This book presents the theory of output-driven maps and provides a fresh perspective on the extent to which phonologies can be characterized in terms of restrictions on outputs. Closely related to traditional conceptions of process opacity, but differing in notable ways, the theory of output-driven maps applies equally to SPE-style ordered rules, Optimality Theory, and other phonological theories. It permits a formally rigorous analysis of the issues in Optimality Theory that is not possible with traditional process opacity.

Also presented is a theory of phonological learning. Building on prior work on learning in Optimality Theory, the learning theory exploits the formal structure of output-driven maps to achieve learning that is far more computationally efficient than comparable prior approaches.

In this book Bruce Tesar, one of the founders of the study of learnability in Optimality Theory, presents fresh perspectives in an accessible way for graduate students and academic researchers.

Bruce Tesar is Associate Professor in the Department of Linguistics and the Center for Cognitive Science at Rutgers University, New Brunswick.
In this series

106 Sharon Inkelas and Cheryl Zoll: Reduplication: doubling in morphology
107 Susan Edwards: Fluent aphasia
108 Barbara Dancygier and Eve Sweetser: Mental spaces in grammar: conditional constructions
110 Marcus Tomalin: Linguistics and the formal sciences: the origins of generative grammar
111 Samuel D. Epstein and T. Daniel Seely: Derivations in minimalism
112 Paul de Lacy: Markedness: reduction and preservation in phonology
113 Yehuda N. Falk: Subjects and their properties
114 P. H. Matthews: Syntactic relations: a critical survey
115 Mark C. Baker: The syntax of agreement and concord
116 Gillian Catriona Ramchand: Verb meaning and the lexicon: a first phase syntax
117 Pieter Muysken: Functional categories
118 Juan Uriagereka: Syntactic anchors: on semantic structuring
119 D. Robert Ladd: Intonational phonology second edition
120 Leonard H. Babby: The syntax of argument structure
121 B. Elan Dresher: The contrastive hierarchy in phonology
122 David Adger, Daniel Harbourn and Laurel J. Watkins: Mirrors and microparameters: phrase structure beyond free word order
123 Niina Ning Zhang: Coordination in syntax
124 Neil Smith: Acquiring phonology
125 Nina Topintzi: Onsets: suprasegmental and prosodic behaviour
126 Cedric Boeckx, Norbert Hornstein and Jairo Nunes: Control as movement
127 Michael Israel: The grammar of polarity: pragmatics, sensitivity, and the logic of scales
128 M. Rita Manzini and Leonardo M. Savoia: Grammatical categories: variation in romance languages
129 Barbara Citko: Symmetry in syntax: merge, move and labels
130 Rachel Walker: Vowel patterns in language
131 Mary Dalrymple and Irina Nikolaeva: Objects and information structure
132 Jerrold M. Sadowski: The modular architecture of grammar
133 Dunstan Brown and Andrew Hippines: Network morphology: a defaults-based theory of word structure
134 Bettelou Los, Corrien Blom, Geert Boooj, Marion Elenbaas and Ans van Kemenade: Morphosyntactic change: a comparative study of particles and prefixes
135 Stephen Crain: The Emergence of Meaning
136 Hubert Haider: Symmetry Breaking in Syntax
137 José A. Camacho: Null Subjects
138 Gregory Stump and Raphael A. Finkel: Morphological Typology: From Word to Paradigm
139 Bruce Tesar: Output-Driven Phonology: Theory and Learning

Earlier issues not listed are also available
OUTPUT-DRIVEN PHONOLOGY
THEORY AND LEARNING

BRUCE TESAR
Rutgers University, New Brunswick.
CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Published in the United States of America by Cambridge University Press, New York

Cambridge University Press is part of the University of Cambridge.
It furthers the University’s mission by disseminating knowledge in the pursuit of
education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107001930

© Bruce Tesar 2014

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2014

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

Tesar, Bruce.
Output-driven phonology : theory and learning / Bruce Tesar, Rutgers University, New Brunswick.
  pages cm. – (Cambridge studies in linguistics)
Includes bibliographical references and index.
ISBN 978-1-107-00193-0 (hardback)
1. Grammar, Comparative and general – Phonology.  2. Optimality theory
   (Linguistics)  3. Language acquisition. I. Title.
P158.42.T49  2013
414 – dc23  2013020376

ISBN 978-1-107-00193-0 Hardback

Cambridge University Press has no responsibility for the persistence or accuracy of
URLs for external or third-party internet websites referred to in this publication,
and does not guarantee that any content on such websites is, or will remain,
accurate or appropriate.

Every effort has been made to secure the necessary permissions to reproduce copyright
material in this work, though in some cases it has proved impossible to trace or contact
copyright holders. If any omissions are brought to our notice, we will be happy to
include appropriate acknowledgments on reprinting, and/or in any subsequent edition.

© in this web service Cambridge University Press

www.cambridge.org
This book is dedicated to Heidi and Amanda, for all they have done to inspire my work and my life.

“Who do we ask for help when we don’t know which way to go? The map!”

Dora the Explorer
# Contents

<table>
<thead>
<tr>
<th>List of figures</th>
<th>page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acknowledgments</td>
<td>xix</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1</th>
<th>Characterizing surface orientedness in phonology</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Surface orientedness</td>
<td>1</td>
</tr>
<tr>
<td>1.1.1</td>
<td>Surface orientedness in phonology</td>
<td>1</td>
</tr>
<tr>
<td>1.1.2</td>
<td>Formalizing surface orientedness</td>
<td>3</td>
</tr>
<tr>
<td>1.2</td>
<td>Surface orientedness in Optimality Theory</td>
<td>6</td>
</tr>
<tr>
<td>1.2.1</td>
<td>Markedness violations as opacity</td>
<td>7</td>
</tr>
<tr>
<td>1.2.2</td>
<td>Markedness constraints cause disparities</td>
<td>10</td>
</tr>
<tr>
<td>1.2.3</td>
<td>Something more general</td>
<td>11</td>
</tr>
<tr>
<td>1.3</td>
<td>Formalizing surface orientedness: output-driven maps</td>
<td>13</td>
</tr>
<tr>
<td>1.4</td>
<td>Output drivenness and Optimality Theory</td>
<td>15</td>
</tr>
<tr>
<td>1.5</td>
<td>Output drivenness and learning</td>
<td>16</td>
</tr>
<tr>
<td>1.6</td>
<td>The relationship between learnability and linguistic theory</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2</th>
<th>Output-driven maps</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The main idea</td>
<td>20</td>
</tr>
<tr>
<td>2.1.1</td>
<td>Terminology: candidates and correspondence</td>
<td>20</td>
</tr>
<tr>
<td>2.1.2</td>
<td>Inputs of greater similarity yield the same output</td>
<td>21</td>
</tr>
<tr>
<td>2.1.3</td>
<td>Unifying surface orientedness</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Relative similarity</td>
<td>26</td>
</tr>
<tr>
<td>2.2.1</td>
<td>Relating the disparities of two candidates</td>
<td>26</td>
</tr>
<tr>
<td>2.2.2</td>
<td>Individuating disparities</td>
<td>28</td>
</tr>
<tr>
<td>2.2.3</td>
<td>Relative similarity is a relational notion</td>
<td>30</td>
</tr>
<tr>
<td>2.2.4</td>
<td>The importance of input–output correspondence</td>
<td>33</td>
</tr>
<tr>
<td>2.3</td>
<td>Output drivenness is not process opacity</td>
<td>37</td>
</tr>
<tr>
<td>2.3.1</td>
<td>One map, multiple generalizations</td>
<td>37</td>
</tr>
<tr>
<td>2.3.2</td>
<td>Epenthesis and assimilation in Lithuanian</td>
<td>39</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Closeness with processes</td>
<td>42</td>
</tr>
<tr>
<td>2.4</td>
<td>Formal analysis with segmental IO correspondence</td>
<td>44</td>
</tr>
</tbody>
</table>
2.4.1 Maps from inputs to candidates 44
2.4.2 The internal structure of candidates 46
2.4.3 Relating candidates to each other 48
2.4.4 The non-uniqueness of input–input correspondence 53
2.4.5 Removing disparities by changing the input 54
2.4.6 The identical disparity requirement and surface orientedness 55
2.4.7 Individuating disparities (again) 57

2.5 Expanding to other representational theories 59
2.5.1 Non-identical corresponding representational elements 59
2.5.2 Non-unique correspondence 60
2.5.3 Autosegmental representation 60

2.6 The map 61

3 Output-driven maps in Optimality Theory 63
3.1 Background: ERC entailment in Optimality Theory 64
3.1.1 Elementary ranking conditions 64
3.1.2 Single ERC entailment: L-retraction and W-extension 66
3.1.3 Joint ERC entailment: fusion 67
3.2 Relating output-driven maps to Optimality Theory 69
3.2.1 Output-driven maps and optimization 69
3.2.2 A designated competitor: aoy 70
3.2.3 Relationships among the disparities 74
3.2.4 As goes bpy, so goes aoy 78
3.2.5 Output-driven maps and constraints 81

3.3 Sufficient conditions for output-driven maps 83
3.3.1 Properties of Gen: correspondence uniformity 84
3.3.2 Properties of constraints: output-driven preserving 86
3.3.3 Proof of sufficient conditions for output-driven maps 87

3.4 Basic constraints: overview of the results 88
3.4.1 Terminology: faithfulness and input-referring constraints 88
3.4.2 Markedness constraints 89
3.4.3 Value-independent input-referring constraints 89
3.4.4 Value-restricted input-referring constraints 90

3.5 Analysis of relationships between disparities 91
3.5.1 The set-up 91
3.5.2 Deletion disparities 93
3.5.3 Insertion disparities 94
3.5.4 Identity disparities 96
3.5.5 Comments/discussion 101

3.6 Output-driven preserving constraints: the proofs 101
3.6.1 Outline of the proof structure 101
3.6.2 Max 102
3.6.2.1 Partition of the deletion disparities 102
<table>
<thead>
<tr>
<th>Contents</th>
<th>xi</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.6.2.2</td>
<td>Corresponding deletion disparities for (aoy) and (bpy) 103</td>
</tr>
<tr>
<td>3.6.2.3</td>
<td>Corresponding deletion disparities for (akx) and (bmx) 103</td>
</tr>
<tr>
<td>3.6.2.4</td>
<td>Non-corresponding deletion disparities for (aoy) and (akx) 104</td>
</tr>
<tr>
<td>3.6.2.5</td>
<td>(\max(bpy) &lt; \max(bmx)) entails (\max(aoy) &lt; \max(akx)) 104</td>
</tr>
<tr>
<td>3.6.2.6</td>
<td>(\max(bpy) = \max(bmx)) entails (\max(aoy) \leq \max(akx)) 105</td>
</tr>
<tr>
<td>3.6.3</td>
<td>(\text{Dep}) 105</td>
</tr>
<tr>
<td>3.6.3.1</td>
<td>Partition of the insertion disparities 105</td>
</tr>
<tr>
<td>3.6.3.2</td>
<td>Corresponding insertion disparities for (aoy) and (bpy) 106</td>
</tr>
<tr>
<td>3.6.3.3</td>
<td>Corresponding insertion disparities for (akx) and (bmx) 106</td>
</tr>
<tr>
<td>3.6.3.4</td>
<td>Non-corresponding insertion disparities for (aoy) and (akx) 106</td>
</tr>
<tr>
<td>3.6.3.5</td>
<td>Proof: (\text{Dep}(bpy) &lt; \text{Dep}(bmx)) entails (\text{Dep}(aoy) &lt; \text{Dep}(akx)) 107</td>
</tr>
<tr>
<td>3.6.3.6</td>
<td>Proof: (\text{Dep}(bpy) = \text{Dep}(bmx)) entails (\text{Dep}(aoy) \leq \text{Dep}(akx)) 107</td>
</tr>
<tr>
<td>3.6.4</td>
<td>(\text{Ident}[F_{in} \in V]) 108</td>
</tr>
<tr>
<td>3.6.4.1</td>
<td>Partition of the identity disparities 108</td>
</tr>
<tr>
<td>3.6.4.2</td>
<td>Identical corresponding identity disparities for (aoy) and (bpy) 109</td>
</tr>
<tr>
<td>3.6.4.3</td>
<td>Corresponding identity disparities for (akx) and (bmx) 110</td>
</tr>
<tr>
<td>3.6.4.4</td>
<td>Non-Corresponding and non-identical corresponding identity disparities for (aoy) and (akx) 110</td>
</tr>
<tr>
<td>3.6.4.5</td>
<td>Proof: (\text{Ident}<a href="bpy">F_{in} \in V</a> &lt; \text{Ident}<a href="bmx">F_{in} \in V</a>) entails (\text{Ident}<a href="aoy">F_{in} \in V</a> &lt; \text{Ident}<a href="akx">F_{in} \in V</a>) 110</td>
</tr>
<tr>
<td>3.6.4.6</td>
<td>Proof: (\text{Ident}<a href="bpy">F_{in} \in V</a> = \text{Ident}<a href="bmx">F_{in} \in V</a>) entails (\text{Ident}<a href="aoy">F_{in} \in V</a> \leq \text{Ident}<a href="akx">F_{in} \in V</a>) 111</td>
</tr>
<tr>
<td>3.6.5</td>
<td>(\text{Ident}[F_{out} \in V]) 112</td>
</tr>
<tr>
<td>3.6.5.1</td>
<td>Partition of the identity disparities 113</td>
</tr>
<tr>
<td>3.6.5.2</td>
<td>Corresponding identity disparities for (aoy) and (bpy) 113</td>
</tr>
<tr>
<td>3.6.5.3</td>
<td>Corresponding identity disparities for (akx) and (bmx) 114</td>
</tr>
<tr>
<td>3.6.5.4</td>
<td>Non-corresponding identity disparities for (aoy) and (akx) 114</td>
</tr>
<tr>
<td>3.6.5.5</td>
<td>Proof: (\text{Ident}<a href="bpy">F_{out} \in V</a> &lt; \text{Ident}<a href="bmx">F_{out} \in V</a>) entails (\text{Ident}<a href="aoy">F_{out} \in V</a> &lt; \text{Ident}<a href="akx">F_{out} \in V</a>) 114</td>
</tr>
<tr>
<td>3.6.5.6</td>
<td>Proof: (\text{Ident}<a href="bpy">F_{out} \in V</a> = \text{Ident}<a href="bmx">F_{out} \in V</a>) entails (\text{Ident}<a href="aoy">F_{out} \in V</a> \leq \text{Ident}<a href="akx">F_{out} \in V</a>) 115</td>
</tr>
</tbody>
</table>

3.7 The map 116
4 Analysis of constraint behavior 117
4.1 Non-ODP constraints and non-output-driven maps 117
4.2 Illustrating the three ways of non-ODP 119
4.2.1 Distinction only at lesser similarity 119
4.2.2 Distinction only at greater similarity 122
4.2.3 Distinction conflict 125
4.3 Relating non-ODP constraint behaviors to non-ODM map patterns 127
4.4 Faithfulness conditioned on output context 129
4.4.1 Positional faithfulness I: position-specific Dep 129
4.4.2 Positional faithfulness II: position-specific Ident[F] 132
4.4.3 Conjoined markedness and faithfulness 134
4.4.4 Value-restricted Dep 137
4.4.5 Summary: independent context 138
4.5 Faithfulness conditioned on input context 139
4.5.1 Value-restricted Max 139
4.5.2 Relation to output-conditioned faithfulness 142
4.6 Multiply conditioned faithfulness 143
4.6.1 Joint input–output value restrictions 143
4.6.2 Conditioning on disparities 144
4.7 Conditioned antifaithfulness 145
4.7.1 Material implication constraints 145
4.7.2 F → M as output-conditioned antifaithfulness 146
4.8 Reference to other forms: sympathy 148
4.9 Eventual idempotency 150
4.9.1 Classical OT grammars are eventually idempotent 151
4.9.2 Absolute vs. relative satisfaction of constraints 154
4.9.3 Derived environment exchanges 155
4.9.4 Absolute vs. relative characterizations of similarity 159
4.10 The role of relative similarity 161
4.10.1 Distinguishing faithfulness from faithfulness constraints 161
4.10.2 The rotation theorem isn’t about faithfulness 162
4.10.3 Relative similarity links grammars to maps 166
4.11 More on relating non-ODP behaviors to non-ODM patterns 167
4.12 Maps and grammars 170
4.13 The map 171

5 Learning phonotactics 173
5.1 An overview of the learning problem 174
5.2 Stress and length: a system for illustration 175
5.3 Constructing a constraint hierarchy from winner–loser pairs 177
5.3.1 Recursive Constraint Demotion 177
5.3.2 Stratified constraint hierarchies 181
5.3.3 Constraint conflict and ranking relations 184
5.4 Selecting winner–loser pairs 185
  5.4.1 Error detection 186
  5.4.2 Production-directed parsing with stratified hierarchies 187
    5.4.2.1 Mark pooling 188
    5.4.2.2 Conflicts Tie 191
    5.4.2.3 Variationist EDCD 194
  5.4.3 MultiRecursive Constraint Demotion (MRCD) 195
  5.4.4 MRCD step by step 197
    5.4.4.1 First pair 198
    5.4.4.2 Second pair 198
    5.4.4.3 Third pair 198
    5.4.4.4 Fourth pair 199
    5.4.4.5 Fifth pair 199
    5.4.4.6 Last pair 200
  5.4.5 Limitations of loser production via stratified hierarchies 200
  5.5 Assessing computational requirements 203
    5.5.1 The computational complexity of MRCD 203
    5.5.2 Grammar space vs. language space 205
  5.6 Restrictiveness biases 207
    5.6.1 Restrictiveness in learning 207
    5.6.2 Phonotactic learning 209
    5.6.3 Language subsets vs. grammar subsets 211
    5.6.4 Biased Constraint Demotion 213
      5.6.4.1 RCD and restrictiveness 213
      5.6.4.2 Estimating the restrictiveness of grammars: the r-measure 215
      5.6.4.3 A restrictiveness bias for RCD 216
    5.6.5 Enforcing restrictiveness in phonotactic learning 220
      5.6.5.1 Enforcing restrictiveness with BCD 220
      5.6.5.2 Restrictiveness with “hierarchy only” learning 221
    5.6.6 Implicit representation of phonotactic restrictions 222
  5.7 Phonotactic contrast 224
    5.7.1 Contrast and the nature of phonotactic learning 224
    5.7.2 A canonical form for phonotactic contrast 226
  5.8 Phonotactic information underdetermines languages 233
  5.9 The map 236
  5.10 Appendix: the Stress/Length typology 237

6 Learning with paradigmatic information 246
  6.1 Paradigmatic information 247
  6.2 The explosive growth of lexical hypothesis spaces 248
    6.2.1 Now that’s big 248
    6.2.2 The basic alternant constraint 248
### Contents

6.2.3 Selected prior work 250  
6.2.4 Combinatorics of the Stress/Length linguistic system 252  
6.3 An aside on methodology 254  
6.4 Inconsistency detection 256  
6.5 Setting underlying features via inconsistency detection 258  
6.5.1 Feature setting 259  
6.5.2 Setting a single unset feature 260  
6.5.3 Multiple unset features 263  
6.5.4 Multiple words and local lexica 266  
6.6 Non-phonotactic ranking information 269  
6.6.1 Ranking information in local lexica 269  
6.6.2 The join operation 270  
6.6.3 Extracting shared ranking information 272  
6.7 The Contrast Pair and Ranking (CPR) algorithm 274  
6.8 Computational issues for CPR 276  
6.9 The map 277  

7 Exploiting output drivenness in learning 279  
7.1 Contrast with Richness of the Base 280  
7.1.1 Contrastive for an input 280  
7.1.2 Contrastive for a morpheme 282  
7.1.3 Contrast in output-driven maps 283  
7.2 Relative similarity lattices 284  
7.3 Limiting lexical search in output-driven maps 287  
7.4 Phonotactic contrast and underlying feature values 291  
7.5 Morphemic alternation and non-phonotactic ranking information 294  
7.6 Contrast pairs 299  
7.6.1 When one word isn’t enough 299  
7.6.2 A disjunction of disparities 302  
7.6.3 Restricted rankings create inconsistencies 305  
7.6.4 The roles of alternation and contrast in contrast pairs 308  
7.6.5 Multiple words and relative similarity 310  
7.6.6 Another illustration: setting the stress feature for r1 316  
7.6.7 Setting environment morpheme features 317  
7.7 Beyond error-driven learning 321  
7.7.1 Uncertainty in the lexicon 323  
7.7.2 Uncertainty about the ranking 328  
7.7.3 Single form learning 332  
7.7.4 Contrast pair learning 334  
7.8 The Output-Driven Learner (preliminary) 335  
7.9 Learning language L20 336  
7.9.1 Phonotactic learning 336  
7.9.2 Initial single form learning 338
Contents

7.9.2.1 r1s1 338
7.9.2.2 r1s4 340
7.9.2.3 r2s1, r4s1 343
7.9.2.4 r3s1, r1s3 344
7.9.3 Contrast pair learning 345
7.9.4 Second single form learning 348
  7.9.4.1 r3s3 348
  7.9.4.2 r4s3, r1s3, r2s3 349
  7.9.4.3 r1s2, r2s4 350
  7.9.4.4 The final learned grammar for L20 351
7.10 The map 352

8 Paradigmatic subsets 353
8.1 The phenomenon: paradigmatic subsets 353
  8.1.1 Language L8, the subset language 353
  8.1.2 Language L7, the superset language 355
  8.1.3 L8 is a paradigmatic subset of L7 355
8.2 The problem: attempting to learn L8 358
8.3 Paradigmatic restrictiveness and the lexicon 360
8.4 The solution: Fewest Set Features 363
  8.4.1 The Fewest Set Features procedure 363
  8.4.2 Algorithmic details of Fewest Set Features 370
  8.4.3 The relation to maximum likelihood 372
  8.4.4 Summary: restrictiveness in the lexicon 376
8.5 Evaluating the Output-Driven Learner 376
  8.5.1 The Output-Driven Learner (revised) 376
  8.5.2 Simulation results 377
     8.5.2.1 No contrast languages 378
     8.5.2.2 Single form learning, no contrast pairs 378
     8.5.2.3 Contrast pairs 379
     8.5.2.4 Fewest Set Features 380
     8.5.2.5 A little perspective 381
  8.5.3 Issues 381
8.6 The map 383

9 Linguistic theory and language learnability 385
9.1 Contrast and the final lexicon 385
9.2 Forms of restrictiveness enforcement 390
9.3 Evaluation metrics 392
9.4 What is stored by the learner (and why) 394
  9.4.1 Ranking information 394
  9.4.2 Lexical information 395
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.4.3 Structural ambiguity and multiple hypotheses</td>
<td>396</td>
</tr>
<tr>
<td>9.5 Beyond output drivenness in learning</td>
<td>398</td>
</tr>
<tr>
<td>9.6 What has been accomplished</td>
<td>399</td>
</tr>
</tbody>
</table>

### References

401

### Index

410
Figures

2.1 Relative similarity relation (upward is greater internal similarity). page 27

2.2 Candidates \(akx\) and \(bmx\), and their input–input correspondence. 51

3.1 Candidates \(akx\), \(bmx\), \(aoy\), and \(bpy\). 73

3.2 Relationships among the disparities. Corresponding disparities (same output) are connected with double-sided arrows; analogous disparities (same input) are connected by the square-angled lines. 75

3.3 Candidates \(akx\), \(bmx\), \(aoy\), and \(bpy\). 80

3.4 Relationships among the disparities when \(out_y = [tebig]\). 81

7.1 Relative similarity lattice for output \([paká]\) (higher in the diagram means greater internal similarity). 285

7.2 Relative similarity lattice for output \(paká\): (feature matrix version). 286

7.3 Setting \(s4\) to +long. This is the relative similarity lattice for the output of \(r1s4\), \([paká]\). The shaded sublattice contains all candidates with \(s4\) underlyingly –long. 289

7.4 The relative similarity lattice for \(r3s4\) (nonviable candidates, with \(s4\) underlyingly –long, are marked with shaded diamonds). 296

7.5 The relative similarity lattice for \(r1s1\). The shaded diamond nodes have \(r1\) underlyingly +long, counter to what has already been set in the lexicon, and so are not viable candidates. 311

7.6 The relative similarity lattice for \(r1s3\). The shaded diamond nodes have \(r1\) underlyingly +long and/or \(s3\) underlyingly +long, counter to what has already been set in the lexicon, and so are not viable candidates. 311

7.7 Relative similarity lattice for \(r1s1\) (viable candidates only). 312
Figures

7.8 Relative similarity lattice for r1s3 (viable candidates only). 312
7.9 The joint relative similarity order for the contrast pair r1s1 and r1s3. Each node is labeled with underlying forms for r1s1 and r1s3, in that order. The left suborder contains all local lexica in which r1 is –stress underlyingly, while the right suborder contains all local lexica in which r1 is +stress underlyingly. 313
7.10 Testing the stress feature for s1. The shaded nodes are the local lexica with +stress assigned to s1 underlyingly. 315
8.1 Relative similarity lattice for r1s1, where s1 has been set to –long. 366
8.2 Viable sublattice for r1s1. 367
8.3 Viable sublattice for r1s1 with inconsistent inputs shaded. 368
The work in this book was conducted over a period of several years. During that time, I have benefitted significantly from conversations with a number of people, including Crystal Akers, Eric Baković, Karen Campbell, Paul de Lacy, Jane Grimshaw, Fernando Guzman, Brett Hyde, Gaja Jarosz, John McCarthy, Nazarré Merchant, Alan Prince, Jason Riggle, and Paul Smolensky. I also received useful feedback from several audiences who kindly listened to various incarnations of pieces of this work: the Rutgers Optimality Reading Group (RORG), the Mathematics Association of America Seaway Section and New York Association of Two Year Colleges Fall 2007 Meeting, several meetings of the Northeast Computational Phonology Circle, the Learnability Meets Acquisition Workshop at the 31st Jahrestagung der Deutsche Gesellschaft für Sprachwissenschaft in Osnabrück, NELS39, the Seventh North American Phonology Conference, the Linguistics Department at SUNY Stony Brook, and the Linguistics and Cognitive Science Department at the University of Delaware. Useful comments were also provided by three Cambridge University Press reviewers. Exceptional patience was exhibited by the students of two incarnations of my Seminar on Learnability and Linguistic Theory (Ling670), as well one incarnation of Phonology III (Ling522).

All factual errors and bizarre-sounding conceptions are the sole responsibility of the author.

Special thanks are due to Alan Prince for suggesting the term “output-driven.”

I wish to thank, for institutional support, the Department of Linguistics and the Rutgers Center for Cognitive Science (RuCCS), at Rutgers, the State University of New Jersey, New Brunswick.

I would like to thank Helen Barton, of Cambridge University Press, for her support of this project and her assistance in getting this work into print.

I wish to express gratitude to my wife, Esther, and my daughters, Heidi and Amanda, for their support and exceptional tolerance of the countless hours I spent in my basement office, working on “the book.” Special thanks to Heidi and Amanda for loving me despite the fact that I do not like ketchup.