

The Transmission of Language: models of biological and cultural evolution

**Kenneth Smith
M. A. (Hons), M. Sc.**

**A thesis submitted in fulfilment of requirements for the degree of
Doctor of Philosophy**

**to
Theoretical and Applied Linguistics,
University of Edinburgh**

March 2003

© Copyright 2003

by

Kenneth Smith
M. A. (Hons), M. Sc.

Abstract

Theories of language evolution typically attribute its unique structure to pressures acting on the genetic transmission of a language faculty and on the cultural transmission of language itself. In strongly biological accounts, natural selection acting on the genetic transmission of the language faculty is seen as the key determinant of linguistic structure, with culture relegated to a relatively minor role. Strongly cultural accounts place greater emphasis on the role of learning in shaping language, with little or no biological adaptation.

Formal modelling of the transmission of language, using mathematical or computational techniques, allows rigorous study of the impact of these two modes of transmission on the structure of language. In this thesis, computational models are used to investigate the evolution of symbolic vocabulary and compositional structure. To what extent can these aspects of language be explained in terms of purely biological or cultural evolution? Should we expect to see a fruitful interaction between these two adaptive processes in a dual transmission model?

As a first step towards addressing these questions, models which focus on the cultural transmission of language are developed. These models suggest that the conventionalised symbolic vocabulary and compositional structure of language can emerge through the adaptation of language itself in response to pressure to be learnable. This pressure arises during cultural transmission as a result of 1) the inductive bias of learners and 2) the poverty of the stimulus available to learners. Language-like systems emerge only when learners acquire their linguistic competence on the basis of sparse input and do so using learning procedures which are biased in favour of one-to-one mappings between meanings and signals. Children acquire language under precisely such circumstances.

As the second stage of inquiry, dual transmission models are developed to ascertain whether this cultural evolution of language interacts with the biological evolution of the

language faculty. In these models an individual's learning bias is assumed to be genetically determined. Surprisingly, natural selection during the genetic transmission of this innate endowment does not reliably result in the development of learning biases which lead, through cultural processes, to language-like communication – there is no synergistic interaction between biological and cultural evolution. The evolution of language may therefore best be explained in terms of cultural evolution on a domain-general or exapted innate substrate.

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Kenneth Smith

Acknowledgements

Acknowledgements sections are usually over-long and toe-curlingly awful. I've managed to keep this one short, but have singularly failed to remove the cringe factor. Sorry.

Thanks firstly to my wife Becky, who has supported me with good grace and charm through my interminable stint as a student. You are the best.

For financial support, my parents deserve a mention, as do the Carnegie Trust for the Universities of Scotland. The department of Theoretical and Applied Linguistics also stumped up several sums of cash so I could go off to conferences and summer schools. I hope everyone feels they've got their money's worth.

On an academic level, thanks first of all to Simon Kirby and Jim Hurford for being excellent supervisors. Thanks are also due to Henry Brighton, who has worked closely with me on some of the models outlined in this thesis, in particular the model in Chapter 5. Andrew Smith also deserves a mention, for the helpful feedback he's given me over the years, for providing an enjoyable working environment in the Junior LEC office, and for not objecting too strongly to his nickname.

Thanks to Cedric McMarty and the rest of the computing support team. I've been thrashing every processor in the department for the past three years, and I couldn't have done so without your help.

Finally, no thanks to Scotland (for being useless at football, but not quite useless enough to stop me caring), UEFA (for running the Champions' League every year) and FIFA (I could have done without World Cup 2002, to be honest).

Contents

Abstract	iii
Declaration	v
Acknowledgements	vii
Chapter 1 Introduction	1
1.1 The uniqueness of language	2
1.1.1 Design features of language	3
1.1.2 Design features of animal communication	4
1.1.3 Oliphant's taxonomy of communication systems	6
1.1.4 Refining the taxonomy	8
1.2 Theories of Language	9
1.2.1 The Nativist paradigm	9
1.2.2 The Empiricist paradigm	11
1.3 Evidence for language evolution	13
1.3.1 Archaeological evidence	13
1.3.2 Current-day evidence	16

1.4 Theories of language evolution	19
1.4.1 Non-adaptationist accounts	20
1.4.2 Adaptationist accounts	29
1.5 Formal models	33
1.6 Guide to the thesis	35
 Chapter 2 Models of cultural transmission	 37
2.1 General and linguistic models	38
2.1.1 A general model	38
2.1.2 Linguistic models	39
2.2 Transmission and Cultural Traits	47
2.2.1 Cultural traits and transmission in the general model	47
2.2.2 Cultural traits and transmission in linguistic models	49
2.3 Forces acting on cultural transmission	56
2.3.1 Natural selection of cultural variants	58
2.3.2 Guided variation	66
2.3.3 Directly biased transmission	69
2.3.4 Indirectly-biased transmission	76
2.3.5 Frequency-dependent bias	79
2.3.6 Transmission through a bottleneck	81
2.4 Summary of the Chapter	91
 Chapter 3 The cultural evolution of communication	 93
3.1 Models of the evolution of vocabulary	94

3.2	The communication model	99
3.3	Model 1: a feedforward network model	100
3.3.1	The communicative agent	101
3.3.2	The Iterated Learning Model	104
3.3.3	Network architecture, learning bias and natural selection	106
3.3.4	Summary	113
3.4	Model 2: an associative network model	114
3.4.1	Communicative agents	115
3.4.2	Acquisition of an optimal system	117
3.4.3	The Iterated Learning Model	119
3.4.4	Maintenance of an optimal system	120
3.4.5	Construction of an optimal system	121
3.4.6	Summary: The classification hierarchy	122
3.5	The Key Bias	124
3.5.1	The key bias in the associative network model	125
3.5.2	The key bias in the feedforward network model	128
3.5.3	The key bias in other models	132
3.6	Biases in vocabulary acquisition in humans and non-humans	135
3.6.1	Biases against synonymy in humans	136
3.6.2	Biases against homonymy in humans	138
3.6.3	Biases in non-human animals?	143
3.7	Summary of the Chapter	146

4.1	Modelling genetic transmission	149
4.1.1	Natural selection on genetic transmission	150
4.1.2	Models of the genetic transmission of communication	151
4.2	The dual inheritance model	155
4.2.1	The genetic transmission of direct bias	156
4.3	The Evolutionary Iterated Learning Model	158
4.3.1	The evolutionary iterated learning of vocabulary	158
4.4	Model 1: Adding genetic transmission to the feedforward network model	161
4.4.1	Genotypes, phenotypes and the genotype-phenotype mapping . .	162
4.4.2	Reproduction	162
4.4.3	The EILM algorithm	163
4.4.4	Emergence of a communication system	164
4.4.5	Maintenance of an optimal system	167
4.4.6	Summary	169
4.5	Model 2: Adding genetic transmission to the associative network model	170
4.5.1	Genotypes, phenotypes and the genotype-phenotype mapping . .	170
4.5.2	Reproduction	171
4.5.3	The EILM algorithm	171
4.5.4	Main result: optimal communication rarely emerges	174
4.5.5	Varying the speed of convergence by varying e	180
4.5.6	Varying the speed of convergence by varying cultural population size	180

4.5.7	Summary	184
4.5.8	Discussion	188
4.6	Summary of the Chapter	193
Chapter 5 The cultural evolution of compositionality		195
5.1	Models of the cultural evolution of linguistic structure	196
5.1.1	Varying the learner model	196
5.1.2	Varying the transmission bottleneck	198
5.1.3	Varying the structure of the meaning space	198
5.2	An Iterated Learning Model	200
5.2.1	Languages and communication	200
5.2.2	Communicative agents	202
5.2.3	The Iterated Learning Model	207
5.2.4	Environments	209
5.2.5	Measuring compositionality	210
5.3	The impact of transmission bottleneck and environment structure	213
5.3.1	Linguistic evolution in the absence of a bottleneck	213
5.3.2	Linguistic evolution in the presence of a bottleneck	220
5.3.3	Summary	229
5.4	Exploring the impact of learning bias	230
5.4.1	Acquisition of a compositional system	230
5.4.2	Maintenance through a bottleneck	232
5.4.3	Construction through a bottleneck	233

5.4.4 The classification hierarchy	235
5.5 The key bias	235
5.5.1 An overview of the learning biases	235
5.5.2 The two parts of the bias	241
5.5.3 The bias in other models	248
5.5.4 Summary	251
5.6 One-to-one biases and the acquisition of linguistic structure	251
5.6.1 One-to-one biases in general: clarity and isomorphism	252
5.6.2 One-to-one biases in morphology	254
5.6.3 One-to-one biases in syntax	256
5.7 Summary of the Chapter	260
 Chapter 6 The evolution of compositionality in populations	263
6.1 Models of the evolution of linguistic structure in populations	264
6.1.1 Cultural evolution in populations	264
6.1.2 Gene-culture coevolution in populations	265
6.2 Languages, communication and communicative agents	270
6.3 Cultural evolution in populations	271
6.3.1 Linguistic evolution in the absence of a bottleneck	274
6.3.2 Linguistic evolution in the presence of a bottleneck	277
6.3.3 Summary	281
6.4 The evolution of learning biases for compositional language	282
6.4.1 Genotypes, phenotypes and reproduction	282

6.4.2	The EILM	283
6.4.3	The environment	285
6.4.4	A negative result	285
6.4.5	A positive result: the evolution of learning biases for compositional language	286
6.4.6	Summary	293
6.5	Discussion	294
6.6	Summary of the Chapter	294
Chapter 7	Conclusions	297
Appendix A	Mathematical models of transmission	309
A.1	Models of cultural transmission	309
A.1.1	Basic cultural transmission models	309
A.1.2	Pressures acting on cultural transmission	312
A.2	Genetic transmission and natural selection	322
A.3	Dual transmission and direct bias	324
Appendix B	Published papers	329
References		437

List of Tables

1.1	Trends in brain size	14
2.1	The loss of phonemic distinctions in Czech	42
3.1	Production and reception in feedforward neural networks	103
3.2	The imitator learning bias	108
3.3	The obverter learning bias	108
3.4	The imitator random bias	109
3.5	The obverter random bias	110
3.6	The classification of weight-update rules	123
3.7	Learning biases of the weight-update rules	129
5.1	Sensitivity to initial conditions	218
5.2	Comparison of the stability of compositional language across environments	229
5.3	A compositional language	231
5.4	The classification of weight-update rules	236
5.5	Values of g for various analyses	237
5.6	Values of g for various analyses	240
5.7	Values of g for various analyses	242
5.8	Average internal compositionality for all rules	243

5.9	Values of g for holistic analyses	245
5.10	Values of g for compositional analyses	247
6.1	Summary of results for the population ILM	278

List of Figures

1.1	Oliphant's taxonomy of communication systems	6
1.2	An extended taxonomy of communication systems	9
1.3	Jackendoff's schedule for the evolution of language	31
2.1	A general model of cultural transmission	39
2.2	The transmission of language from generation to generation	40
2.3	The Arena of Use in the cycle of linguistic transmission	45
2.4	The Expression/Induction cycle	45
2.5	The Negotiation Model	47
2.6	The Iterated Learning Model	48
2.7	Language as a mapping between three spaces	55
2.8	Tomasello's cultural ratchet	67
2.9	Hutchins & Hazelhurst's autoassociator network	72
2.10	Exemplars in Batali's model	87
2.11	Argument maps in Batali's model	90
3.1	Imitator and obverter network architectures	102
3.2	Communicative accuracy in imitator and obverter populations	107

3.3	Communicative accuracy in imitator populations, with natural selection of cultural variants	111
3.4	Communicative accuracy in obverter populations, with natural selection of cultural variants	111
3.5	Communicative accuracy in imitator populations, with natural selection of cultural variants and noise on cultural transmission	112
3.6	Communicative accuracy in obverter populations, with natural selection of cultural variants and noise on cultural transmission	113
3.7	The associative network	116
3.8	Production in the associative network	117
3.9	Learning in the associative network	118
3.10	Maintenance of an optimal system	121
3.11	Construction of an optimal system	123
3.12	The hierarchy of weight-update rules	124
3.13	The hierarchy of weight-update rules restated	128
3.14	The bias of imitator networks	130
3.15	The bias of obverter networks for ambiguous mappings	131
3.16	The bias of obverter networks for unambiguous mappings	131
3.17	The learning bias of Oliphant & Batali's obverter	136
4.1	A simple model of genetic transmission	150
4.2	The dual transmission model	156
4.3	The generational EILM	159
4.4	The gradual EILM	159
4.5	Genotype-phenotype mapping in the feedforward network model	162

4.6	Communicative accuracy in imitator populations, with natural selection on genetic transmission	165
4.7	Communicative accuracy in imitator populations, with natural selection on genetic transmission	166
4.8	Communicative accuracy in obverter populations, with natural selection on genetic transmission	168
4.9	Slumping in imitator populations	169
4.10	Selection pressure arising from tournament selection	173
4.11	Selection pressure on cultural variants arising from tournament selection .	174
4.12	Evolution of learning biases leading to optimal communication	175
4.13	The three phases of a successful run	177
4.14	Relative communicative accuracy of constructors	178
4.15	Relative communicative accuracy of maintainers	178
4.16	Relative communicative accuracy of learners	179
4.17	Relative communicative accuracy of non-learners	179
4.18	The impact of learning	181
4.19	The impact of spatial organisation	183
4.20	Time-space diagrams for cultural spatialisation	185
4.21	Time-space diagrams for genetic spatialisation	186
4.22	Time-space diagrams for combined spatialisation	187
5.1	The structured associative network	203
5.2	Learning in the structured associative network	205
5.3	Production in the structured associative network	206
5.4	Analysis pairs as parse trees	207

5.5	Unstructured environments	210
5.6	Structured environments	210
5.7	I-compositionality for sparse environments, with no bottleneck	214
5.8	E-compositionality for sparse environments, with no bottleneck	215
5.9	E-compositionality for medium density environments, with no bottleneck	216
5.10	E-compositionality for dense environments, with no bottleneck	217
5.11	Sensitivity to initial conditions	218
5.12	Compositionality in sparse environments, where $c = 0.8$	222
5.13	Compositionality in medium density environments, where $c = 0.4$	222
5.14	Compositionality in medium density environments, where $c = 0.6$	223
5.15	Compositionality in medium density environments, where $c = 0.8$	223
5.16	Compositionality in dense environments, where $c = 0.25$	224
5.17	Compositionality in dense environments, where $c = 0.4$	224
5.18	Compositionality in dense environments, where $c = 0.6$	225
5.19	Compositionality in dense environments, where $c = 0.8$	225
5.20	Maintenance of a compositional language	233
5.21	Construction of a compositional language	234
5.22	Connection weights after learning	238
6.1	Compositionality in sparse environments, with no bottleneck	275
6.2	Compositionality in medium density environments, with no bottleneck . .	276
6.3	Sample runs of the population ILM	279
6.4	Sample runs of the population ILM	286

6.5	Evolution of learning bias leading to communicatively optimal, compositional language	287
6.6	Early stages of evolution of learning bias leading to communicatively optimal, compositional language	288
6.7	Learning bias, communicative accuracy and compositionality	289
6.8	Relative communicative accuracy for +constructor, +ic-preserved weight-update rules	290
6.9	Relative communicative accuracy for +maintainer, -constructor, +ic-preserved weight-update rules	290
6.10	Relative communicative accuracy for +maintainer, \pm constructor, -ic-preserved weight-update rules	291
6.11	Relative communicative accuracy for -maintainer weight-update rules . .	291

