

Colour Terms, Syntax and Bayes Modelling Acquisition and Evolution

Mike Dowman

February 2004

PhD Thesis,
School of Information Technologies,
Faculty of Science,
University of Sydney

Supervised by Associate Professor Judy Kay

Colour Terms, Syntax and Bayes: Modelling Acquisition and Evolution

Mike Dowman

Abstract

This thesis investigates language acquisition and evolution, using the methodologies of Bayesian inference and expression-induction modelling, making specific reference to colour term typology, and syntactic acquisition. In order to test Berlin and Kay's (1969) hypothesis that the typological patterns observed in basic colour term systems are produced by a process of cultural evolution under the influence of universal aspects of human neurophysiology, an expression-induction model was created. Ten artificial people were simulated, each of which was a computational agent. These people could learn colour term denotations by generalizing from examples using Bayesian inference, and the resulting denotations had the prototype properties characteristic of basic colour terms. Conversations between these people, in which they learned from one-another, were simulated over several generations, and the languages emerging at the end of each simulation were investigated. The proportion of colour terms of each type correlated closely with the equivalent frequencies found in the World Colour Survey, and most of the emergent languages could be placed on one of the evolutionary trajectories proposed by Kay and Maffi (1999). The simulation therefore demonstrates how typological patterns can emerge as a result of learning biases acting over a period of time.

Further work applied the minimum description length form of Bayesian inference to modelling syntactic acquisition. The particular problem investigated was the acquisition of

the dative alternation in English. This alternation presents a learnability paradox, because only some verbs alternate, but children typically do not receive reliable evidence indicating which verbs do not participate in the alternation (Pinker, 1989). The model presented in this thesis took note of the frequency with which each verb occurred in each subcategorization, and so was able to infer which subcategorizations were conspicuously absent, and so presumably ungrammatical. Crucially, it also incorporated a measure of grammar complexity, and a preference for simpler grammars, so that more general grammars would be learned unless there was sufficient evidence to support the incorporation of some restriction. The model was able to learn the correct subcategorizations for both alternating and non-alternating verbs, and could generalise to allow novel verbs to appear in both constructions. When less data was observed, it also overgeneralized the alternation, which is a behaviour characteristic of children when they are learning verb subcategorizations. These results demonstrate that the dative alternation is learnable, and therefore that universal grammar may not be necessary to account for syntactic acquisition. Overall, these results suggest that the forms of languages may be determined to a much greater extent by learning, and by cumulative historical changes, than would be expected if the universal grammar hypothesis were correct.

Preface

The aim of this thesis is to show how linguistic phenomena which have been identified by other researchers can be explained. The results reported in this thesis were obtained by running simulations on computers, and comparing the outputs of these simulations to data reported in published sources. Therefore, the research involved no empirical data collection. This thesis attempts to provide explanations of a number of phenomena which have been identified as a result of empirical investigations, and subsequently reported in the literature. However, while this thesis contains discussions of many empirical findings, and of analyses of empirical data, it is not concerned with these analyses in themselves, or with the correctness of the data on which they are based. For example, much of this thesis relies heavily on Kay and Maffi's (1999) analysis of the data of the World Colour Survey, which is discussed in detail in section 2.1. Kay and Maffi's analysis remains controversial (Levinson, 2001), but I do not attempt to analyse the primary data myself. That would go beyond the scope of this thesis, so instead I simply tried to create a computer model which would account for the emergence of colour term systems corresponding to those which Kay and Maffi reported were attested in the World Colour Survey.

While criticisms have been made of Kay and Maffi's work (many of which are discussed in section 2.1 below), a wealth of empirical data supports their general conclusions. The basis of most of the criticisms seems to be that Kay and Maffi's

analysis does not take account of the full complexity of colour term systems, especially in that it excludes non-basic colour terms, and that it ignores the secondary connotations of basic colour terms. However, neither of those factors would seem to in any way invalidate the results which Kay and Maffi did report. Hence it would seem that, even if some of the details of Kay and Maffi's analyses turn out to be incorrect, that is unlikely to greatly affect the validity of the work presented in this thesis. Many other issues relating to this thesis remain controversial, such as exactly what neurophysiological mechanisms underlie colour vision, and whether the dative alternation can be explained in terms of semantic regularities. However, like with the controversies surrounding colour term typology, resolving these issues concerning empirical findings goes beyond the scope of this thesis. Further empirical findings may necessitate a revision of this work, but at present it is only possible to work on the basis of the results which have been published up to this point.

While the research for this thesis relied entirely on computer modelling, its actual subject matter lies within linguistics, and to some extent related disciplines such as psychology. Hence the thesis has been written so that it should be, as far as possible, understandable by a general linguistic audience. At the same time, I have tried to explicate linguistic terminology whenever possible, for the benefit of non-linguists. In general, where specific technical issues arise, I have tried to make them as accessible as possible, but a description of some aspects of the research necessitates the use of mathematical concepts and related notation which are likely to be unfamiliar to many readers interested in the subject matter of the thesis. This thesis makes considerable use of techniques from statistics and machine learning, in particular Bayesian inference and minimum description length. However, while some attempt is made to justify why these techniques were chosen, actually demonstrating why they are

effective machine learning techniques, or why, for example, Bayes' rule is correct, goes beyond the scope of this thesis. It can be said that, while this thesis uses machine learning, it is not about machine learning itself, and hence I have tried to restrict discussion of the machine learning literature to the minimum necessary.

All of the thesis is my own original work, except where explicit reference is made to other work. However, portions of the thesis have been published, and I have presented results at a number of conferences, and given some other talks. A full list of this research activity appears below (on all of which I am the sole author, except where noted otherwise). In general, the publications do not correspond exactly to particular parts of this thesis, but Chapter 9 is essentially an expanded version of the 2000 Cognitive Science Society conference paper, while parts of the other five papers are reproduced in Chapters 1 to 8.

Publications

2003	Modeling Language as a Product of Learning and Social Interactions. <i>Cognitive Systems</i> , Volume 6, issue 1.
2003	Explaining Color Term Typology as the Product of Cultural Evolution using a Multi-agent Model. In <i>Proceedings of the Twenty-Fifth Annual Meeting of the Cognitive Science Society</i> . Cognitive Science Society.
2002	Modelling the Acquisition of Colour Words. In B. McKay and J. Slaney (eds.) <i>Advances in Artificial Intelligence</i> . Berlin: Springer-Verlag.
2001	A Bayesian Approach to Colour Term Semantics. <i>Lingu@scene</i> , Volume 1.
2001	<i>A Bayesian Approach to Colour Term Semantics</i> . Technical Report Number 528. Sydney: Basser Department of Computer Science, University of Sydney. (This is a more technical version of the lingu@scene paper above.)
2000	Addressing the Learnability of Verb Subcategorizations with Bayesian Inference. In Gleitman, L. R. & Joshi, A. K. (Eds.) <i>Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society</i> . Mahwah, New Jersey, USA: Lawrence Erlbaum Associates.

Other Conference talks and presentations

2002	Modelling the Evolution of Basic Colour Terms. Talk given as part of the <i>Language in Time Symposium</i> , University of Western Australia, 25-27 June.
2002	A Bayesian Model of Colour Term Acquisition and Typology. Talk given at <i>Australian Linguistics Society Conference</i> , Macquarie University, Sydney, 13-14 July.
2002	Bayesian Learning of Linguistic Categories. Talk given at <i>European Society for the Study of Cognitive Systems Special Workshop on Multidisciplinary Aspects of Learning</i> , Paris, 17-19 January.
2002	<i>Explaining Language Typology using a Multiple Agent Model</i> . Talk given at Department of Computer Science and Software Engineering, University of Western Australia, 28 June.
2000	A Bayesian Model of Syntactic Acquisition. Talk given at <i>Australian Linguistics Society Conference</i> , University of Melbourne, 7-9 July.
2000	<i>Learning Verb Subcategorizations - A Case Study for the Acquisition of Syntax</i> . Departmental Seminar, Department of Linguistics, University of Sydney, 3 March.
1999	A Minimum Encoding Inference Approach to Theoretical Syntax. Talk given at <i>Australian Machine Learning Workshop</i> , Australian National University, Canberra, 22-23 November. Co-authored with Jon Patrick.
1999	Syntax: Accounting for Acquisition and Change. Talk given at <i>Postgraduate Students' Conference</i> , Department of Linguistics, University of Sydney, 23 September.

During the course of my Ph.D. I have received help from an awful lot of people, and I'd like to thank them all. Some people have helped me out with relevant literature, others have given me feedback on my papers or have had helpful discussions with me about my research. Others have just given me much needed encouragement along the way. I will probably will forget to mention some people who should be here, but I would like to thank Judy Akinbolu, Stephen Anthony, Brett Baker, Tony Belpaeme, Nils Bruin, Cassily Charles, Garry Cottrel, Gerhard Dalenoort, David Dowe, Mark Dras, Michelle Ellefson, Jeff Elman, Eva Endrey-Walder, Bill Foley, Alexander Francis, Roslyn Frank, Yukari Fujiwara, John Goldsmith, Paul Green-Armytage, Tom Griffiths, Joost van Hamel, Catherine Harris, Timo Honkela, Jim Hurford, Simon Kirby, Bill Labov, Johan Lammens, Darren Ler, Andrew Lum, Sinead Lyle, Rob MacLaury, Que Chi Luu, Chris Manning, Vladimir Novikov, Adam Blaxter Paliwala, Jon Patrick, Sam Pickering, David Powers, Hong Liang Qiao, Christine Rakvin, Jacqueline van der Schaaf, Jane Simpson, Kenny Smith, Josie Spongberg, Anders Steinvall, Kimie Takahashi, Tania Tsatralis. I have had three supervisors during the course of my Ph.D., but Judy Kay deserves special thanks as she has supervised most of it, and I probably would never have completed without her help. I should also give special thanks to Mark Ellison, who supervised my undergraduate dissertation (Dowman,1998), and without whom I probably would never have got interested in minimum description length, Bayesian inference, or any other form of computer modelling. And I could not have done some of the hard maths without the help of Emmanuel Letellier. Of course any errors or omissions remain my own responsibility (and I have not always taken all the advice I was given).

My Ph.D. was supported by a scholarship funded by a grant from the Australian Research Council, an International Postgraduate Research Scholarship from the

Australian government, and an International Postgraduate Award from the University of Sydney.

Contents

Chapter 1 Introduction	1
Chapter 2 Background	6
2.1 Colour Terms across Languages	12
2.2 Modes of Explanation in Linguistics	32
2.3 Psycholinguistic and Neurophysiological Findings	45
2.4 Expression-Induction Models of Language	54
Chapter 3 A Bayesian Model of Colour Term Acquisition	67
3.1 Bayesian Inference	67
3.2 Axioms of the Acquisitional Theory.....	76
3.2.1 A Conceptual Colour Space	78
3.2.2 Evidence from which Children Learn.....	82
3.2.3 Unreliability of Data	83
3.2.4 Salience, Memorability and Location of Unique Hue Points	84
3.2.5 Possible Colour Term Denotations	86

3.2.6	Bayes' Optimal Classification	88
3.3	The Bayesian Model	89
3.3.1	Calculating Probabilities.....	89
3.3.2	More than One Colour Term.....	99
3.3.3	Deriving Fuzzy Sets	102
3.4	Implementing the Model.....	103
3.4.1	Calculating the Probability that a Colour is within the Denotation of a Colour Term.....	104
3.4.2	Derivation of the Integrals	105
3.4.3	Applying the Equations.....	127
3.5	Computer Implementation	127
3.6	Learning the Denotation of English Colour Terms from Examples	128
3.7	Learning with Unreliable Data.....	134
3.8	Robustness of the Model to Random Noise.....	137
	Chapter 4 Comparing Acquisitional and Evolutionary Simulations.....	142
4.1	Learnable Colour Term Systems	142
4.2	Evolvable Colour Term Systems	146
4.3	Implications of the Results.....	153

Chapter 5 Adding Unique Hue Points to the Model	159
5.1 Specification of the New Model	159
5.2 Predictions of the Acquisitional Model	166
Chapter 6 Simulating Colour Term Evolution.....	170
6.1 The Evolutionary Model	170
6.2 Emergent Colour Term Systems	172
6.3 Number of Basic Colour Terms	175
6.4 Typological Analyses.....	178
6.5 Investigating the Effect of Unique Hue Points	193
Chapter 7 Adding Random Noise to the Evolutionary Model.....	205
Chapter 8 Implications and Future Directions	212
Chapter 9 Bayesian Acquisition of Syntax.....	236
9.1 Bayesian Grammatical Inference	239
9.2 Computational Models of Syntactic Acquisition.....	246
9.2.1 Description of Model	252
9.2.2 Results	259
9.3 Learning Verb Subcategorizations	260
9.3.1 Data Used for Learning.....	261

9.3.2	Results	263
9.4	Discussion	268
9.5	Conclusion	290
Chapter 10 The Nature of Language		292
References		304
Appendix A Source Code for Model with Continuous Colour Space		323
Appendix B Source Code for Model with Discrete Colour Space		324
Appendix C Results Obtained with Discrete Colour Model.....		326
Appendix D Ditransitive Verb Corpus.....		330
Appendix E Source Code for the Syntax Model.....		334

List of Figures

Figure 2.1. Berlin and Kay's (1969) Implicational Hierarchy.	16
Figure 2.2 The Main Line of Kay and Maffi's Evolutionary Trajectory.....	24
Figure 2.3. Chomsky's Conceptualization of Language Acquisition.	34
Figure 2.4 Hurford's Diachronic Spiral.....	38
Figure 2.5 Interacting Constraints on Possible Languages	45
Figure 3.1. The Conceptual Colour Space.	79
Figure 3.2. Indexing Colours in the Colour Space.....	90
Figure 3.3. A Hypothesis as to the Denotation of a Colour Term.	91
Figure 3.4. A Hypothesis as which Crosses the Origin.	91
Figure 3.5. Hypothesis Probabilities with Erroneous Data.....	97
Figure 3.6. The Phenomenological Colour Space with Observed Colour Term Examples.....	113
Figure 3.7. The Fuzzy Denotation of English Basic Colour Terms after 5 Examples.	129

Figure 3.8. The Fuzzy Denotations of English Colour Terms after 20 Examples	133
Figure 3.9. The Fuzzy Denotations Learned for <i>Green</i>	136
Figure 3.10 Accuracy of Learning with Noisy Data	139
Figure 4.1. A Learnable Colour Term System of a Type which is Unattested Typologically.	144
Figure 4.2. Outline of the Evolutionary Algorithm.	149
Figure 4.3 A Colour Term System which Emerged in an Evolutionary Simulation.	151
Figure 5.1. Learned Denotations for Urdu Colour Terms.....	168
Figure 6.1. The Basic Colour Term Systems of Four Artificial People from the same Simulation.	173
Figure 6.2. Relationship of Number of Frequency of use of Colour Terms to Number of Basic Colour Terms in Emergent Languages.	177
Figure 6.3. Denotations of Basic Colour Terms for all Adults in a Community.	179
Figure 6.4. Percentage of Colour Terms of each type in the Simulations and the World Colour Survey.	184
Figure 6.5. Scatter Graph Showing Relationship between the Frequencies of Colour Terms in the World Colour Survey and the Simulations.	189
Figure 6.6. Locations of Colour Term Prototypes.	194
Figure 6.7. Frequency Distribution of 10,644 WCS Colour-term Foci across the Hue Columns of the Ethnographic Munsell Array.	195

Figure 6.8. The Distribution of Unique Hue Points between Colour Terms.	198
Figure 6.9. Percentage of Colour Terms of each Type.	200
Figure 9.1. A Structure Assigned by the Learned Grammar	265
Figure 9.2. Architecture of Allen's Model.....	281
Figure C.1 The Basic Colour Term Systems of all the Artificial People from the Simulation Reported in Section 6.2.	327-329

List of Tables

Table 3.1. Axioms of the Acquisitional Theory.....	78
Table 3.2. Summary of Symbol's Meanings.....	106
Table 3.3. Summary of the conditions under which each equation applies.....	127
Table 3.4. The Ranges of the Denotations for the English Colour Terms Taught to the Model.	130
Table 3.5. Means and Standard Deviations Showing Precision and Accuracy of Learning with Noisy Data.	140
Table 6.1. Means and Standard Deviations of the mean number of Basic Colour Words in Emergent Languages.....	177
Table 6.2. Frequencies of Colour Terms of each type in the Simulations and the World Colour Survey.	183
Table 6.3. The Most Common Colour Term Systems Emerging in the Simulations.	191
Table 6.4. Classification of How the Languages in the Simulations Diverge from the Attested Evolutionary Sequences.....	192
Table 6.5. The Frequency of Each Type of Colour Term in Simulations without Unique Hues.....	201

Table 6.6. The Most Common Types of Emergent Colour Term Systems in Simulations without Unique Hues.	202
Table 6.7. Classification of Emergent Languages.	203
Table 9.1. Data for English.	253
Table 9.2 Grammar Describing English Data.	253
Table 9.3. Form of Initial Grammars.	256
Table 9.4. Examples of Changes that Could be Made to an Initial Grammar.	258
Table 9.5. Evaluations for English Grammar.	260
Table 9.6. Evaluations for Ditransitive Verbs Data.	263
Table 9.7. Grammar Learned from Ditransitive Verbs Data.	264
Table 9.8. Evaluations for Ditransitive Grammars with <i>sent</i> in Irregular Class or Regular Class.	266
Table C.1. The Number of Examples Remembered by Each Artificial People in the Simulation Reported in Section 6.2.	327