerful relation. Recognition of its role in grammar is a major contribution to generative grammar. As with types, subsumption defines a natural hierarchy from very general to very specific descriptions.

Subsumption provides a natural way in which to express Kiparsky's Elsewhere Principle (Kiparsky 1973). This principle holds that where two rules both apply to a representation with conflicting demands for structural change, and where the structural description of one rule refers to a specific instance of the structural description of the other, the more specific rule has priority over the more general rule. For example, both of the following
redundancy rules are natural, as voiced obstruents are marked, whereas sonorant consonants (which do not normally enter into the voicing opposition) are usually voiced:

\[(5.17)\]
\[
a. [+consonant, +sonorant] \rightarrow [+\text{voice}]
\]
\[
b. [+\text{consonant}] \rightarrow [-\text{voice}]
\]

Informally, these rules can be read 'sonorant consonants are voiced, but otherwise consonants are generally voiceless'. When contexts are added to such rules, sets of rules with related structural descriptions can be glossed 'in context X, rule Y applies, but elsewhere, rule Z applies'; hence the name 'Elsewhere Principle'.

Within the terms of transformational phonology, Kiparsky and his intellectual successors interpreted this as a principle of rule ordering, and placed more general rules later in the grammar. However, as stated here, it can be seen as a statement about prioritisation which applies to adjudicate multiple possible rule assignments. In the trivial example of rules (a) and (b) it can be seen that if (a) is not prioritised over (b), in a declarative grammar it can never do any work, since if (b) has priority, all consonants unmarked for [voice] will be given the feature [−voice], to which (a) can then not apply, because of the prohibition against removing or rewriting parts of representations. The Elsewhere Principle hardly needs to be explicitly stated in Declarative Phonology, for there would be little point in including more specific constraints in the grammar if they could be scuppered by a more general constraint. In such a case, the more specific constraint would do no work, and ought to be eliminated.

Any plausible phonological theory needs to have a mechanism for adjudicating between the demands of conflicting constraints in such cases. Every theory needs to have a way of expressing exceptions to generalisations, that is, a way of granting priority to the more specific constraints. In derivational generative phonology this is done in a very inessential way, using exception features. In that approach, the Elsewhere Principle did not appear to follow from more general principles of grammar, such as subsumption. In Declarative Phonology, we hypothesise that besides silent conjunction and explicit disjunction of rules, 'elsewhere' prioritisation is the only principle of rule prioritization.

5.5.8 Syntagmatic structure

The fact that feature-value categories are partial functions means that they can also be represented as a particular kind of graph: a directed, acyclic
graph, or dag. Dags are like trees, except in three respects: (i) dags are unordered; (ii) nodes may have more than one mother; (iii) therefore, unlike trees, dags may have more than one root. I shall argue that these three properties hold of phonological representations, too. First, order is in general redundant in structured phonological representations (Cheng 1971). This may sound surprising in view of the fact that order is built into practically all other approaches to phonological representation, but this practice follows from the inappropriate use of 'phonetic' transcriptions as phonological representations. Consider the examples from Japanese introduced in chapter 3. As noted in chapters 3 and 4, Japanese words generally consist of sequences of consonant–vowel dyads called moras, for instance ka, shi, tsu and so on. The order of consonant and vowel is completely predictable in a phonological representation of these objects, and is thus redundant. In languages with more complex syllable structures, such as English, the order of consonants and vowels is predictable if syllables are given hierarchical representations such as figure 5.8. In such a representation, the order of nucleus /i/ with respect to coda /t/ is completely predictable: since codas always follow nuclei, this redundant information should be expressed as a general ordering principle (nucleus < coda), rather than encoded in each phonological representation. Likewise, onsets always precede rimes, and so the relative order of /p/ and /it/ is completely predictable. Even in consonant clusters and diphthongs, the order of the parts follows general principles (see 5.6.5, 7.1.9 and 7.1.11 below), and thus their order has no more place in a phonological (lexical) representation than does representation of, say, [+voice] with every occurrence of [+nasal], instead of the general statement that nasality is generally accompanied by voicing in English. Thus in general unordered graphs are more
suitable for phonological representation than ordered graphs (with the addition of ordering statements only in those parts where order is distinctive, such as the distinction between /sC/ and /Cs/ clusters in English codas). In Declarative Phonology, unordered phonological structures may be defined by immediate dominance rule-schemata (as in the formalisation of metrical parameters in chapter 4). Ordering generalisations may be defined by linear precedence constraints.

The second property of dags, that some nodes (called **re-entrant** nodes) have more than one mother, is also applicable to phonological representations. For instance, the phenomenon of ambisyllabic consonants (consonantal forms which occur only at the junction between two syllables, such as the [ɹ] in the pronunciation of words like butter [b̪əɹ] or Betty [bɛɹ]) in some English dialects is most felicitously analysed in a Declarative theory as being constituents of two syllables at once. They are neither the coda of the first syllable (since isolation pronunciations such as [b̪əɹ] for 'but' or [bɛɹ] for *bet* are not found), nor are they the onset of the second syllable, since [ɹ] is never otherwise found syllable-initially. In addition, in their phonetic interpretation, they have a transition into the tap like a coda, and a transition out of the tap like an onset. The simplest Declarative analysis is that they are simultaneously both coda of one syllable and onset of the next. (An alternative analysis proposed in Kiparsky (1979) and Nespor and Vogel (1986) is that such elements are in one place in structure at one step in the derivation, and at another at a later step in the derivation. Such 'movement' or 'resyllabification' rules, however, are not expressible in this monotonic framework, which is more constrained, and contains no structure-changing operations. Some linguists regard 'improper bracketing' with horror, but providing the number of mothers a node can have is finitely bounded, it is nowhere near as damaging to the goal of producing a constrained grammar formalism as the use of structure-changing rules.) The representation of this phenomenon in figure 5.9 assigns more than one mother to such nodes. 'Sharing' or spreading of features across adjacent non-sisters, as in cases of assimilation, admits of a similar representation (figure 5.10). Having more than one root in a phonological representation may also be desirable. Semitic languages, for instance, have a non-concatenative inflectional morphology in which vowels and consonants form disjoint, but intercalated, morphemes. As is well known, Arabic consonantal morphemes such as *k-t-b* may be intercalated with different vowel patterns that distinguish singular from plural, active from passive, and so on. The separate consonant and vowel morphemes may be appropriately
Figure 5.9 Ambisyllabicity represented by a re-entrant graph (butter)

Figure 5.10 Assimilation represented by a re-entrant graph (it put)

represented with disjoint graphical structures with separate roots and intersective fringes:

inflection

k i t a b

stem

(Bird and Blackburn (1991) present a Declarative analysis of Arabic morphology which employs conventional single-rooted prosodic structures.)

The two types of representation discussed above (columns of feature-
value pairs and directed, acyclic graphs) are fully equivalent. A column of feature-value pairs can be represented as an equivalent unordered, hierarchical, tree-like diagram as follows. Each feature-value pair is represented as an arrow, the body of which is labelled with the name of the feature, and the head of which is labelled with the value. Thus [−grave] may be represented:

(Some of the features used in my phonological analyses are no longer widespread in other varieties of generative phonology, especially Jakobsonian features. However, the declarative hypothesis does not rest on any particular choice of feature-names or feature-structures, and Declarative Phonology makes no substantive claims in this regard. Arguments for my specific choice of features are presented in the analyses in chapters 6 and 7.) Following customary practice, I shall orient such diagrams so that all arrows point down the page: this convention enables the arrowheads to be omitted, as in conventional tree-diagrams, if it is convenient or desirable to do so. Sets of feature-value pairs are graphed by arrows with shared tails. For instance, a consonantal place-of-articulation category such as [−grave, −compact] can be represented as in figure 5.11. In Declarative Phonology, some features, such as [consonantal] (location of a consonantal obstruction), take other categories, rather than simple + or − values, as their value. These are represented by rooting a category at the head of the category-valued arrow, as in figure 5.12, for example. The proposals of Clements (1985), Hayes (1990) and others concerning the internal structure of segments are thus provided with a formal instantiation (see also Bird 1991, 1995: 73–5; Broe 1992). Also, the use of category-valued features allows some constituents of a structure to be the value of a feature [head]. This is one way in which headed structures can be represented. Root-to-frontier
paths through a category are expressed using : as an associative infix operator. For instance, in figure 5.12, consonantal:grave = − and consonantal:compact = −.

In the fragment of English phonology in chapter 7, structured categories are extensively used. For example, every simple phonological constituent is composed of up to three category-valued features: [consonantal], [vocalic] and [source], corresponding roughly to the category-valued features [Place], [Dorsal] and [Laryngeal] in feature-geometry theory (see Broe 1992). The feature [consonantal] takes a consonantal place-of-articulation category as its value, [vocalic] takes a vocalic place-of-articulation category as its value, and [source] takes a combination of various other features, such as [voice], [continuant] and [nasal] as its value. For instance, the dark lateral /l/ is represented equivalently as in (5.18) or figure 5.13. (Although conven-
tional segmental categories have no status in the present theory, I shall
employ letters in slashes, such as /H/, as shorthand abbreviations for feature-
structures.)

\[
\begin{align*}
\text{consonantal:} & \quad \begin{cases}
\text{compact:} & - \\
\text{grave:} & -
\end{cases} \\
\text{vocalic:} & \quad \begin{cases}
\text{grave:} & + \\
\text{compact:} & +
\end{cases} \\
\text{source:} & \quad \text{[nasal:} & - \text{]}
\end{align*}
\]

Other analyses using feature-structures are possible, of course. In
Dependency Phonology (see Lass 1984: 290), segments are analysed into
three ‘gestures’: categorial, initiatory and articulatory. The feature-struc-
ture formalism developed in Unification Grammar is equally suitable for
the formal representation of such alternative proposals. The existence of
[consonantal] and [vocalic] nodes in my analyses of English and Japanese
is defended on language-internal phonological grounds in chapters 6 and
7. (The feature [source] is simply an \textit{ad hoc} place-holder node under which
to group together all features not included in [consonantal] and [vocalic].)

A number of details of figure 5.13 should be noted. (i) Use of category-
valued features [consonantal] and [vocalic] allows the features [grave] and
[compact] to be used with different values regarding the location of primary
and secondary articulation. For instance, the primary articulation of /H/ is
alveolar [−grave, −compact], whereas the secondary articulation is velar
[+grave, +compact]. (ii) Some features used elsewhere in the phonology,
such as [round] or [voice], do not occur in this category, since they are not
distinctive in the structural positions in which this category may occur (in
English). In other words, categories may be underspecified. (iii) Liquids
such as /l/ and /r/ are those categories which distinctively bear both [conso-
nantal] and [vocalic] subcategories. Obstruents do not distinctively bear a
[vocalic] subcategory in English, though in languages with distinctive sec-
ondary articulation in obstruents, such as Russian and Irish, obstruents
will have [vocalic] features. Vowels do not bear a [consonantal] subcategory,
and /h/ bears neither [consonantal] nor [vocalic], only [source]. In other
proposals regarding phonological classification, it is not possible for a
feature to have more than one value anywhere in a structure, a more restric-
tive claim than the unification-based theory of category structure enforces.
Since primary and secondary articulations may indeed have ‘opposite’
values in some distinctive dimensions (such as ‘back-velar closure with palatal resonance’, or ‘rounded consonant coarticulated with an unrounded vowel’, as in the onset of English she or ray, or ‘velarized palatal closure’), the employment of the ‘same’ [grave], [compact] and possibly [round] features as subcategories of both [consonantal] and [vocalic] is preferable to theories in which such restricted cases of ‘contradictory values’ cannot be expressed. For example, the use of quite different features for place of articulation of consonants and vowels (including ‘secondary articulations’ such as palatalisation) in Chomsky and Halle’s (1968) feature-system fails to capture cross-categorial relations such as /v/ ~ /u/ (e.g. revolve/revolution, solve/solution), /l/ ~ /j/ (e.g. garden/yard) or /l/ ~ /w/ (e.g. guard/ward). In Clements (1985), the category-valued feature DORSAL, which does partially cross-classify consonantal and vocalic place of articulation, cannot express palatalised velars (which would require ‘palatalised’ [dorsal: [back: -]] to be combined with ‘velar’ [dorsal: [back: +]]) or velarised palatals, because no attention is paid to the distinction between the largely independent primary and secondary (or consonantal and vocalic) aspects of articulation. (This problem is ameliorated in Clements (1991) by the establishment of a separate V-place node.) Thus, my use of the category-valued features [consonantal] and [vocalic] is primarily motivated by the need to assign conflicting values to the same features ([grave] and [compact]) within a single category. Evidence that the same features need to be used with different values to categorise consonantal obstruction and vocalic resonance is provided by the work of Jakobson (see Jakobson, Fant and Halle 1952: 28, 30; Jakobson and Waugh 1979: 92–111, especially 105–7, where McCawley’s criticisms of Jakobson’s features, which formed part of the motivation for Chomsky and Halle’s revised system of features, are addressed). Chomsky and Halle’s (1968: 307) observation that Jakobson’s feature-system needs to use separate features such as [sharp] and [flat] to characterise secondary palatality, velarity and pharyngeality is addressed by my proposal to use [grave] and [compact] within independent [consonantal] and [vocalic] subcategories. That is, Jakobson’s [flat] will be represented as [vocalic: [grave: +]] and [sharp] will be represented [vocalic: [grave: -]]. This proposal addresses Chomsky and Halle’s main criticism of Jakobson’s features, while remaining consistent with Jakobson’s arguments that the same features cross-categorise consonantal and vocalic articulation, as well as resolving the problems with Chomsky and Halle’s features raised by Campbell (1974). More recent evidence for independent [consonantal] and [vocalic] subcategories may be
5.5 Formalism for phonological representation


Category structure, syllable structure, higher levels of phonological structure, and other levels of structure, such as morphological and syntactic structure, are fully integrated in this framework. Thus, as a coda, \( /H/ \) is represented as in (5.19) or figure 5.14.

(5.19) \[
\begin{align*}
\text{coda:} & \quad [\begin{array}{c}
\text{consonantal:} \\
\quad [\begin{array}{c}
\text{compact: } - \\
\text{grave: } - \\
\end{array}] \\
\text{vocalic:} & \quad [\begin{array}{c}
\text{grave: } + \\
\text{compact: } + \\
\end{array}] \\
\text{source:} & \quad [\text{nasal: } -]
\end{array}] \\
\end{align*}
\]

The syllable full can be (partially) represented as in figure 5.15. The following details of figure 5.15 should be noted: (i) [voice] only occurs in the onset:source. This is because it is distinctive at this ‘place in structure’, but not in nuclei or /H/. (ii) The [vocalic] value of the nucleus is shared with the [vocalic] value of the onset (nucleus:vocalic = onset:vocalic). This is because ‘true consonants’ such as voiceless fricatives bear no [vocalic] specification of their own in English, but are coarticulated with the nucleus.
with which they are tautosyllabic, by sharing its [vocalic] specification. Such coarticulation is in part phonological, not phonetic, since it varies from language to language and therefore cannot be phonetically determined (though when it does occur, it has an obvious articulatory basis). A finer level of mechanical or ‘hard’ coarticulation, which is not linguistically controllable, is also modelled in the phonetic interpretation component.

The nodes of these graphs are labelled with complex phonological categories, which are themselves directed, acyclic graphs. This development enables phonological oppositions to be expressed between constituents of any category, not just over terminal symbols. So, for instance, it is possible to express contrasts between rounded vs. unrounded syllables, nasal vs. non-nasal rimes, voiced vs. voiceless clusters, as well as the more conventional oppositions such as high vs. low vowels or velar vs. alveolar consonants.

Each position in structure, in other words each node in the phonological graph, is the locus for systems of oppositions between distinct utterances. Phonological oppositions may be expressed between constituents of any type. For instance, in morphophonological representations, phrasal stress is located at phrasal nodes; compound stress is represented at the com-
pound word node; the stress of morphemes of Greek or Germanic origin is represented at the Greco-Germanic Stress (level 2) node; Latinate main and secondary stresses over the Latinate Stress domain (level 1); nasality, front-ness, rounding and voice are onset- and rime-level features; place of articulation is a terminal-domain feature; and vowel-height is a nucleus-domain phonematic feature.

The 'sharing' of a category by two or more category-valued features is denoted using integer indices. For instance, the syllable in figure 5.15 can be partially represented as in (5.20).

(5.20)

\[
\begin{array}{l}
\text{onset:} \quad \text{consonantal:} \ldots \\
\quad \quad \text{vocalic:} [ ] \\
\text{rime:} \quad \text{nucleus:} [ \text{vocalic:} [ \text{height: close} ] ] \\
\quad \quad \text{coda:} \ldots
\end{array}
\]

In (5.20), the label [ ] is used to index the category [height: close, +grave]. Example (5.20) is exactly equivalent to (5.21):

(5.21)

\[
\begin{array}{l}
\text{onset:} \quad \text{consonantal:} \ldots \\
\quad \quad \text{vocalic:} [ \text{height: close} ] \\
\text{rime:} \quad \text{nucleus:} [ \text{vocalic:} [ ] ] \\
\quad \quad \text{coda:} \ldots
\end{array}
\]

Examples (5.20) and (5.21) are not equivalent to (5.22), however, in which two different tokens of the category [vocalic: [height: close, +grave]] occur, not one shared token.

(5.22)

\[
\begin{array}{l}
\text{onset:} \quad \text{consonantal:} \ldots \\
\quad \quad \text{vocalic:} [ \text{height: close} ] \\
\text{rime:} \quad \text{nucleus:} [ \text{vocalic:} [ \text{height: close} ] ] \\
\quad \quad \text{coda:} \ldots
\end{array}
\]
Although dags are somewhat more complex than trees, in that their nodes may be multiply dominated, established techniques for describing or constructing trees and parts of trees may be extended to the description and construction of dags. Dags differ from trees in three ways: their constituents are unordered, their nodes may have more than one mother, and they may have more than one root. To be set against these properties, however, are the following observations. (i) Constituent order in trees is merely conventional in linguistics. It is not explicitly notated (it is implicit in the left-to-right order of nodes) and in graph theory trees are not necessarily ordered. Furthermore, pictures of dags must be conventionally ordered in practice if they are drawn on the page or represented in a computer program, for instance as Prolog terms. (ii) Both trees and dags primarily denote the relation 'is a constituent of'. (iii) A collection of trees (a forest), like a dag, also has more than one root.

Dags may be represented and constructed using a small extension of the techniques used to represent and construct trees. Specifically, distinguishing between trees and ordered trees, we can state that a dag is formed from one or more unordered trees, plus possibly some 're-entrancy arcs' for those nodes dominated by more than one mother. This can be summarised in the informal formula:

(5.23) \( \text{dag} = \text{unordered forest} + \text{re-entrancy arcs} \)

More formally, a dag can be analysed as the join of its depth-first spanning forest with its back-edges (see Gibbons 1985: 20), where the back-edges are the sharing arcs. Observe that the dag in figure 5.16 is equivalent to the join of the two dags in figure 5.17 at XY. This is done by unifying X and Y, by adding the equation X = Y to the descriptions of the two dags. Since each of the two component dags is an unordered tree, the re-entrant dag of figure...
5.5 Formalism for phonological representation

5.16 can be represented as the combination of two trees and an equational constraint, as in figure 5.18. This type of sharing is employed in the construction of larger phonological structures from smaller ones. Phrase-structure rules are used to build trees, and equations are used to join them together into re-entrant dags. Local (1992) argues that this type of representation is appropriate in the analysis of assimilation. As another example, figure 5.19 is a representation of the disyllabic word cousin, which illustrates syllable structure, (minimal) metrical structure, features at non-terminal nodes, and sharing of constituents (ambisyllabicity).

5.5.9 Heads and feature propagation

A commonly employed distributional test of headship is obligatory occurrence. Constituents which must occur are heads, whereas optional constituents are dependents. But in some cases, it is not immediately clear which constituent of a particular structure is the head. The various diagnostics of headship may be mutually contradictory. For example, in English, short vowels do not occur in isolated stressed, open syllables. This fact could be
Figure 5.19 Part of a non-segmental phonological representation of the word cousin

accounted for in several ways. For instance, we might consider that the nucleus and coda of short, stressed syllables are both heads. If both are obligatory, short, stressed syllables without a coda would be ill formed. (Two-headed structures have been proposed in some linguistic analyses. For instance, the GPSG account of conjunction (Gazdar et al. 1985: ch. 8) has both conjuncts as heads, an analysis which is prompted by the manner in which head features are propagated in that framework.) However, this account of short-vowel distribution does not explain why coda constituents are heads when the nucleus is short. An alternative mechanism for accounting for this distribution is to encode a feature agreement constraint between nucleus and coda. For instance, if short vowels are marked [+checked], and empty codas are marked [-checked], a constraint on rime structure which required nucleus and coda to agree in the value of the feature [checked] would prohibit short vowels from co-occurring with an empty coda, that is, in open rimes.

One of the early motivations for proposing the head/non-head distinction was to allow for expression of the observation that an object which is
5.5 Formalism for phonological representation

composed of smaller constituents is characterised by the same sort of properties as some, but not all, of those constituents. For example, in syntax, noun-phrases may bear grammatical properties characteristic of simple nouns, such as number, person and gender, while verb-phrases may bear typical verb characteristics such as tense. In X-bar syntactic theories, phrasal and their head daughter categories also share the categorial properties [± N] and [± V] (Chomsky 1972: 52, Gazdar et al. 1985: 50). Defining one daughter of each category to be the head distinguishes the relationship between that daughter and the mother as special, and paves the way for general principles of feature propagation such as the Head Feature Convention (Gazdar, Pullum and Sag 1982). Propagation can be regarded as an informal procedural interpretation of the effect of feature agreement constraints holding between an element and one (or more) of its constituents. Feature propagation mechanisms may be used to encode long-domain phonological phenomena, such as feature-spreading. For example, applying the constraint Coda Features = Rime Features to the structure of figure 5.20 forces [Rime Features] and [Coda Features] to be identical. If the coda features are evaluated before the rime features, the constraint produces the effect of percolation of the coda features up to the rime node. Conversely, if the rime features are determined before the coda features, the constraint gives rise to the effect of the rime features trickling down to the coda node. If both sets of features are determined together, the constraint appears to check that they agree, whereas before either set has been determined, the constraint ensures that when one of them is evaluated, the other will automatically be filled in through either percolation or trickling. If the temporal interpretation of the mother node is of greater extent than that of the daughter node, percolation can also be regarded as bringing about the extension in time of the relevant features. In the case of codas and rimes, for instance, percolation of coda features
to the rime node can be interpreted as extending the exponents of the coda features in time to fill the temporal domain of the entire rime, covering also the temporal domain of the nucleus. It should be emphasized that the feature-spreading or copying which derives from the interaction of vertical feature-sharing constraints and temporal constraints is merely an effect (albeit a useful one) which arises solely from the procedural interpretation of a declarative formalism.

In the phonology of English, it can be argued that the opposition of voicing holds of onsets and rimes, not consonants and vowels. (Voicing is not distinctive in vowels even in segmental generative phonology, and the distinctive voicing of a cluster is the voicing of the head of the cluster.) The phonetic exponents of the feature [voice] are long-domain (Firth 1935b, Hockett 1955, Yoshioka 1981, Browman and Goldstein 1986: 227), an interpretation which is to be expected if the phonological feature [voice] is non-terminal, but which requires special explanation (e.g. in terms of feature-spreading) otherwise.

In a segmental representation, however, [voice] is of course specified for obstruents. In order to express its prosodic role, the value of this feature would have to be percolated to the onset or coda node dominating the obstruent. If this is done, all instances of [voice] on the path between the onset or rime node and the terminal obstruent node are redundant. By percolating longer-domain features to higher domains in this way information is distributed more evenly over the entire phonological structure, and not concentrated in the terminal nodes. Alternatively, one could abandon the use of segmental representations and represent such features in the lexicon at non-terminal nodes (Anderson 1986, Vincent 1986, Goldsmith 1990), after the fashion of Firthian Prosodic Analysis. In Prosodic Analysis, features located at non-terminal nodes are termed prosodic features or prosodies. Those located at terminal nodes are termed phonematic features. This distinction is the principal way in which non-segmental phonology differs from structured segmental phonology. (Anderson (1986) makes some recent proposals towards the inclusion of prosodic (non-terminal) features in Dependency Phonology. Selkirk (1980: 569) and Pierrehumbert and Beckman (1988: 21) do likewise for Metrical Phonology.)

Similarly, feature propagation can be used to account for the spread of the liprounding feature [round], and [nasal] too is often considered to be a rime- or syllable-domain feature, rather than a segmental feature (e.g. Browman and Goldstein 1986: 231). Kelly and Local (1986) present evi-
dence that the phonetic exponents of gravity in the liquids /l/ and /r/ are distributed across entire metrical feet, an observation which could be modelled by feature propagation. This is not to say that all cases of featurer-spreading should be attributed to feature propagation in this theory. Distributional criteria remain the best basis on which to determine the domain of each distinctive feature.

Because feature-sharing constraints are tied to the parse tree, some constraints on sequences of consecutive terminal symbols are expressed with greater ease than others. In particular, constraints which hold of sisters are especially easy to express. It is possible to express constraints holding between non-sisters, but additional feature-passing is required to do this. Also, constraints holding of an entire substring are more easily stated if the substring is a well-defined constituent. Constraints on substrings that span a constituent boundary may be stated, at the expense of requiring more feature-sharing constraints.

In contrast, if a grammar has a mechanism for stating constraints on sequences of terminal symbols without regard to constituent structure, spreading of features over non-constituent domains may be expressed as easily as over constituent domains. For example, a constraint against sequences of obstruents with different values of the feature [voice] could be stated as follows:

\[(5.24) \quad \left[ +\text{obstruent} \right] \left[ +\text{obstruent} \right] \]

Pierrehumbert and Nair (1995) argue that sequence constraints such as (5.24) are necessary, because they may cut across the constituents of the syllable tree, applying, for example, across the coda of one syllable and the onset of the next. Such constraints are simple to implement by intersecting regular expressions with the frontier of the prosodic tree. This does not affect the power of the grammar, since regular and context-free languages are closed under intersection with a regular language. There is, nevertheless, a tension between the tree-based feature-sharing constraint mechanism described above and a string-based mechanism, in that in the tree-based approach it is easier to express constraints on constituent substrings than on non-constituent substrings. This asymmetry seems reasonable to me: when the coanalysis technique of section 5.6.2 is taken into account, there is little reason to add the ability to add constraints on arbitrary substrings to the basic context-free or regular phonological grammar.
5.5.10 Simultaneous and parallel rule interaction in Declarative Phonology

The preceding sections set out most of the formal resources which Declarative Phonology employs to express linguistic generalisations as constructive constraints. By a judicious use of disjunction and implication, rules employing deletion, feature-changing or restructuring can be recast in monotonic terms. The effect of this is that extrinsic ordering can be eliminated. In this section I shall review and spell out how constraints in Declarative Phonology interact in a simultaneous or parallel fashion.

1 Implicit disjunction

Extrinsic, sequential rule application employs a form of conjunction, 'and then'. Some pairs of rules, however, are inherently disjoint, as they could never both apply to the same part of a representation (though they could both apply to different parts of a representation). For example, consider a pair of rules such as (5.25):

(5.25) a. C[+voice, -continuant] → [+continuant] / V —> V
b. C[+voice, -continuant] → [-voice] / V

These two rules apply in different contexts: (a) lenites intervocalic voiced stops after stress, and (b) devoices stops before a stressed vowel. Since (a) and (b) are disjoint in their application, there is no need, and no point, in ordering (a) before (b) or vice versa. SPE phonology makes provision for such cases by claiming that rule ordering is not total, only partial. As far as it goes, this position is correct. Another example can be seen in the disjunctive application of a set of rules that in SPE notation is ‘compacted’ into a single rule using the angle bracket convention (Chomsky and Halle 1968: 77). Disjunctive ordering is easily overlooked in expositions of derivational phonology, in which rule ordering looms large over disjunctive rule application. However, it is a transparent case of parallel or sequential rule application that should not be neglected.

Disjunction is also implicit in the way in which context-free grammars work. Consider a grammar such as (5.26):

(5.26) a. PrWd → Foot
    PrWd → Footw Foot_s
b. Foot → Syll_{heavy}
    Foot → Syll_{heavy} Syll_{light}
c. Syll_{heavy} → Onset Rime_{heavy}
    Syll_{light} → Onset Rime_{light}
5.5 Formalism for phonological representation

The grammar can be divided into blocks of rules that have the same symbol on the left-hand side, as in (a) and (b). One rule in each of the blocks (a), (b) and (c) must do some work in the prosodic analysis of every word. The relationship between the blocks, therefore, is a conjunctive one. (This does not mean that they have to be applied in a particular order, however. Since all the blocks must do some work, as they relate to different levels of structure, they could all be applied together, simultaneously, in parallel or in any order.) Within each block, the rules are disjoint, since their right-hand sides characterise incompatible structures. It is natural to think of rules within each block as applying simultaneously or in parallel to each word. Only one of the rules in each block ends up contributing to the analysis, but all of them must be tried. SPE phonology did not make use of context-free grammars, and rules were rarely and only contingently context-free. But with the incorporation of prosodic phonology into the generative mainstream, it is timely to examine the rôle that such a development plays in simplifying the derivational rule-ordering hypothesis.

2 Explicit disjunction
From this uncontroversial starting-point, we can advance the scope of parallel or sequential rule application by casting other rule interactions as disjunctive. The examples of counter-feeding rule ordering and the selection of two possible rules by a parameter-setting mechanism were two other cases discussed in 5.5.2 above. We also saw that other uses of explicit disjunction, such as in lexical representations, advance the possibilities for declarative analysis.

3 Implications
As shown above, redundancy rules exemplify logical implication rather than rewriting. Implications do not map one string onto a different string (which would give rise to a derivation step), but relate one part of a string onto another part. Apparently structure-changing rules such as word-final place assimilation in English coronal consonants can be brought under the same umbrella as redundancy rules by employing underspecification. If coronal consonants are left unspecified for place, the acquisition of place features from the initial consonant of the next word is a feature-filling, rather than feature-changing, operation, and can hence be regarded as true of a single level of representation. As in redundancy rules, spreading and sharing rules relate one part of a representation to another, not one representation to a
different one at the next step in a derivation. In the isolation form of word-final coronals, the underspecified value could be filled in by a default or redundancy rule – again, an implication – under the Elsewhere Principle, another way of eliminating extrinsic rule ordering that requires no further comment.

4 Stratification

'Stratification' refers to the use of hierarchical structure in order to avoid the ordering of blocks of rules (strata) found in e.g. Lexical Phonology. An example of this will be given in section 6.4.3, where the four strata of lexical rules in Halle and Mohanan's (1985) analysis of English segmental phonology are enabled to apply disjointly, by building four separate suprasegmental domains ('\(W_1\)', '\(W_2\)', 'Compound' and 'Word' – see example (7.141)). This approach arose out of work in Lexical Phonology (e.g. Booij and Rubach 1984) which correctly observed that the domains to which blocks of lexical rules apply in many cases do not correspond to morphological constituents. Several such cases are noted in Cohn (1989), Inkelas (1989) and Cole and Coleman (1993), all of whom conclude that a lexical hierarchical structure distinct from morphological structure (though built on morphological categories) should be recognised. The non-terminal nodes in this structure define a hierarchy of levels or strata – parts of the word to which sets of Lexical Phonological rules apply. Since the rules can now be related to independent strata, the blocks of rules in Lexical Phonology can be applied simultaneously to their particular domains.

5 Conjunction

All other instances of rule interaction are conjunctive. In the discussions of context-free grammars and stratification above, it was shown that where two or more rules both apply to a representation, they can often be applied simultaneously. This is obvious in the case where the two rules do not affect each other. Like the cases of inherent disjunction, pairs of rules relating to different parts of a representation without any interaction are also inherently unordered. Because conjunction is commutative ('A & B' is true under the same conditions as 'B & A'), parallelism of conjunction is also quite general. (The identification of instances of and-parallelism and or-parallelism is the basis of concurrent programming.)

The problematic cases of conjunctive rule application are those in which there is an interaction or dependency between two rules. These arise in SPE solely in cases of destructive rules. For example, a context-sensitive deletion
rule A may remove the context under which some other rule B would otherwise apply. Hence, the order 'A, then B' will yield a different result from the opposite order. Similarly, a feature-changing rule may alter the context in which some other rule would otherwise apply, and a structure-changing rule may remove structure that would otherwise trigger another rule. In each case, order is significant, as different rule orders may yield different results. In no other circumstances is rule ordering significant. Consequently, if deletion, feature-changing and restructuring are eliminated from grammars, using the techniques of declarative analysis discussed above, rule ordering becomes unnecessary. All cases of 'and then' rule ordering reduce to simple, commutative 'and'.

5.5.11 Restrictiveness of Declarative Phonology
Context-free grammars, which are central to much work in Declarative Phonology, are more restrictive than the powerful kinds of grammars used in derivational phonology, as shown in chapter 3. In fact, if context-free grammars are used to define level-ordered hierarchical structures without recursion, the languages which they define are regular, that is, a proper subset of the context-free languages. However, some of the formal resources introduced in this chapter are potentially damaging to this increased restrictiveness. Note that the use of complex categories (e.g. features at terminal and non-terminal nodes) does not affect the weak generative capacity of the grammar. Nor do the feature-checking constraints between mother and daughters in a rule, by which feature propagation is implemented.

Multiple domination does introduce the potential for unbounded expressive power if the number of mothers that a node may have is unbounded, as shown in chapter 4. This problem is avoided here by limiting re-entrancy to a maximum of two mothers per node.

A more serious problem is presented by the richness of feature-structures as described so far. Using category-valued features, for instance, it is possible to count, using a version of Peano arithmetic (see Hofstadter 1979: 216–27). For example, let [succ] (an abbreviation for ‘successor of’) be a category-valued feature. Let the empty category be denoted Z (for ‘zero’). With this single feature, we can build a complete set of integers:

0 can be represented as Z
1 can be represented as [succ: Z]
2 can be represented as [succ: [succ: Z]]
etc.
We shall refer to such feature-structures as Peano feature-structures. More formally, define an alphabet of terminal feature-structures as follows:

- 0 abbreviates $N[\text{num}: Z]$
- 1 abbreviates $N[\text{num}: [\text{succ}: Z]]$
- 2 abbreviates $N[\text{num}: [\text{succ}: [\text{succ}: Z]]]$
  
  etc.

The terminal symbol ‘E’ can be read ‘equals’. Using such a notation for integers, we can incorporate some powerful counting and equality-checking operations. Equality of Peano feature-structures can be checked using context-free rules, co-indexing, and a couple of extra features, as follows:

\[(5.27)\]

a. $X[\text{left:\{\}, right:\{2\}] \rightarrow X[\text{left:\{succ:\{\}, right:\{succ:\{2\}]})$

b. $X[\text{left:\{\}, right:\{2\}] \rightarrow N[\text{num:\{\} E N[\text{num:\{2\}]})$

These two rules, and the start symbol $X[\text{left:\{} \text{right:\{}}$ define a language whose strings are just ‘0 E 0’, ‘1 E 1’, ‘2 E 2’, and all others in which the first and third symbols are identical, the middle symbol being ‘E’. To see why this is so, consider the derivation of the string ‘2 E 2’ (recall that ‘2’ is an abbreviation):

\[
\begin{align*}
X[\text{left:\{} \text{right:\{}} & \rightarrow \, N[\text{num:\{\} E N[\text{num:\{2\}]}) \\
X[\text{left:\{succ:\{\}, right:\{succ:\{\} & \rightarrow \, X[\text{left:\{succ:\{\}, right:\{succ:\{\} \\
X[\text{left:\{\}, right:\{\} & \rightarrow \, X[\text{left:\{\}, right:\{\} \\
N[\text{num:\{\} & \rightarrow \, N[\text{num:\{\} \\
E & \rightarrow \, E \\
N[\text{num:\{\} & \rightarrow \, N[\text{num:\{\} \\
\end{align*}
\]

Each application of (5.27a) reduces the Peano representation of the numbers that are values of the features [left] and [right] by exactly one: that is, it is a cancellation rule. The only way to get from the root to the frontier is through a derivation in which the values of [left] and [right] are equal. Equality is checked by cancellation and the fact that the start symbol requires both [left] and [right] to be zero.

This mechanism is extremely powerful. It is easy to see how a feature [mid] could be added to [left] and [right] in order to check the equality of three successive symbols. If this is done, it becomes easy to implement the strictly context-sensitive language $a^n b^n c^n$, and other more exotic sets of strings, using only context-free rules, feature-structures and co-indexing.

The great power of this system arises from the unboundedness of the feature structures. It is obvious that in order to implement arithmetic or unbounded cross-serial dependencies requires unboundedly deep feature
structures to be available in principle. Similar techniques could be used, for instance, to encode an arbitrarily large tree as a feature-structure, and then pass it to another part of the linguistic representation using feature propagation. Arbitrary tree transformations capturing elaborate dependencies could be implemented in this way.

The purpose of this discussion is not to undermine the use of unification grammars, but to address a common (but irrelevant) objection. Because of examples like the above, it is sometimes claimed that unification grammars are no better than unrestricted rewriting systems, because they are just as powerful. In general, this claim is correct. But it is irrelevant to Declarative Phonology (and to most syntactic theories based on unification grammar, including GPSG), because those theories are careful to require feature-structures to be of bounded depth (Gazdar et al. 1985: 36). This ensures that the set of legal feature-structures is finite. The arbitrary packaging of constituent structure into feature-structures that unbounded depth permits is thus prevented. There is no scope for counting. Not even indexed grammars, which lie between context-sensitive and context-free grammars, can be defined if feature-structures are bounded. A grammar formalism which uses a finite set of feature structures in context-free rules is equivalent to classical context-free grammar, not more powerful.

Now let us consider a practical example of a phonological phenomenon that Declarative Phonology predicts cannot occur. After all, the discussion so far has centred on how Declarative Phonology can deal with the phenomena addressed by derivational phonology. The most straightforward example is that in a declarative grammar it is not possible to change feature-values or delete information. For example, Hayes (1984) presents an analysis of Russian phonology in which the preposition [v] (lexically, /w/) in some circumstances is devoiced by a rule of final devoicing, and is later revoiced by a rule of sonorant revoicing. Thus the /w/ at the end of the first word in Zdorov li (‘Is he healthy?’) has a derivation [+voice] => [-voice] => [+voice]. Such an analysis is not possible in Declarative Phonology. Either of the devoicing or revoicing rules could be implemented declaratively using underspecification and implication, as either [] D [-voice] or [] D [+voice]. Let us suppose that the devoicing step is implemented in this way. Now, the revoicing step could be regarded as phonetic, rather than phonological: that is an option in Declarative Phonology, as in derivational phonology, of course. But in this case, it is the only option available: alteration of [-voice] is not possible in Declarative Phonology. Moreover, that forces
certain consequences: if revoking is phonetic, we expect it to behave phonetically, that is, perhaps in a gradient fashion; not subject to lexical exceptions etc. An alternative that is quite consistent with Declarative Phonology is that neither devoicing nor revoking are phonological, but both are phonetic (postlexical). This possibility, in fact, is suggested by Kiparsky (1985).

Phonological analyses in which the value of a feature is changed twice are very rare in the phonological literature. I know of no analyses in which a feature is changed three times (excluding SPE stress rules). This observation has no explanation at all in terms of derivational phonology, in which there should be no reason to expect multiple feature changes to be uncommon, given the length of derivation supposed in most transformational phonological analyses. The absence of such feature-changing chains has a simple explanation under the declarative hypothesis, which thus wins support from the lack of such derivations.

5.6 Phonetic interpretation, 1: temporal interpretation

In this section I shall present a theory of phonetic interpretation of non-segmental, hierarchical, graphical structures described above. The theory of phonetic interpretation is based on a parametric, dynamic model of phonetics. The distinction between ‘head’ and ‘non-head’ constituents is central to the phonetic interpretation model. This model is formally explicit, and has been computationally implemented in a novel speech synthesis system, ‘YorkTalk’ (Coleman 1990a, b, 1991; see also Dirksen and Coleman, 1997). The phonological representations employed in that system are structures of the kind discussed in this and the next chapters. The model of phonetic interpretation employed in the system has led to the general proposals regarding phonetic interpretation which I shall now present.

For expository reasons, I shall discuss phonetic interpretation in two parts. Firstly, I discuss temporal interpretation – that is, the general principles by which the phonetic exponents of phonological units are ordered in time with respect to each other. Secondly, I shall briefly discuss the parametric interpretation of consonant–vowel transitions. No implication is intended by this order of presentation regarding the actual order in which these two aspects of interpretation might be determined in human speech production, in a model of human speech or in a speech synthesis system.
5.6 Temporal interpretation

5.6.1 Temporal interpretation of concatenation in logic grammars

In one common variety of Unification Grammar, logic grammars (a class which includes Definite Clause Grammar; see Pereira and Warren 1980), phrase-structure rules are interpreted as well-formedness constraints expressed in first-order predicate logic. Strings are represented in what is called position notation (Pereira and Shieber 1987: 33), meaning that the start and end of the string, and every inter-word space, are numerically indexed. A phrase-structure rule such as:

\[(5.28) \quad S \rightarrow NP \ VP\]

is interpreted as a shorthand form of the logical description:

\[(5.29) \quad \forall xy \ NP(x) \ & \ VP(y) \ \supset \ S(x \ -. \ y)\]

Dropping the universal quantification and representing strings in position notation:

\[(5.30) \quad NP(p_0, p_1) \ & \ VP(p_1, p) \ \supset \ S(p_0, p)\]

where \(p_0\) denotes the beginning of the string, \(p\) the end, and \(p_1\) some intermediate position. Applied to the indexed string \(0 \ Mary \ 1 \ runs \ 2\), constraint satisfaction can be used to determine that \(p_0 = 0, p_1 = 1, p = 2\), \(Mary\) is an NP, \(runs\) a VP and \(Mary \ runs\) an S. (The indices can be inserted automatically, of course.) By casting the phrase-structure rule in position notation, the order of the NP term and the VP term is no longer significant, since conjunction is commutative. Expressions (5.30) and (5.31) are equivalent:

\[(5.31) \quad VP(p_1, p) \ & \ NP(p_0, p_1) \ \supset \ S(p_0, p)\]

These formal devices are also applicable in phonetic interpretation. If we were to develop a concatenative speech synthesis system in which utterances were simply context-free sequences of prerecorded words with a fixed suprasegmental intonation contour, then the logic grammar model would be entirely appropriate. For instance, the beginning and end point indices could be actual times in milliseconds denoting the time at which each prerecorded word is to begin and end. The temporal structure assigned to the string \(Mary \ saw \ John\) would be figure 5.21.

As well as defining the start and end time of each word in the sequence, the constraint satisfaction mechanism also defines the start and end time of every phrasal constituent, including the S. It would be possible to set the start and end times of a separate, perhaps rule-generated, intonation contour, so that it was precisely synchronised with the timing of the sequence of words. Constraint satisfaction can thus be used to impose a
self-regulating, sequential and parallel temporal structure on a structured linguistic representation. However, this formalism is not sufficient for the temporal interpretation of phonological representations because of the existence of overlap between constituents in phonology and phonetics. In the next section a case of phonological overlap is considered and an appropriate formal extension of phrase-structure theory proposed to deal with it. The issue of overlap in the phonetic interpretation of syllable constituents is pursued at length below.

5.6.2 An Extension to the ID/LP formalism for overlapping constituents
In the preceding sections it was explained how logic grammars factor temporal sequence out of linguistic representations and express it as separate constraints on otherwise unordered (i.e. conjoined) constituent structures. This is similar to the ID/LP mechanism of Generalised Phrase-Structure Grammar (Gazdar et al. 1985). In classical Chomskyan phrase-structure grammars, a phrase structure rule such as:

(5.32) $S \rightarrow NP \ VP$

represents two independent kinds of linguistic information: information about the constituency relations between categories (the ‘is a constituent of’ relation), and information about the linear order of daughter categories (NP precedes and is concatenated to VP).

In Generalised Phrase-Structure Grammar, classical phrase-structure rules such as the above are factored into two components, expressing the immediate dominance (constituency) information, and the linear precedence information separately. Grammars expressed in this way are said to be in ID/LP format. The chief motivations for ID/LP format are:

Constituency and order are linguistically independent relations; a linguistic description which reflects this fact will be more
parsimonious and able to express independent generalisations about constituency and order more felicitously.

ID/LP grammars have the formal property that LP rules define an exhaustive, constant partial ordering of the terminal categories of all strings defined by the grammar. This property, the ECPO property, is not a property of all phrase-structure grammars, and therefore ID/LP format incorporates a restrictive empirical hypothesis about ordering in natural-language grammars.

For example, in the ID/LP format, the classical rule (5.32) is factored into two parts: an order-free constituency statement:

(5.33) \( S \rightarrow NP, VP \)

and a statement of linear order:

(5.34) \( NP < VP \)

The categories employed in LP statements are logically independent of the categories which may be mentioned in the ID rules, which gives added motivation for ID/LP format. For instance, in Gazdar et al. (1985), most of the facts of word order in English are expressed by just two LP statements, one of which effectively states that lexical heads precede their complements, the other of which effectively states that non-subcategorising categories (such as complementisers, determiners, conjunctions and prepositions) precede subcategorising categories (such as nouns and verbs).

The particular case of non-concatenative constituency I shall now consider concerns non-congruent parallel structures, in which a given string functions simultaneously as two independent kinds of constituent. When English words are assigned a morphological structure, showing the division into the various strata of prefixes, suffixes and stems, and also a metrical structure, showing the division into feet and syllables, there is no guarantee that the morphological and the metrical structures of a given word will always be congruent. In particular, since the morphological and metrical structures of level 1 (Latinate) words are not congruent, they must be described using separate subsets of constituency rules. Since Latinate words may themselves be constituents of larger words (for instance, they may combine with level 2 (e.g. Germanic) prefixes, or form part of a compound word), the grammar must contain a unitary way of referring to level 1 words (stress and morphology combined). Level 1 metrical structure and level 1 morphological structure are two non-congruent, intersecting constituents of level 1 words. In other words, part of the phonological structure of
level 1 words is described by an unordered constituency rule of the following kind:

(5.35) \( \text{Word}_i \rightarrow \text{Stress}_i, \text{Morphology}_i \)

If this rule were to be interpreted in the fashion of classical phrase-structure rules according to the interpretive conventions of logic grammars, it would receive the interpretation:

(5.36) \( \text{Stress}(t_0,t_1) \& \text{Morphology}(t_0, t_1) \supseteq \text{Word}(t_0,t) \)

This interpretation is clearly wrong, however, because it is a concatenative interpretation. The required interpretation is the non-concatenative interpretation:

(5.37) \( \text{Stress}(t_0,t_1) \& \text{Morphology}(t_0, t_1) \supseteq \text{Word}(t_0,t) \)

in which Stress and Morphology are simultaneous, overlapping constituents of Word, dominating the same terminal objects, with coextensive phonetic exponents.

We want to enforce the following temporal constraint:

(5.38) \( \text{Stress}_i \preceq \text{Morphology}_i \)

where \( a \preceq b \) denotes that \( a \) and \( b \) are coextensive. As a piece of convenient notation, then, let us introduce rules of the form (5.39) to denote (5.40).

(5.39) \( X \rightarrow A \approx B \approx \ldots C \)

(5.40) \( A(t_0,t) \& B(t_0,t) \& \ldots C(t_0,t) \supseteq X(t_0,t) \)

in which \( A, B, \ldots C \) are coextensive in time.

5.6.3 Temporal interpretation and overlap
Temporal relations in phonetic representations have been formalised and extensively studied by Bird and Klein (1990). Bird and Klein's system employs intervals as its elementary temporal units, and allows no reference to be made to points of time, even phonetically salient times such as endpoints of an interval, maxima, minima or inflections of a particular phonetic parameter. Although they regard this restriction as a strength of their formalism, reference to salient points in the time-course of a phonetic event is necessary for explicit phonetic description using current speech synthesis models. Common to YorkTalk, Browman and Goldstein's and Bird and Klein's models of the phonetics-phonology interface is the formal inclusion of (at least) two temporal relations, precedence and overlap. (YorkTalk contains two types of temporal overlap: coproduction
5.6 Temporal interpretation

**Figure 5.22** Concatenative interpretation (b) of syllable-structure (a)

of *different* phonetic parameters, and overlap of the parameters of one constituent on top of the *same* parameters of another constituent, a distinction which parallels in some respects the distinction which Browman and Goldstein (1990) make between within-tier *blending* and between-tier *hiding.*

In structured phonological representations, the phonetic exponents of smaller constituents overlap the phonetic exponents of larger constituents. The *concatenative* temporal interpretation of a structure such as figure 5.22a is therefore figure 5.22b, in which the onset exponents and the rime exponents overlap (indeed are part of) the syllable exponents, and the nucleus exponents and the coda exponents overlap the rime exponents and the syllable exponents. The temporal constraints which encode this pattern of sequence and co-occurrence are:

\[(5.41)\] syllable start time = onset start time

\[(5.42)\] syllable end time = rime end time

\[(5.43)\] onset end time = rime start time

\[(5.44)\] rime start time = nucleus start time

\[(5.45)\] rime end time = coda end time

\[(5.46)\] nucleus end time = coda start time

If phonetic representations were made up exclusively of well-ordered sequences of elementary phonetic objects such as allophones, then such a set of temporal constraints would suffice to divide up and group together phonetic subsequences into phonological constituents. This, it seems, is the
strategy offered by generative phonology in its original form and in more modern versions, such as Dependency Phonology.

As I argued in chapter 2, however, it makes little sense to consider speech as the concatenation of phonetic segments, as the exponents of one constituent may overlap those of another. This is a well-established observation. For instance, Öhman (1966) found that in $V_1CV_2$ ‘sequences’, the transition from $V_1$ to $V_2$ (i.e. the point at which the coarticulatory influence of $V_1$ on $C$ gives way to the coarticulatory influence of $V_2$ on $C$) does not necessarily occur at the $V_1C$ boundary (figure 5.23b), nor the $CV_2$ boundary (figure 5.23c), but depending on various factors, may fall either somewhat before (figure 5.23a) or somewhat after the $C$, an observation which suggested to Öhman, Fowler (1983) and myself the existence of an ‘underlying’ vowel-to-vowel movement, upon which the $C$ is overlaid. Such a model of articulation has been proposed by Liberman et al. (1967), Perkell (1969), Mermelstein (1973), Gay (1977), Bell-Berti and Harris (1979) and Mattingley (1981), and proposed as a basis for non-linear phonological organisation by Firth and Rogers (1937) and Griffen (1985). It has not been generally adopted in generative phonology, however, and even Autosegmental Phonology presents phonological representations constructed from well-ordered sequences of consonants and vowels, with no account of their temporal synchronisation or coarticulation.
5.6 Temporal interpretation

The 'overlay' account of VCV articulatory timing can also be applied to the interpretation of onsets (figure 5.23a and b) and codas (figure 5.23c). In the case of codas, temporal constraint (5.46), which means that the coda follows the nucleus, must be replaced by (5.47), which ensures that the coda and the nucleus end at the same time.

(5.47) nucleus end time = coda end time

Taking this together with the contingent fact that the exponents of codas are (generally) shorter than the exponents of nuclei, we can illustrate the temporal interpretation of figure 5.24a as figure 5.24b. This is done by augmenting figure 5.24a with temporal constraint (5.47). It follows from constraints (5.44), (5.45) and (5.47) that the exponents of rimes are coextensive with the exponents of nuclei (figure 5.25). Onsets come in for similar treatment in the 'overlaying' model. Rather than prefixing the exponents of the onset to the exponents of the rime, we state that the exponents of onsets are overlaid on the exponents of rimes starting at the beginning. This is done by replacing (5.43) by (5.48):

(5.48) onset start time = rime start time

Since rimes and nuclei are coextensive, the consequence of this is that the exponents of onsets are also overlaid on the exponents of nuclei.
From constraints (5.41), (5.42) and (5.48) it follows that the exponents of syllables are coextensive with the exponents of rimes, and that the temporal interpretation of figure 5.26a is figure 5.26b. Combining figures 5.25 and 5.26 gives figure 5.22a with the temporal interpretation in figure 5.27, in which syllable, rime and nucleus are coextensive in time. (This does not mean that the phonological distinctions and the non-temporal phonetic distinctions between syllable, rime and nucleus have been lost: the categories syllable, rime and nucleus have not become conflated.)

Compare figure 5.27 with figure 5.22b. In figure 5.26 the onset exponents are overlaid on the nucleus by virtue of the facts that the onset is overlaid on the rime, by constraints (5.41), (5.42) and (5.48) and the rime is coextensive with the nucleus, by constraints (5.44), (5.45) and (5.47). This is despite the fact that onset and nucleus are not sisters in the phonological structure, and yet it is exactly the desired result, according to the phonetic model. The fact that exactly the right phonetic interpretation for onset timing arises from the interaction of all six temporal constraints, simultaneously according with the established view of English syllable structure and the ‘overlay’ model of timing, suggests that this is a very promising approach to temporal interpretation of phonological structure, maintaining the appropriate degree of independence of phonetics from phonology.
5.6.4 Heads

The fact that syllable, rime and nucleus are coextensive leads naturally to consideration of whether they are in fact separate objects at all. For syllables without onset or coda, a theory such as Dependency Phonology answers this question in the negative, proposing that they are just three ways of looking at a single object, the vowel (figure 4.16). With onsets and codas, however, Dependency Phonology maintains the concatenative interpretation (figure 4.17).

In Dependency Phonology, the head of a structure is the object written directly below the mother of that structure in the dependency graph. Since the rime is the obligatory part of the syllable, and the nucleus is the obligatory part of the rime, the nucleus node is written below the rime node and the rime node below the syllable node. Onset and coda are optional. The previous section showed that the distinction between syllable, rime and nucleus, on the one hand, and onset and coda, on the other, is manifested not only in the phonological domain (obligatory vs. non-obligatory), but also phonetically: syllable, rime and nucleus are all coextensive. Since these constituents are special, we shall adopt conventional practice in referring to them as heads. Onset and coda we call margins.

Armed with this distinction, we can make three generalisations about temporal interpretation: (i) the temporal domain of the head of a constituent is coextensive with the temporal domain of the whole constituent; (ii) the temporal domain of the modifier of a constituent begins or ends at the same point as the temporal domain of the whole constituent, but is shorter; (iii) therefore, the temporal domain of the modifier constituents overlap the temporal domains of their head sisters. These three principles are not sufficient to replace the temporal constraints discussed above, but they are sufficient to direct the flow of control in a procedural interpretation of phonological structure. The principle is: 'at each level of structure, the head is interpreted first', as described in greater detail in section 5.7 below.

5.6.5 Clusters

In 1987, while modelling consonant clusters in YorkTalk, Adrian Simpson proposed that obstruent+glide clusters also have an 'overlaid', rather than concatenative, temporal structure. Simpson's proposal modelled the timing of obstruent+glide clusters in a similar fashion to onset+rime structures, as in figure 5.28. Mattingley's (1981) proposed synthesis model also has this analysis of clusters. This analysis is borne out by a detailed examination of the phonetic exponency of the obstruent+glide cluster /tr/, compared with,
Declarative Phonology

Glide exponents
Obstruent exponents

is analogous to

Rime exponents
Onset exponents

Figure 5.28 Simpson's analogy

Labiality
Retroflection
Closure  Aspiration

Figure 5.29 Labiality and retroflection of /r/ are present throughout /tr/

/r/ exponents
/s/ exponents  /p/ exponents

Figure 5.30 Labiality and retroflection of /r/ are present throughout /spr/

Onset
Obstruent  Glide

Figure 5.31 Structure of obstruent + glide onsets

say, /pl/. In /tr/, the strong labiality which is a characteristic exponent of /r/ is present throughout the cluster, and the 'spooning' of /r/ is audible throughout the burst of /t/, suggesting the temporal orchestration illustrated in figure 5.29.

In the cluster /spr/, compared with /spl/, the two /s/s have audibly different qualities, a distinction which can be modelled with the temporal organisation in figure 5.30. The phonological structure of binary branching onsets like /pr/ is illustrated in figure 5.31. In this case, neither the
5.6 Temporal interpretation

obstruent nor the glide is obligatory, so headship cannot be determined on phonological grounds alone. But the hypothesis that phonetic interpretation is head-first requires one daughter to be taken as the head, and the other to be dependent. Experimentation with the YorkTalk model has confirmed that in this theory, the glide must be taken to be the head. This accords with the observation that the glide is more sonorous than the head: the obstruent–sonorant relation is parallel to the onset–rime relation. The phonological structure of /spr/-like onsets could be either that of figure 5.32a (as in Kiparsky 1979: 432 or Cairns and Feinstein 1982: 200), or that of figure 5.32b (as in Cairns 1988). In the analysis of English in chapter 7, the doubly branching structure will be used.

5.6.6 Elision

Further evidence for the foregoing account of syllable organisation is provided by cases of apparent vowel deletion in unstressed prefixes in words such as prepose, propose and suppose. In Declarative Phonology, no material is removed or lost in these examples, so no deletion rules or processes are required. In my account of this phenomenon, the phonological representation of these items is invariant, but may receive various possible temporal interpretations, depending on phonological factors (e.g. metrical structure) and/or phonetic factors (e.g. speech rate). (Browman and Goldstein 1990 present a very similar account to this in the analysis of consonant elision in casual-speech renditions of phrases such as perfec(t)
The 'unreduced' form of these prefixes have the temporal organisation shown in figure 5.33. In the putatively reduced forms, the nucleus duration can be reduced up to the point at which the end of the nucleus coincides with the end of the onset (fig. 5.34). This analysis is supported by the fact that the exponents of onset are appropriate to the presence of the exponents of the nucleus in each case, and that the qualities of the release or glide differ in the expected way in the 'reduced' forms of prepose and propose. These observations present serious problems to procedural segmental phonology: deletion involves the removal of a segment; coarticulation demands its presence, so the supposedly mechanical coarticulation process would have to precede a phonological deletion rule, as we saw in section 3.4.4. The overlay analysis suffers from no such problems. (As well as complete onset-nucleus overlap, complete nucleus-coda overlap is also found, for instance in the pronunciation of (i)mportant, (o)f and (a)n etc.)
5.7 Phonetic interpretation, 2: parametric interpretation.

Parametric phonetic interpretation is a relation between phonological categories (feature structures) at places in structure (i.e. nodes in the syllable tree) and sets of parameter sections. A parameter section is a sequence of ordered pairs, each of which represents the value of that parameter at a particular (salient) time. For instance, the parameter section \([(t_1, v_1), (t_2, v_2), (t_3, v_3)]\) denotes value \(v_1\) at time \(t_1\), value \(v_2\) at time \(t_2\) and value \(v_3\) at time \(t_3\) for some parameter. The times may be constants, denoting absolute times in milliseconds, or they may be functions of the start and end times of the intervals determined by the temporal interpretation of the constituent being evaluated, denoting relative times. For instance (Onset End, Value) might denote the value of some parameter at the notional ‘transition point’ that marks the boundary between onset and nucleus. Likewise, (Coda End – 100, Value), (Rime End – 100, Value) and (Syllable End – 100, Value) all denote the same parameter-value, 100ms from the end of the syllable, given that the temporal constraints presented above hold. Although they are extensionally equivalent, they are intensionally distinct, a circumstance which sometimes requires close attention. For example, if labiality is observed throughout the entire extent of a syllable utterance, it is necessary to determine whether the phonological representation of labiality is located at the syllable node, or whether, for instance, it is located at a lower node,
the phonetic exponents of which happen to be coextensive with the phonetic exponents of the syllable node. The first analysis takes labiality to be a syllable-level phonological object, and its syllable-length extent follows 'for free'; the second analysis requires the 'excessive' duration of labiality to be explicitly specified in the phonetic interpretation component by temporal constraints. Thus the determination of the phonological domain of a phonetic phenomenon cannot be found simply by measurement and comparison of its extent with other phenomena of similar extent.

The phonetic interpretation relation, which is sometimes called phonetic exponency, is simply a (large) set of ordered pairs of the form \((\text{node(Category, [Start Time, End Time]}), \text{Parameter Sections})\) which can also be represented:

\[(5.49) \quad \text{node (Category, [Start Time, End Time]) } \rightarrow \text{Parameter Sections}\]

Compare this with the discussion of parametric phonetic interpretation in Ladefoged (1980: 495). Ladefoged proposes rules such as:

\[(5.50) \quad [+\text{velar}, +\text{stop}] \rightarrow P_{\text{front raising}} -1.0\]
\[(5.50) \quad [+\text{velar}, +\text{stop}] \rightarrow P_{\text{back raising}} +3.0\]

The kinds of object on the left-hand side and right-hand side of these rules are qualitatively different. They are not rewrite rules of the conventional form; the arrows denote a (many-to-many) mapping between two dissimilar domains of description, linguistic feature structures, representing distinctions and similarities between phonological objects, and phonetic parameters, representing the `controls' for the speech model. The only difference between the YorkTalk format and Ladefoged's format is that the YorkTalk phonetic interpretation statements include time.

Klatt, in Allen, Hunnicutt and Klatt (1987: 113–15) presents a method for the calculation of onset formant-frequency transitions which takes account of the formant-frequency motions of the nucleus, implicitly recognizing that the nucleus is the head of the syllable and governs the onset. Klatt's method is as follows, using the second-formant frequency motion of the syllable \(\text{go}\) by way of example (fig. 5.35). Firstly, the motion for the isolation form of the nucleus \(F_2\) is determined. Klatt models this motion as: (i) a constant 1100kHz for the interval from 200ms to 230ms; (ii) a linear fall from 1100kHz at 230ms to 850kHz at 380ms; (iii) a constant 850kHz from 380ms to 460ms. (The absolute values of the specific times mentioned in this description are arbitrary, and could all be increased or decreased if required.) As illustrated in fig. 5.36, the onset part of the second-formant
motion is then evaluated over the overlapping interval from 200ms (the time at which the velar closure is released) to 280ms (the time at which the underlying formant motion for the nucleus appears independent of the onset). The value of the second-formant frequency at 200ms is a function of three variables: a backward-extrapolated locus $L$ for the $F_2$ value during the closure, the value $V$ of the second formant at the 280ms consonant–vowel boundary, and a theoretical measure of coarticulation $C$ between the (extrapolated) consonant locus and the (observed) vowel value. This function is:

$$F_{2,\text{burst}} = L + C(V - L)$$

Further examples of the exponency function are discussed in Coleman (1992). In this model, parametric phonetic interpretation of the syllable tree is performed top–down and head-first (i.e. heads are interpreted before non-heads). The motivation for this flow-of-control régime is so that the parameters of the nucleus are evaluated before onset and coda, both of which are dependent on the nucleus. In this way, the observation that onset
and coda coarticulate with the nucleus is instantiated. Furthermore, this hypothesis does not require the notion of look-ahead to be invoked in order to implement onset–nucleus coarticulation, because the nucleus has already received its interpretation by the point at which the onset is interpreted.

In addition, this mode of phonetic interpretation is directly parallel to Gazdar et al.'s (1985: 223–4) account of the semantic interpretation of syntactic structures: 'it is clear that many features also contribute to the semantic interpretation of the structures in which they occur . . . The question arises as to where in the tree such features should make their contribution to the interpretation. Our answer is: at the highest point of occurrence.'

5.8 Conclusion

In this chapter, I have sketched the principal characteristics of Declarative Phonology and a theory of phonetic interpretation. I have shown how the maintenance of a rigid distinction between phonological structure (which is unordered) and phonetic interpretation (which includes timing) enables many phonological phenomena, including coarticulation and 'elision', to be modelled without the use of transformations or strings of segments. The price of this benefit is that phonological representations must have more structure than strings, a detail which is, however, borne out by many phonological and phonetic considerations.

In the final two chapters I shall explore the theoretical claims and examples of this chapter by presenting analyses of two fairly extensive fragments of Japanese and English phonology.
6 \hspace{1em} \textit{A declarative analysis of Japanese words}

6.1 Introduction

In the preceding chapters, I have shown how the formal resources of declarative grammar can be applied to the description of various phonological phenomena. As in other approaches to phonological theory, the success of the declarative approach to phonology was evaluated by a consideration of particular, partial analysis of fragments of several unrelated languages. In this chapter and the next, I shall illustrate the utility and coherence of the declarative approach by presenting and defending more complete (lexical) phonological analyses of the two unrelated languages, Japanese and English, which have been a focus for discussion and exemplification in earlier chapters. For both of these languages I shall present similar grammars of phonological structures up to the level of the word. In each case I shall discuss the phonological oppositions employed in those languages, the various types of features, the well-formed feature-structures and the phonological constituent structures. For the Japanese fragment, the systems of phonological oppositions were discussed in chapter 3 above. Consequently, this chapter attends in particular to the development of a phrase-structure grammar of the structures in which those systems are brought together. For the English fragment, presented in chapter 7, both systems and structures are described in detail. The grammar of English phonological structure has been implemented as part of a computer program for text-to-speech synthesis (Coleman 1991), and has consequently been tested for consistency and completeness more rigorously than the Japanese grammar.

Both this chapter and the first part of the next are, in Firthian style, descriptive \textit{statements}, with a minimum of theoretical argumentation. This style can be contrasted with the more discursive style characteristic of less exhaustive studies, such as McCawley (1968). Nevertheless, the criteria by which my statement is to be judged are the same criteria that must
be applied to any other approach to phonological representation, such as observational adequacy (agreement with a reasonably large body of data), descriptive adequacy (parsimony, generality, symmetry, and accordance with other such 'theoretical aesthetic' constraints, such as consistency) and explanatory adequacy (formal restrictiveness of the theory). The emphasis on attainment of descriptive and explanatory adequacy in generative phonology has led in the past, with few exceptions (such as Jones 1967, Kohler 1967, Chomsky and Halle 1968, Halle and Mohanan 1985 and Cairns 1985), to a widespread neglect of observational adequacy. But a disparate collection of fragmentary analyses does not collectively constitute a consistent theory (grammar), however elegant each such analysis may be. The likelihood is that a collection of fragments will contain internal contradictions. It is for this reason that I have attempted to present analyses which may (like any others) be improved upon, but which at least (unlike some others) meets the minimal standard, observational adequacy.

6.2 Japanese syllable-structure

6.2.1 Introduction
The analysis of Japanese phonological structure in this chapter draws on both first- and second-hand observations. Detailed first-hand face-to-face phonetic observations and impressionistic transcriptions of one native speaker of standard Japanese (a female student from Osaka) were made in weekly sessions over a period of several months. The utterances recorded in this way were elicited sentences which were constructed to be suitable minimal contexts for a large number of lexical items, e.g. 'that's a(n) ___', 'I ___ed a(n) ___' etc.

In addition, a number of digital tape recordings of the same speaker and two others reading written Japanese word- and phrase-lists were subsequently made in a recording studio. During the same period, I made a number of more casual observations of a native-speaker teacher of Japanese, both in the classroom setting and in conversation. I have also drawn extensively on a number of grammars and reference works on Japanese (see the papers and books on Japanese phonology referenced throughout the previous chapters). The following, in particular, have been indispensable to this chapter: Takehara (19-), Japan Foundation (1986), Clark and Hamamura (1981) and Hinds (1986).
6.3 Moras

As mentioned in chapter 3, Japanese and western linguistic traditions alike view Japanese words as being constructed from suprasegmental, but subsyllabic phonological elements called *moras*. As well as being usual in analyses of Japanese, these units have also been employed in the analysis of other languages in Hyman (1985) and more recent works, such as Hayes (1989). There are three different mora-structures in Japanese, which may be mnemonically labelled CV, V and C. The class of consonantal articulations found in C moras is quite different from those found in CV moras, C mora consonants being either mora nasals or mora obstruents, unspecified for place of articulation. These are therefore traditionally labelled as N and Q respectively, to distinguish them from the Cs of CV moras. The constituent structures of these three kinds of moras are:

\[
\begin{align*}
\text{a.} & \quad \mu \quad \text{b.} \quad \mu \quad \text{c.} \quad \mu \\
C & \quad V & V & \quad C
\end{align*}
\]

e.g. /si/ e.g. /i/ e.g. /Q/ or /N/

Relationship between CV and C moras

The three-way distinction between CV, V and C moras hides certain generalisations regarding the distributions of these three subcategories of moras. For example, CV moras alternate with C moras under certain morphological operations, such as /iti/ (‘one’) ~ /iQ+pon/ i.e. [ippon] (‘one [classifier]’), in which the CV mora /ti/ permutes with the C mora /Q/. As well as CV ~ C alternations in which the C mora is the mora of gemination, examples such as /nom+u/ (‘drink’, infinitive) ~ /noN+de/ (‘drink’, progressive), in which C is the nasal mora, are not uncommon. Because both CV and C moras occur with a similar (though not identical) distribution, to wit after CV moras, the possibility of regarding them as two instances of some more abstract category of moras will be considered below.

Relationship between CV and V moras

However, in addition to CV ~ C alternations, there are also some CV ~ V alternations, such as /kaku/ (‘write’, infinitive) ~ /kai+ta/ (‘write’, perfective) (believed to derive historically from /kakita/; see Poser 1986: 174). If one of the possible instantiations of the C constituent of CV moras could
be the empty category, CV and V moras could be regarded as variants of a single mora-structure, which might be labelled (C)V.

Relationship between V and C moras
As well as alternating with CV moras, V moras also alternate with C moras (sometimes with a morphological function e.g. /ka-u/ 'to buy' vs. /ka-t-te/'buying'), since either C or V may follow a CV mora to form the two-mora sequence CVC or CVV respectively. Since every VX sequence is possible in Japanese (where X is a C or V mora), there is no language-internal basis for preferring the C.VX partition that the ‘onset/rime’ theory proposes over the CV.X partition that the mora-based analysis proposes. Although the distribution of V moras partially overlaps the distribution of C moras, it is not feasible to regard V moras as a subcategory of C or vice versa, since they are distributionally different in some respects. For example, V moras alternate with CV moras (as shown above), and they may be iterated to form V+ sequences. In both these respects they are unlike C moras. Consequently, the analysis of V moras as a subcategory of (C)V moras with empty C is no more well motivated than treating V moras as a subcategory of C moras: if V moras are analysed as (C)Vs, then C moras could by the same logic be analysed as C(V)s.

CyV moras will be analysed as CV moras with a palatalised C, though the possibility that they are CiV sequences will be examined below, and the affricates [ts], [dz], [tʃ] and [dʒ] will be analysed as forms of /t/, /d/ and /z/ in CV moras containing close vowels. /dy/ and /zy/ are not phonetically distinct, the phonetic interpretation of both being [dʒ].

6.4 Syllables and moras
Although the structure of Japanese words can be analysed in terms of sequences of moras, like other languages, Japanese utterances may also be analysed as being constructed from syllables. If the conventional theory of syllable structure is insisted upon, the division of syllables into onsets and rimes, on the one hand, and moras, on the other, will impose non-congruent parses on many words. For example, /seNsei/ ('teacher') must be parsed /s-eN-s-ei/ into onsets and rimes, but /se-N-se-i/ into moras. I shall show that a phrase-structure grammar which expresses Japanese syllable-structure in terms of moras, rather than the onset/rime division, is well motivated, before proceeding to consider the relationship between moras and the syllable constituents ‘onset’, ‘nucleus’ and ‘coda’.

In Japanese, syllables consist of between one and four moras. I shall con-
6.4 Syllables and moras

Consider the structure of these in order of increasing size. The account of syllables which follows is based on two criteria:

1. **Phonetic: the sonority cycle.** A syllable extends from one sonority minimum (or obstruction maximum) to the next sonority minimum.

2. **Distributional: the delimitation of words.** A word consists of an integral number of syllables, so that the shortest possible isolated word is one syllable long. Non-syllables (such as C moras) may not form isolated words.

### 6.4.1 One-mora syllables

Of the three types of mora (CV, V and C), only two (CV and V) may form a syllable (and thus a word) by themselves. (The dictionary entries /N/, an interjection or a contracted form of /no/, will be discounted because they are orthographic representations of an interjection and a fast-speech, not isolation, form. If these exceptional cases are excluded, the generalisation can be made that C moras do not occur in isolation.) Most of the CV and V moras are meaningful words in Japanese, for instance /te/ ‘hand, arm’, /si/ ‘city’, /me/ ‘eye’, /le/ ‘picture’, /lo/ ‘tail’, /lu/ ‘cormorant’. C moras only occur in syllables in combination with a preceding V or CV mora. Since a V or CV mora is an obligatory constituent of syllables, I shall refer to them as syllable-head moras, and I shall describe C moras as non-head moras. One-mora syllables thus have the following structures:

\[
\begin{array}{c}
\sigma \\
\mu \\
C & V
\end{array}
\]

\[
\begin{array}{c}
\sigma \\
\mu \\
V
\end{array}
\]

E.g. /ki/ ‘a tree’ e.g. /hi/ ‘a well’

### 6.4.2 Two-mora syllables

The set of the three basic mora-types yields the following set of nine potential two-mora syllable patterns. Not all of the logically possible combinations are allowed in Japanese phonology, however. In the chart which follows, those marked with a ? are phonetically monosyllabic variants of disyllabic sequences, arising from ‘devoicing’ or ‘eclipse’ of one of the Vs. Such ‘eclipsed’ vowels are indicated with an apostrophe. The entries marked with a * do not occur. Those marked with a ?? are perhaps instantiated by
one or two examples which are nevertheless exceptional in various respects. Those marked with a + may only occur if a CV mora follows, and are thus not isolated word-forms.

<table>
<thead>
<tr>
<th>2nd</th>
<th>CV</th>
<th>V</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st CV</td>
<td>?/des/ 'to be'</td>
<td>CVV</td>
<td>CVC</td>
</tr>
<tr>
<td>V</td>
<td>/is/ 'a chair'</td>
<td>VV</td>
<td>VC</td>
</tr>
<tr>
<td>C</td>
<td>??/tte/ 'going'</td>
<td>*C-V</td>
<td>*CC</td>
</tr>
</tbody>
</table>

Of these nine possible syllable patterns, two (CVCV and VCV) are phonetically monosyllabic variants of two 'full' syllables, and three (CCV, C-V and CC) are highly exceptional, if not completely unattested. There are consequently only four two-mora syllable patterns in the 'productive' part of the grammar: CVV, CVC, VV and VC. If the mora-based analysis of syllable-structure is pursued, these four syllable patterns may be assigned the following constituent structures:

(6.3)

\[
\begin{array}{ll}
\sigma & \sigma \\
\mu & \mu \\
C & V & V \\
\end{array}
\]

e.g. /koo/ 'thus'

e.g. /ii/ 'good'

\[
\begin{array}{ll}
\sigma & \sigma \\
\mu & \mu \\
C & V & C \\
\end{array}
\]

e.g. /saN/ 'three'

/haN/ 'yes Sir!'

\[
\begin{array}{ll}
\sigma & \sigma \\
\mu & \mu \\
V & C \\
\end{array}
\]

e.g. /oN/ 'a sound'

/eQ/ 'eh?'
The syllable-structures I have proposed here are rather different from the more usual division of syllables into onset and rime widely proposed in contemporary phonological analyses of other languages (and often assumed for Japanese, too; see e.g. Abe 1986) and exemplified by figure 5.7 in the previous chapter. However, I shall show that the Japanese pattern is either a stereoisomer or a minor variant of the more common analysis of syllables. The first possibility, stereoisomerism, is predicated on the observation that the left-branching syllable-structure presented above for syllables such as /tii/ and /siN/ is the mirror-image of the right-branching onset/rime structure of the standard theory of syllable-structure. If the analysis of Japanese into moras which I propose is found to be convincing, and is exemplified in other languages too, it might be proposed that the standard theory of syllable-structure should be revised along the lines of X-bar theory or the ID/LP proposal, so that the linear precedence of head and non-head constituents of the syllable is set by a binary parameter, although the asymmetrical branching pattern of dominance in the onset/rime analysis and the mora-based analysis might still be regarded as language-universal.

The second possibility (not incompatible with the first), is that although the structure of Japanese syllables is rather different from the usual pattern, the constituents which occur in Japanese syllables, with the exception of the rime, are the same as those which occur in the conventional syllable structure, that is, onset, nucleus and coda. Correspondences (if not outright identities) can be drawn between the onset/rime model and the mora-based model for three out of four of the traditional syllable constituents. For instance, according to the analysis of two-mora syllables presented above, CV moras contribute the onset and nucleus to a syllable, or, if the V of a CV mora is 'devoiced or 'eclipsed', theoda (e.g. /des'/) or part of the onset (e.g. /s'ki/). V moras contribute the nucleus peak when in first-mora position, and the vocalic offglide of diphthongal nuclei when in second mora position, or, exception-ally, a V mora may contribute a glide onset, as in /i/u/ [ju:] 'to say' (i-l'say', cf. progressive /itte/; infinitive inflection /-u/). C moras contribute the coda to a syllable. If we attempt to reconcile the standard and the mora-based models of syllable-structure, the Japanese syllable template can be represented:

(6.4)
The terms 'antirime', 'coda', 'onset' and 'nucleus' will be used below as functional terms, like 'subject' and 'object' in syntax. They will not be used as formal categories of Japanese phonological structure. This syllable-structure can be defined by a context-free phrase-structure grammar. A first attempt at formulating such a grammar, which can be viewed equivalently either as a set of phrase-structure rules or as a set of local constraints on trees (partial structural descriptions), are the following six rules:

(6.5) \[
\sigma \\
\mu \quad \text{or } \sigma \rightarrow \mu
\]

(6.6) \[
\sigma \\
\mu \quad \mu \quad \text{or } \sigma \rightarrow \mu \mu
\]

(6.7) \[
\mu \\
V \quad \text{or } \mu \rightarrow V
\]

(6.8) \[
\mu \\
C \\
V \quad \text{or } \mu \rightarrow CV
\]

(6.9) \[
\mu \\
N \quad \text{or } \mu \rightarrow N
\]

(6.10) \[
\mu \\
Q \quad \text{or } \mu \rightarrow Q
\]

There are several respects in which these six rules are not adequate as a mechanism to define all and only the one- and two-mora syllables of Japanese:

1. Rule (6.6) does not distinguish between subcategories of mora. It will be necessary to do so, since C moras may only occur syllable-finally, and if the second mora is CV, C must be voiceless and V close. I shall refer to N, Q, non-syllabic V and C[-voice]V[+high] (e.g. /s(ʊ)/, /h(ʊ)/, /s(ɨ)/ etc.) as 'marginal' moras, since they occur in syllable margins, and syllabic V and sequenced CV moras (e.g. /se/ , /to/ etc.) as 'nuclear', since they contain a syllable nucleus. The above grammar sanctions all types of mora in both first and second place in a syllable, and thus overgenerates.

2. Rule (6.8) does not determine the phonological function (onset,
antirime or coda) of different types of CVs, or the marginal vs.
uclear distinction between different types of CV mora.

3 Similarities in the distribution and behaviour of the nasal and
obstruent moras are not accounted for by (6.9) and (6.10).
4 The contextual restriction on the distribution of the obstruent
mora to word-internal position is lacking from (6.10).

In order to rectify these deficiencies, I shall now propose some unification-
based modifications to the classical context-free phrase-structure grammar
given above. With reference to (6.6), I shall first describe the distribution of
consonantal moras in branching syllables in more detail. The second mora
of such syllables must be non-branching (i.e. either V or C), whereas the first
mora must contain a V, and may or may not have an initial C. To encode the
distributional distinction between moras that may form a syllable by them-
selves (CV and V) from those that may not form a syllable by themselves (C),
I shall add the feature [±head] to rules (6.5)–(6.10). X[+head] should be
read as ‘X can be a head’, not ‘X is a head’. C moras are never syllable heads
(since they may not form syllables by themselves) and are thus categorised
as [−head]. V moras may occur as syllable heads, as in the first mora of /ai/
or the second mora of /ie/, or as non-heads, such as the second mora of /ai/
or the first mora of /ie/. V moras will consequently not be specified for [head]
by the phrase-structure rules. Whispered CV moras will be categorised as
[−head], to ensure that they only occur in marginal positions in the syllable.
Therefore, whether CV moras are [+head] or [−head] depends on other fea-
tures, such as the voicelessness of the C or the closeness of the V. If the V of
a CV mora is not whispered, the CV mora in which it occurs is the head of
the syllable. Non-whispered CV moras will therefore be categorised as
[+head]. The [head] feature will also serve to distinguish the two moras of
a branching syllable from each other. The revised form of (6.5) is:

\[
(6.11) \begin{array}{c}
\sigma \\
\mu \\
\mu \\
\end{array}
\begin{array}{c}
[+head] \\
[+head] \\
\end{array}
\begin{array}{c}
or \\
\rightarrow \\
\rightarrow \\
\end{array}
\begin{array}{c}
\sigma \\
\mu \\
\mu \\
\end{array}
\begin{array}{c}
[+head] \\
[−head] \\
[+head] \\
[−head] \\
\end{array}
\]

Rule (6.6) will be replaced by the following:

\[
(6.12) \begin{array}{c}
\sigma \\
\mu \\
\mu \\
\end{array}
\begin{array}{c}
[+head] \\
[−head] \\
[+head] \\
[−head] \\
\end{array}
\end{array}
\begin{array}{c}
or \\
\rightarrow \\
\rightarrow \\
\rightarrow \\
\end{array}
\begin{array}{c}
\sigma \\
\mu \\
\mu \\
\mu \\
\end{array}
\begin{array}{c}
[+head] \\
[−head] \\
[+head] \\
[−head] \\
\end{array}
\]

The generalisation expressed by (6.12) expresses the fact that the few cases
of two-mora syllables in which the first mora is [−head] and the second
mora is [+head], such as /iu/ and /s'ki/ are exceptional. Second-language learners of Japanese must learn the phonetic interpretation of /iu/ as a special case, and in C₁V₁C₂V₂ syllables, the number of attested instantiations of Cs and Vs by which V₁ is whispered and V₂ the head is much less than the number of attested instantiations of Cs and Vs by which V₂ is whispered and V₁ is the head. This is because V₂ is whispered if it falls after a voiceless C₂, whereas for V₁ to be whispered, not only C₂ but also C₁ must be voiceless. In order to include words such as /s'ki/ in this grammar, it would be necessary to add the rule σ → μ[−head] μ[+head] together with some mechanism to ensure that μ[+head] is a CV mora in which the C is [−voice]. This could be done by marking the entire mora as [−voice], a possibility argued for in greater detail below.

If /iu/ is analysed as a syllable which consists, exceptionally, of a [−head] mora followed by a [+head] mora, a similar analysis might also be proposed for syllables containing palatalised onsets such as /kyu/. Such syllables could be regarded as consisting of a [−head] initial CV mora, in which V is /i/, followed by a [+head] mora such as /u/ or /yu/. Thus, using the diacritics ~ to mark the vowel of [−head] moras, and , to mark the vowel of [+head] moras, syllables such as /kyu/ could be analysed as /kiu/. Although there are no synchronic CiV ~ CyV alternations to support this, this proposal accords with what is believed about the history of Japanese (Miller 1967, Poser 1986), in that, for example, contemporary /kyoo/ and /myooto/ are recorded as /keu/ and /meoto/ respectively in a Korean transcription of Japanese dating from 1495 (cf. the contemporary alternation /-mei/ (name suffix) (cf. Ch. ming) ~ /myooji/ (surname + nominal) (cf. Ch. mingzi). Furthermore, the syntagmatic phenomenon of palatalisation of initial coronal consonants before /i/ also occurs to coronal consonants in CyV syllables.

In order to ensure that the subcategories of moras in the syllable rules are respected by the mora structures which may occur in those syllable positions, antirime CV moras must be augmented with the feature [+head] and coda moras with [−head]. Rule (6.8) must be modified as follows:

\[
\begin{align*}
\mu & \rightarrow [\text{+head}] \\
\text{C} & \rightarrow V \\
\text{[+voice]} & \rightarrow [\text{+voice}]
\end{align*}
\]

where the phonetic exponents of C and V are not completely coextensive.

Whispered C[−voice] V[+high] moras and nasal and obstruent moras, must bear the feature specification [−head]:
where the phonetic exponents of C and V are coextensive in time.

The analysis of /Cy/ onsets as /Ci/ moras mentioned above would require the following similar, but not identical structure to be added to the grammar:

\[
\begin{align*}
(6.15) & \quad \mu \\
& \frac{[\text{-head}]}{C} \frac{V}{\text{or } \mu[\text{-head}] \rightarrow C \quad V} \\
& \frac{[\text{+high}]}{[\text{-voice}]} \quad \frac{[\text{-low}]}{[-\text{back}]} \\
\end{align*}
\]

The similarity between the rules (6.14) and (6.15) could be captured by collapsing them into the single rule:

\[
\begin{align*}
(6.16) & \quad \mu \\
& \frac{[\text{-head}]}{C} \frac{V}{\text{or } \mu[\text{-head}] \rightarrow C \quad V} \\
& \frac{[\text{+high}]}{[\text{-voice}]} \quad \frac{[\text{-low}]}{[-\text{back}]} \\
\end{align*}
\]

Here, I have also promoted [back] and [voice] to be mora features rather than V features. Rule (6.7) does not need to be altered, because V moras can be either first or second (or [+head] or [−head]) mora of a syllable. In order to restrict C moras to coda position, (6.9) and (6.10) can be modified as follows:

\[
\begin{align*}
(6.17) & \quad \mu \\
& \frac{[\text{-head}]}{N} \quad \frac{\mu \rightarrow N}{[\text{-head}]} \\
\end{align*}
\]
Declarative analysis of Japanese words

(6.18) \[ \mu \\
| \quad \text{[−head]}
\]
\[ \quad \text{Q} \\
\quad \text{or} \\
\quad \mu \rightarrow \text{Q} \\
| \quad \text{[−head]}
\]

These revisions provide an almost complete characterisation of the distribution of the nasal and obstruent moras. Rules (6.17) and (6.18) could be eliminated from the grammar if C moras are marked with the feature [+consonantal], to generalise across the categories N and Q. The categories N and Q can then be analysed solely in terms of features, without referring to any constituents of \( \mu \), as follows:

(6.19) N moras:
\[ \mu \\
| \quad \text{[−head]}
\]
\[ \quad \text{[+cons]} \\
\quad \text{[+nasal]}
\]

Q moras:
\[ \mu \\
| \quad \text{[−head]}
\]
\[ \quad \text{[+cons]} \\
\quad \text{[−nasal]}
\]

If the [+consonantal] feature is employed, the [−head] specifications in the above two definitions can be predicted by a constraint, as both kinds of C moras are [−head]:

(6.20) \( \mu [+\text{consonantal}] \supseteq [\text{−head}] \)

If the feature [consonantal] is also employed to distinguish C and CV moras from V moras, the CV mora-structure can be represented:

(6.21) \[ \mu \\
| \quad \text{[+head]} \\
\quad \text{[+cons]} \\
\quad \text{[+cons]} \\
\quad \text{[−cons]}
\]

Even though the mora node in this structure bears the feature [+consonantal], constraint (6.20) will not apply because the mora node is marked [+head]. Rule (6.21) takes priority over (6.20) because it is more specific.

Accordingly, V moras would be categorised \( \mu [−\text{consonantal}] \). This reformulation also permits restrictions on the distribution of moras within words to be stated extremely simply, since [+head] moras may constitute a syllable or word, and may appear word-initially, whereas [−head] moras only ever occur after (or at least, always in combination with) a [+head]
mora. [−nasal, −head] moras, in addition, only ever occur before a [+head] mora, that is, between two [+head] moras.

The categorisation of CV, V and C moras in terms of the two features [head] and [consonantal] is summarised as follows:

(6.22)  
CV moras: μ [+consonantal]
V moras: μ [−consonantal]
C moras: μ [+consonantal, −head]

Similarities of distribution between CV and C moras can be expressed by reference to the feature [+consonantal], similarities of distribution between CV and V moras can be expressed by reference to the feature [+head], and the relationship between C moras and offglide/length V moras can be expressed by reference to the feature [−head].

6.4.3 Three- and four-mora syllables
In addition to one- and two-mora syllables, the possibility of adding V moras to V or CV moras to form syllables with diphthongal (two-mora) or even triphthongal (three-mora) nuclei can be combined with the addition of a final C or whispered CV coda mora to make three- and four-mora syllables. Hinds (1986: 404) states that 'any combination of vowels may occur, and in theory there is no limit to the length of such combinations'. In fact, the range of legal three- and four-mora syllables is only a tiny fraction of the range of possible three- and four-mora sequences, as I shall show in this section, and there is only a handful of rather unusual words containing three- or four-vowel sequences.

The set of possible three-mora syllables is illustrated in the following chart. (The symbols ?, ??, *, + and − have the same interpretations as in the chart of one- and two-mora syllables above.)

<table>
<thead>
<tr>
<th>3rd</th>
<th>V</th>
<th>CV</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st and 2nd</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV</td>
<td>??VVV</td>
<td>??VVV</td>
<td>??VVV</td>
</tr>
<tr>
<td></td>
<td>??/aoi 'blue'</td>
<td>??/ais' 'ice'</td>
<td>??/ais' 'ice'</td>
</tr>
<tr>
<td>CVV</td>
<td>??CVVV</td>
<td>??CVVV</td>
<td>??CVVV</td>
</tr>
<tr>
<td></td>
<td>??/baai 'event'</td>
<td>??/baai 'event'</td>
<td>??/baai 'event'</td>
</tr>
<tr>
<td>VC</td>
<td>*VC-V</td>
<td>*VC-V</td>
<td>*VC-V</td>
</tr>
<tr>
<td></td>
<td>*CVCCV</td>
<td>*CVCCV</td>
<td>*CVCCV</td>
</tr>
<tr>
<td>CVC</td>
<td>*CVC-V</td>
<td>*CVC-V</td>
<td>*CVC-V</td>
</tr>
<tr>
<td></td>
<td>*CVCCV</td>
<td>*CVCCV</td>
<td>*CVCCV</td>
</tr>
<tr>
<td></td>
<td>??/paNts 'pants'</td>
<td>??/paNts 'pants'</td>
<td>??/paNts 'pants'</td>
</tr>
<tr>
<td></td>
<td>??/hoN-te/ '(someone says) that books'</td>
<td>??/hoN-te/ '(someone says) that books'</td>
<td>??/hoN-te/ '(someone says) that books'</td>
</tr>
</tbody>
</table>
Declarative analysis of Japanese words

Of the twelve three-mora combinations, not all are possible syllables. \([\mu(C)V, [\mu(C)]_\mu, [\mu V]_\mu\) combinations are ill formed. Adding \(V\) to \(N\) always results in a disyllable, \([\sigma(C)V N, [\sigma V]_\sigma\) since \(N\) can only constitute a coda. A \(V\) mora cannot be appended to \(Q\), since \(Q\) codas only occur before \(CV\).

Tautosyllabic triphthongs are limited to the pattern ‘onglide–peak–offglide’, where the onglide and offglide must be closer in quality than the peak. In addition, the distinction between /i/ and /e/ and the distinction between /u/ and /o/ is neutralised in ‘offglide’ position, which I shall indicate by writing ‘i/e’ and ‘u/o’. The possible triphthongs are thus:

<table>
<thead>
<tr>
<th>Opening–closing triphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>onglide</td>
</tr>
<tr>
<td>i, u</td>
</tr>
<tr>
<td>i, u, e, o</td>
</tr>
</tbody>
</table>

More formally, ‘i/e’ and ‘u/o’ could be represented as the categories [-low, -back] and [-low, +back] respectively. Since in this section the precise representation of vowel height is not relevant, I shall use the informal terms ‘close’, ‘mid’ and ‘open’. However, I shall argue below that the three vowel heights should be analysed using two binary features [high] and [low], and not a three-valued [height] feature, as in the description of English in the next chapter, for two reasons. First, there is no evidence of cyclic shift in Japanese vowel height, which will be used to warrant a single three-valued feature in the English analysis. Second, there is a vowel ‘fusion’ phenomenon in Japanese in which the phonological combination of a high vowel with a low vowel results in a mid vowel. For example, Okawa (1982: 428) reports the fast-speech vowel coalescence phenomenon /itaï/ \(\Rightarrow\) /itee/ ‘painful’, and Hasegawa (1979a: 129) reports that in men’s vulgar speech such forms as /semai/ \(\Rightarrow\) /semeel/ ‘narrow, small’ are common. A related, though different, phenomenon is represented by forms such as /sugoi/ \(\Rightarrow\) /sugeel/, in which mid (grave) /o/ followed by (high), non-grave /i/ forms mid, non-grave /ee/. The /tal/ \(\Rightarrow\) /ee/ phenomenon is susceptible to a unification-based analysis in which mid vowels are specified as [+high, +low], the unification of [+high] and [+low]. In order for this analysis to work it is necessary to categorize mid vowels as [+high, +low], rather than as [−high, −low]. contra Chomsky and Halle’s (1968: 305) claim that [+high, +low] is beyond ‘the phonetic capabilities of man’. While simultaneously raising and lowering the tongue may be physically impossible, classifying close and mid vowels as [+high] and mid and open vowels as [+low] does not give rise to any (phono)logical paradox, if dis-
tinctive features are interpreted relationally as referring to natural classes, rather than according to a simplistic phonetic theory. Even if these Japanese examples are regarded as marginal, languages with comparable but more productive open/close vowel coalescence may be found to support this criticism of Chomsky and Halle's claim. (The names of these features [high] and [low] will be replaced by [compact] and [diffuse] below, for reasons which will be spelt out later.) Three-vowel sequences in which the 'onglide' and 'offglide' are no less close than the peak constitute a larger set of potential triphthongs, enumerated as follows. (Unattested sequences are marked with an asterisk. ‘S-J’ indicates Sino-Japanese words, and ‘Ch.’ denotes their Modern Standard Chinese cognates, which are presented for comparison. ‘(i-adj.)’ denotes adjectives formed using the adjectival suffix /-i/, and ‘(u-inf.)’ denotes infinitive verbs formed using the infinitival suffix /-u/. ‘(n.)’ denotes nouns, where it is necessary to disambiguate them from (i-adj.) and (u-inf.) words.)

Closing triphthongs

Open–mid–close: aei (/yaei/ 'bivouac'), aoi (/ao+i/ 'blue–green' (i-adj.)), *aeu, *aou.
Open–mid–mid: *aee, aoe (Aoe, a name), *aeo, aoo (/kaoo/, a toilettry brand name).
Open–open–close: aai (/ba+ai/ 'occasion, event'; the /ai/ is S-J, cognate with Ch. hé), *aau.
Open–close–close: *aii, aui (/aya+u+i/ 'critical' (i-adj.)), *aiu, *auu.
Mid–mid–close: eei (S-J /see+i/, /sei+i/ 'sincerity' Ch. chéngyi), oei (/to+ei/ 'made in Tokyo'), *eoi, ooi (/oo+u+i/ 'many, numerous' (i-adj.)), *e eu, *o eu, eou (/seou/ 'be burdened with'), oou (/oo+u/ 'to hide' (u-inf.)).

Opening triphthongs

Declarative analysis of Japanese words


Level triphthongs


Open–open–open: *aaa.

Opening–closing triphthongs


6.4 Syllables and moras

6.4.4 Absence of triphthongs within Japanese morphemes

The enumeration of possible triphthongs given above demonstrates that, far from being potentially unlimited, as the quotation from Hinds (1986) above states, even three-vowel sequences are extremely restricted in Japanese, because words are not just sequences of consonants and vowels: those sequences are structured by syllables and moras. While most possible two-vowel sequences are attested in the Japanese lexicon, most three-vowel sequences are unattested. The foregoing list shows that those which are attested can be divided into four classes:

1 Native Japanese words containing a morpheme boundary which may be a syllable boundary, for instance /-i/ adjective inflection or /-u/ infinitive inflection on stems which end in a diphthong. For example: /ao+i/ 'blue-green', /oo+i/ 'many, numerous', /too+i/ 'far', /nio+u/ 'give off a smell', /yo+u/ 'gets drunk/seasick', /kayo+u/ 'commute', /kayu+i/ 'itchy', /yo+i/ 'good', /nio+i/ 'a smell' (n.), /nia+u/ 'suit, match well'. Also of this kind are words containing a non-inflectional morpheme boundary, for instance /siri+ai/ 'acquaintance', /yamu+oe+nai/ 'unavoidable', /yamu+oe+zu/ 'unavoidably' (from /yamu/ 'stop'), /iki+oi/ 'energy, power', /boo+eki/ 'trade'. While morpheme boundaries do not always coincide with syllable boundaries (or even mora boundaries, e.g. /kak+u/ 'write' infinitive), if a morpheme boundary between two V-units is analysed as a phonological syllable boundary, these examples support the statement that syllables in underived, monomorphemic words do not contain more than two successive V units.

2 Sino-Japanese words containing a syllable boundary that is a morpheme boundary, for instance /sui+ei/ 'swimming' = /sui/ 'water' + /ei/ 'swimming', /hi+yoo/ 'daily expenses', /dyuuyoo/ 'important', /tyuu+i/ 'attention, caution', /syuu+i/ 'circumference, surroundings', /koo+en/ 'public park', /oo+en/ 'assistance', /si+ai/ 'match, game'. Long vowels and diphthongs are common in Sino-Japanese morphemes, respectively deriving from VN and VV rimes in Chinese. The general one-to-one correspondence between syllables and morphemes in Chinese precludes Sino-Japanese words from containing more than two vowels within each syllable.

3 One or more of the Vs in VVV may be analysed as a glide C (i.e. /y/ or /w/), for example /yo+u/ 'gets drunk/seasick', /kayo+u/...
Declarative analysis of Japanese words

‘commute’, /kayu+i/ ‘itchy’, /iya/ ‘disagreeable’, /iwa/ ‘crag’, /niwa/ ‘garden’, /yoo/ ‘business, engagement’, /iyo-yo/ ‘more and more’, /yue/ ‘reason’, /yu-uu/ ‘evening’, /yo+i/ ‘good’, /huyu/ ‘winter’, /wai/ ‘short’, /kyuu/ ‘nine’. Although it is an analytical decision whether to treat glides as Cs or Vs, the former option allows the restriction on the length of V sequences to no more than two to be maintained. In addition, if /yl/ and /w/ are analysed as Cs, not Vs, then there are no four-vowel sequences either, since the only candidates, words like /yaei/ ‘bivouac’ and /iwai/ ‘congratulate, celebrate’, contain a glide.

4 Monomorphemic words. The number of cases of morpheme-internal three-vowel sequences which remain when agglutination, inflection, Sino-Japanese syllable/morpheme boundaries, and glides are taken account of is very small indeed, aei (/yaei/ ‘bivouac’), iie (iie/ ‘no’), uai (/uai/ ‘condition, health’) being perhaps the only examples. The most restrictive analysis of within-syllable vowel sequences that is consistent with the data presented above is that the above words are disyllabic (e.g. /ba-aai/, /ya-ei/, /ii-e/ or /i-ye/ and /gu-aai/), and that Japanese syllables contain no more than two vowels in the nucleus. (There may be whispered vowels in the onset or coda, however.) Thus the maximal lexical syllable structure is the three-mora sequence CV–V–C(V). More complex onsets and codas may be derived by concatenation of CV moras containing a whispered or elided V, for example /tskue/ (from /tsukue/), /ski/ (from /suki/), /kuttsk/ (from /kuttsuku/), but since these always have alternative pronunciations with the full vowel, it is not necessary to regard such forms as containing complex onsets or codas in the lexicon. I shall return to this matter of apparent consonant clusters at the end of this chapter, after developing a declarative analysis of devoicing and voiceless vowel ‘elision’ which will help account for such cases.

The structure of the maximal three-mora syllable is defined by the following phrase-structure rule:

(6.23) $\sigma \rightarrow \mu [+\text{head}] \mu [+\text{head}, -\text{consonantal}] \mu [-\text{head}]$

Having examined the composition of syllables in terms of various kinds of moras as well as their relationship to the conventional syllable constituents, onset, nucleus and coda, in the following section I shall examine CV moras containing close vowels, in order to develop a unified declarative account of a number of related phenomena initially discussed in chapter 3:
coarticulation, nucleus devoicing, fricative nuclei, nucleus elision and regressive palatalisation.

6.5 Structure of CV moras

In addition to the constituent structure of syllables, one category of mora, the CV mora, has an internal constituent structure and internal constraints on the temporal interpretation of C and V. It is this structure and its temporal interpretation that I shall now examine, before turning to the particular case of CV moras with close vowels which are involved in nucleus devoicing, fricative nuclei, nucleus elision and regressive palatalisation.

In table 3.1, I listed the combinations of consonants and vowels that may form a CV mora. The categorisation of consonants and vowels in that table follows the traditional order of the kana writing systems. The entries in that table are minimal 'systematic phonetic' representations in which I gave only as much phonetic (i.e. allophonic) detail as is usually recognised and accounted for in phonemic analyses. In section 3.3.1, I described the variability in the pronunciation of the C and V constituents of some CV moras. Since CV moras with close vowels are of special relevance to the phonotactic structure of Japanese, I presented some more detailed phonetic observations of such moras, taken from my first-hand observations, in table 3.2. To account for the secondary articulation of C with V, for example, it could be argued in a segmental, derivational account that the vocalic quality features of the vowel are (i) copied to the consonant as palatal secondary articulation (velar secondary articulation in the case of 'back-nucleus' moras), before (ii) the vowel is deleted. The apparent movement of features from the V segment to the preceding C segment which results from these two operations can be portrayed as follows:

(6.24)

```
\begin{array}{c}
\text{C} \\
[\beta \text{ strident}] \\
[\gamma \text{ continuant}] \\
[\text{\text{-nasal}]}
\end{array}
\quad \rightarrow \quad
\begin{array}{c}
\text{V} \\
[+\text{high}] \\
[-\text{low}] \\
[\alpha \text{ back}]
\end{array}
```

However, if the analysis of nuclear friction includes the proposal that a segment or timing unit is deleted, as in the transformational analysis in
chapter 3, the material which remains after such deletion ought to be of shorter duration than a CV mora, given the relation between segments and timing-units assumed in segmentally based theories. I showed that this is not the case, however: the duration of the period of friction in fricative-nucleus moras is greater than in their whispered and close vocalic counterparts, so if a vowel deletion rule is proposed, it is also necessary to include a rule or principle in the grammar of Japanese that results in the reassignment of syllabicity to the friction period, or a compensatory lengthening rule to account for its increased duration. Unlike compensatory lengthening in other languages (see Wetzels and Sezer 1986 for an overview), this would be onset lengthening, not coda lengthening. Perhaps in ‘antirimes’ this type of compensatory lengthening could be expected, however, since codas in ‘rime languages’ and onsets in ‘antirime’ languages are directly dominated by σ in both cases. In the discussion which follows, I shall develop a declarative analysis of these phenomena which does not involve copying, deletion or compensatory lengthening.

In section 6.4, the distribution of voiceless high vowels was partly accounted for in terms of syllable-structure. It was shown that CV moras may occur as either first or second mora in a syllable. In first-mora position, all the legal CV combinations of table 3.1 are found. But in second-mora position, in lexical representations, the CV must act as the syllable coda. Only CV moras containing a voiceless C and a close vowel can do this, because it is only that type of CV mora which can have a ‘whispered’ or ‘eclipsed’ V, and thus not contribute another nucleus to the syllable. Thus in the present analysis, ‘devoicing’ is not regarded as conditioned by sequences of consonants and vowels, but voicelessness, the consonant/vowel distinction, and the distribution of moras in syllables will be treated in an integrated fashion as aspects of syllable-structure. In the remainder of this chapter I shall examine the distribution of the features of voicing and palatality, and their interaction with the categorial distinctions between Cs and Vs, and stops and fricatives.

6.6 Regressive palatalisation

The first aspect of the data in table 3.2 that I shall consider is the palatalisation of /s/ and /t/ before /i/ which results in the phonetic interpretations [ japonese] and [t̚ japonese] respectively. By employing the mora as a phonological category, the (tautomoraic) palatalisation of coronal onsets in moras with close vowels can be encoded simply by adding the ‘long-domain’ or ‘prosodic’ feature-cluster [+high, -low, -back] to the mora node, that is,
6.6 Regressive palatalisation

(6.25)

\[
\begin{array}{c}
\mu \\
\begin{array}{l}
[+\text{high}] \\
-\text{low} \\
-\text{back}
\end{array} \\
C \\
\begin{array}{l}
[-\text{voice}] \\
[+\text{coronal}] \\
-\text{sonorant}
\end{array} \\
V
\end{array}
\]

e.g. /si/, /ti/ ([fi], [ʃi]).

Such factorised features are distinctive properties of the mora as a whole, rather than of the onset or nucleus independently. While they might be regarded as features that trickle to both daughters, trickling is not necessary to characterise the set of well-formed phonological representations in Japanese, nor to provide them with a phonetic interpretation, if the account of phonetic interpretation of non-segmental phonological representations in chapter 5 is applied to the analysis of Japanese presented here. Moras with distinctively palatalised onsets, such as /tya/ ([tʃa]) and /kya/ will be represented, like /si/ ([ʃi]) and /ti/ ([tʃi]), with the 'vocalic' features [+high, -low, -back] borne by a single, unbranching C onset.

Although the apparent 'spreading of palatality from /i/ to coronal onsets' can be analysed declaratively by treating the features [+high, -low, -back] as a mora-level prosodic unit, the 'spreading of palatality' to a C mora in a preceding syllable as in /matti/ ([matʃi]) is more problematic, since the (onset) C and the preceding (coda) C-mora are not sisters. Sharing the onset C ambisyllabically with the coda of the preceding syllable as in (6.26) would yield an account of palatalisation in which the second onset inherits palatality from the antirime of the second syllable, and the first coda inherits palatality from the ambisyllabic C unit.

(6.26)

\[
\begin{array}{c}
\text{Syllable} \\
\begin{array}{c}
\text{Antirime} \\
(C) \\
\text{Coda} \\
(V) \\
\text{Antirime}
\end{array} \\
\text{m} \\
\text{a} \\
\text{t} \\
\text{t} \\
\text{i} \\
\text{g} \\
\text{e} \\
\text{n} \\
\text{k} \\
\text{i}
\end{array}
\]

\[
\begin{array}{l}
[+\text{high}] \\
-\text{low} \\
-\text{back}
\end{array}
\]

\[
\begin{array}{c}
\text{Syllable} \\
\begin{array}{c}
\text{Antirime} \\
(C) \\
\text{Coda} \\
(V) \\
\text{Antirime}
\end{array} \\
\text{m} \\
\text{a} \\
\text{t} \\
\text{t} \\
\text{i} \\
\text{g} \\
\text{e} \\
\text{n} \\
\text{k} \\
\text{i}
\end{array}
\]

\[
\begin{array}{l}
[+\text{high}] \\
-\text{low} \\
-\text{back}
\end{array}
\]
This analysis is not satisfactory, however, as it would require a branching coda rule such as ‘Coda → C C’ to be constrained to apply only if the second C is ambisyllabic. In preference to this, we need only to ensure that the group of vocalic features denoting palatality, [+high, –back, –low], are shared between the coda and the onset:

(6.27)  Word → σ[coda: C₁] σ[antirime: [C: α high, β back, γ low]]

Although it is not associated with an independently motivated level of syllable structure, the shared unit C₁ is a prosodic unit because it is an autonomous, non-terminal, non-segmental, feature-structure. This configuration has been extensively studied by Itô (1986) and Yoshida (1990).

A Unification-based Phrase-Structure Grammar can thus represent the ‘spread’ or domain of palatality and its localisation to specific limits, though it is more verbose to state constraints allowing sharing between non-sisters than between sisters (as noted in section 5.5.9 above). Since the propagation of prosodic features is restricted by the phrase structure of the representations, the unification-based analysis accounts for this localisation, whereas in a transformational ‘copying’ analysis, no such limits are enforced by the formalism, and the localisation of regressive palatalisation would be arbitrary.

If representations such as (6.25) are sanctioned by the grammar formalism, other features than just [high] and [back] could be analysed prosodically too. For instance, [round] can be specified at the mora level in Japanese, as its value can be specified for the whole mora, since the C and V of a CV mora never have conflicting values for [round]. (In the mora /wa/ there is no contradiction, as /w/ is not rounded, but is a spread bilabial or labiodental frictionless continuant.) This removes the need for a regressive assimilation rule, but more importantly, leads to the following consequences: (i) when the grammar formalism contains a natural definition of ‘domain’, as in this case, the notion of a copying ‘process’ is dispensable; (ii) since all the vowel-features can be specified for entire moras, the need for V elements as (pre)terminal categories at all is called into question. The second proposition has also been defended at length by Hyman (1985), and now has considerable support both within Autosegmental Phonology (see Prince 1984) and in other non-linear frameworks, such as Government Phonology and Dependency Phonology.

This discussion has now reached a position in which the traditional partition between ‘consonantal’ and ‘vocalic’ features which I assumed to begin...
with is no longer necessary. The features [high], [low], [back] and [round] are treated as prosodic mora features relevant to the phonology of both ‘consonant’ systems (e.g. palatalised vs. non-palatalised) and ‘vowel’ systems. There remain, however, some features which are specific to the consonant system alone (such as [nasal] and [strident]). C and V features can thus be partitioned into two disjoint sets, since V features are prosodic (mora-level) and C features are terminal features. Since the C and V features may be partitioned, C features may also be associated in lexical representations with the mora nodes, as their trickling to the daughter C node is completely predictable. Under this circumstance, establishment of C and V nodes in lexical representations would be redundant in two respects – ordering and lexical categorisation – and therefore contrary to the aim of non-redundant lexical representations to which generative phonology aspires. A minimally redundant representation of the structure of moras such as /si/ and /ti/ is the following, in which there are no C and V labels, the C features [−voice, +coronal, −sonorant] can be predicted to trickle to one of the daughter nodes, there are no ordering redundancies among the features, [−consonantal] is predictable from [−sonorant], and, being incompatible with [−consonantal], the vocalic feature [−consonantal], if it needs to be specified at all, can be predicted to trickle to the other daughter node. The daughter nodes are unordered, since the relative order of C- and V-type nodes can be predicted on the basis of the features which they bear. The CV moras /si/ and /ti/, for example, can be represented as follows:

\[
\mu
\begin{array}{l}
+\text{head} \\
+\text{high} \\
-\text{low} \\
-\text{back} \\
-\text{voice} \\
+\text{coronal} \\
-\text{sonorant}
\end{array}
\]

\[
\begin{array}{l}
X \\
X
\end{array}
\]

Although it represents a departure from contemporary phonological orthodoxy (but see Hirst 1985), this solution has several attractions. First, there are no C or V units, since this distinction is predictable from the distinctive features. Second, vowel and consonant features are specified for the entire mora: that is, the secondary articulation features of initial consonants derives not from regressive spreading of features from the nucleus, but from prosodic features of the mora. This is not only more logically
satisfactory and simpler than segmental analyses (being free of the need for
copying rules), but it is also phonetically preferable, as we shall see.

In lexical representations, each mora is represented as a redundancy-free
structure of distinctive features. In CV moras, each mora’s feature-structure
can be partitioned into two substructures. One of these subsets contains the
category-valued feature [vocalic], the other [consonantal]. In the case of
consonantal moras Q and N, only the [consonantal] category-valued
feature is lexically specified (regression of palatality may add [vocalic] fea-
tures); in V moras, only the category value of [vocalic] is lexically specified.
The features in the value of [consonantal] define phonological distinctions
in the stricture parameters of speech production. This set of features
includes specification of the place and degree of stricture, but not the
‘manner of articulation’. The [vocalic] feature-structure is complementary
to the [consonantal] feature-structure, and its features partially encode dis-
tinctions in the phonetic resonance parameters of speech production. In
order to capture generalisations that are pertinent to both the stricture and
the resonance exponents of a mora, I shall employ some features from
Jakobson, Fant and Halle (1952). In the main, the customary articulatory
and acoustic interpretations of these features are supposed, although their
segmental status is not. Furthermore, the greater structural expressiveness
of the current phonological framework obviates the need for some of the
SPE features, namely [delayed release] and the vowel features [high], [back]
and [low], which are replaced by [vocalic: [compact]], [vocalic: [grave]] and
[vocalic: [diffuse]] respectively.

The feature [grave] has been employed here, contra Chomsky and Halle
(1968: 303) and Goldstein (1995), as it is well motivated in a number of
respects (see Campbell 1974): (i) the Jakobsonian interpretation of [+grave]
expresses the relationship between backness and labiality that cuts across
traditional primary distinctions of vowel and consonant; (ii) likewise, the
interpretation of [−grave] captures the relationship that holds between front
vowels and palatal and palatalised consonants; (iii) thus, [grave] expresses
the distributional relationships between ‘allophonic’ distinctions of frica-
tives in Japanese, and the vowels with which they are combined.

These three points are exemplified by the group of Japanese moras tradi-
tionally analysed as /ha/, /hi/, /hu/. Omitting much detail, these may be
represented phonetically as:

/ha/: [ha] – grave stricture, grave resonance.
/hu/: [fu] – grave stricture, grave resonance.
Place of articulation is further distinguished by the [compact] feature,labial and alveolar articulation being [-compact], palatal and velar articulation [+compact]. Compactness is also used to differentiate open resonance, which is [+compact], from non-open (i.e. mid or close) resonance, which is [-compact]. Thus, [ha] is grave and compact throughout, whereas [ϕu] is grave and non-compact throughout.

The distributional, phonological and acoustic phonetic naturalness of the compactness contrast in Japanese is further supported by the fact that items with non-compact resonance must be distinctively specified for the feature. As shown in tables 3.1 and 3.2, gravity is not distinctive with compact resonance (i.e. /a/). Compact resonance is redundantly specified with the gravity of the associated stricture; compare grave [ha] and [ka] with non-grave [sae] and [tae], all of which are compact throughout. This detail of the pronunciation of open vowels is subject to great variability, as would be expected for an opposition that is not lexically distinctive. Hinds (1986: 397–8) claims that in certain Tokyo pronunciations, ‘fronting’ of /a/ is lexically idiosyncratic, occurring in /atası/ ‘I’, but not in /anata/ ‘you’. To make his analysis more symmetrical, McCawley (1968) ‘hypothesises’ a gravity opposition in compact vowels, and then justifies the validity of his hypothesis with a discussion about the history of the Japanese vowel system. In the present analysis, attention to phonetic detail confirms McCawley’s hypothesis. Likewise, diffuseness is not distinctive in nongrave stricture (e.g. /t/, /s/) with nongrave, noncompact resonance (i.e. /i/, /e/), but the diffuseness of nongrave strictures is that of the associated resonance. This constraint can be represented by value-sharing:

\[
\begin{array}{l}
\text{consonantal:} \\
\quad \begin{array}{l}
\text{grave: -} \\
\text{diffuse: } -
\end{array} \\
\text{vocalic:} \\
\quad \begin{array}{l}
\text{grave: -} \\
\text{compact: -} \\
\text{diffuse: } [\hat{\imath}]
\end{array}
\end{array}
\]


According to this proposal, a partial representation of the mora phonemically analysed as /ti/ is presented in (6.30). (For clarity, this representation does not include any specification for [voice] or [nasal].) Compactness, diffuseness and gravity are specified prosodically for the entire mora. This will not necessarily be the case for all moras, as moras
with compact, non-diffuse obstruence and non-compact, diffuse resonance, such as /ki/ also occur. I have simply analysed all features common to both obstruence and resonance 'gestures' as prosodic. Continuance is not specified among the resonance features, since it can be predicted as in (6.37) below.

\[
(6.30) \quad \mu
\]

\[
\begin{array}{c}
\text{[-grave]}
\text{[-compact]}
\text{[+diffuse]} \\
\text{[consonantal: \text{\textbackslash}j\text{\textbackslash}]} \\
\text{[-continuant]}
\end{array}
\quad [\text{vocalic: \text{\textbackslash}i\text{\textbackslash}}]
\]

Since the [vocalic] daughter is the head of the mora, the prosodic features [-grave, -compact, +diffuse] can be shared with the [vocalic] constituent by a general principle, the Head Feature Convention of Gazdar et al. (1985). Consequently, the [vocalic] feature does not need to be subscripted \(\text{\textbackslash i}\), though the fact that [vocalic] is the head of the mora feature-structure would need to be encoded in some other way. Compactness, diffuseness and gravity are closely connected, and in some sense codependent, since they all phonetically denote aspects of the articulatory displacement of the tongue and lower jaw and the acoustic resonance characteristics of the vocal-tract transfer function. Vocalicity and continuance, however, are of a different order, since they relate to articulatory envelope. This distinction is reflected in the fact that [compact], [diffuse] and [grave] may be prosodic features which are not associated with either stricture or resonance in particular, whereas [vocalic] and [continuant] are specifically associated with (respectively) the head and non-head daughters of the mora node in the phonological representation which I have proposed. Both stricture and resonance are defined in terms of interdependent features of gravity, compactness and diffuseness in (6.30). Because [vocalic] and [consonantal] are not temporally ordered in this analysis, (6.31a) is exactly equivalent to (6.31b):

\[
(6.31) \quad \begin{array}{lll}
a. & \mu & [\text{vocalic}] [\text{consonantal}] \\
\quad & \mu & [\text{consonantal}] [\text{vocalic}]
\end{array}
\]

6.7 Phonetic interpretation of CV moras

In the preceding discussion I considered the phonological representation of palatalisation of coronal consonants in combination with /i/, phonological
aspects of coarticulation and a kind of assimilation, regressive palatalisation, in a declarative account of Japanese phonotactics. I proposed to represent each mora as an unlinearised structure of distinctive features and to describe the distribution of phonological categories in syllables by suprasegmental phonotactic constraints, such as phrase-structure rules and feature co-occurrence restrictions. These analytical decisions have a number of significant consequences: (i) palatalisation is not treated as a directed (regressive) assimilatory process; (ii) ordering redundancies are removed from lexical representations; (iii) much syllable-internal structural information need not be represented in the lexicon. On this basis, I shall argue in this section that fricative nuclei do not result from vowel deletion, as claimed by the transformational analysis in chapter 3, but are an epiphenomenon of the temporal interpretation of lexically unlinearised structures. Affrication will be shown to be an epiphenomenon of the temporal interpretation of coronal stops in combination with /l/. I shall also show how whispered vowels are accounted for in this framework, and how the four phenomena - affrication, fricative nuclei, whispered vowels, and temporal interpretation - are non-arbitrarily related to each other.

The traditional claim, discussed in chapter 3, that nucleus whispering or friction is specific to voiceless environments is not wholly correct. On a number of occasions, I have observed nucleus whispering and friction before lax, voiced consonants (often the first consonant of the following word). One of my transcriptions, which records nucleus friction in accelerated speech and whisper in slower speech, before a lax, voiced consonant in both cases, is (6.32):

(6.32) It's an insect (mushi desu).

\( \text{muf(i)des'} \)

Note: J(i) is a complete fricative in faster speech.

In (6.32), the first mora of mushi, /mu/, is louder than the second mora /si/, it is voiced throughout, and may bear a high rising-falling pitch-contour. The voicelessness or friction of /l/, then, appears by contrast with the first syllable to be an exponent of syllable weakness, this despite the fact that the second mora is accented (cf. /músil/, 'selflessness', with accent on the first mora). Since nucleic whisper or friction can occur before lax, voiced consonants, it would be more accurate to state that nucleic whisper or friction occurs more readily before tense, voiceless consonants than before lax, voiced consonants. Similarly, although it is typically only close vowels which devoice, on occasions this phenomenon extends to mid vowels too.
In my records, however, nucleic friction or whisper only occur after tense, voiceless onsets in unaccented moras, in agreement with other descriptions of Japanese phonology. The dependency between the voicing of a mora’s onset and the possibility that its nucleus will also be voiceless is apparently not accidental, but neither is it simply a case of mechanical inertia, since it is accent-governed, as I shall now establish. The traditional segmental analysis of whispered or devoiced vowels is predicated on the twin assumptions that (i) vowels are by default voiced – that is, voiceless vowels are the exception rather than the rule; and (ii) voicing is specified segment by segment. According to my observations, plain voicing is rather rare in Japanese. This is reflected in Shapiro’s (1974) analysis, which proposes that tenseness, and not voicing, is the primary paradigmatic distinction between tenues and mediae in Japanese. Indeed, various types of tense excitation, including aspiration, voiceless oral friction, pharyngeal friction, breathy voice, creaky voice and whisper are much commoner in my records than voice or voicelessness. The so-called voiced or lax (non-nasal) onsets are actually only very rarely phonetically voiced; they are virtually restricted to accented moras, reduplication forms and compound words; and they are specially marked in the orthography: moras with voiced non-nasal onsets are written using the symbol for the mora with the corresponding voiceless onset, together with a voicing diacritic mark called ‘dakuten’ or ‘nigori’. So /ga/ is written just like /ka/, but with the addition of the voicing diacritic. According to the assumptions presented in the previous paragraph, this is an idiosyncrasy of the Japanese writing system. Alternatively, though, ‘nigori’ can be viewed not just as marking the laxness of the onset, but also as encoding the information that the nucleus is never devoiced. According to this idea, phonological voicing operates in Japanese as a mora-domain (rather than a segmental-domain) feature, and phonetic voicing in the nucleus of moras with voiceless onsets is an exponent of accent. The unmarked state of affairs is:

(6.33) a. 

\[
\begin{array}{c}
\mu \\
[+\text{voice}] \\
(\text{Phonetically}) \\
(\text{Phonetically}) \\
\text{voiced onset} \\
\text{voiced nucleus}
\end{array}
\]

b. 

\[
\begin{array}{c}
\mu \\
[-\text{voice}] \\
(\text{Phonetically}) \\
(\text{Phonetically}) \\
\text{voiceless onset} \\
\text{voiceless nucleus}
\end{array}
\]
In addition to these, there is also the marked case:

\[
\begin{array}{c}
\mu \\
[-\text{voice}] \\
[+\text{accent}]
\end{array}
\]

(Phonetically) voiceless onset (Phonetically) voiced nucleus

I propose, therefore, that the feature [voice] is a mora-level feature, in line with *kana* orthographic practice, and that the grammar should define not when vowels are 'devoiced', but rather when voiced vowels occur in voiceless moras (the marked case). Now I shall bring fricative nuclei into this analysis of voicing and accent. As with whispered vowels, rather than being an all-or-nothing assimilation phenomenon, the present analysis views fricative nuclei as one point in a tempo- and accent-dependent 'strength' scale:

- Voiced vocalic nuclei
- Whispered vocalic nuclei
- Fricative nuclei

Accented (strongest)
Unaccented
Unaccented/accelerated speech (weakest)

This scale is phonetically quite natural. In traditional descriptive articulatory terms, the distinction between fricatives and vowels is a difference in the degree of approximation of the active and passive articulators: close approximation in the case of fricatives, and open approximation in the case of vowels. This traditional analysis is oversimple, however, because frictionless continuants (semi-vowels) cannot be accommodated in this simple account: they cannot have a stricture of close approximation, for then they would not be frictionless, yet they cannot be of open approximation either, for that would make them vowels. The *articulatory* definition of a fricative is circular, for a fricative is an articulation produced with close approximation, and close approximation is that degree of stricture which produces audible friction (or, in 'pure' articulatory terms, that degree of stricture which forms fricatives, which is a tautology). If a sustained frictionless continuant such as \[\text{j} \] is produced, and then the voice is 'turned off', the resulting sound is the voiceless fricative \[\text{k} \], despite the fact that the tongue is held in the same position *vis-à-vis* the palate. Clearly, then, the distinction between fricatives, frictionless continuants and close vowels is not dependent solely on the traditional articulatory distinction between open and close approximation.
Declarative analysis of Japanese words

(see Catford 1977: 120). A phonetic relationship between voiceless friction and whispered vowels is not problematic. These observations call into question the traditional hard-and-fast phonetic division of speech events into 'consonants' and 'vowels', which complements the arguments raised above against C and V units in Japanese phonological representations.

CV combinations of stricture or resonance features can be represented as a dag or feature structure rather than as a tree or ordered graph in order to eliminate temporal ordering from lexical phonological representations, except where it is distinctive. For example, moras in which an obstruence maximum phonetically precedes a resonance maximum, as in [tji], can be considered as the phonetic interpretation of the same abstract unordered phonological representation as the corresponding phonetic forms with fricative nuclei, such as [tj:], in which the obstruence maximum co-occurs with the resonance maximum. In the first case, the resonance and obstruence envelopes (i.e. the way in which resonance and obstruence change in prominence over time) are 'out of phase', whereas in the second case they are 'in phase'. The difference between the two phonetic interpretations is thus a distinction of phasing, which is one of the operations involved in the temporal interpretation of abstract phonological representations (Abercrombie 1964, Kelly and Local 1989: 86–9). Under this analysis, deletion of the vowel need not be invoked:

\[(6.35) \quad \text{[tji]} \quad \text{vs.} \quad \text{[tj:]}\]

The relationship between accent and nucleic friction is thus analysed as a chain of logical consequences: (i) one of the phonetic exponents of phonetic unaccentedness is short duration of the phonetic interpretation of the mora; (ii) because of this short duration, the coarticulation of the stricture and resonance components of the phonetic interpretation of the mora is more pronounced, and they are produced more 'in phase' with each other; (iii) the coincidence of friction and resonance which results is interpreted in segmental analyses as a 'fricative nucleus'. Although this 'phasing' account explains the relationship between unaccentedness and fricative nuclei, it does not explain why coronal stops are affricated – that is, have a fricative phase – which is a prerequisite for the above account to hold for affricated stops as well as fricatives.

As an example, I shall now present and discuss the representation of the
mora with exponents [tʃi] ~ [tʃː]. The stricture feature [continuant] distinguishes partial from complete closure, or in other words a long envelope from a short envelope. Long envelopes are characterised by the possibility of being freely extended, subject to respiratory restrictions. (Catford (1977: 128) calls this property ‘prolongability’ and ‘maintainability’.) Short envelopes, on the other hand, may not be extended. (Catford (1977: 129) calls this property ‘momentariness’.) A non-continuant has a duration, albeit very brief, which has a lower potential maximum than that of continuants. Since even complete closure can be freely maintained for relatively long periods, as in ‘syllabic stops’ (see Hoard 1978), such articulations would in this analysis be linguistically characterised as continuants. An example from the current analysis is the mora of gemination, Q, which is complete closure that nevertheless continues in time for longer than a non-continuant stop (e.g. an onset).

The analysis of fricative nuclei assumes this definition of continuance, and the account of affrication presented below embodies the interaction of long-envelope resonance with short-envelope stricture, resulting in an extension of the release portion of the stricture envelope:

(6.36) Degree of obstruction: max.

\[
\begin{array}{c}
\text{stricture component} \\
\text{resonance component} \\
\text{extension of release}
\end{array}
\]

In this example, [continuant] is used to distinguish the stricture phase of moras such as /ṭi/ ([tʃi]) and /ṭe/ ([te]), which are [consonantal: [−continuant]], from /sil/ ([ʃi]) and /sel/ ([ʃe]), which are [consonantal: [+continuant]]. To generalise over fricative and vocalic nuclei, the set of resonance features is held to (predictably) include [+continuant]. This implication could be represented by a feature co-occurrence constraint:

(6.37) [vocalic: ANY] ⊇ [vocalic: [+continuant]]

In a phonetically interpreted theory of phonology, not only must interdependencies between different parameters be modelled, such as the
phasing of obstruence and resonance envelopes discussed above, but so
must the process by which units in the phonetic interpretation of unordered
representations are assigned temporal order. According to the parametric
phonetic viewpoint, the process of speech production involves two types of
synchronisation. The ordering of events within individual parameter
'streams', is termed sequencing by Kelly and Local (1989: 132). The second
aspect of temporal interpretation concerns the phasing relationships
between units in independent streams, which determine co-ordination of
the speech organs.

(6.38) a. In vocalic nuclei: [consonantal] $\prec$ [vocalic]
'Consonants costart with vowels.'

\[
\begin{array}{c}
\text{[vocalic] exponents} \\
\text{[consonantal] exps.}
\end{array}
\]

b. \[
\begin{array}{c}
\text{[vocalic] exponents} \\
[\text{consonantal}] \quad \text{exponents} \\
[(-\text{continuant})]
\end{array}
\]

c. In fricative nuclei: [consonantal: [+continuant]] $\approx$ [vocalic: [+diffuse]]
'Fricatives may completely overlap vowels.'

d. \[
\begin{array}{c}
\text{[vocalic] exponents} \\
\begin{array}{c}
\text{consonantal:} \\
(-\text{continuant})
\end{array}
\end{array}
\]

The relation 'A costarts with B' implies that 'A does not finish at the same
time as B' (in the absence of the stronger claim 'A is coterminous with B').
I shall assume the convention that 'A costarts with B' entails that 'A finishes
before B'.

Examples of ordering statements are (6.38a) and (6.38c). These specify
that in moras with both stricture and resonance phases, the stricture
maximum never follows the resonance maximum, although it may precede
and/or coincide with the resonance peak; and [-continuant, -grave] stric-
ture is affricated in [+diffuse] moras. The temporal interpretation of the
phonological representation in (6.39a) (e.g. [tʃi]) according to (6.38) is
impressionistically portrayed in the 'timing diagram' (6.39b).

(6.39) a. \[
\begin{array}{c}
\mu \\
\begin{array}{c}
\text{[+diffuse]} \\
\text{consonantal:} \\
\begin{array}{c}
\text{[−continuant]} \\
\text{[−grave]}
\end{array}
\end{array}
\begin{array}{c}
\text{vocalic:} \\
\begin{array}{c}
\text{[+continuant]} \\
\text{[−compact]}
\end{array}
\end{array}
\end{array}
\]
6.8 Consonant 'clusters'

b. 

<table>
<thead>
<tr>
<th>[±-continuant] exponents</th>
</tr>
</thead>
<tbody>
<tr>
<td>[±-continuant] exponents</td>
</tr>
<tr>
<td>[vocalic] exponents</td>
</tr>
<tr>
<td>[+diffuse] exponents</td>
</tr>
<tr>
<td>[consonantal] exps.</td>
</tr>
</tbody>
</table>

The nodes in (6.39a) are not sequenced: temporal organisation is only a property of its parametric phonetic interpretation (6.39b). Representing affrication in an unsequenced structure affords an explanation of the relationship between fricative-nucleus and vocalic-nucleus affricate-structure moras as a difference in phasing of the stricture and resonance envelopes. In moras with vocalic nuclei, the interpretation of the [consonantal] feature-structure costarts with the interpretation of [vocalic], but finishes before it (i.e. [consonantal] |< [vocalic]), whereas for fricative nuclei [consonantal: [±-continuant]] is coterminous with [vocalic] (cf. (6.36)).

6.8 Consonant 'clusters' in onset and coda

The foregoing analysis of moras with whispered or 'elided' nuclei enables the grammar of Japanese syllable-structure presented above to be completed. Although the grammar defines maximal syllables of three-moras, that is, CV.V.C(V) it is possible to concatenate voiceless CV moras to the onset C or coda C(V) to create complex consonant 'clusters'. The following comprehensive chart of examples is compiled from a variety of sources, including Hinds (1986: 401–3), Ueda (1976) and my own notes. The abundance of loan-words in this selection is deliberate, so that Japanese syllable-structure may be compared to English. It should be clear, however, from the substantial number of native Japanese words in this chart that these 'clusters' are not restricted to loan-words.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Examples</th>
<th>Glosses</th>
<th>Examples</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>sk</td>
<td>s'kebe</td>
<td>oversexed, wolfish</td>
<td>mas'k'</td>
<td>mask</td>
</tr>
<tr>
<td>s'ki</td>
<td></td>
<td>(1) to like; (2) plough (n.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) crack, opening, opportunity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sky</td>
<td>s'kyandaru</td>
<td>scandal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ss</td>
<td>s'sumeru</td>
<td>advance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sf</td>
<td>s'fi</td>
<td>sushi</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Onset 'clusters' Coda 'clusters'
Declarative analysis of Japanese words

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Examples</th>
<th>Gloses</th>
<th>Cluster</th>
<th>Examples</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>st</td>
<td>s’tando</td>
<td>stand</td>
<td>stf</td>
<td>s’tjuades’</td>
<td>stewardess</td>
</tr>
<tr>
<td></td>
<td>s’tekki</td>
<td>walking-stick</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s’toraiki</td>
<td>strike</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s’tobu</td>
<td>stove</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s’teki</td>
<td>wonderful</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s’t’eefon</td>
<td>station</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s’topp’</td>
<td>stop</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sp</td>
<td>s’pein</td>
<td>Spain</td>
<td>sp</td>
<td>s’poots’</td>
<td>sports</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fk</td>
<td>f’ka</td>
<td>deer</td>
<td>fk</td>
<td>f’kaku</td>
<td>square; qualification, competency</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fky</td>
<td>f’kyuu</td>
<td>womb</td>
<td>fky</td>
<td>f’kyuu</td>
<td>specutation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fs</td>
<td>f’saku</td>
<td>speculation</td>
<td>fs</td>
<td>f’saku</td>
<td>speculation</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fj</td>
<td>f’jamo</td>
<td>smelt</td>
<td>fj</td>
<td>inof’j’c</td>
<td>boar</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ft</td>
<td>f’ta</td>
<td>bottom, below, under; tongue; did</td>
<td>ft</td>
<td>f’taku</td>
<td>preparation, arrangement</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ftj</td>
<td>f’tyuu</td>
<td>stew</td>
<td>ftj</td>
<td>f’tyuu</td>
<td>stew</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fk</td>
<td>f’ka</td>
<td>shark</td>
<td>fk</td>
<td>waf’k’c</td>
<td>Japanese clothes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fk</td>
<td>f’kai</td>
<td>deep</td>
<td>fk</td>
<td>f’kai</td>
<td>deep</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fj</td>
<td>f’fin</td>
<td>depression</td>
<td>fj</td>
<td>f’fin</td>
<td>depression</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ft</td>
<td>f’ta</td>
<td>top, lid</td>
<td>ft</td>
<td>f’ta</td>
<td>top, lid</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ftj</td>
<td>f’tjash</td>
<td>rise and fall</td>
<td>ftj</td>
<td>f’tjash</td>
<td>rise and fall</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ftj</td>
<td>f’tsuu</td>
<td>ordinary</td>
<td>ftj</td>
<td>f’tsuu</td>
<td>ordinary</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ff</td>
<td>f’fuku,</td>
<td>dissatisfaction</td>
<td>ff</td>
<td>f’fuku,</td>
<td>dissatisfaction</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kk</td>
<td>k’kanai</td>
<td>do not listen</td>
<td>kk</td>
<td>k’kanai</td>
<td>do not listen</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ks</td>
<td>k’sa</td>
<td>grass, weed</td>
<td>ks</td>
<td>k’sa</td>
<td>grass, weed</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ksai</td>
<td>k’sai</td>
<td>smelly, stinking; fishy, questionable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k’sari</td>
<td>k’sari</td>
<td>chain</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k’sen</td>
<td>k’sen</td>
<td>a habit</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k’sou</td>
<td>k’sou</td>
<td>dung, excrement</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>k’suri</td>
<td>k’suri</td>
<td>medicine</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kf</td>
<td>k’fami</td>
<td>a sneeze</td>
<td>kf</td>
<td>k’fami</td>
<td>a sneeze</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kt</td>
<td>k’tunai</td>
<td>dirty</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ktf</td>
<td>k’tfjigai</td>
<td>strange</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
6.9 Concluding remark

The canonical maximal (CV.V.C) syllables were defined above by the phrase-structure rule (6.23), repeated here for convenience:

\[(6.40) \quad \sigma \rightarrow \mu[+\text{head}] \mu[+\text{head, –consonantal}] \mu[–\text{head}]\]

Recall that devoiced CV moras are represented as \(\mu[+\text{head, –voice}]\). If onset and coda consonant clusters containing devoiced CV moras are to be incorporated into the syllable template, (6.40) would have to be amended to:

\[(6.41) \quad \sigma \rightarrow \mu[+\text{head, –voice}] \mu[+\text{head}] \mu[+\text{head, –consonantal}] \mu[–\text{head}] \mu[+\text{head, –voice}]\]

To ensure that ‘devoiced’ CV moras are only permitted in voiceless onsets, it is necessary to ensure that if \(n > 0\), the mora before \(\mu[+\text{head, –consonantal}]\) in the above rule must also be [–voice].

6.9 Concluding remark

Few of the components of the above analysis are really new. For instance, in a short study of Japanese pronunciation for teachers, Daniels (1958: 58) made the following observation:
Declarative analysis of Japanese words

Consonants are said to be 'prefixed to' the vowels rather than to 'precede' them because... in Japanese it is necessary to put the speech organs into the position for the vowel... before producing the consonant — so far, that is to say, as it is possible to do this and still produce 'the consonant'. There is therefore more or less difference in almost all cases between the ways in which 'the same' consonant is produced when prefixed to different vowels.

Abstraction of features that pertain to complete moras, rather than their distribution into individual segments, is a common principle of Firthian prosodic analysis. The use of a classical generative notation for phonological features in prosodic analysis is novel, however, as is their new interpretation in certain cases.

In the next chapter, I shall present another Declarative Phonology fragment: an analysis of English monosyllabic words, with some more programmatic remarks on the phonology of polysyllabic words.
A declarative analysis of English words

7.1 A grammar of English monosyllables

7.1.1 Introduction

In chapters 3 and 5 it was shown that declarative grammars have the desirable property of being independent of any particular procedural implementation, so that a given grammar can be used with a variety of different parsing algorithms, or can just as easily be employed for generating sets of strings and structures. The grammar of English monosyllables which I will describe here has been implemented and thoroughly tested as part of a speech synthesis system (Coleman 1990b, 1992). The grammar described below was employed both for parsing input strings presented to the system, and for enumerating the set of input strings admitted by the grammar. In the normal operation of the speech synthesis program, the grammar is interpreted as a parsing program – a Prolog Definite Clause Grammar (DCG; Pereira and Warren 1980) – to determine the phonological structure of an input string. The structures determined in this way are then used to evaluate a compositional phonetic interpretation, without using rewrite rules. In order to ensure that the set of strings defined by the grammar was complete, the grammar was used generatively and its yield automatically tested against an on-line dictionary.

The phonotactic grammar of English will be discussed in two parts: syllable-level and word-level. The first part of this chapter (section 7.1) concentrates on the syllable-level subgrammar, and the second part, section 7.2, describes the word-level subgrammar.

The data for these subgrammars were in part drawn from numerous English grammars and descriptions of English phonology, although the collection, formalisation and integration of those observations into a complete grammar fragment is original. The syllable-level subgrammar makes formally explicit observations on English phonology which are also noted in Fudge (1969, 1987), Gimson (1962), Selkirk (1984b), Clements and Keyser (1983), Jones (1918) and Trnka (1935). Gimson (1962) and Fudge
Declarative analysis of English words (1987) contain particularly useful tabulations of the combinatorial possibilities in English, but their analyses do not make explicit the underlying principles which govern these combinatorics, and they do not discriminate sharply between productive and exceptional cases. Clements and Keyser (1983) discuss onsets alone, and Selkirk (1984b) concentrates on syllable-structure. Her use of a fixed integer-valued feature [sonority] is no more explanatory than a taxonomic partition of phonemes into classes labelled 'stop', 'fricative', 'sonorant', 'vowel' etc. The word-level subgrammar is based on the work of Selkirk (1984a), Church (1985) and Lass (1987).

7.7.2 A sketch of the grammar formalism

The monosyllable grammar is a Unification Phrase-Structure Grammar of phoneme strings. As discussed in chapter 5, this type of grammar differs from classical phrase-structure grammars in three main ways:

1. The symbols of the grammar are feature-structures (partial functions, directed acyclic graphs) rather than atomic symbols.

2. Each phrase-structure rule may be augmented by constraints: logical statements which restrict the applicability of a rule, and consequently allow exceptions and restrictions to be felicitously incorporated in the grammar, thus complementing the generalisations expressed in the rule. As in a DCG all constraints other than phrase-structure rules are linked to a particular phrase structure rule, expressing the domain in which the constraint holds. Constraints are not independent, as in the conception of Scobbie (1993) and Bird (1995), but are linked into groups, with the phrase-structure rules having a privileged status, as the example grammar below illustrates.

3. Matching a symbol in the left-hand side of one rule to the right-hand side of another rule requires only that the symbols can be unified, not exactly the same in all features.

For example, rule (7.1) states that a unit of category σ (syllable) consists of an O (onset) and an R (rime), and that the syllable and its rime must bear the same value for the feature [heavy]. In other words, heavy syllables are those with heavy rimes, and light syllables are those with light rimes.

(7.1) \[ \sigma \rightarrow O \ R ; \sigma: \text{heavy} = R: \text{heavy} \]

Equivalent formulations of this rule, which exploit the uniqueness of logical variables and the fact that the symbols are feature-structures, are the following:
7.1 Grammar of monosyllables

(7.2) \( \sigma[\alpha \text{heavy}] \rightarrow O \ R[\alpha \text{heavy}] \)

or more concretely:

(7.3) \([\sigma: \text{heavy: } \alpha] \rightarrow [O: []] [R: \text{heavy: } \alpha]\)

Since constraints are associated with particular rules, every constraint in the grammar is local to a specific domain, the local tree of the grammar rule. There are no global constraints, such as the Feature Co-occurrence Restrictions or Head Feature Convention of a GPSG, other than constancy of terms (that is, all tokens of a constant have the same value everywhere in the grammar, and all tokens of a variable have the same value everywhere within a rule). This is dictated largely by the practicality of computationally implementing the grammar in Prolog as a Definite Clause Grammar (Pereira and Warren 1980) or a Functional Logic Grammar (Boyer 1988), rather than being a theoretically principled standpoint.

Determining an appropriate combination of rules and constraints is guided by two complementary principles:

1. The maximal generalisation principle: there should be a minimal number of rules, without disjunctions where possible, for each generalisation concerning English phonotactics. For example, there should not be numerous slightly different expansions of a category if a more general formulation of the same information can be found. Each legitimate generalisation need not by itself be attested by all the cases over which it holds: a generalisation may be subject to certain exceptions, which are governed by:

2. The minimal exception principle: there should be a minimal number of negative constraints (e.g. filters) to encode legitimate exceptions to the rules. Legitimate exceptions can be motivated by exceptional phonology or phonotactics, such as spelling pronunciations (e.g. kiln), proper nouns, loan-words (e.g. vraic, voodoo) including forms borrowed from other dialects (e.g. vixen), archaisms (e.g. tierce), or phonotactic oddities (e.g. traipe). The filter construction *[ . . . ] may be used in combination with phrase-structure rules to indicate an exception to the generalisation which the rule expresses (e.g. examples (7.16) and (7.90) below). In this grammar, exceptions are preferably only filtered if the addition of a single filter accounts for more than one gap in the combinatory possibilities. Many historically anomalous developments are encoded in this way. In many cases where such anomalies inform our understanding of the synchronic
phonotactic patterns, I have noted some etymologies. This is not to imply that contemporary speakers are aware of the historical factors giving rise to the current patterns, of course. Such details are provided to demonstrate the appropriacy of the constraints I propose.

7.1.3 The phoneme alphabet
Writing a phonotactic grammar of a set of phonemic representations is somewhat different from writing a grammar of sentences, since for any language, there is no unique phonemic analysis (Chao 1934), and thus the set of terminal symbols is not a 'given'. For example, to represent the distinction between so-called tense and lax nuclei (as in *peat* vs. *pit*) it is reasonable either to use two different symbols, such as /i/ and /I/ (Jones 1918), or to employ a 'modifier' or 'length phoneme', /iː/ vs. /i/ (Gimson 1962); or a combination of the two techniques could be employed to represent both the length and quality difference, thus /iy/ vs. /i/ (Bloch and Trager 1942, Chomsky and Halle 1968).

The first of the above three notations fails to represent the fact that the English vowels do not form an unstructured set. There are short vowels and long vowels, for instance (Lass 1976: 21–5). Some of the long vowels (perhaps all, depending on the analysis) are phonologically diphthongs. Both the second and third solutions encode the long/short distinction, but they partition the long vowels in two different ways. For instance, in the third notation, /iy/ and /oy/ are related by having the same second element, whereas in Gimson's notation they are not related in this way (/iː/ vs. /oː/).

I have chosen a form of phonemic representation which is designed to make the grammar as simple as possible. For example, since it is possible to give a structural account of syllable weight, which is important in lexical stress placement (see the grammar of English words in section 7.2 below), I have not employed the first of the above methods for representing vowels. For a combination of reasons that are spelt out in the next subsection, I have chosen the third type of notation.

7.1.4 The vowel systems
The analysis of vowel systems in this grammar is a generative adaptation of the largely unpublished Prosodic Analysis of English developed and taught by Mrs E. Whitley to a generation of students at the School of Oriental and African Studies, and passed on in the present day by some of those students, through written notes and lectures (Whitley, mss.). Sprigg (1984,
1986) has suggested some alternatives to parts of Whitley's analysis in unpublished papers. A fragment of this analysis, pertaining largely to vowels, is given in Albrow (1975). In that analysis, vowels are divided into three systems: short vowels, rising diphthongs and centring diphthongs. Each of these systems generally has no more than five or six members, and can be arranged in a $2 \times 3$ pattern.

Short vowels are distributionally distinct from diphthongs in the following ways, amongst others:

1. Short vowels may not occur in stressed open monosyllables: the syllable must be closed, for instance /pey/ and /pet/ are possible stressed monosyllables, but /pe/ is not.
2. Long vowels and diphthongs are metrically 'heavy': the presence of a long vowel or diphthong in a syllable allows that syllable to bear final or penultimate lexical stress in Latinate words: for instance, *inert* and *reprisal* are well formed, whereas in *topic* and *typical*, the light vowels of *-ic* and *-al* may not bear main stress, which thus falls on the preceding syllable.

Each of the vowel systems in the Prosodic Analysis of English can be categorised in terms of a three-way vowel height distinction, and a two-way front/back distinction. These two dimensions are historically and morphologically relevant as the dimension of the synchronic reflexes of the Great Vowel Shift and I-mutation of short back vowels (see Campbell 1959: 71), respectively.

Vowel height and the Great Vowel Shift

The Great Vowel Shift is a historical phenomenon in which, stated oversimply, long vowels changed quality in the following way (ignoring predictable details of diphthong quality and final inflection):

1. Open vowels became mid, e.g. /naːm/ became /neːm/ (*name*) and /haːl/ became /hoːl/ (*whole*).
2. Mid vowels became close, e.g. /feːt/ became /fiːt/ (*feet*) and /toːθ/ became /tuːθ/ (*tooth*).
3. Close vowels (more specifically, the first element of the long vowel) became open, e.g. /biːt/ became /bait/ (*bite*) and /hauːs/ became /haus/ (*house*).

For further discussion of the diachronic developments involved in the Great Vowel Shift see Lass (1976), Chomsky and Halle (1968), Jones (1989:
Declarative analysis of English words

140–1, Kökeritz (1953) and Wolfe (1972). I shall ignore the shift of Middle English /e:/ ~ /e:/ distinction (Jespersen 1909/54: 242–4, Kökeritz 1953: 194–204, Wolfe 1972: 156–8) as not directly relevant to the synchronic grammar. Thus the *leap*/leapt* distinction is now isomorphic to the originally distinct *keep*/kept* distinction. (Scholars of historical English phonology will regard this discussion with some disapproval: I ask them to regard it as an account of synchronic phonology, informed by, though not slavishly encapsulating, historical developments.)

The result of this cyclic permutation would not be particularly relevant to the synchronic phonology were it not for the fact that in Old and Middle English, vowel length could have a morphological function, for instance /bi:t/ (*bite*) is the present-tense stem form, whereas /bit/ (*bit*) is the stem of the past tense (preterite) form. In permuting only the height of long vowels, the Great Vowel Shift replaced what was formerly a length opposition by a composite length-and-vowel-quality opposition, such as /bait/ (present tense) vs. /bit/ (past). In order to understand the vowel patterns of contemporary English, then, it is necessary to take account of the synchronic consequences of the Great Vowel Shift. Chomsky and Halle (1968) achieved this using abstract 'underlying' vowel qualities and a synchronic transformational description of the historical changes involved in the Great Vowel Shift. The present grammar employs a three-valued vowel height feature [height], but does not express phonological generalisations about morphological relations derivationally. However, employing a three-valued vowel height feature permits phonological generalisations about these morphological relations to be expressed. Suppose we define a function, shift(…), which pairs each value of the vowel height feature with its successor, modulo 3: that is,

\[
\text{shift}([\text{height}: n]) = [\text{height}: (n+1)_3] \\
\text{shift}([\text{height}: 0]) = [\text{height}: 1] \\
\text{shift}([\text{height}: 1]) = [\text{height}: 2] \\
\text{shift}([\text{height}: 2]) = [\text{height}: 0]
\]

The conventional names of numerals or phonological feature-values are, from the point of view of the human language faculty, completely arbitrary. The names that phonologists use to refer to the values of features are purely conventional. In a modulo 3 arithmetic system, we could equally well employ the names 'open', 'mid' and 'close' instead of 'zero', 'one' and 'two' respectively. I shall take the liberty of doing so where it helps to make my account less baroque in appearance.
7.1 Grammar of monosyllables

Morphosyntax and semantics:

Verb
[- tense: past]
[- semantics: past (bite)]

Verb
[- tense: present]
[- semantics: present (bite)]

Phonological representations:

σ

b V t

long: -

height: 2

grave: -

σ

b V t

long: +

shift ([height: 2])

grave: -

Figure 7.1 Relationships between morphosyntactic, semantic and morphophonological information

In the lexicon, the phonological representation of *bit* will include the feature description [height: 2]. The phonological representation of *bite* will include the feature description shift([height: 2]), which is extensionally equivalent to [height: 0] (ignoring length and the predictable diphthongisation which also results from the Great Vowel Shift). Thus 'shifted' close vowels will receive the same phonetic interpretation as unshifted open vowels (as in e.g. *aisle*). (The ‘shift’ relation should not be confused with the older ablaut relations of Old English and other Germanic languages, from which the *sit* ~ *sat* opposition derives. I shall not attempt an account of ablaut relations here, though clearly the three-valued [height] feature which I have proposed will provide a sound basis for a paradigmatic, relational account of ablaut.)

My analysis of Great Vowel Shift relations, then, is not derivational, but paradigmatic (Calder and te Lindert 1987). The relationship between shifted and unshifted vowels in allomorphs of a single morpheme, such as /kiyp/ ~ /kept/, /gəw/ ~ /gon/, /bayt/ ~ /bit/, /indiws/ ~ /indəktə/, /luwz/ ~ /lost/, is a paradigmatic, lexical relation (sometimes called ‘via rules’; see Lass 1984: 224). The relationship between some of the features of the vowels in /bit/ vs. /bayt/ is illustrated in figure 7.1. The distribution of those features in the syllable structure will be described after a fuller consideration of the analysis of vowels and syllable nuclei.

As noted in section 6.4.3 above, the representation of vowel height in my analyses of English and Japanese are fundamentally different. The English analysis has a three-valued [height] feature, which is motivated in large part
by the cyclic nature of the Great Vowel Shift, whereas the Japanese analy-
sis has two binary height features, motivated by consonant–vowel cross-
classification à la Jakobson, Fant and Halle (1952), and the fusion of close
and open vowels to produce mid vowels. This difference between the two
analyses is a specific counter-example to universalist theories of phonolog-
ic categories. This implies that no language has both cyclic shift of vowel
height and height fusion.

Diphthongs, length and the Great Vowel Shift
The foregoing account of vowel height and vowel shift is incomplete, since
it did not address the issue of the representation of vowel length, which is a
formal prerequisite for the Great Vowel Shift, or the representation of diph-
thongs, which is a consequence of the Great Vowel Shift. The simplified
description given above represents the shifted open and shifted mid vowels
as monophthongal long vowels, that is /neim/ name, /hoːl/ whole, /fiʃt/ feet,
/tuːt/ tooth, and the shifted close vowels as diphthongal, /bayt/ bite and
/haws/ house. While some dialects do have fairly monophthongal shifted
open and shifted mid vowels, in most dialects of English truly monoph-
thongal long-vowel qualities are rare. I shall now argue that the proper
representation of the 'long' vowels is as structurally bipartite nuclei (i.e.
diphthongs) and that the dialects with monophthongs can be analysed as a
monophthongal phonetic interpretation of phonological diphthongs, for
instance /iːl/ = [iː], /eːl/ = [eː], /oːl/ = [oː] etc. I shall thus revise my pre-
liminary description of the changes induced by the Great Vowel Shift as
follows:

1. Open to mid: /aːl/ > diphthongal /eːl/ (e.g. name); /aːl/ > diphthongal
   /oːl/ (e.g. whole).
2. Mid to close: /eːl/ > diphthongal /iːl/ (e.g. feet); /oːl/ > diphthongal
   /uːl/ (e.g. tooth).
3. Close to open: /iːl/ > diphthongal /aːl/ (e.g. bite); /uːl/ > diphthongal
   /aːl/ (e.g. house).

If we represent these nuclei as binary-branching structures containing
two V elements, which in the contemporary phonological representations
are the peak (V₁) and offglide (V₂) of the diphthong (except in /iw/ = [ju],
which is discussed below), it is possible to state the fact that the shift opera-
tion defined above applies not to the nucleus as a whole, but only to the
peak (V₁) of the diphthong. (The quality of V₂ is defined by constraints dis-
cussed below.) Thus:
Open V₁ to mid V₁: /aːl > /eɪl (e.g. name); /aːl > /awl (e.g. whole).

Mid V₁ to close V₁: /iːl > /iy/ (e.g. feet); /oʊl > /uw/ (e.g. tooth).

Close V₁ to open V₁: /iːl > /ay/ (e.g. bite); /uu/ > /aw/ (e.g. house).

The binary-branching analysis of the supposed monophthongal long vowels in the 'input' to the Great Vowel Shift is, in the case of /ee/ and /oo/, consistent with the spellings meet, feet, foot, book etc., which reflect the older phonology despite the historically changed vowel qualities. The quality of V₁ in the 'shifted' nuclei is predictably the front offglide /y/ when V₁ is front and non-open; and the back offglide /w/ when V₁ is back and non-open. This generalisation can be expressed by feature agreement as follows:

\[
(7.5) \quad \text{Nucleus} \\
\quad \text{V₁} \quad \text{V₂} \\
\quad \text{[grave: } α \quad \text{[height: } \neg \text{open]} \quad \text{[grave: } α \quad \text{[height: close]}\]
\]

In the open diphthongs /ay/ and /aw/, the backness of V₂ is the same as the backness of the V₁ of its historical ancestor, but this backness is not preserved in the contemporary V₁, which in many dialects has the opposite backness from its historical ancestor. Thus, the contemporary reflex of /huus/, with back V₁ and V₂, is /haws/, with back V₂ still but front V₁, whereas the contemporary reflex of /biit/, with front V₁ and V₂, is /bayt/, with front V₂ still, but back V₁. This is a relatively recent development. The /a/ part of these diphthongs is lexically neither back nor front: the front/back distinction is encoded in the glide. We could describe these generalisations with an alpha-switching (dissimilation) constraint, such as:

\[
(7.6) \quad \text{Open Diphthong Dissimilation} \\
\quad \text{Nucleus} \\
\quad \text{V₁} \quad \text{V₂} \\
\quad \text{[grave: } -α \quad \text{[height: open]} \quad \text{[grave: } α \quad \text{[height: close]}\]
\]

But this representation does not account for the fact that it is the backness of V₂ that is historically preserved, not the backness of V₁. In a non-deriva- tional, monostratal analysis, it is not possible to appeal to the 'underlying backness' of open vowels, since there are no 'underlying' representations. In order to make the generalisation that the front quality of the offglide in
Declarative analysis of English words

*bite* is the same as the front quality of the vowel in *bit*, it is necessary to treat the front/back distinction as property of the nucleus as a whole. It is then possible to account for the quality of the offglide of these diphthongs independently of the height of $V_1$:

\[
\begin{array}{c}
\text{Nucleus} \\
\left[\text{grave: } \alpha \right] \\
V_1 \\
\left[\text{grave: } \alpha \right. \\
\left. \text{height: close} \right]
\end{array}
\]

The idiosyncratic, dissimilatory quality of $V_1$ in open diphthongs is now stated by constraint (7.6).

I-mutation and the front/back vowel opposition

I-mutation is a characteristic of the phonology of Germanic languages which, like the vowel-shift morphophonology of the above examples, is still preserved in some corners of the English lexicon. For instance, the singular–plural opposition in a number of common Germanic words is expressed as a phonological front/back opposition in vowels, for instance /guws/ *(goose)* vs. /giys/ *(geese)*, /maws/ *(mouse)* vs. /mays/ *(mice)*. In mid vowels this dimension is also evidenced, albeit employed with different functions, such as ablaut and borrowings from other dialects for instance,

1. morphological ablaut: /get/ *(present tense)* vs. /got/ *(past tense)*;
2. lexical doublets: /nəwzl/ *(nose)* vs. /neyz/ *(naze, nas(al))*; /hwəwl/ *(whole)* vs. /heyl/ *(hale)*;
3. register/stylistic variants: /nəwl/ *(no)* vs. /neyl/ *(nay)*;
4. dialect variation: /həwml/ *(home)* vs. Northern and Scottish /heyl/ *(hame)*.

Despite the use of the traditional term ‘vowel’ in this section, the grammar does not actually employ the category ‘vowel’. Vowel letters are phonologically analysed as three different kinds of object in this grammar: short vowels, unchecked nuclei and vocalic rimes.

A short vowel bears the [vocalic] place-of-articulation feature, whose value is a category representing both the height of the vowel on the three-point scale (0 or open, 1 or mid, 2 or close), and its gravity (back-roundedness or labiovelarity). For example, the place of articulation of the short vowel /i/ is represented:
The place-of-articulation features of the short-vowel system are given in the following chart:

<table>
<thead>
<tr>
<th>[height: close]</th>
<th>[-grave]</th>
<th>[+grave]</th>
</tr>
</thead>
<tbody>
<tr>
<td>/i/</td>
<td></td>
<td>/a/</td>
</tr>
<tr>
<td>/e/</td>
<td></td>
<td>/o/</td>
</tr>
<tr>
<td>/a/</td>
<td></td>
<td>/A/</td>
</tr>
</tbody>
</table>

This chart oversimplifies the phonology of the short-vowel systems in three ways. First, the close grave short-vowel category is a composite of two historically distinct categories: historically unaltered close grave short vowels, as in *push* and *pull*, and historically shortened close grave long vowels, as in *wool* and *book*, which, judging by the spelling, apparently derive from mid, grave long vowels by the Great Vowel Shift. Second, the contemporary distinction between close and open short grave vowels (e.g. *look* vs. *luck*) is a historical residue: the open category is the regular reflex in Received Pronunciation (RP) of former close short grave vowels (Jespersen 1909/54: 332), but some close short grave vowels, especially in labial environments, escaped the march of time (Jespersen 1909/54: 333–5, Lass 1987). It might be conjectured that the shift from /u/ to /A/, the only case of Great Vowel Shift within the short vowels, was caused by the shift of /o:/ to /u:/ and its subsequent shortening to /u/ in words such as *look*. However that may be, the resulting situation does not seem regular enough to be described by a simple rule. Third, the distinction between grave and non-grave open vowels implied in the chart above is lexically but not morphophonologically distinctive. A better case could perhaps be made for analysing some instances of [aː] in the grave, open short vowel ‘slot’, since it distributes like other short vowels. For example, unlike ‘true’ long vowels it can occur before coda consonant clusters, such as [maɪsk]. In many other dialects, such words indeed have a phonetically short vowel [mask]. The one phonetic category [aː] has two different phonological representations in different words, /ar/ in open syllables such as /bar/, but /a/ in heavy rimes such as /mask/. (Note that RP English is not rhotic.) The problems presented by historical and dialectal peculiarities such as these can be overcome by maintaining a careful separation between phonological representation and phonetic interpretation, between structures and systems, and between categories and their distribution.

My categorisation of English vowels is summarised in table 7.1. Further
Table 7.1 Summary of the vowel categories

<table>
<thead>
<tr>
<th></th>
<th>Short Front</th>
<th>Back</th>
<th>Rising to front</th>
<th>to back</th>
<th>Centring Front</th>
<th>Back</th>
<th>Central</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>/i/</td>
<td>/u/</td>
<td>/iy/</td>
<td>/uw/</td>
<td>/iə/</td>
<td>/ur/</td>
<td></td>
</tr>
<tr>
<td>Mid</td>
<td>/e/</td>
<td>/o/</td>
<td>/ey/ /oy/</td>
<td>/aw/</td>
<td>/εr/</td>
<td>/ɔr/</td>
<td></td>
</tr>
<tr>
<td>Open</td>
<td>/a/</td>
<td>/ʌ/</td>
<td>/aɪ/</td>
<td>/aw/</td>
<td>/aə/</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Details concerning the representation of each vowel category are given in the next subsection.

7.1.5 Vowel categories

The short-vowel symbols of a phonemic representation correspond to the categories of the present analysis as follows:

(7.9)  \[
\text{\textit{\textit{i}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]

(7.10) \[
\text{\textit{\textit{e}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]

(7.11) \[
\text{\textit{\textit{a}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]

/\textit{a}/ is not specified for gravity, in order to generalise over the front [\textit{a}] and long, back [\textit{u}] variants of /\textit{ar}/, as well as give the same head vowel to /\textit{ay}/ and /\textit{aw}/, differing predictably according to the dissimilation constraint (7.6).

(7.12) \[
\text{\textit{\textit{u}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]

(7.13) \[
\text{\textit{\textit{u}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]

For instance in \textit{shook, look}, etc.

(7.14) \[
\text{\textit{\textit{u}}} = \left[ \begin{array}{c} \text{vocalic:} \\ \text{\textit{\textit{v}}} \end{array} \right]
\]
7.1 Grammar of monosyllables

(7.15) \[ /\alpha / = [\text{vocalic: } \checkmark \text{ grave: } + \text{ shift([height: close])}] \]

\(/\alpha /\) is categorised as shift([height: close]), and not as the extensionally equivalent [height: open], since all instances of \(/\alpha /\) correspond by vowel shift to \(/u/\) in more conservative dialects, as most speakers are aware.

There is no short schwa nucleus (though \(/\alpha /\) does occur in \(/\alpha r / = [\alpha i ]\)), because it does not occur distinctively in English monosyllables, being dependent on stress placement in polysyllabic words. Short vowels cannot occur in stressed open monosyllabic words and, unlike long vowels, they can occur before certain clusters, such as \(/lt/\). These two facts can be collapsed into one constraint by defining the notion of checked and unchecked syllables: open syllables are all unchecked, as are some closed syllables, such as those ending in single consonants, such as \(/l/\), \(/l/\) or \(/\beta /\); see (7.113) below. We shall require that \(N:\text{checked} = C:\text{checked}\). Closed syllables ending in single consonants which may not be combined with long vowels, such as \(/g/\) (see (7.112) below) and clusters such as \(/lt/\), however, are checked. Short vowels can only occur in checked syllables.

All the 'vowels' defined above may occur in checked syllable nuclei, with the exception of \(/u/\):

(7.16) \[ N \rightarrow \begin{cases} \text{long} \\ +\text{checked} \end{cases} * [\text{vocalic: } \checkmark \text{ grave: } + \text{ height: close}] \]

Because of its special history, shifted \(/u/\) (orthographic oo) occurs with a similar distribution to long vowels. Like long vowels, it cannot combine with [+checked] codas such as \(/lt/\). \(/u/\) thus has a special nucleus rule to itself (7.17), to ensure that a nucleus which contains only:

* [\text{vocalic: } \checkmark \text{ grave: } + \text{ height: close}] is [-checked], not [+checked] as in (7.16). Otherwise \(/u/\) could be generated before coda clusters such as \(/ld/\) which are generally prohibited from combining with long vowels.

(7.17) \[ N \rightarrow \begin{cases} \text{long} \\ -\text{checked} \end{cases} \text{vocalic: } [\checkmark \text{ grave: } + \text{ height: close}] \]

Rules (7.16) and (7.17) are complementary, and if desired could be collapsed into a single rule, with the absence or presence of [grave: +, height: close] related to the value of [checked].
Diphthongs are represented phonemically as a sequence of a vowel symbol followed by one of the symbols /y/, /w/ or /r/. These 'offglide' symbols correspond to three phonological categories, which again can be expressed using just the features [grave] and [height], although there is no open semi-vowel category, nor is there a distinction between front and back mid-height glides:

Not every sequence of vowel and offglide that agrees with the above rules and constraints is a well-formed phonological category. Certain deviant sequences (e.g. /uy/, /Ay/, /ay/, /ew/, /AW/, /AT/) escape the strictures of the constraints I have proposed. (Perhaps /Ay/ should be non-distinct from /ay/, if the gravity opposition is neutralised in open vowels. Likewise, perhaps /ew/ should be non-distinct from /ew/, if [grave] is a property of the diphthong as a whole, so that the single representation:

parses both /ew/ and /aw/.) It is not clear what generalisations over these exceptional forms could be made, or how the rules and constraints proposed above should be modified to prevent them.

In this grammar, each of the three types of diphthong (rising to front, rising to back, centring) is described by a separate phrase-structure rule. There are three rules, rather than a single, more general rule, because the exceptions to each diphthong type are different in each case. Since exceptions are dealt with by putting constraints on rules, there must be three separate rules for diphthongs. The first two rules define /Vy/ and /Vw/ diphthongs respectively:

(7.18) \[ N \quad [+ grave] \rightarrow V \quad [vocalic: [grave: - [height: close]]] \]

(7.19) \[ N \quad [+ grave] \rightarrow V \quad [vocalic: [grave: + [height: close]]] \]
I have included /iw/ among the symbols for diphthongs in view of its historical status as the modern reflex of a former front, close rounded vowel category (Old English and Norman French /ü/), as well as Middle English /əw/, thus accounting for the apparently defective distribution of vowel categories after syllable-initial [Cy] (see 7.1.9 below).

Irrespective of the treatment of inadmissible vowel-offglide sequences, the first V of the centring diphthong rule which follows must be constrained to exclude /ir/, /ur/ and /er/:

\[(7.20)\] 
\[
\begin{array}{c}
\text{N} \\
\text{long: +} \\
\text{checked: -} \\
\text{weight: heavy}
\end{array} \rightarrow 
\begin{array}{c}
\text{V} \\
\text{vocalic: } \{\text{height: close}, \text{height: mid}\} \\
\text{grave: -}
\end{array} \rightarrow 
\begin{array}{c}
\text{V} \\
\text{vocalic: } \{\text{height: mid}\}
\end{array}
\]

It is not that /ir/, /ur/ and /er/ do not occur at all in English; rather, they are not nuclei – they occur only in open syllables, and thus permute with whole rimes. They are thus assigned directly to the category R, not N:

\[(7.21)\] 
\[
\begin{array}{c}
\text{N} \\
\text{long: +} \\
\text{checked: -} \\
\text{weight: heavy}
\end{array} \rightarrow 
\begin{array}{c}
\text{V} \\
\text{vocalic: } \{\text{height: close}, \text{height: mid}\} \\
\text{grave: -}
\end{array} \rightarrow 
\begin{array}{c}
\text{V} \\
\text{vocalic: } \{\text{height: mid}\}
\end{array}
\]

The exceptions to this treatment of /ir/ and /er/ are beard, weird, bierce, pierce, fierce, tierce, kirsch (German), cairn (Gaelic), laird (Scots), bairn, and scarce – quite a few, but hardly enough to confound the analysis, especially considering their origins. The rimes /ird/, /irs/ and /ern/, and the words /kirʃ/, /lerd/ and /skers/ may be listed separately from the core syllable grammar, reflecting their anomalous status.

I shall conclude this discussion of vowel categories by considering the way in which the relationships between the nuclei in cases like /bit/ ~ /bayt/ are to be encoded, given the categories, structures and constraints presented above. The lexical representations of the nuclei of these words would be:

\[(7.22)\] 
\[
\begin{array}{c}
a. /i/ \\
\text{N} \\
\text{[long: -]} \\
\text{[grave: -]} \\
\text{V} \\
\text{[vocalic: [height: close]]}
\end{array} \\
\begin{array}{c}
b. /ay/ \\
\text{N} \\
\text{[long: +]} \\
\text{[grave: -]} \\
\text{V} \\
\text{[vocalic: [height: close], [shift(height: close)]]}
\end{array}
\]

Application of constraint (7.18) – it must be (7.18) and not (7.19) to agree with constraint (7.7), by which nucleus and offglide nodes must agree in the
value of [grave] – to the lexical representation of /ay/ (7.22b), and evaluation of the shift function, results in the addition of a /y/ offglide:

\[(7.23)\]
\[
\begin{array}{c}
\text{N} \\
\text{long: +}
\end{array}
\longrightarrow
\begin{array}{c}
\text{V} \\
\text{vocalic: [height: open]} \\
\text{grave: -} \\
\text{weight: heavy}
\end{array}
\]

The dissimilation constraint (7.6) adds the information that the first V has the opposite value of the feature [grave] to the second V. The result of applying (7.6) to (7.24) is:

\[(7.24)\]
\[
\begin{array}{c}
\text{N} \\
\text{long: +}
\end{array}
\longrightarrow
\begin{array}{c}
\text{V} \\
\text{vocalic: [height: open]} \\
\text{grave: -} \\
\text{weight: heavy}
\end{array}
\]

7.1.6 Consonant categories

Consonant categorisation presents fewer intricacies than vowel categorisation, and the following features are adequate. [consonantal] denotes consonantal place of articulation, or location of consonantal stricture. Its values are feature-structures made up of the two features [grave], which distinguishes labial and velar place of articulation ([+grave]) from (alveo-)palatal and alveolar ([−grave]), and [compact], which distinguishes velar and palatal place of articulation ([+compact]) from labial and alveolar ([−compact]). In consonants, [compact] is equivalent to Chomsky and Halle’s (1968) [anterior]: [+anterior] corresponds to [−compact] and [−anterior] corresponds to [+compact]. In terms of contemporary theories of feature-geometry (Clements 1985, Broe 1992), [+compact] denotes dorsal articulations. Together, [grave] and [compact] define four place-of-articulation categories:

\[
\begin{array}{c}
[+grave] \\
[−grave] \\
[+compact] \\
[−compact]
\end{array}
\]

\[
\begin{array}{c}
\text{Velar} \\
\text{(Alveo-)Palatal} \\
\text{Labial} \\
\text{Alveolar}
\end{array}
\]

There are also four manner-of-articulation features: [voice], [nasal], [continuant], which distinguishes fricatives and approximants from stops and nasals, and [strident], which distinguishes strident fricatives such as /s/ from mellow fricatives such as /θ/, and strident stops (i.e. affricates) such as /č/ from mellow stops such as /t/ (Jakobson, Fant and Halle 1952: 24). This analysis of affricates thus differs from the Japanese analysis presented in
chapter 6, in which I analysed [tʃ] as the delayed release exponent of an alveolar stop in palatal contexts.

These features are primarily (morpho)phonological and relational. They have no intrinsic phonetic interpretation. For example, not all articulations made with vocal-cord vibrations are classified as [+voice], and not all exponents of [+voice] constituents are phonetically voiced. Voiced nuclei are an example of the former, since voicing is not distinctive in vowels, liquids and nasals. Utterance-initial [+voice] stops (which are often voiceless or late-voiced; Ladefoged 1971: 9, Kelly and Local 1989: 87, Lass 1990: 1063) exemplify the latter. The phonological features are based primarily upon distribution and functional oppositions, and only secondarily on considerations regarding phonetic exponency. Bearing these comments in mind, the following definitions should be transparent.

(7.25) \[ \begin{array}{c}
\text{C} \\
\begin{array}{c}
\text{voice: } - \\
\text{consonantal: } \begin{array}{c}
\text{grave: } + \\
\text{compact: } -
\end{array}
\end{array} \\
\text{nasal: } - \\
\text{strident: } - \\
\text{continuant: } -
\end{array} \] \rightarrow /p/

(7.26) \[ \begin{array}{c}
\text{C} \\
\begin{array}{c}
\text{voice: } - \\
\text{consonantal: } \begin{array}{c}
\text{grave: } - \\
\text{compact: } -
\end{array}
\end{array} \\
\text{nasal: } - \\
\text{strident: } - \\
\text{continuant: } -
\end{array} \] \rightarrow /l/

(7.27) \[ \begin{array}{c}
\text{C} \\
\begin{array}{c}
\text{voice: } - \\
\text{consonantal: } \begin{array}{c}
\text{grave: } + \\
\text{compact: } +
\end{array}
\end{array} \\
\text{nasal: } - \\
\text{strident: } - \\
\text{continuant: } -
\end{array} \] \rightarrow /k/

(7.28) \[ \begin{array}{c}
\text{C} \\
\begin{array}{c}
\text{voice: } + \\
\text{consonantal: } \begin{array}{c}
\text{grave: } + \\
\text{compact: } -
\end{array}
\end{array} \\
\text{nasal: } - \\
\text{strident: } - \\
\text{continuant: } -
\end{array} \] \rightarrow /b/
Declarative analysis of English words

(7.29) \[
\begin{array}{c}
\text{C} \\
\text{voice: +} \\
\text{consonantal: grave: -} \\
\text{nasal: -} \\
\text{strident: -} \\
\text{continuant: -}
\end{array}
\Rightarrow /d/
\]

(7.30) \[
\begin{array}{c}
\text{C} \\
\text{voice: +} \\
\text{consonantal: grave: +} \\
\text{nasal: -} \\
\text{strident: -} \\
\text{continuant: -}
\end{array}
\Rightarrow /g/
\]

(7.31) \[
\begin{array}{c}
\text{C} \\
\text{voice: -} \\
\text{consonantal: grave: -} \\
\text{nasal: -} \\
\text{strident: +} \\
\text{continuant: -}
\end{array}
\Rightarrow /\&/
\]

(7.32) \[
\begin{array}{c}
\text{C} \\
\text{voice: +} \\
\text{consonantal: grave: -} \\
\text{nasal: -} \\
\text{strident: +} \\
\text{continuant: -}
\end{array}
\Rightarrow /\&/
\]

(7.33) \[
\begin{array}{c}
\text{C} \\
\text{voice: -} \\
\text{consonantal: grave: +} \\
\text{nasal: -} \\
\text{strident: +} \\
\text{continuant: +}
\end{array}
\Rightarrow /l/
\]

(7.34) \[
\begin{array}{c}
\text{C} \\
\text{voice: -} \\
\text{consonantal: grave: -} \\
\text{nasal: -} \\
\text{strident: -} \\
\text{continuant: +}
\end{array}
\Rightarrow /0/
\]
(7.35) \[ \begin{align*} 
\text{voice: } &- \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &- \\
\text{compact: } &- 
\end{cases} \\
\text{nasal: } &- \\
\text{strident: } &+ \\
\text{continuant: } &+ 
\end{align*} \rightarrow \text{/sl/} \]

(7.36) \[ \begin{align*} 
\text{voice: } &- \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &- \\
\text{compact: } &+ 
\end{cases} \\
\text{nasal: } &- \\
\text{strident: } &+ \\
\text{continuant: } &+ 
\end{align*} \rightarrow \text{/šl/} \]

(7.37) \[ \begin{align*} 
\text{voice: } &+ \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &+ \\
\text{compact: } &- 
\end{cases} \\
\text{nasal: } &- \\
\text{strident: } &+ \\
\text{continuant: } &+ 
\end{align*} \rightarrow \text{/vl/} \]

(7.38) \[ \begin{align*} 
\text{voice: } &+ \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &- \\
\text{compact: } &- 
\end{cases} \\
\text{nasal: } &- \\
\text{strident: } &- \\
\text{continuant: } &+ 
\end{align*} \rightarrow \text{/ðl/} \]

(7.39) \[ \begin{align*} 
\text{voice: } &+ \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &- \\
\text{compact: } &- 
\end{cases} \\
\text{nasal: } &- \\
\text{strident: } &+ \\
\text{continuant: } &+ 
\end{align*} \rightarrow \text{/zl/} \]

(7.40) \[ \begin{align*} 
\text{voice: } &+ \\
\text{consonantal: } &\begin{cases} 
\text{grave: } &+ \\
\text{compact: } &- 
\end{cases} \\
\text{nasal: } &+ \\
\text{strident: } &- \\
\text{continuant: } &- 
\end{align*} \rightarrow \text{/ml/} \]
Declarative analysis of English words

The symbol /n/ represents all non-labial nasals, since \(-\) [grave: +, compact: -] unifies with [grave: -, compact: -] (alveolar), [grave: -, compact: +] (palatal) and [grave: +, compact: +] (velar). The range of [consonantal] values for nasals is slightly different in onsets and codas, as too is the value of [voice] in the two positions. In onsets, nasals may have labial or alveolar place of articulation ([consonantal: [compact: -]]), whereas in codas, nasals may have labial /m/, alveolar /n/, alveo-palatal (only in clusters such as /pɛ/ or velar /ŋ/ place of articulation. In onset nasals, the value of [voice] is either \(-\) or unspecified (section 7.1.9 below). In coda nasals, the value of [voice] must be + in order for syllable termination voicing agreement to work, by which the alternation cat[s] vs. dog[z] is determined (section 7.1.8). The phonemic categories /h/, /ŋ/ and /ʒ/ do not appear in the above list, because there are no apparent generalisations regarding the features of /h/ and the few consonants with which it permutes (/c, j, y/), so the restriction of /h/ to non-branching onsets is determined by a separate 'h/rule' (7.87), in which /h/ has no features. /ŋ/ is analysed as a cluster, and /ʒ/ is a non-native category which does not generally occur in monosyllables. The omission of these three categories is deliberate and motivated.

The definitions of each consonant symbol given above can be made less verbose by using category 'macros' or 'templates' (see Shieber 1986: 55). Suppose the following types are defined:

\[
\text{(7.42) } \text{Stop} = C \begin{array}{c}
-\text{nasal} \\
-\text{strident} \\
-\text{continuant}
\end{array}
\]

\[
\text{(7.43) } \text{Fricative} = C \begin{array}{c}
-\text{nasal} \\
+\text{continuant}
\end{array}
\]

\[
\text{(7.44) } \text{Nasal} = C \begin{array}{c}
+\text{nasal} \\
-\text{strident} \\
-\text{continuant} \\
+\text{voice}
\end{array}
\]
7.1 Grammar of monosyllables

(7.45) Affricate =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
-\text{grave} \\
+\text{compact}
\end{array} \right]
\]
-\text{nasal} \\
+\text{strident} \\
-\text{continuant}
\]

(7.46) Voiced =
\[
C \left[ \begin{array}{c}
-\text{nasal} \\
+\text{voice}
\end{array} \right]
\]

(7.47) Voiceless =
\[
C \left[ \begin{array}{c}
-\text{nasal} \\
-\text{voice}
\end{array} \right]
\]

(7.48) Strident =
\[
C \left[ +\text{strident} \right]
\]

(7.49) Mellow =
\[
C \left[ -\text{strident} \right]
\]

(7.50) Labial =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
-\text{grave} \\
+\text{compact}
\end{array} \right]
\]

(7.51) Non-labial =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
-\text{grave} \\
-\text{compact}
\end{array} \right]
\]

(7.52) Labiovelar =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
\text{grave: +} \\
-\text{compact: -}
\end{array} \right]
\]

(7.53) Alveolar =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
\text{grave: -} \\
-\text{compact: -}
\end{array} \right]
\]

(7.54) Palatal =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
\text{grave: -} \\
+\text{strident} \\
-\text{compact: +}
\end{array} \right]
\]

(7.55) Velar =
\[
C \left[ \begin{array}{c}
\text{consonantal:} \\
\text{grave: -} \\
-\text{compact: +}
\end{array} \right]
\]

By unifying such representations with each other, rules (7.25)–(7.41) can be replaced by the following:

(7.56) Voiceless U Labial U Stop → /p/

(7.57) Voiceless U Alveolar U Stop → /t/

(7.58) Voiceless U Velar U Stop → /k/
Declarative analysis of English words

(7.59) Voiced $\cup$ Labial $\cup$ Stop $\rightarrow /b/$
(7.60) Voiced $\cup$ Alveolar $\cup$ Stop $\rightarrow /d/$
(7.61) Voiced $\cup$ Velar $\cup$ Stop $\rightarrow /g/$
(7.62) Voiceless $\cup$ Affricate $\rightarrow /t/$
(7.63) Voiceless $\cup$ Affricate $\rightarrow /d/$
(7.64) Voiceless $\cup$ Labial $\cup$ Fricative $\rightarrow /f/$
(7.65) Voiceless $\cup$ Mellow $\cup$ Alveolar $\cup$ Fricative $\rightarrow /\theta/$
(7.66) Voiceless $\cup$ Strident $\cup$ Alveolar $\cup$ Fricative $\rightarrow /s/$
(7.67) Voiceless $\cup$ Palatal $\cup$ Fricative $\rightarrow /s/$
(7.68) Voiceless $\cup$ Labial $\cup$ Fricative $\rightarrow /l/$
(7.69) Voiceless $\cup$ Mellow $\cup$ Alveolar $\cup$ Fricative $\rightarrow /\theta/$
(7.70) Voiceless $\cup$ Strident $\cup$ Alveolar $\cup$ Fricative $\rightarrow /\theta/$
(7.71) Labial $\cup$ Nasal $\rightarrow /m/$
(7.72) Nonlabial $\cup$ Nasal $\rightarrow /n/$

7.1.7 Frictionless continuant (‘glide’) categories
There are four glide categories in this analysis of English, /l/, /r/, /w/ and /y/. They are distinguished from vowels and obstruents by bearing both non-empty [consonantal] and [vocalic] specifications. Both [consonantal] and [vocalic] properties are needed because glides are phonologically similar to consonants (the alternations ward/guard and garden/yard, for instance), but phonetically similar to vowels. The similarities of /y/ and /w/ to both consonants and vowels may be relatively uncontroversial, but including /l/ and /r/ in this class demands a little further explanation. First, phenomena such as /l/-vocalisation and structurally conditioned clear vs. dark secondary articulation of /l/ are well known, and show the phonological similarity of /l/ to vowels. Kelly and Local (1986) argue that /r/ as well as /l/ is distinctively associated with ‘clear’ or ‘dark’ secondary articulation in a dialect-variable manner (see also Olive, Greenwood and Coleman 1993). And just as /y/ and /w/ are customarily analysed as the non-syllabic counterparts of /i/ (e.g. benefic[y]al vs. benefic[i]ary) and /u/, /r/ can be analysed as the non-syllabic counterpart of /a/ (Jakobson, Fant and Halle 1952: 22). Glides are thus characterised as follows:

(7.73) Glide =
\[
\begin{array}{c}
\text{[consonantal: ANY]} \\
\text{[vocalic: ANY]}
\end{array}
\]
Glide symbols are categorised by the following correspondences:

(7.74) \[\text{Alveolar U Glide U}\] [vocalic: \[\begin{array}{c}
\text{grave: +} \\
\text{height: close} \\
\text{round: -}
\end{array}\]\] \[\rightarrow /l/\]

(7.75) \[\text{Alveolar U Glide U}\] [vocalic: \[\begin{array}{c}
\text{grave: +} \\
\text{height: mid} \\
\text{round: +}
\end{array}\]\] \[\rightarrow /rl/\]

(7.76) \[\text{Labiovelar U Glide U}\] [vocalic: \[\begin{array}{c}
\text{grave: +} \\
\text{height: close} \\
\text{round: +}
\end{array}\]\] \[\rightarrow /w/\]

(7.77) \[\text{Palatal U Glide U}\] [vocalic: \[\begin{array}{c}
\text{grave: -} \\
\text{height: close} \\
\text{round: -}
\end{array}\]\] \[\rightarrow /y/\]

Observe that [grave] distinguishes palatal /y/ and clear /rl/ from labiovelar /w/ and dark /l/. (According to Kelly and Local 1986, the values of [grave] in /l/ and /rl/ might be reversed in other dialects, but would nevertheless be opposite to each other in all the dialects of English they have examined.) [height] is assigned to /y/, /w/ and /rl/ equal to the [height] of their syllabic counterparts. /l/ is assumed to be [height: close], on the basis of its vocalised form in, for instance, ‘milk’ [mɪlk]. Although two binary-valued features should be sufficient to distinguish four categories, the particular assignment of values of [grave] and [height] in (7.74)–(7.77) is not sufficient to categorise all four glides in terms of [vocalic] features alone, as /l/ and /w/ are both [grave: +, height: close], so a third feature, [round] is included in the value of [vocalic]. The value of [round] is assigned according to simple phonetic observation: [+round] for /w/ and /rl/ (Jones 1918: 195, 205; Gimson 1962: 177; Brown 1981: 67), [−round] for /l/ and /y/.

7.1.8 Syllable structure

The structure of monosyllabic words is described by the following rule:

(7.78) \[W \rightarrow \sigma (T)\]

T (termination) is the category of productive inflectional suffixes. In monosyllables, these must also be tautosyllabic with the last syllable of the word: that is, the regular verbal third person singular, nominal plural or nominal genitive suffixes /s/ and /z/ (but not /iz/, which forms a separate syllable), the weak past-tense suffixes /t/ and /d/ (but not /id/) and the nominalising suffix /θ/. In most accounts of English syllable structure (e.g. Fudge 1969, 1987, Fujimura and Lovins 1978, Borowski 1989), these elements are put in a
syllable-final position, after the coda. But in my analysis, they are not within-syllable constituents, as they occur only in word-final position (see section 7.2.6). Since words may have a complex internal structure that separates a free monosyllabic stem from such inflectional suffixes by several levels of intervening structure, these units are treated most simply as independent of regular syllable structure. A fuller account of this view of word structure is given in section 7.2. The set of \( T \) constituents is enumerated by the following rules:

\[
(7.79) \quad T[+\text{voice}] \rightarrow /l/z/ \\
(7.80) \quad T[-\text{voice}] \rightarrow /s/ \\
(7.81) \quad T[+\text{voice}] \rightarrow /d/ \\
(7.82) \quad T[-\text{voice}] \rightarrow /l/ \\
(7.83) \quad T[-\text{voice}] \rightarrow /\theta/ \\
\]

The termination and the last constituent of the syllable must agree in voicing. The last constituent may be the coda, if there is one, or the rime, otherwise. However, it may be only the last part of the coda or rime if the coda or rime is branching. Since ‘the last constituent of the syllable’ is not a well-defined primitive concept in the grammar formalism developed so far, it is necessary to introduce a new refinement: the addition of the category-valued feature [last] to every category. In every rule, the value of [last] of the mother is the category of the rightmost daughter constituent. In every structure this feature serves as a pointer to the rightmost terminal constituent, enabling adjacency constraints over non-sister constituents to be modelled, by percolating the relevant information up to the level of phonological structure at which the two constituents have sister ancestors. The voicing agreement constraint can then be expressed:

\[
(7.84) \quad o:\text{last}:\text{voice} = T:\text{voice} \\
\]

(Since /T/ does not alternate with /\theta/, it is not specified for [voice], so nominalisations such as ‘breadth’ and ‘width’ do not violate this constraint.)

A similar technique can be used to encode constraints on coda+termination sequences: */s\theta/ , */c\theta/ , */j\theta/ , */c\theta/, */s\theta/, */zz/ , */\theta\theta/ , */tt/ and */dd/. Ambisyllabic elements can be treated in like fashion, as the unification of the value of [last] in one syllable with the value of [first] of the next, where [first] is defined in a similar way to [last], as a pointer to the left-most terminal constituent.
The binary-branching structure of the syllable is expressed by the following phrase-structure rule:

\[(7.85) \quad \sigma \rightarrow O \ R\]

In order to express the facts that onsets are not involved in the determination of stress, and that syllable weight is determined solely by rime weight, the constraint \(\sigma\):weight = \(R\):weight is added to the above rule (see section 4.4.3).

A large number of filters could be placed on the syllable rule to avoid overgeneration. A complete grammar could include all of the following:

* /pli/, /snl/, /twol/, /dwal/, /dwo/, /zol/, /dol/, /vul/, /zul/, /z\alpha/, /smul/, /prul/, /pr\alpha/, /sprul/, /stru/, /skru/, /plul/, /klu/, /tru/, /dru/, /grul/, /frul/, /vru/, /slu/, /Cwu/, /C\alpha/ (where \(C\) = any consonant), * /\theta u/, * /\delta u/, * /\j u/, * /dwa/. Many of the large number of unattested combinations of onset and rime are accidental, unsystematic gaps and therefore do not demand to be filtered by the grammar. By testing the grammar against an on-line dictionary, I found that all these filters were necessary in practice in order to avoid extensive overgeneration. (By employing underspecification and unification to capture generalisations over this set, the number of filters that need to be added to the syllable rule is much smaller than the number of entries in this list.) The many restrictions involving /u/ in the above list reflect the fact that all instances of /u/ are historically exceptional. The other restrictions generally have some lexical exceptions (e.g. plinth, dwarf, voodoo), but these are highly idiosyncratic.

7.1.9 Onsets

Affricates and /h/ do not distribute like other obstruents (stops and fricatives). They cannot be preceded by /s/ or followed by a glide, and thus form onsets by themselves. Historically, affricates derive primarily from velar stops before close front vowels (which gives rise to doublets such as kirk/church), a development which occurred in both Old English (whence most contemporary voiceless affricates) and in Romance (especially voiced affricates) (Jespersen 1909/54: 51). The requirement that the velar stop preceded a vowel explains why such affricates are never followed by a glide. (I argue below that the only glides permitted in the onset are /rl/, /l/ and /w/, not /yl/.) The voiced velar stop /g/ cannot be prefixed by /s/, and although /k/ can be prefixed by /s/, the regular reflex of Old English /sk/ is /\sl/, contemporary /sk/ deriving mainly from Old Norse /sk/, for example doublets such as skirt (from Old Norse) vs. shirt (from Old English), and from more
Declarative analysis of English words

recent loans from Dutch, such as schooner. It is therefore justifiable to define affricate onsets by an independent phrase-structure rule:

\[(7.86) \ O \rightarrow \text{Affricate}\]

'Affricate' is defined in (7.45) above. The distribution of /h/ is similarly specific:

\[(7.87) \ O \rightarrow /h/\]

/s/+ stop and /s/+ nasal clusters distribute just like stops and nasals respectively, and so both are assigned to a superordinate category, Cl (closure). The analysis of onsets is thus (contra Clements and Keyser 1983: 41–51 and Cairns 1988) binary branching with two levels of structure:

\[
\begin{array}{c}
O \\
/ \ \\
Cl \ (\text{Glide}) \\
/ \ \\
/s/ \ C
\end{array}
\]

The upper part of this structure is defined by rules (7.88) and (7.89):

\[(7.88) \ O \rightarrow \text{Cl}\]

Only closure which does not contain a nasal, however, can be combined with a glide:

\[(7.89) \ O \rightarrow \text{Cl}[-\text{nasal}] \text{Glide}\]

The non-terminal feature [-nasal] prohibits the occurrence of /mr/, /snl/ etc. Voiced fricatives are historically separate from other obstruents, and do not occur in obstruent–glide clusters. This is described by the filter, which relates to rule (7.89):

\[(7.90) \ 
\begin{array}{c}
\text{Cl} \\
[+\text{voice} \ \\
+\text{continuant}] \\
\end{array}
\]

The only vowel which may occur after C+[y] word ‘beginnings’ is [u:]. Inclusion of /iw/ in the set of nuclei, coupled with prohibition of the glide /yl/ in onset clusters, ensures the required distribution. /yl/ can occur before other vowels, for instance in words such as yes, yacht etc., but not in clusters. Thus the following rule is needed to describe solitary /yl/:

\[(7.91) \ O \rightarrow /yl/\]

A number of other filters must be placed on rule (7.89):
7.1 Grammar of monosyllables

1 */stw/ and */skl/ (sclerosis is a lexical exception).

2 */pw/, */bw/ and */fw/ require a ‘*labial–labial’ filter. According to the categories of the present analysis, this would be a prohibition on the sequence * [+ grave, − compact][+ grave]. If the feature [labial] were employed, *[+ labial][+ labial] would be a simpler expression of this filter. The exceptions, such as pueblo, bwana, fuego, foie (gras), are all loans. Pierrehumbert and Nair (1995) point out that this sequence constraint is not limited to branching onsets, but applies across syllable boundaries too.

3 /gw/ only occurs in loans, especially from Welsh, but the common occurrence of /kw/ precludes extension of the *[+ labial][+ labial] filter to include velars as well (so that *[+ grave][+ grave] is not an appropriate formulation).

4 */tl/, */dl/, */θl/. This restriction appears to belie the ‘neutralisation’ of alveolar and velar place of articulation in stop+/l/ clusters, since [dl] occurs in some varieties of English as a token of /gl/, according to Firth (1935b/1957a: 43) and Jespersen (1909/54: 353). The Survey of English Dialects (Orton and Dieth 1962) records [tl] in all the test-words beginning with /kl/ and [dl] in both the /gl/- test-words for some speakers in Westmoreland, Cumberland, County Durham, Northumberland, Lancashire and Yorkshire. This suggests that stop+/l/ clusters should be analysed as [− labial][− labial] sequences. Both /d/ and /g/ being [− labial], they are not functionally distinct at this place in structure.

5 */sr/, */šl/, */šw/. These gaps suggest that compactness is not distinctive in non-grave fricatives before glides. Cairns (1988) categorises the glides /l/, /w/ and /y/ as [+ high] and /t/ as [− high] (equivalent to [height: close] vs. [height: mid], as I have done), /s/ as [− high] and /š/ as [+ high]. Thus /st/ would be [− high][− high], and /šl/ and /šw/ would be [+ high][+ high]. Viewed in this light, the prohibition against them is another instance of dissimilation, and can be expressed in the same manner as the labial–labial filter above. Many other dissimilatory phonotactic constraints are introduced below, such as gravity dissimilation constraints on rimes and pansyllabic constraints on onset and codas. The prevalence of such dissimilatory constraints would seem to call into question the unmarked status accorded to assimilation (i.e. spreading) in theories such as Autosegmental Phonology and Government-based
Declarative analysis of English words


Non-branching onsets
A single stop or nasal can be an onset:
(7.92) \( \text{Cl} \rightarrow \begin{array}{c} C \\ [-\text{strident}] \\ [-\text{continuant}] \end{array} \)

providing that \( C \) is not a glide, and that palatal and velar nasals are prohibited:
(7.93) \( *C \)

[\text{vocalic: ANY}]
(7.94) \( *C \)

[\text{+nasal} \\ \text{consonantal: [+compact]}]

Continuants can also occur in the Cl position, but glides are excluded by prohibiting consonants with a non-empty value of the feature \([\text{vocalic}]\). In other words, of the continuants, only fricatives can occur in the Cl position:
(7.95) \( \text{Cl} \rightarrow \begin{array}{c} C \\ [+\text{continuant}] \\ [-\text{vocalic: ANY}] \end{array} \)

Branching closure
Stop and nasal consonants can be prefixed by /s/:
(7.96) \( \text{Cl} \rightarrow /s/ \begin{array}{c} C \\ [-\text{voice}] \\ [-\text{strident}] \\ [-\text{continuant}] \end{array} \)

The following two filters on this rule might be avoided, but are included here for completeness. As in closure without an initial /s/, constraint (7.94) excludes tokens of /sn/ with a palatal or velar feature-specification of /n/.

The constraint:
(7.97) \( *C \)

[\text{-nasal} \\ [+\text{voice}]]

excludes /sb/, /sd/ and /sg/ from the right-hand daughter position in rule (7.96), though it is arguably only a matter of convention that the stops in /s/+stop clusters are represented using /p/, /t/, /k/ and not /b/, /d/, /g/ (cf.
Italian or Scots Gaelic spelling of these clusters, for example). If it were not for the [+voice] feature on nasal categories, which is redundant in onset position anyway, it would be possible to eliminate this filter by altering rule (7.96) to:

(7.98) \[ \text{Cl} \rightarrow /s/ \]

\[-\text{voice}]\]

\[-\text{strident} \]

\[-\text{continuant} \]

\[-\text{voice} \]

In onsets, therefore, /n/ should either be [-voice] or unspecified for [voice], whereas in codas, /n/ must be [+voice] for the correct T allomorphs to be selected. It might be proposed that nasals are not lexically marked [+voice], but that coda nasals alone acquire [+voice] by an implicational constraint on feature-specification.

The Cl position may be empty, since glides may occur by themselves in onset position:

(7.99) \[ \text{Cl} \rightarrow \emptyset \]

(where \( \emptyset \) denotes the empty string).

7.1.10 Rime structure

Rimes are defined by the following rule, augmented by a number of constraints specific to English:

(7.100) \[ \text{R} \rightarrow \text{N} \text{Co} \]

Syllable weight is determined by rime weight. Although weight is determined by the shape or structure of the rime, it is not enough to describe the determinants of syllable weight in terms of rime structure alone, as is done in most Metrical analyses, such as Borowsky (1989): information about the weight of the rime needs to be available at the syllable node, as the grouping of syllables into feet depends on syllable weight (see section 4.4.3). I shall therefore introduce a feature [weight] with values 'heavy' and 'light' to encode the distinction determined by the structure of the rime.

A rime is heavy if either its nucleus or its coda is heavy; otherwise it is light. Rather than invoking final consonant extrametricality, we assume that syllables closed by one consonant are light and those closed by a cluster are heavy. Note that in this analysis, most single intervocalic consonants are codas as well as onsets, i.e. amabisyllabic, and thus a single consonant in coda position is insufficient to make a rime heavy. The value of the syllable [weight] feature cannot simply be determined by requiring
rime, nucleus and coda to agree in the value of [weight] by setting $R:\text{weight} = N:\text{weight} = C_0:\text{weight}$, because it is possible to have opposite weights in nucleus and coda, for instance light nucleus + heavy coda: /int/, heavy nucleus + light coda: /eyt/. So branching codas and nuclei will be marked [weight: heavy]. Unbranching codas and nuclei are not marked [weight: light]; they are left unmarked. (Weight is thus a privative feature in codas and nuclei.) Rime weight can then be determined by the following two constraints:

(7.101) $(C_0:\text{weight} = \text{heavy} \lor N:\text{weight} = \text{heavy}) \supset R:\text{weight} = \text{heavy}$

Otherwise:

(7.102) $R:\text{weight} = \text{light}$

(The ‘otherwise’ clause might be dropped, making $R:\text{weight}$ privative too.)

The main rime constraints refer to the subconstituents of branching nuclei or codas. Constraint (7.103) expresses the generalisation that if the coda branches in an unchecked rime, the second $C$ of the coda must be alveolar. In other words, in $V_1 V_2 C_1 C_2$ rimes, $C_2$ must be alveolar:

(7.103) $(C_0:\text{left} \neq C_0:\text{right} \land R[-\text{checked}]) \supset C_0:\text{right} = \text{Alveolar}$

Exceptions to this restriction are the rimes:

1. /arskl/ (e.g. mask), /arsp/ (e.g. clasp) and /arnčl/ (e.g. branch), which are only apparent exceptions. The nucleus of these rimes is a phonetically long, back exponent of the phonological short-vowel category /a/. These rimes could therefore be represented phonologically as /ask/, /asp/ and /anč/. The phonetic interpretation of these rimes can be described by exponent rules such as /asC/ $\rightarrow$ [asC] or the equivalent in the /anč/ case.

2. /ornčl/, which is found in only paunch, haunch and launch.

3. /orlk/ in baulk. The usual contemporary form of this rime is /ork/ as in chalk, walk etc. or /alk/ or /olkl/, as in baulk and caulk. /orlk/ is apparently a spelling pronunciation at odds with the general pattern of English phonology, so it is reasonable to account for this oddity by a specific phrase-structure rule, if necessary.

4. /eynj/ in range and mange, and /awnj/ in ‘lounge’.

There are four closely related constraints on the ‘crossing’ diphthongs /aw/, /oy/ and /ay/, which I shall group together under the general label of
Gravity dissimilation. A possible generalisation over them is that open back diphthongs may not co-occur with velar closure codas, nor open front diphthongs with alveo-palatal closure. However, the details of the pattern prevent such a simple formulation, and the four cases will be stated separately:

(7.104) Gravity dissimilation

\[
\begin{array}{c}
\ast R \\
N \quad \text{Co} \\
a \quad w \quad \text{C}
\end{array}
\]

[consonantal: [+grave]]

This prohibits the rimes /awg/, /awk/, /awb/, /awp/, /awf/ and /awv/.

(7.105)

\[
\begin{array}{c}
\ast R \\
N \quad \text{Co} \\
o \quad y \quad \text{C}
\end{array}
\]

[consonantal: [+grave]]

This prohibits the rimes /oyg/, /oyk/, /oyb/, /oyp/, /oyf/ and /oyv/. (The exception coif is a loan.)

(7.106)

\[
\begin{array}{c}
\ast R \\
N \quad \text{Co} \\
a \quad y \quad \text{C}
\end{array}
\]

[consonantal: [-grave ] [+compact]]

This prohibits /ayj/, /ayč/ and /ayš/.

(7.107)

\[
\begin{array}{c}
\ast R \\
N \quad \text{Co} \\
o \quad y \quad \text{C}
\end{array}
\]

[consonantal: [-grave ] [+compact]]

This prohibits /oyj/, /oyč/ and /oyš/.

Two final rime constraints concern unchecked alveolar nasal+stop
cluster codas. Firstly, /unt/, /iynt/ and /uwnt/ are ill formed, although /int/ (which is [+checked]) is well formed (e.g. lint, mint, dint, hint).

(7.108) *R

N
[height: close]

Co
[−checked]

nt

Secondly, the only branching nuclei with which coda /nd/ may combine are /ay/ and /aw/ (fiend is an exception):

(7.109) R

N
[height: close]

Co

V

V

n
d

a

[voice: [height: close]]

7.1.11 The coda systems

There is a general constraint on codas which can actually be stated in the rime rule (7.100), since it applies to all codas, not just to particular sub-categories of coda: /l/ is the only glide which can occur in codas, /r/, /w/ and /y/ being considered part of the nucleus (i.e. the 'offglide').

Simple Codas

Since codas are obligatory in this grammar, in order to keep the rime rule simple, empty codas must be permitted, to allow for open syllables. Voicing agreement between rime and termination (constraint 7.84) in e.g. key[z], rowe[d], requires that empty codas be [+voice].

(7.110) Co → 0
[+voice]

Non-empty, non-branching codas are described by the rule:

(7.111) Co → C

This rule must be augmented by three filters. First, as in onsets, only alveolar and labial nasals are permitted in non-branching clusters. Palatal and velar nasal non-branching codas must be excluded, as they only occur in the clusters /nɔl/, /ŋɔl/, /ŋɔk/ and /ŋɔgl/. This constraint is the same as (7.94).

Second, /g/ does not occur after unchecked nuclei (i.e. long vowels or /u/). Enforcing agreement in the value of the feature [checked] in both nucleus
and coda (i.e. N:checked = Co:checked), this restriction can be expressed
locally to the coda constituent:

(7.112)        Co
              [+checked]  
               /g/  

The exceptions *league, fugue, brogue, drogue, rogue, morgue, erg* and *berg*
are all loans.  Third, although *with* has /ð/ as coda, /ð/ is so exceptional after checked
nuclei that it may be filtered out to avoid overgeneration:

(7.113)        Co
              [-checked]  
               /ð/  

The highly restricted distribution of /ð/ is perhaps an artefact of its almost
exclusive occurrence in function words, such as *this, that, the, other, whether, rather*. /ð/ also occurs intervocalically in content words such as
*heather, leather, weather, mother* etc. In these cases, /ð/ arose historically
through voicing assimilation/spread (cf. *heath*), or through lenition of an
intervocalic /ld/. The phonetically voiced coda in *with* could be analysed as
a morphosyntactically conditioned special phonetic interpretation of /θ/,
since *with* is pronounced [wðθ] in Scots and some Northern dialects (see also
Jespersen 1909/54: 201), in which case /ð/ is not required in the analysis of
monosyllabic words, because of the implications:

(7.114) Content word: θ ⊇ [-voice] 
Function word: θ ⊇ [+voice]  

Branching codas I, sonorant clusters

In English, all branching codas are heavy, and thus bear the feature [weight:
heavy]. Geminates are prohibited:

(7.115) Co:left ≠ Co:right 

Sonorant +consonant sequences are common:

(7.116) Co → {Glide} C  
        \Nasal/  

The disjunction in this rule could be eliminated by using the feature
[+sonorant] to characterise glides and nasals. This general phrase-structure rule must be augmented by a number of further constraints, however.
First, C cannot be a glide (i.e. /l/). Thus, /nl/ and /ml/ are prohibited, as is /ll/, though that is also excluded by the more general constraint on geminate codas described above:

\[(7.117) \text{C} \neq \text{Glide}\]

Second, if the left-hand daughter of Co is Nasal, it must be homorganic with C:

\[(7.118) \text{Nasal:consonantal} = \text{C:consonantal}\]

/mb/ does not occur in codas, except in iamb, rhomb (a non-native, learned word), rhumb (a contraction of Spanish and Portuguese rumbo), zimb, an Ethiopian fly (Jespersen 1909/54: 217) and, for some speakers, jamb.

Third, there are a number of restrictions on nasal+voiceless fricative codas. Few speakers have the codas /ns/ or /ls/. (For most speakers, lexicalisation of phonetic [t]-epenthesis has made these codas /nc/ and /lc/.) They must consequently be filtered out. Since, in sonorant+consonant clusters, the sonorant can only be /l/, /n/ or /m/ according to the previous constraints, and since /ms/ violates the homorganic nasal constraint (7.118), all that is needed is a constraint that ensures Co:right $\neq /s/$. The exception, welsh, is historically a contraction of the two-syllable, two-morpheme Old English word welisc, i.e. Wale+ish. Welsh has an alternative pronunciation /welc/, with lexicalised epenthetic [t], just as the contemporary pronunciation /frenç/ developed from earlier /frenʃ/. /nθ/ must also be excluded. The two exceptions are plinth, a Classical loan-word, and month, which was originally two syllables monath and two morphemes món (moon) + suffix -ath.

The cluster /ns/ is excluded after unchecked nuclei (the exceptions being rimes [ans] and /awns/, of which the former is analysed as /ans/ in section 7.1.5 above). This constraint is enforced locally to the coda category in the sonorant cluster rule (7.116) by reference to the coda's value of the feature [checked], which must agree with that of the nucleus:

\[(7.119) \text{Co}_{+\text{checked}}\]

\[
\begin{array}{c}
\text{n} \\
\text{s}
\end{array}
\]

Fourth, nasal+voiced fricative codas such as /mv/, /nz/ and /nθ/ are prohibited:

\[(7.120) \ast \text{Co}\]

\[
\begin{array}{c}
\text{Nasal} \\
\text{Voiced} \sqcup \text{Fricative}
\end{array}
\]
Fifth, glide + voiced alveolar fricative codas such as /lz/ and /lo/ are prohibited:

\[(7.121) \quad \ast \text{Co} \]
\[
\quad \text{Glide} \quad \text{Voiced} \quad \text{Alveolar} \quad \text{Fricative}
\]

In monosyllables, /nz/ and /lz/ are always coda + termination, not a branching coda.

Sixth, /l/ + voiceless consonant sequences do not occur after unchecked nuclei. Again this is enforced locally to codas by exploiting the agreement of N:checked with Co:checked enforced by the rime rule:

\[(7.122) \quad \text{Co} \quad [+\text{checked}] \]
\[
\quad \text{Glide} \quad \text{Voiceless}
\]

The exception is the coda of wolf, one of the few cases of a historically and morphophonologically short /u/.

Seventh, /n/ is generally prohibited in codas after /l/:

\[(7.123) \quad \ast \text{Co} \]
\[
\quad \text{Glide} \quad \text{Alveolar} \quad \text{Nasal}
\]

(the sole exception is kiln, a spelling pronunciation according to Jespersen 1909/54: 208) as are the voiced grave stops /g/ and /bl/ (exceptions: bulb and alb):

\[(7.124) \quad \ast \text{Co} \]
\[
\quad \text{Glide} \quad \text{Labiovelar} \quad \text{Stop}
\]

Branching codas II, obstruent clusters

In stop+/s/ syllable endings, /s/ is usually the nominal plural, nominal genitive or third person singular verbal termination, rather than a constituent of the coda. However, in some cases, including those which are spelt with an x or tz, the stop+/s/ sequence is a tautomorphic coda cluster. These codas are always [+checked], and thus cannot occur in a rime with a long vowel or /u/ as its nucleus. The sole exception is 'traipse', which according to OED is either etymologically possibly a verb form /treyp/ (Middle Dutch and Middle Low German trappen 'to tread, trample') plus the third person
singular verbal termination /sl/, or a reduced disyllable (Old French *trapper, trapesser, trepasser* > *trespass*). The structure of stop+/s/ coda clusters is thus:

\[(7.125) \quad \text{Co} \rightarrow \text{Voiceless} \sqcup \text{Stop} /s/\]
\[\quad [+\text{checked}]\]

The structure of fricative + stop codas is defined:

\[(7.126) \quad \text{Co} \rightarrow \text{Voiceless} \sqcup \text{Strident} \sqcup \text{Fricative} \quad \text{Voiceless} \sqcup \text{Stop}\]

A constraint is added to this rule to enforce restrictions on its two subcases: either Fricative is /s/, or, providing that the nucleus is checked (enforced by stating Co[+checked]), Co is /ft/. (An exception to the checked nucleus restriction is the rime [aft], which can be analysed as /aft/, according to the argument in section 7.1.10 above.) /ft/ codas occur mainly in words which are etymologically past-tense forms of verbs ending with /v/, but which no longer have a transparent past-tense morphology. The restriction of these codas to checked rimes can be attributed to the vowel quantity opposition in such pairs as *leave ~ left*, and *weave ~ weft*. Other examples of /ft/ codas are words such as *neft*, a dialect form of *newt*. Historically, and in the dialects in which this pronunciation occurs, [ef] is actually an exponent of the nucleus /iw/ (for similar examples in Tyneside English, Fang and Swahili, see Kelly and Local 1989: 135–9, 150–3; Kelly 1991), the coda being just /t/. However, for speakers who are not conscious of the etymology of these forms, /ft/ must be considered the coda, as in *lift* or (adjective) *left*. The constraint can be expressed as follows:

\[(7.127) \quad \text{Co[left: Alveolar]} \lor \text{Co[+checked, left: Labial, right: Alveolar]}\]

Closely related to /ft/ codas are stop+/t/ codas, that is, /pt/ and /kt/:

\[(7.128) \quad \text{Co} \rightarrow \text{Voiceless} \sqcup \text{Stop} /t/\]
\[\quad [+\text{checked}]\]

/\tt/ is prohibited by the more general constraint against geminate branching codas (7.115). Despite the great similarity between (7.128) and (7.125), they cannot be collapsed as (7.127) applies to (7.128), but not to (7.125).

### 7.1.12 Pansyllabic constraints

In addition to constraints on the concatenation of sisters, such as onset and rime, or at least adjacent constituents, such as the coda and syllable termination, there are prohibitions against the repetition of certain consonants in certain onset and coda positions. These are described in Cairns
The first pansyllabic constraint that Cairns describes is that in syllables which begin with an /sC/ onset, the first constituent of the coda may not also be /s/. This constraint prohibits syllables such as */stask/, */spast/, */skusp/, */strask/, */splast/ and */skrusp/. The second constraint Cairns notes is that in syllables which begin with an /sC/ onset (as before), the C segment may not be repeated in the head of the coda if it is the first coda constituent and if the head consonant of the onset closure is marked for phonological content. This constraint prohibits such forms as */spup/, */skak/, */snon/, */smam/, */splup/ and */skrak/, without preventing pup, crack, none, skulk, smarm or stet, strut (in which /t/ is said to be contentless).

The final pansyllabic constraint considered by Cairns is that a segment in onset glide position may not appear anywhere else in the same syllable. This constraint excludes */klilt/, */krark/ and */klul/ while allowing click, crock (glide is not repeated), lilt and lull (initial /l/ is in head, not glide position). Quote does not fall foul of this constraint because Cairns analyses initial /kw/ as a single consonant, rather than a stop+glide sequence.

An autosegmental paradox
These pansyllabic constraints are formally different from the constraints on phrase-structure rules of the type employed above, since they are not restricted to sisters in the syllable-structure (or even adjacent constituents). Cairns suggests (somewhat tentatively) that since they all prevent the repetition of segments in certain structural configurations, they can be regarded as structure-dependent restrictions on the Autosegmental spreading of a segment across two positions. For instance, the second constraint mentioned above, which prohibits forms such as /spup/, could be expressed as a constraint against a branching association between a segmental feature specification, a branching onset position and a coda position:

(7.129) O Co Syllable structure plane
     *X X... X
     [F] Feature plane

The other two pansyllabic constraints could be formulated similarly. Ingenious though this proposal is, Autosegmental formulations such as (7.129) will be mutually contradictory. Let us consider why this is so.
The first constraint prohibits an /s/ from being shared between the syllable positions which I have called O:Cl:left and Co:left in the same syllable. The features anchored to O:Cl:right must be on a different tier from the features which are shared between O:Cl:left and Co:left in order to avoid crossing lines. That is:

(7.130) *Sharing*
O:Cl:left features and Co:left features are coplanar.

(7.131) *Non-blocking*
O:Cl:right features are on a different tier from O:Cl:left features and Co:left features.

Now consider the second constraint. A complete Autosegmental version of this restriction (7.129) states that a segment may not be shared between O:Cl:right and Co:right if O:Cl:left is filled, Co:left is empty and Co:right is non-empty. If Co:left is filled, the constraint is blocked. The Autosegmental account of this fact is that the presence of an association line between Co:left and its features blocks the sharing of an autosegment between O:Cl:right and Co:right, rendering the filter (7.129) inapplicable. This is achieved by putting Co:left features on the same tier as O:Cl:right features and Co:right features. The occurrence of features in Co:left position blocks the sharing of another feature specification on the same tier between O:Cl:right and Co:right, since such a configuration would have crossing lines. Thus:

(7.132) *Sharing*
O:Cl:right features and Co:right features are coplanar.

(7.133) *Blocking*
Co:left features are on the same tier as O:Cl:right features and Co:right features.

For the Autosegmental account of the second constraint to work, then, O:Cl:right and Co:left must lie on the same tier (7.133), which flatly contradicts (7.131). Taken individually, the Autosegmental accounts of these constraints, exemplified in (7.129), are simple and appealing, constructed as they are on fundamental and universal mechanisms: tiered representations, the Obligatory Contour Principle and the No Crossing Constraint. Taken together, however, they are mutually contradictory. This contradiction presents a deeper challenge to Autosegmental Phonology, for it shows that a number of attractive and apparently disjoint parts of an Autosegmental analysis may have unforeseen, mutually contradictory consequences.

Since the depth of syllable-structure is finitely bounded, however, it is
possible to express these constraints using feature-structures. The three relevant filters on syllable-structure are as follows:

(7.134) \[ \sigma: \begin{cases} O: \left[ \text{left: /s/} \right] \\ \text{head: C} \\ R: \left[ \text{Co: [left: /s/]} \right] \end{cases} \]

(7.135) \[ \sigma: \begin{cases} O: \left[ \text{left: /s/} \right] \\ \text{head:[C: */t/*]} \\ R: \left[ \text{Co: [glide: 0]} \right] \end{cases} \]

(7.136) \[ \sigma: \begin{cases} O: \left[ \text{glide} \right] \\ R: \left[ \text{C: [glide]} \right] \end{cases} \]

By using subscripts to denote equality of content of non-adjacent nodes does not suffer from the paradox that the No Crossing ‘Constraint’ brings to Cairns’s Autosegmental analysis of these constraints.

To conclude the grammar of monosyllabic words, I shall list the sets of sub-strings defined by the rules and constraints given above for onset, nucleus and coda.

**Onsets:** č, ğ, ž, h, y, n, m, sn, sm, θ (the empty string), sp, st, sk, p, t, k, b, d, g, f, θ, s, š, r, l, w, spr, spl, str, kw, pr, pl, tr, tw, kr, kl, kw, br, bl, dr, dw, gr, gl, fl, gl, sw, šr.

**Nuclei:** i, e, a, o, ʌ, iy, ey, ay, oy, iw, aw, uw, əw, ar, or, ər.

**Codas:** l, m, p, t, k, b, d, g, f, v, s, š, θ, z, č, j, ld, ts, ks, ps, ft, sp, st, sk, pt, kt, n, ng, ð, ps, ks, mp, mf, nt, nd, nč, nj, ns, nk, lp, lt, lk, lc, lj, lm, lf, lv, lθ, ls.

Having a reliable, accurate and explicit grammar of syllables is an essential prerequisite for developing a declarative account of English metrical phonology, which will require concepts such as ‘vowel length’, ‘closed vs. open syllable’, ‘checked’ and ‘syllable weight’ to be formally well defined. In the remainder of this chapter, I shall show how a declarative account of English metrical structure in conformity with the accounts of lexical and metrical structure discussed in section 4.4 above can be built on the grammar of English monosyllabic words presented in this section.
7.2 A grammar of English words

Mr. William Archer, after a long list of seemingly arbitrary accentuations in the English language (*America To-Day*, p. 193), goes on to say: 'But the larger our list of examples, the more capricious does our accentuation seem, the more evidently subject to mere accidents of fashion. There is scarcely a trace of consistent or rational principle in the matter.' It will be the object of the following pages to show that there are principles, and that the 'capriciousness' is merely the natural consequences of the fact that there is not one single principle, but several principles working sometimes against each other. (Jespersen 1909/54: 160)

7.2.1 Introduction

The grammar presented in the preceding section did not define words of more than one syllable, although it included a rudimentary monosyllabic word-structure rule (7.78), which said that a word could be a single syllable or a syllable with a termination. But in words of more than one syllable, a new factor must be considered in phonological descriptions: the representation of stress, or relative prominence of each syllable compared to the others. In this section, I outline a declarative account of English word stress which has also been implemented as a parser which is used for, amongst other things, syllabification and stress assignment in a text-to-speech system. It consists of a phrase-structure grammar which specifies how words are made up of stress domains containing feet, which are in turn made up of syllables. The division of a word into syllables is thus no different in kind from the division of syllables into smaller units such as onsets, nuclei and codas. On the basis of such a grammar, standard parsing techniques can be used to analyse words into metrical units if a procedural interpretation of the grammar is required. It is not necessary to write ad hoc, special-purpose heuristic 'syllabification' or 'stress assignment' algorithms, as is usually done in text-to-speech systems.

The grammar of English words described below extends the grammar of monosyllables described in the previous section to higher levels of structure. Certain additional constraints (such as an account of 'reduced' vowels and syllabic consonants) and relaxations of the monosyllabic word constraints (such as the prohibition against short open syllables enforced by the feature [+checked] in rule (7.16) which does not apply if another syllable follows) are needed to extend the monosyllable grammar to disyllables, but since they are not central to the discussion in this section, I shall not detail them.

7.2.2 Abstract stress

Elementary phonetics and phonology texts (as well as some scholarly works too, such as Jakobson and Halle 1956:34 and Mohanan 1986: 165) often
describe stress as a phonetically simple phenomenon, a composite of the
‘prosodic’ parameters of amplitude, duration and pitch. From the standpoint
of phonetic theory, however, this view of stress is excessively simplistic. It is
easily demonstrated that stress is a phonological category (Pierrehumbert,
Beckman and Ladd 1996) which may have completely contrary phonetic
exponents in different dialects of English (Wells 1988: 24). Furthermore, the
phonetic exponents of stress may include parameters usually regarded as seg-
mental, including voicing, aspiration, labiality and vowel quality. A striking
example of this (due to Simpson 1988, and also noted by Selkirk 1972/80:
22–40) is demonstrated by a comparison of the phonetic exponents of phrasal
stress in the pronunciation of preposition + pronoun sequences such as to
him, to her, for him and for her. If stress falls on the preposition in to him
/tuim/ or to her /tuwl/, the exponents of stress are (amongst others) labial-
ity and closeness of the nucleus /uw/, and the exponents of non-stress on the
pronoun include absence of /h/, as well as ‘traditional’ exponents of stress
such as vowel length (stressed, long /uw/ in to vs. unstressed short /ə/ in her).
When the stress falls on the pronoun, we have /təhim/ and /təhar/, without a
labial or close quality in the nucleus of the preposition, but with /h/ and a long
nucleus in the pronoun. Other preposition + pronoun sequences display
similar behaviour: compare /forim/ with /fəhim/ and /fəar/ with /fəar/.
These cases do not demonstrate that stress is an unusually difficult phenom-
enon to describe phonetically, simply that it must be regarded as an abstract
phonological category (like any other) with various phonetic interpretations
in different structures, dialects and lexical subsystems.

7.2.3 Lexical structure
Until the publication of the ground-breaking analyses of Chomsky, Halle
and Lukoff (1956), Chomsky and Miller (1963: 315) and Chomsky and
Halle (1968), English lexical stress was considered to be such a complicated
phenomenon that it was regarded as too arbitrary to be described by a few
general principles, as taxonomic descriptions such as Kingdon (1958)
testify. In such taxonomic studies, English is described as if it were a ‘free
accent’ language, in contrast to languages such as Polish or Czech, in which
accent placement is fixed.

If the reasons for the apparently anarchic nature of English word stress
are examined, however, it is apparent that the variety of different word stress
patterns in English can in part be attributed to the fact that the English
lexicon contains words which have been borrowed from many different lan-
guages, as a result of which English has several different patterns of lexical
stress. If we are careful to distinguish words which are, or which pattern as
if they are, of different origin, stress placement can be reliably determined
Declarative analysis of English words

according to a small number of general principles of stress placement of the various stress subsystems. As in the syllable grammar, the language is regarded as a system of several simple, interacting systems.

An extensive number of attested stress systems are defined by the parameters of metrical structure described in section 4.4.3. I shall argue below that one choice of settings determines one of the English stress subsystems, but does not provide a complete account of lexical stress. This initially problematic situation is remediated by a recognition and separation of several hierarchically distinct levels of prosodic constituency, corresponding to Latinate, Germanic, Compound and Inflectional morphophonology, the four lexical strata of Halle and Mohanan's (1985: 57–8, 64) version of lexical phonology.

Although most people are unaware of the history of English, of course, speakers behave as if operating with different principles of stress assignment for Germanic and Latinate words, and for lexical and phrasal stress. (The term 'Latinate' covers loans directly from Latin, especially scientific and ecclesiastical terms, as well as words introduced into English from Norman French. Later loans from French and other Romance languages are often 'exceptional' loans, just like loans from less familiar languages: compare déjá vu with vindaloo. Some Greek affixes, such as hetero-, behave like Germanic affixes, although Greek stems often behave like Latin stems.) The grammar of English words will implicitly encode information about the different language systems which historically contributed to its present form. This is the extent of people's implicit acquired knowledge of etymological factors in stress assignment. The same distinction of source language systems helps to account for the ordering and grouping of prefixes, stems and suffixes in complex words such as unacceptable. In this word, un- is a Germanic prefix, ac- a Latinate prefix, -cept a Latinate stem, and -able a Germanic suffix. In mixed Germanic/Latinate polymorphemic words, the order 'Germanic prefixes, Latinate prefixes, Latinate stems, Latinate suffixes, Germanic suffixes' is quite regular (see Selkirk 1982: 77–124). The fact that Latinate morphemes are sandwiched between Germanic affixes and have the same distribution as Germanic stems, suggests that words have a branching structure of the following general shape:

(7.137)

```
Word
  /   \
Germanic prefixes  Stem  Germanic suffixes
    /   \
  Latinate prefixes  Latinate stem  Latinate suffixes
```
Aronoff (1976) discusses a number of apparent violations of this affix-ordering generalisation. For example, the suffixes -able, -ment and -ise do not shift stress (7.138a) and may attach to Germanic stems, revealing them to behave like Germanic suffixes (7.138b). Historically, however, they are Latinate morphemes (7.138c). (In order to avoid unwanted historical connotations, I shall refer to Latinate, inner or stress-shifting morphemes as level 1 morphemes and Germanic, outer or non-stress-shifting morphemes as level 2 morphemes. For example, *photograph* is etymologically Greek, but since it is a stem, behaves synchronically like etymologically Latinate stems, and is hence level 1. Greek prefixes, on the other hand, are often stressed like Germanic prefixes, in a separate stress from the stem, and are thus level 2. Subscripts 1 and 2 are used to show the level of a category.)

(7.138)  
\[ \text{a. patent-able, govern-ment, protestant-ise} \]
\[ \text{b. wearable, washable, shipment, wonderment, womanise, lionise} \]
\[ \text{c. viable, definable, testament, parliament, finalise, utilise} \]

However, these suffixes may be followed by a Latinate, stress-shifting morpheme (7.139):

(7.139)  
\[ \text{patent-abil-ity, govern-mént-al, protestant-is-ation} \]

In view of such examples, we must recognise the possibility of recursion of the structure in (7.137). For instance, the words in (7.139) have the structure shown in (7.140).

(7.140)  
\[ \text{W}_1 \]
\[ \text{W}_2 \]
\[ \text{Suffix}_1 \]
\[ \text{W}_1 \]
\[ \text{Suffix}_2 \]
\[ \text{Stem}_1 \]
\[ \text{(Suffix}_2\text{)} \]
\[ \text{e.g patent} \]
\[ \text{able} \]
\[ \text{ity} \]
\[ \text{govern} \]
\[ \text{ment} \]
\[ \text{al} \]
\[ \text{protest} \]
\[ \text{ant} \]
\[ \text{ise} \]
\[ \text{ation} \]

Note that this yields a classical example of cyclic stress assignment. Latinate stress assignment on the inner \(W_1\) constituent assigns secondary stress to
the initial syllables in (7.139), as well as assigning primary stress to the final foot on the outer $W_1$ cycle. We discuss the interaction between morphological structure and stress in section 7.2.6 below.

Compound words, such as *blackbird*, can be added to this scheme too, as can termination morphemes, which do not really belong in the syllable grammar. The following contrived example shows the structure of the invented but grammatical compound word *underinertness-gradings*. The fact that even such an artificial word has a morphological structure is demonstrated by the fact that other permutations of the various prefixes, stems and suffixes are all ungrammatical viz. *inunderertness-gradings* (prefixes not properly nested), *underinert-gradingsness* and *underinertness-gradesing* (suffix placed after inflectional termination). *Underinertnesses-grading*, however, is also well formed, and accords with the grammar behind (7.141).

\[
(7.141) \quad W
\]

\[
\begin{array}{c}
W_2 \\
W \\
W \\
W
\end{array}
\]

\[
\begin{array}{c}
T \\
-s \\
-gradings
\end{array}
\]

\[
\begin{array}{c}
Prefix_2 \\
Prefix_1 \\
in-
\end{array}
\]

\[
\begin{array}{c}
W_2 \\
W_1 \\
-s
\end{array}
\]

\[
\begin{array}{c}
Stem_1 \\
-ness
\end{array}
\]

\[
\begin{array}{c}
Suffix_2 \\
Suffix_1
\end{array}
\]

This theory of word-structure is adapted from Church (1985: 251), who presents the following context-free grammar schema:

\[
(7.142) \quad \text{word} \rightarrow \text{level}3 \ (\text{regular-inflection})^* \\
\quad \text{level}3 \rightarrow (\text{level}3-\text{prefix})^* \ \text{level}2 \ (\text{level}3-\text{suffix})^* \\
\quad \text{level}2 \rightarrow (\text{level}2-\text{prefix})^* \ \text{level}1 \ (\text{level}2-\text{suffix})^* \\
\quad \text{level}1 \rightarrow (\text{level}1-\text{prefix})^* \ (\text{syll})^* \ (\text{level}1-\text{suffix})^*
\]

Church's grammar is based on Halle and Mohanan's (1985) version of lexical phonology, which is idiosyncratic in postulating four levels. (Previous work and most subsequent work on English employs only two
levels.) Halle and Mohanan's four levels derive some additional support from the fact that they are also said to be the domains of morphophonological phenomena, such as CiV Lengthening (e.g. caucásus → caucásian) at level 1, *-lowering (e.g. abound ⇒ ab\[A\]ndance) at level 2, dialect C Vowel Tensing (e.g. citi hall) at level 3, and /-Resyllabification (e.g. seal\[a\] office) at level 4. Rather than attempt to translate the many transformations proposed in analyses such as Halle and Mohanan (1985) into declarative terms, however, I shall concentrate on stress alone in the remainder of this section. For a fuller discussion of declarative lexical phonology, see Coleman (1995).

I shall in fact employ only two etymological levels, levels 1 and 2, approximately corresponding to the \([±\text{Latinate}]\) distinction in Chomsky and Halle (1968), or the \([±\text{cyclic}]\) distinction in Halle and Kenstowicz (1991). Inflection is not superordinate to compounding, as examples such as underinertnesses-grading demonstrate. I shall leave it to the correct formulation of morphological rules to determine that Ts may not be spuriously multiplied, or affixed to words of an inappropriate category. Also, 'compound word' is not superordinate to level 1 and level 2 words. Besides level 2 compounds such as taxi-driver there are level 1 compounds such as Graeco-Roman or electromagnetism, in which, as might be expected, the two elements are more tightly bound to each other than the components of a level 2 compound. Graeco-, aero- etc. do not occur in isolation. Although in some respects they are like level 1 prefixes, they are metrically quasi-independent. For example, the level 1 compound element aero- in aero-mechanic has an unreduced second vowel, in contrast to the reduced second vowel of the level 1 prefix aero- in (British English) aeroplane [εərəploen]. Compare compound electrotherapy with prefix+stem electrolytic for another example.

7.2.4 Metrical structure of English words

One of the challenges presented by stress, especially for the declarative approach, is the way in which it apparently moves about. For instance, adding *-er to photograph /ʃə\w\tə\grəf/ (with stress on the first syllable), results in a pronunciation which is phonemically /fə\tə\grəf/, in which the vowel quality of every syllable is different from in the isolated stem form, and the stress falls on the second syllable. These variations appear as if they might present a challenge to the Principle of Compositional Interpretation, which states that the meaning and form of an utterance is a function of the meaning and form of the parts of that utterance. The word photographer is quite clearly composed of photograph and *-er, so simply listing two separ-
Declarative analysis of English words

ate entries in the lexicon is an abandonment of the Principles of Economy and Compositionality. Such an account would fail to capture generalities of stress effects, and since photography contains the same stem allomorph as photographer, there are, at least, some generalisations to be made regarding the stress effects of -y and -er suffixation. The Principle of Compositionality does not require that the function which combines the parts of an utterance into the whole cannot alter those parts (see e.g. Montague 1970/4: 238, 1973/4: 252, Dowty, Wall and Peters 1981: 179), although that restriction does apply if a monostratal, monotonic grammar formalism is desired. So either the Principle of Compositionality must be abandoned, or monotonicity given up, or a different way of looking at the problem must be found.

A compositional, monotonic analysis of word stress is possible. Ignoring for a moment stress placement, I shall first consider the analysis of the vowel quality alternations. These are due to the different stress placement in the two words: in the analysis of Chomsky and Halle (1968: 28 etc.) taken up in Selkirk (1972/80: 78–179) and Halle and Mohanan (1985: 83), it is claimed that vowels in syllables bearing primary or secondary stress display their ‘unreduced’ vowel quality, whereas vowels in unstressed syllables are ‘reduced’ to /ə/. Photograph bears primary stress on the first syllable, and secondary stress on the last syllable, so these syllables have their ‘unreduced’ vowel qualities /aw/ and /a/, whereas the second syllable is unstressed, so its vowel is reduced to /ə/. Photographer, on the other hand, is stressed only on the second syllable, so its vowel has its ‘unreduced’ quality /əl/, whereas all the other syllables are unstressed, and thus their vowels are all /ə/. This is a neat account so far, but it has two major problems. First, it requires a mechanism to change vowel qualities in a way which is destructive of phonological information (distinctions), which again contradicts the aim of keeping the grammar formalism monostratal and monotonic. Second, it requires the representation of photograph at some point in the derivation to be an amalgamation of the two surface forms, with all of its vowels ‘unreduced’: /fɔwtogəf/. If the units of phonological representations such as this are phonetic objects, as classical generative phonology maintains, then this representation is a phonetic form which is never uttered. This is a strange, if not intrinsically contradictory, basis for phonological representation, and many phonologists have found this abstract standpoint difficult to accept (e.g. Hooper 1976). Its only strength, which has maintained its widespread acceptance, is the fact that a better non-derivational analysis has not yet been proposed, even in Metrical...
Phonology, which has replaced so much of Chomsky and Halle's (1968) account of stress assignment.

Part of the problem with the above analysis of *photograph* is that the phonological representation is also a phonetic representation, in that it is a string of phonetic units. If the view that phonetic and phonological representations are completely distinct is held, as proposed in chapter 5, /ˈfəwtogərəf/ (with unreduced vowels throughout) could be regarded as a purely phonological representation, providing that it is not regarded as a phonetic representation. Such a view would fit in with the aim of developing a monostratal, monotonic grammar, since instead of turning /əw/ into /ə/ and so on by rewriting, the phonological object /əw/ could be interpreted as phonetic [ə] in unstressed syllables. A partial working-out of this idea is explored in Dirksen and Coleman (1997). This style of analysis is quite similar in some ways to classical phonemics and to the two contemporary phonological theories called 'two-level phonology' (Karttunen 1983, Koskenniemi 1983). Not only is stress an abstract category, but so is even phonological vowel quality. In the monosyllable grammar discussed above, for instance, phonologically 'mid' vowels, those between 'close' and 'open' in the vowel systems, are all of category [height: mid], irrespective of their phonetic vowel height: compare the vowels of *debt* and *day*, both of which are categorised as [height: mid], but with slightly different phonetic interpretations arising from the different systems and contexts in which they occur.

Thus the striking vowel quality differences of *photograph* and *photographer* can be accounted for in a declarative fashion, providing that the differences of stress placement are accounted for. The analysis of stress placement developed below is broadly in line with the standard Metrical analysis developed by Liberman and Prince, Kiparsky, Hayes and others (see also 4.4.2 and 4.4.3 above), except in the following respects:

1. Being declarative, restructuring of metrical structure is prohibited.
2. Rather than employing extrametrical syllables, as in Hayes (1982), ternary feet are permitted, as in Lass (1987) and Burzio (1994). Hayes has to employ additional machinery, including metrical-structure adjustment rules, to account for words like *abracadabra* which contain a word-internal ternary foot in the absence of a morpheme boundary. Rather than employing extrametrical consonants, a single consonant in coda position is not regarded as sufficient to make the syllable in which it occurs heavy. Intervocalic
consonants are typically parsed as ambisyllabic, rather than as onsets, so that only a coda cluster or branching nucleus will make a syllable heavy (see 7.1.10 above).

For level 1 words, the general descriptive principles of main stress placement are these (Lass 1987):

1. Only the last three syllables are relevant to primary stress placement.
2. If the last syllable is heavy, it is stressed.
3. Otherwise, if the next-from-last syllable is heavy, it is stressed.
4. Otherwise, the second-from-last syllable is stressed.
5. If the word has only one syllable, it is stressed.
6. If the word has only two syllables, and the last syllable is light, the first syllable is stressed.

A procedural algorithm for stress assignment based on these principles could inspect the last three syllables of a word, check if the last syllable is heavy, if not, inspect the next-from-last syllable, and so on. The alternative approach to be set out below defines a grammar which encodes this information declaratively, in a non-derivational manner, as we shall see.

The first of the above six principles is implemented in part by regarding Latinate words as composed of two metrical parts: a pretonic ‘beginning’, of arbitrary length, and a final ‘main-stress domain’, consisting of the tonic (i.e. most prominent) foot:

(7.143)  
\[ W_j \]

Pretonic  \hspace{1cm}  Tonic foot  
(no more than three syllables)

e.g. *disestablishment*  \hspace{1cm}  -arian

This structure instantiates the word rule (4.17), repeated here for convenience:

(7.144)  
\[ W_j \rightarrow \Sigma_s, \Sigma_w^p \]

where \( W_j \) is the Latinate word-stress domain, \( \Sigma_s \) the tonic foot and \( \Sigma_w^p \) the pretonic. In order to ensure that the pretonic precedes the tonic, parameter \( P1 \) is set to ‘Right’ (4.19): \( \Sigma_w^p < \Sigma_s \). Examples of very long Latinate words such as *antidisestablishmentarian*, with several secondary stress-feet in the pretonic, suggest that parameter \( P10 = No \) (feet are iterative), that is \( p \geq 0 \).

A foot consists of a stressed ‘ictus’ syllable followed by up to two unstressed, light, ‘remiss’ syllables. The main-stress domain is the last foot
of the level 1 part of the word, and secondary stresses fall on the heavy syllable of each foot before the last. English words rarely have enough syllables for more than one or at most two pretonic feet, although there may be one or two stray light syllables at the beginning of the word when the remainder of the word is partitioned into feet (e.g. *parásitical*); see the discussion of parameter P3 in section 4.4.3. These syllables are unstressed, and form a foot without an ictus or a two-syllable remiss in a foot with an empty silent ictus. Cases of two non-final unstressed light syllables such as this provide further evidence in support of the possibility that ternary feet can occur in English other than word-finally through extrametricality. The general structure of level 1 words is thus:

\[ (7.145) \quad (\Sigma_w) \quad \Sigma^* \quad \Sigma_s \]

In (7.145), I use superscript L as shorthand for [weight: light]. Likewise, superscript H is shorthand for [weight: heavy]. The tonic foot has three possible forms: (i) one syllable (cases 1 and 4 below); (ii) two syllables (cases 2 and 5); or (iii) three syllables (case 3). I will account for the five cases using five different structures, defined parametrically. Whether or not the tonic foot is just a part of a word or the whole of the word can be determined by employing a feature which encodes whether or not the tonic is word-initial, such as [±initial] and the pair of rules:

\[ (7.146) \quad W, \rightarrow \Sigma_w^\alpha [\alpha \text{ initial}] \Sigma_w [-\alpha \text{ initial}] \]
\[ (7.147) \quad \Sigma_w^\alpha [-\text{initial}] \rightarrow \emptyset \]

Thus, if the pretonic is empty, it will be marked [−initial], and the tonic foot will consequently be marked [+initial], through alpha-switching in (7.146). Non-empty pretonics must be marked [+initial]. Alternatively, the tonic could be marked [−initial] when \( p = 0 \).

The possible structures of the tonic foot are as follows:

\[ (7.148) \quad \text{Case 1} \]
\[ \Sigma_s \]
\[ [-\text{initial}] \]
\[ \sigma^H \]

STRESSED

This (and case 4) instantiate the foot rule (4.20) \( \Sigma \rightarrow \sigma_s \sigma_w^n \) in the case where \( n = 0 \), with the additional requirement in this case that \( \Sigma \) is [−initial].
Since the single syllable must be heavy to occur in this structure, the Latinate stress system is quantity-sensitive, so parameter P5 = Yes (4.30), i.e. [+heavy] \( \supseteq \sigma_s \). In addition, since this structure and that of case 4 instantiate a strong (i.e. stressed) branch of a foot which does not itself branch, parameter P7 = No; for example, in(ert) (parentheses denote foot boundaries).

(7.149) **Case 2**

\[
\Sigma_s \\
[-\text{initial}] \\
\sigma^H \\
\sigma^L \\
\text{STRESSED}
\]

This and case 5 instantiate the foot rule (4.20) \( \Sigma \rightarrow \sigma_s \sigma_w^0 \) with \( n = 1 \). For example, in(ertia).

(7.150) **Case 3**

\[
\Sigma \\
\sigma \\
\sigma^L \\
\sigma^L \\
\text{STRESSED}
\]

Examples of (7.150) are (a)bo)(riginal) and (a)braca)(dabra).

(7.151) **Case 4**

\[
\Sigma_s \\
 [+\text{initial}] \\
\sigma \\
\text{STRESSED}
\]

Examples: (hit), (héat), (hint). This case is very similar to case 1, except that the tonic foot is marked [+initial], and the single syllable of the word need not be heavy in order to be stressed (e.g. hit). (Burzio (1994) argues that unbranching feet are unnecessary: since they are always word-final, he analyses them as binary feet in which the second syllable is empty. This ingenious proposal has the merit that stress clashes can never arise, because all feet are at least binary.) This is in accordance with the parameter setting P7 = No, which determines the omission of the constraint \( \sigma_s \leftrightarrow [+\text{heavy}] \) from the grammar. In English, therefore, although [+heavy] syllables are strong (i.e. stressed), not all stressed syllables are heavy. A light syllable may be stressed by virtue of its position, for example as the first syllable of a ternary foot, or in this case, as the sole syllable of the word. In the foot rule
\[ \Sigma \rightarrow \sigma_s \sigma_w^n \], the \( \sigma_s \) is obligatory. It is only the \( \sigma_w^n \) part of the rule that can be 'suppressed' by setting \( n = 0 \) in order to parse a foot in a word of one syllable. The same consideration applies in case 5, to ensure that if there are only two syllables in the word, the first of them will be stressed, irrespective of its weight, in accordance with the setting \( P4 = \text{Left} \):

(7.152)  
\[
\begin{array}{c}
\Sigma_s \\
[+\text{initial}] \\
\sigma \\
\sigma_l \\
\text{STRESSED}
\end{array}
\]

Examples of case 5 are (mätter), (mëter). As the five cases 1–5 are mutually orthogonal, it is not necessary for them to be listed in the grammar or applied to the analysis of a word in any particular order, and a context-free phrase-structure grammar will be sufficient to define them, instead of an ordered sequence of transformations such as Chomsky and Halle (1968) propose or a derivational theory of foot construction as in contemporary Metrical Phonology.

In short, the metrical structures defining Latinate stress placement in English are defined by the declarative account of parameter setting presented in section 4.4.3, with parameter settings \( P1 = \text{Left}, P2 = \text{Binary}, P3 = \text{Right}, P4 = \text{Left}, P5 = \text{Yes}, P6 = \text{Rime}, P7 = \text{No}, P8A = \text{Yes}, P8 = \text{Right} \) and \( P10 = \text{Yes} \). (In this section, I employed ternary feet instead of parameters \( P8 \) and \( P8A \), that sanction extrametricality.)

### 7.2.5 Level 2 and level 3 stress

Germanic morphemes, unlike Latinate words, do not group syllables into feet right to left from the end of the word, as in the Latinate stress system (parameter setting \( P1 = \text{Right} \)). Germanic words, according to Lass (1987), bear main stress on the first syllable of the stem. (The Old English origins of this principle are examined in Halle and Keyser 1971; a Lexical–Metrical account is presented in Suphi 1988.) On monosyllabic stems in words without suffixes, such as under-stánd, this will be the last syllable of the word, so that level 2 stress placement coincides with level 1 stress placement in examples such as these, but that is coincidental. The majority of level 2 cases are of this kind, leading mainstream Metrical Phonology to analyse Germanic words with a single set of stress rules, those of level 1. Cases in which the two systems assign stress to the same syllable need to be treated with caution, since they obscure the division between the two systems. As
Declarative analysis of English words

evidence that the Germanic stress systems survive in present-day English, consider the stress of monomorphemic stems such as threshold, Brobdingnag, zeitgeist, cashmere, ashlar, mushroom, phosgene, husband, Asgard, Wesley and zombie in which the first syllable, even if light, is stressed in preference to a later heavy syllable. Single-syllable (Germanic) prefixes such as un-, and be- are unstressed, whereas polysyllabic prefixes such as under- or (Greek) hetero- are stressed according to the normal foot-structure rules presented in the previous section.

A consequence of this analysis, if it is correct, is that the parameter settings for Germanic metrical structures must be different from the Latinate parameter settings discussed above, with the consequence that the language cannot simply be described by a single set of parameter settings. Different settings are appropriate to different levels of lexical structure. The hierarchical organisation of different levels of lexical structure in (7.141) above, however, facilitates the separation of systems that is needed to retain a poly-systemic version of the parametric metrical theory.

Regular level 2 stress, as in threshold, mushroom or husband, appears to apply to many recent loans or coinages, such as phosgene, zeitgeist, cashmere, or zombie, but not to all. The well-known examples of American place-names such as Apalachicola, Winnepesaukee, Topéka, Lüxipalilla and Ökefenökee have level 1 stress, like other polysyllabic proper nouns such as Nebuchadnezzar or Kilimanjaro. Why non-native words should be stressed in two different ways (Asgard vs. Topéka) is a mystery to me, but a way of integrating level 1 and level 2 stress is suggested by the analysis of words with initial ternary stress like Winnepesaukee, Lüxipalilla, Ökefenökee in Hammond (1989) and Halle and Kenstowicz (1991). Hayes (1980) analysed such forms as follows. First, binary feet are assigned from right to left: a) (Win)(nepe)(saukee). Then, a stress clash between the first and second syllables was alleviated by a rule of post-stress destressing which tore down the secondary stress foot: (b) (Win)nepe(saukee). Finally, the second and third syllables, now stray, were adjoined to the preceding syllable, to make an initial ternary foot: (c) (Winnepe)(saukee), the correct outcome. Hayes's analysis excludes ternary feet except in special cases as a matter of principle, leading him to propose (i) an abstract binary footing which is unfaithful to the surface; (ii) a stress clash; (iii) a destructive operation on metrical trees to remediate the clash, as in (b); (iv) unusually, non-initial stray syllables adjoin to the preceding syllable, as in (c), not to the following syllable, as in Topéka. As Burzio (1994) points out, all this machinery is born of Hayes's unwillingness to number ternary feet among the core foot templates.
Hammond proposed that the derivation of forms with initial ternary feet is simpler if secondary stressed feet are built not right-to-left as in Hayes’s analysis, but left-to-right. (The tonic foot is still built at the right-hand edge.) Thus, his analysis proceeds by building the tonic foot *Winnepe(säukee)*, and then non-tonic feet from the beginning: *(Winne)pe(säukee)*. This asymmetry of directionality in primary and secondary foot construction is very similar to the asymmetry in headedness of level 1 and level 2 words: level 1 words are right-headed and level 2 words are left-headed. Under Hammond’s analysis, the pretonic of level 1 words is also left-headed: perhaps here we can find a resolution of the two systems. Specifically, suppose we identify the level 1 pretonic with the level 2 stress domain L2:

(7.153)

```
W
  / \   /
L2_w /   \_1
|   |   /
\   /  \_2
| /   |
Σ_s Σ_w
```

e.g. Winnipe  saukee
  anti  dise  stablishment

In level 2 stems such as *husband*, it is natural to suppose that the Σ₁ constituent is empty, in which case the initial stress is the main stress. When Σ₁ is non-empty, it will be the main stress. Stress-neutral prefixes and suffixes are outside the scope of (7.153).

A general account of compound stress is elusive, although a number of subprinciples have been widely reported, such as the fact that stress falls on the first word in noun compounds such as *blackbird*. The prominence of the constituents of compounds depends also on semantic factors, such as whether the second member of the compound is the marked or unmarked (default) member of a semantic field. Compare (unmarked) *Ôxford Strèet* with *Ôxford Rôad* (marked) (Jespersen 1909/54: 154–6, Sproat 1990).

The existence of cases in which Germanic stress happens to coincide with Latinate stress (as in e.g. *understand*) is one of the factors which may have obstructed the development of a successful principled account of English stress for many years. In those instances where Germanic stress assignment and Latinate stress assignment would place stress in different places, it is
necessary to know which case applies. Morphological constituent structure such as (7.141) above helps to provide this information, though morphology does not resolve the different stress placement in *Brobdingnag* and *Topeka*. The reconciliation between morphological and metrical structure proposed in the next subsection allows us to overcome Jespersen's (1909/54: 160) claim that the independent principles of stress assignment sometimes work against each other. On the contrary, with stress systems factored out onto independent levels of structure, each principle holds true in its own domain.

7.2.6 *Reconciling morphological and metrical structure*

I have outlined declarative accounts of two different aspects of English word structure: lexical structure and metrical structure. These two accounts are complementary: compound words are the domain of compound stress placement, Germanic morphological structures are the domain of Germanic stress placement, and Latinate morphological structures are the domain of Latinate stress placement (modulo the oddities of loan-words and proper nouns discussed above):

(7.154)

Although the metrical structures of compound stress and Germanic stress project straightforwardly onto their corresponding morphological domains, the same is not the case for Latinate stress, which once again is more tricky. For example, the morphological constituents of *photographic*
are the level 1 stem *photograph* and the level 1 suffix *-ic*, whereas its metrical constituents are pretonic *photo-* and tonic foot *-graphic*. The two structures are not congruent. The declarative analysis I propose here uses two separate sets of context-free rules to describe the metrical and morphological structure of Latinate morphemes. The intersection of two context-free languages is not in general context-free, but is context-sensitive, although it employs only context-free rules. In the current analysis, I see two ways out of the problem. First, by keeping any recursion in the morphological grammar, and ensuring that the language defined by the phonological grammar is strictly regular, even though it may use context-free rules for their structural appeal, the intersection can be kept context free. This proposal seems feasible, in that the strongest case that can be made for recursive phonological structure (Ladd 1986) involves structures that are built on recursive syntactic structure, which nobody seriously disputes. Alternatively, J. Calder and E. J. Briscoe (personal communications) have suggested to me that structural mismatches such as these may perhaps be better analysed using a categorial grammar, rather than using intersections of context-free phrase-structure grammars as I have described here, as categorial grammars typically have multiple parse trees. For a similar proposal, see Steedman (1991). The integration of metrical and morphological structure which results from treating the strata of Lexical Phonology as levels of representation is simpler and more restrictive than the cyclic metrical analysis of Kiparsky (1979), in that metrical tree transformations are not permitted.

There are a number of affixes with effects on stress placement which add to the principles described above. The analysis of these affixes is by no means yet complete, yet there are a number of fruitful means within a declarative analysis for dealing with these. For example, *-ic* must be lexically marked, to ensure that either it occurs as the first syllable of a foot (as in *electricity*), or as the second syllable of a foot. A consequence of this second possibility is that the syllable which precedes *-ic* must be stressed, since it will be the strong syllable. If third syllables of ternary feet are marked in some way (e.g. [+em]), the fact that *-ic* is a 'strong attractor' will follow from its being marked [-em].
Declarative analysis of English words

l-velarisation, which are covered by the analysis in this chapter, are treated as lexical rules in Halle and Mohanan (1985). Other lexical rules resting principally on parsing considerations are prenasal g-deletion (e.g. sig[n]ature ~ sign), n-deletion (e.g. hym[n]al ~ hymn), l-resyllabification (e.g. seal ~ seal o ffice) and non-coronal deletion (e.g. bon[b]ard ~ bomb). Parsing and alternation with zero are unproblematic within Declarative Phonology, as are spreading rules such as s-voicing (e.g. Malthu[s] ~ Malthu[z]ian) and the English rules of spirantisation (e.g. react ~ reac[j]ion) and palatalisation (e.g. presiden[t] ~ presiden[f]ial). In Coleman (1995) I give detailed declarative analyses of CiV lengthening (e.g. custody ~ custo: dial), s-voicing (e.g. Malthu[s] ~ Malthu[z]ian), trisyllabic shortening (e.g. divi:ne ~ divinity), cluster shortening (e.g. cruci[f]a[y] ~ cruci[f]i:xion), -ic shortening (e.g. co:ne ~ conic) and velar softening. Any approach to generative phonology needs to address this array of lexical morphophonological phenomena, though it will be apparent from the above analysis that the idiosyncrasy of many of these phenomena is such that they are not central to Declarative Phonology, which focusses primarily on phonotactic regularities. Two other well-known phenomena of English phonology, flapping of /l/ and /d/ in some North American English varieties and linking and intrusive /r/, are critically discussed by Coleman (1995) and Bird (1995).

Many of the rules regarded as phonological in Lexical Phonology are considered to be rules of phonetic interpretation in Declarative Phonology. CiV lengthening, s-voicing, shortening rules, vowel reduction, vowel tensing and stem-final lengthening are candidates to be rules of phonetic interpretation of metrical structure. Nasal assimilation, æ-tensing and l-velarisation are candidates to be rules of phonetic interpretation of syllable structure. In Lexical Phonology, all phonetic rules are postlexical, though there is no reason in principle why this should be so. In any case, Lexical Phonology is largely silent on the question of phonetic interpretation. The term ‘postlexical’ embraces two quite distinct aspects: phonetic interpretation, and supralexical phonological structure, such as prosodic phrasing. These are recognised as quite separate in Declarative Phonology.

7.4 Envoi

The analyses in this chapter and the last illustrate the character of declarative phonological analyses of non-trivial language fragments. A number of characteristics of the approach are evident. First, the analyses focus on phonotactic regularities of surface structure, rather than phonological
alternations. Alternations are considered primarily as diagnostic of natural classes to be accounted for in the selection of phonological features and the postulation of aspects of structure that are shared by alternants, rather than as indicative of a single underlying lexical form. Second, regular morphophonological relations are given priority over unproductive morphophonology and also purely allophonic relations. Third, prosodic structure is central, whereas autosegmental structure is not employed. Whereas earlier chapters argued for the desirability of Declarative Phonology, these chapters demonstrate its feasibility. They are, I hope, a useful illustration of how phonological analysis may be conducted without recourse to rewriting rules or loss of generality.
References

(1976) Nasal consonants and the internal structure of segments. Language 52.2, 326–44.
References


(1991) Feature structures and indices. Phonology 8, 137–44.


References


330 References


Fourakis, M. S. (1980) *A Phonetic Study of Sonorant Fricative Clusters in Two
References


References

References


References


McCarthy, J. (1979) *Formal Problems in Semitic Phonology and Morphology*. Ph.D.
References

dissertation. MIT. Published by Indiana University Linguistics Club, Bloomington, in 1982.


(1981) *Everything that Linguists have always Wanted to Know about Logic (but were Afraid to Ask).* Oxford: Blackwell.


References


References


References


References


Whitley, E. (mss.) Lecture handouts and course notes. Firthian Phonology Archive, Phonetics Laboratory, Department of Language and Linguistic Science, University of York.


Index

affrication 53, 73–4, 114–18
allophony 53, 177–8
Alphabet, Organic 22, 28, 38–9
    Roman 20
    Shaw 39
Arabic 69
articulation, secondary 52, 70, 199, 251, 255, 290
Articulatory Phonology 23–4
assimilation 43, 66–7, 178–9, 195–6, 205, 211
automata 48, 79–80
Berber, Tamazight 69
boundaries 79
Chamorro 82
charts 131
Chinese 249
Chomsky hierarchy 79, 121
cocatenation 27–9, 60
column-vector: see feature-matrix
Compositionality, Principle of 29, 175, 313–14
concatenation 27
    vertical: see cocatenation
Constraint, No Crossing 90, 128–46, 306–7
    Structure Preservation 91–2
    True Generalisation 90
constraints, feature-sharing 209
    imperative interpretation of 9
    naturalness 89–91
notational 23
pansyllabic 304–7
semantic 27
sequence 209
structural 91
temporal 221–4
on transformations 89–92
well-formedness 119–20, 122–4, 127–8
Declarative Phonology 7–10, 151–9, 165–325
declarativeness 5, 93, 96
deletion 80–2, 91, 114, 121, 251–2
    declarative alternatives to 181–3
    recoverability of 71–2
denotation 22, 25, 32–4, 37–8, 50, 89
Dependency Phonology 5–6, 23, 146–9, 225
derivations 92–105
devoicing 52, 70, 102, 215, 251–2
diagrams 37, 41–4
dissimilation 68, 173–4, 189, 277, 284, 299
Dutch 293, 303
elision 43, 66, 71, 116, 181–2, 227–9
Elsewhere Principle 191–3
English, Guyanese 143–5
epenthesis 43, 66, 180–3, 302
Fang 304
feature-matrix 28, 39–41, 60, 105
feature-structures 184–5, 189–93, 197–205, 307
features, binary 23
distinctive 22, 58–65
predictable 58, 63, 190
prosodic 207–9
unary 23
formalism 12–14
Gaelic 283, 297
German 283, 303
Government-based Phonology 24, 295
grammars, context-free 8–9, 81, 94–6, 210–11
    context-sensitive 79, 81, 83, 94
    generative 78–87
    graph 121
    ID/LP 87, 126–7, 153–9, 195, 218–20, 239
    monotonic 79, 81
    non-transformational 84
    transformational 84–6
    unrestricted 79, 81, 83, 161
graphs, derivation 97–105
Greek, Classical 310–11
interpretation, categorial 31–3
parametric 229–32
phonetic 168–77, 216–18, 220–4, 258–65
superweak segmental 43–5
temporal 33, 36–8, 70, 220–4
Irish 199
Italian 297
Japanese 4, 50–4, 68, 70–6, 102, 162–3, 233–68
Latin 155, 310
Lexical Phonology 212, 219–20, 310–3, 323–4
logic 186–91
temporal 30
Mayan, Aguacatec 159
metathesis 43, 66–7, 83, 183–4
Metrical Phonology 4, 149–51, 160–4
minimal-pairs test 55
model theory: see semantics, model-theoretic
notation, componential 22
segmental 21
Old English 283, 293, 302, 319
Old French 304
Old Norse 293
oppositions 55–8, 63
palatalisation 53, 74–6, 114–18, 252–4
parsing 87
phonemic phonology 47–65
phrase-marker: see graphs, derivation
Polish 82
Portuguese 302
‘Principles and Parameters’ 4, 151–60, 164, 166
recursiveness 85–7
representations, cognitive 11
resyllabification 195
rule application, cyclic 82, 89, 311
rules, autosegmental 114, 120–2
context-free 125–7, 217
context-sensitive 66–7, 78–9, 151
deletion: see deletion
delinking 114
phonological: see transformations
transformational: see transformations
types 78–83, 92
see also grammars
Russian 68, 199, 215
Sanskrit 112
Scots 283, 301
segmentation 36–7, 41–2
segments 21
semantics 25, 29–38, 64–5
model-theoretic 29–31
Sinhalese 68
Spanish 7–10, 302
SPE phonology 2, 7
stress, English 5–6, 94–7, 149–151, 155, 158, 308–24
subsumption 191–3
Swahili 304
syllabification 7–10, 95–7, 308
syllables 85, 94–5, 147–8, 234–307
symmetry 54
syntax 25–9, 64–5
theorem proving 9–10
theorem, Kuratowski’s 137
transcription 24–6
transformational-generative phonology: see
also SPE phonology 2, 66–77
transformations 2, 83, 85, 89
Turing Machine 79–81
Turkic languages 69
unification 1, 40, 107, 108, 125, 179, 189
variability, phonetic 28, 47, 49–54
vowel harmony 69
Warao 154
Welsh 188, 295