An understanding of language as a complex system helps us to think differently about linguistics, and helps us to address the impact of linguistic interaction. This book demonstrates how the science of complex systems changes every area of linguistics: how to make a grammar, how to think about the history of language, how language works in the brain, and how it works in social settings. Kretzschmar argues that to construct the best grammars of languages, it is necessary to understand the complex system of speech. Each chapter makes specific recommendations for how linguists should manage empirical data in order to form better generalizations about a language and its varieties. The book will be welcomed by students and scholars working in linguistics and English language, especially the study of language variation and the historical development of English.

WILLIAM A. KRETZSCHMAR, JR. teaches English as Harry and Jane Willson Professor in Humanities at the University of Georgia. He also has appointments at the University of Glasgow in Scotland and the University of Oulu in Finland. His recent publications include The Linguistics of Speech (Cambridge, 2009).
Language and Complex Systems

William A. Kretzschmar, Jr.
Turning and turning in the widening gyre
The falcon cannot hear the falconer;
Things fall apart; the centre cannot hold;
Mere anarchy is loosed upon the world,
The blood-dimmed tide is loosed, and everywhere
The ceremony of innocence is drowned;
The best lack all conviction, while the worst
Are full of passionate intensity.

Surely some revelation is at hand;
Surely the Second Coming is at hand.
The Second Coming! Hardly are those words out
When a vast image out of *Spiritus Mundi*
Troubles my sight: somewhere in sands of the desert
A shape with lion body and the head of a man,
A gaze blank and pitiless as the sun,
Is moving its slow thighs, while all about it
Reel shadows of the indignant desert birds.
The darkness drops again; but now I know
That twenty centuries of stony sleep
Were vexed to nightmare by a rocking cradle,
And what rough beast, its hour come round at last,
Slouches towards Bethlehem to be born?

*The Second Coming* (1920), William Butler Yeats
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of tables</td>
<td>xiii</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>xiv</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1 Language and complex systems</td>
<td>5</td>
</tr>
<tr>
<td>Examples of complex systems</td>
<td>7</td>
</tr>
<tr>
<td>Principles of complex systems</td>
<td>11</td>
</tr>
<tr>
<td>Language as a complex system</td>
<td>19</td>
</tr>
<tr>
<td>2 Linguistics, science, the humanities, and complex systems</td>
<td>36</td>
</tr>
<tr>
<td>Linguistics as a science</td>
<td>36</td>
</tr>
<tr>
<td>The epistemology of modern science: Wilson and Gould</td>
<td>38</td>
</tr>
<tr>
<td>An alternative to reductionism</td>
<td>44</td>
</tr>
<tr>
<td>The social history of modern science among the disciplines</td>
<td>47</td>
</tr>
<tr>
<td>The humanities as modern science?</td>
<td>52</td>
</tr>
<tr>
<td>3 Usage-based linguistics and complex systems</td>
<td>56</td>
</tr>
<tr>
<td>Paul Hopper (1987)</td>
<td>57</td>
</tr>
<tr>
<td>Paul Hopper and Elizabeth Traugott (1993)</td>
<td>59</td>
</tr>
<tr>
<td>Joan Bybee (2001)</td>
<td>61</td>
</tr>
<tr>
<td>Janet Pierrehumbert (2001)</td>
<td>64</td>
</tr>
<tr>
<td>Adele Goldberg (2006)</td>
<td>68</td>
</tr>
<tr>
<td>Michael Tomasello (2003)</td>
<td>70</td>
</tr>
<tr>
<td>Nick Ellis and Diane Larsen-Freeman (2009)</td>
<td>73</td>
</tr>
<tr>
<td>Joan Bybee (2010)</td>
<td>76</td>
</tr>
<tr>
<td>Usage-based linguistics and complex systems</td>
<td>78</td>
</tr>
<tr>
<td>4 Grammar and complex systems</td>
<td>81</td>
</tr>
<tr>
<td>Zipf’s Law and the 80/20 Rule</td>
<td>83</td>
</tr>
<tr>
<td>English grammar</td>
<td>86</td>
</tr>
<tr>
<td>Complexity science and grammar</td>
<td>93</td>
</tr>
<tr>
<td>Improved grammars</td>
<td>96</td>
</tr>
<tr>
<td>5 Complex systems and the history of the English language</td>
<td>105</td>
</tr>
<tr>
<td>Clockwork</td>
<td>106</td>
</tr>
</tbody>
</table>
### Contents

**Grammaticalization** 109  
**A-curve frequency profiles** 112  
**Punctuated equilibrium in language change** 121  

**6 Neural networks and complex systems** 131  
- **Models in cognitive science** 134  
- **Neural network simulations** 138  
- **Self-Organizing Maps** 141  
- **Making word choices** 150  

**7 Sociolinguistics, communities, and complex systems** 155  
- **Classic North American sociolinguistics** 155  
- **Speech communities and populations of speakers** 164  
- **New varieties** 167  
- **An experiment on scale-free communities** 170  
- **Measurement of scale-free data** 180  
- **Complex systems for sociolinguists** 192  

**8 Postmodernism and complex systems** 201  
- **Fashionable nonsense** 202  
- **Postmodernism and speech** 207  
- **Last words** 212  

**References** 214  
**Index** 224
Figures

1.1 Ants in a line. “Ant Trail.” Author: Reyham Soenasto, www.myfolio.com (© Reyham Soenasto; licensed under the Creative Commons Attribution-ShareAlike 2.0 Generic License)

1.2 A board showing “gliders” (five cell combinations that begin in the middle and move to the lower right) from Conway’s Game of Life (Gosper Glider Gun. YouTube video by nbidyanta. www.youtube.com/watch?v=pvUiA-Q-3hM)

1.3 Koch island, at different scales (adapted from Mandelbrot 1982)

1.4 Whirlpool as self-organization. “Whirlpool in the ‘Water Garden’ within Alnwick Garden.” Author: Christine Westerback, geograph.org.uk (© Christine Westerback; licensed under the Creative Commons Attribution-ShareAlike 2.0 Generic License)

1.5 Eddies and bubbles in the Regnitz

1.6 Leicester locality ca. 1086 (adapted from Russell 1972)

1.7 Barcelona region ca. 1360 (adapted from Russell 1972)

1.8 “Chaotic” fractal: San Marco Dragon (adapted from Mandelbrot 1982)


1.11 A-curve, dry spell variants

1.12 A-curve, parlor variants

1.13 A-curve for fog vowel realizations

1.14 A-curve for six vowel realizations
List of figures

1.15 Scaling property of A-curve: bureau variants for all, SC/GA/FL, NY, women 30
1.16 Scaling property of A-curve: fog variants for women 31
1.17 Scaling property of A-curve: fog variants for Type II speakers 32
1.18 Perceptual aid from the A-curve 33
2.1 Edward O. Wilson (photo by Beth Maynor Young, from National Park Service www.nps.gov/aboutus/nrawards-2010.htm) 39
2.2 Stephen Jay Gould (photo by Kathy Chapman Online, from Wikipedia under Creative Commons 3.0 license) 39
3.1 Distribution of UGA /u/ realizations 63
3.3 VL and VOL construction frequencies (adapted from Ellis and Larsen-Freeman 2009b) 75
4.1 Zipf’s Law (adapted from PlanetMath, http://planetmath.org/zipfslaw) 84
4.2 The 80/20 Rule 85
4.3 Stubbs, top twenty collocates of undergo (adapted from Stubbs 2001) 94
4.5 Differential frequencies of various constructions in the Longman Grammar. Adapted from figures 2.2–2.3, 8.6, 11.2 in Biber, Douglas, Susan Conrad, and Geoffrey Leech, Longman Student Grammar of Spoken and Written English (London: Longman, 2002) 103
5.1 The Corpus Clock (from http://deetoursbydeepali.wordpress.com/2012/05/24/the-time-eater-of-cambridge, used under Creative Commons Attribution-ShareAlike 3.0 Unported license) 106
5.2 Collocates of clockwork in COCA 107
5.3 Collocates of clockwork in COHA 107
5.4 A-curves at two hypothetical moments in time. From Kretzschmar, William A., Jr., The Linguistics of Speech (Cambridge: Cambridge University Press, 2009), 266. Used by permission 114
List of figures


5.6 Realizations of [æ] in *half*, *past*, *last* from twenty-four New York City speakers in LAMSAS 117


5.9 Geographic speciation modes (prepared by Dana Krempels, in the public domain, from Wikipedia) 124

5.10 Great Vowel Shift (from glottopedia.org, under Creative Commons Attribution-ShareAlike 3.0 Unported license) 125

5.11 Great Vowel Shift, Northern region (adapted from Smith 1996) 127

5.12 Great Vowel Shift, London (I) and the Midlands/South region (II) (adapted from Smith 1996) 128


6.2 Matrix of nodes (10×10) in colored patterns 143

6.3 The Midland and Southern regions, maybe (*Hearth* set) 144

6.4 Raven McDavid as a fieldworker (*Day/Wind* set) 144

6.5 Scattered speakers (*Rain* set) 145

6.6 North Midland region and outliers, maybe (*Land/Yard* set) 145

6.7 Map made with Kleiweg and Nerbonne’s cluster analysis, phonetic data. From www.let.rug.nl/~kleiweg/lamsas/traditional. Used by permission. 146

6.8 Ranked node frequency, 10×10 matrix for *Hearth* set 147

6.9 Ranked node frequency, 10×10 matrix for *House/Inside*, *Land/Yard* sets 148

6.10 Ranked node frequency, 10×10 matrix for *Time*, *Day/Wind*, *Rain* sets 149

6.11 The seven top-ranked nodes 151
xii  List of figures

7.2 Hog pen, top ten responses and chart of all responses 172
7.3 Hog pen, only Age Group 1 172
7.4 A-curves of Distribution Type B and Type C 173
7.5 Distribution Type A order of variants: meadow 175
7.6 Distribution Type B order of variants: cloudburst 175
7.7 Distribution Type B order of variants: andirons 176
7.8 Distribution Type C order of variants: cobbler 177
7.9 Exceptional Distribution Type A variant: bunk for pallet 178
7.10 Exceptional Distribution Type B variant (close to Type A) 179
7.11 Lorenz Curve and Gini Coefficient 181
7.12 Normally distributed data set with 100 tokens 183
7.13 Distribution of UGA /u/ realizations 185
7.14 Four A-curves for different numbers of categories for one data set 186
7.15 Type/Speaker index 191
7.16 Twelve A-curves for subsamples of /i/ in three 194
## Tables

1.1 *Dry spell* adjusted for plurals, inappropriate responses  
   
1.2 LAMSAS Tally for *parlor* ‘place to meet guests’ adjusted for plurals, inappropriate responses  
   
1.3 Variation in the *fog* vowel  
1.4 Variation in the *six* vowel  
3.1 Transcriptions of *ten* from LAMSAS  
3.2 Mothers’ and children’s’ construction frequencies by type (adapted from Tomasello 2003)  
5.1 Middle English long, close *e* class (Modern Standard English [i:, iː]), with speakers from the Survey of English Dialects (adapted from Stockwell and Minkova 1988)  
5.2 Three systems involved in the GVS, ca. 1650 (adapted from Smith 1996)  
6.1 *Hearth* words  
7.1 Counts and coefficients for 148 tokens in different categories  
7.2 American English vowels  
7.3 Variant transcriptions for *æ/ in half*  
7.4 Data from *i/ in three*, by categories of different sizes  
7.5 Averages across all vowels  
7.6 Average over all vowels, simplified phonetics  
7.7 Principles of complex systems applied to speech  
8.1 Decreasing size of speech communities  

*Page 22*  

*Page 23*  

*Page 25*  

*Page 27*  

*Page 67*  

*Page 72*  

*Page 126*  

*Page 128*  

*Page 142*  

*Page 186*  

*Page 187*  

*Page 188*  

*Page 189*  

*Page 190*  

*Page 192*  

*Page 193*  

*Page 210*
Acknowledgements

This book, the sequel to my 2009 book *The Linguistics of Speech*, began before the first book was even published. Chapter 1 began as a talk at a 2008 conference organized by Manfred Krug, himself one of the earliest to write about string frequency. While *The Linguistics of Speech* laid the foundations for the study of speech as a complex system, the chapters in this book apply complex systems to many fields within linguistics. I am grateful to the audiences of the many papers on applications in different fields that I prepared and improved over time. These audiences, some at professional meetings and many at universities where I was invited to speak, were always enthusiastic, always full of questions about how complexity science might help them to understand better the linguistic phenomena in which they were particularly interested.

I would also like to thank Jane Willson, whose generosity to the University of Georgia created the professorship I hold. Resources from the professorship allowed me to keep up a busy schedule of presentations in order to bring the idea of complex systems to new audiences of linguists and people in digital humanities. Without the kind of assistance provided by Jane and her late husband Harry, using resources from their Sunnyland Farms business, new ideas would have a much harder time getting out to be considered in the marketplace of ideas. They were thus benefactors not only of the university but of science more generally. I will continue to give pecans from Sunnyland Farms to friends, and I remain proud, Jane, to be your nutty professor.

Finally, I would like to thank my predecessors on the Linguistic Atlas Project for their friendship and support over many years. Raven and Virginia McDavid brought me into that fold when I was just a graduate student, and Lee Pederson has shared ideas with me for nearly thirty years. Without the Atlas, of course, I would be unlikely to have seen the distributional properties that prove speech to be a complex system. Pederson and the McDavids kept Atlas survey research alive while many others pursued sociolinguistics instead, and so preserved the chance for an independent viewpoint.

WAK