Modelling Language as a Product of Learning and Social Interaction

Mike Dowman

Mike@it.usyd.edu.au

School of Information Technologies,
Madsen Building, F09,
University of Sydney,
NSW 2006
Australia

Abstract

Computational models were constructed to investigate how the meanings of basic colour terms were learned, and to determine why these words have prototype properties, and why they partition the colour space. A Bayesian model of acquisition was able to learn colour term systems with these properties, but could equally well learn colour term systems which did not partition the colour space or have prototype properties, and so it failed to explain the empirical data concerning these words. Computational evolutionary simulations were then conducted by creating a community of artificial people using multiple copies of the Bayesian model. These artificial people then learned colour words from one-another, and colour term systems were allowed to evolve over a number of generations. The emergent colour terms always partitioned the colour space and had prototype properties. These results demonstrate that the Bayesian model is able to account for the properties of colour term systems only when it is placed in a social context and so they provide evidence of the importance of understanding language as a product of both psychology and social interaction.

1. Introduction

By studying wide ranges of languages from throughout the world, linguists have established that languages vary greatly in terms of their grammatical structures, sound systems and semantics, but they have also found that this variation is not without limit. It is possible to identify many properties which are common to all or most languages, and to find implicational universals, which allow some property of a language to be predicted based on the presence of another feature or construction. This paper is concerned with understanding why languages conform to such typological rules, while still showing great variation within the limits that they impose. It is concerned with investigating to what extent such patterns are the product of properties of the human mind and the mechanism which children use to learn language, or whether some such regularities may be best explained as the product of social processes. The focus here is on colour term systems, and on investigating whether the properties of such systems are best explained within a psychological or a social model.

Chomsky has been one of the foremost proponents of the hypothesis that the key to understanding language is to understand individuals and the way they learn language. Chomsky (1972) conceptualised the process of language acquisition as a process of mapping from observed linguistic data to a psychological representation of language in the brain, as is illustrated by Figure 1. The particular language which a person learns will be determined by the linguistic data to which he is exposed, but the nature of the psychological mechanism which children use to learn will also play a role in determining the form of the resultant language. Chomsky (1986) introduced the term *I-language* to refer to knowledge of language in the minds of individual speakers, and he argued that linguistics should be concerned exclusively with the study of I-language. He considered the properties of observed language in the world external to people (*E-language*) to be epiphenomenal, although of course it is partly through the study of E-language that linguists gain their understanding of I-language. Chomsky stated that the range of possible human languages is constrained by the genetically determined properties of the human brain, and so it follows that if we want to understand language universals and typology we should do so by focussing our study on I-language and language acquisition.

Figure 1 Chomsky's Conceptualisation of Language Acquisition

In contrast to Chomsky viewpoint, de Saussure (1959) stressed that language is simultaneously a social and a psychological phenomenon. While the ability to speak and understand language is a psychological ability which must rely on an internalised knowledge of language learned by each language user, language is used primarily for communication, and is hence necessarily a social phenomenon. In order for our knowledge of language to be useful, there must exist a community of speakers who have already adopted the collection of conventions which constitute the language, but every time we speak we produce new linguistic data which may cause other speakers to modify their own I-languages. (For example they may learn a new word, or come to prefer one grammatical construction over another.) This suggests that a better understanding of language may be achieved if we conceptualise language as a system which has both psychological and social components, and focus not only on language in the brain but also on the social processes in which it is used and through which it passes between one generation of speakers and the next.

Hurford (1987, 1990) conceptualises language as a continuous cycle as depicted in Figure 2. People use their internalised knowledge of language to communicate, but this communication takes place in the *arena of language use*. The arena of language use concerns all those factors which influence what people say in practice, and how they say it, and so the properties of this arena will determine exactly what is said to whom. It is the language that is produced in the arena of language use which goes on to form the primary linguistic data from which the next generation of speakers will gain their knowledge of language.. Hence in this model it is not only the language acquisition device which places constraints on the allowable human languages, but also the properties of the arena of language use. From this perspective, the social context in which language is used may be as important in shaping language as any characteristics of individual people.

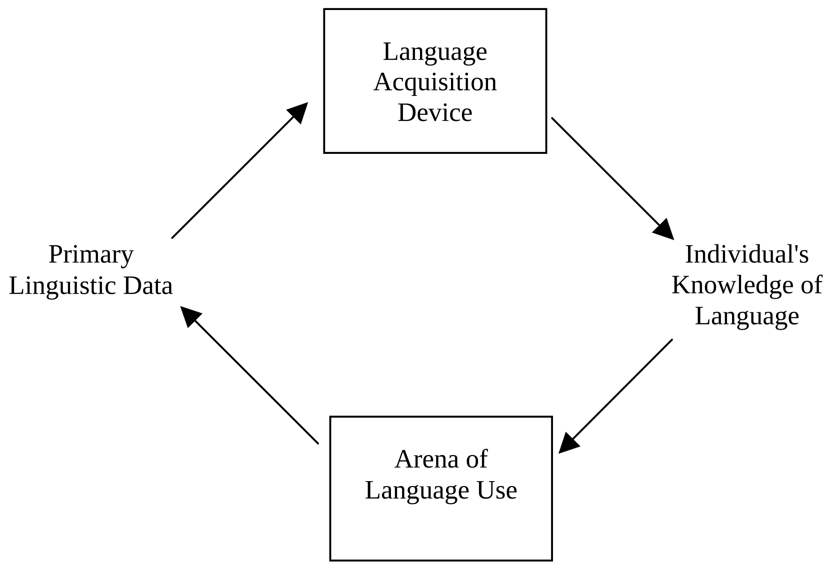


Figure 2 Hurford's Diachronic Spiral

Hurford (1987) constructed a computational model which simulated the diachronic evolution of language. His model was concerned with number constructions, and how individual numerals were combined to express numbers of greater values. He created a population of artificial people who knew a series of words which could express the numbers one to ten, and a syntactic construction in which two of these numbers could be combined to create a new expression, the meaning of which would be the sum of the two numerals which it contained. So, for example, sixteen could be expressed as 'eight eight', or as 'six ten'. The simulation would proceed by first choosing a number between 11 and 20 at random, and then having one person express this number to another person. Initially the artificial people were equally likely to use any combination of numerals which expressed the right value, but once they had heard one numeral used more frequently than others, they would prefer to use that numeral when forming constructions. After a short period of time all speakers in the community would use the numeral ten together with one other numeral for expressing all numbers. This is in accord with typological evidence, which shows that the numbers 11 to 19 are typically expressed with a construction which appears to be based on the word for ten in the language together with another numeral. (For example 'sixteen' in English would appear to be based on the words 'six' and 'ten'.) What is interesting about Hurford's model is that he has shown how a community of speakers can come to develop a shared language which has a universally attested property, even though each individual person in the simulation was capable of learning a language which did not have the property. (For example the artificial people could have learned to always use the numeral 'nine' instead of ten.) Hurford's model was the first such computational evolutionary model, but now his methodology has become firmly established, and has been used to gain an understanding of a wide range of linguistic phenomena as diverse as compositionality (Kirby, 2000) and vowel system typology (de Boer, 2001). This paper applies the same methodology to basic colour terms.

2. Basic Colour Terms

All languages have a small number of highly salient colour terms, the denotations of which are not limited to a subset of the colours denoted by some other colour term, and it is only these words which are considered to be *basic colour terms*. In English there are eleven such words: 'red', 'orange', 'yellow', 'green', 'blue', 'purple', 'pink', 'brown', 'black', 'grey' and 'white'. None of the other English colour terms, such as 'scarlet', 'mauve' or 'buff' are considered to be basic. By conducting a large cross-linguistic survey, Berlin and Kay (1969) determined that all languages have between two and eleven basic colour terms. Regardless of how many basic colour terms a language has , the terms tend to partition the colour space so that for every colour there is a corresponding colour term which can name it. However, the location of the boundaries between colour terms varies between languages, so that, even when there are very similar colour terms in different languages, the exact range of colours which each term denotes will differ.

Another important feature of colours terms is that they have prototype properties (Taylor, 1989). Rather than a word like 'green' simply denoting a range of colours uniformly, some colours within its denotation are better examples of green than are others. We can usually identify a single word which is the best examples of a colour category, which is the prototype, and as colours get more dissimilar to this prototype they become progressively worse members of the category *green*. At the margins of a word's denotation we find *fuzzy boundaries* , where it becomes unclear exactly which colours are members of the category and which are not.

When children learn their first language they must learn which ranges of colour each term denotes. Children are not usually explicitly taught the range of colours which colour terms can be used to identify, so it would seem that the primary source of evidence which children use to learn must be derived by observing other people's speech. In order to learn word meanings, children must observe examples not only of which words are used, but also of what those words are used to refer to. In the case of colour words, these words are used to identify particular colours, and so the evidence from which children learn the meanings of these words must consist of examples of colours which these words have been used to identify. The task of learning will then be to generalise from such examples to determine the full range of colours which come within the word's denotation.

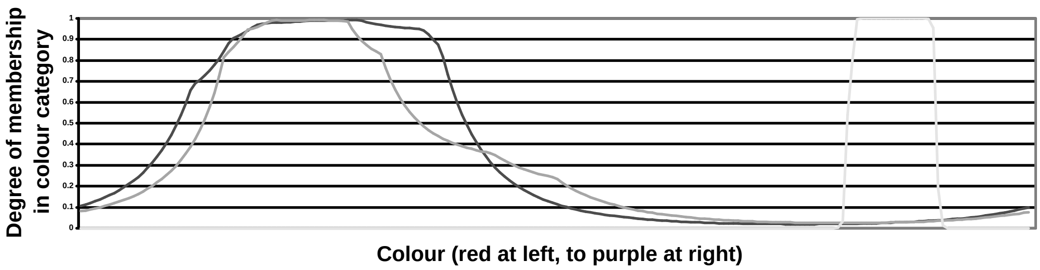
This paper reports work which investigated what properties colour terms learned by a computational model of colour term acquisition would have, and what constraints that model places on the range of possible human languages. This was first investigated in the context of learning from data which was generated artificially in an attempt to create the same kind of data which a child learning a language would encounter. The second simulation used multiple copies of the acquisitional model to simulate a whole community of speakers, allowing social interactions between speakers to be modelled, and the cumulative effect of the evolution of language over several generations to be determined.

1.1 Modelling the Acquisition of Colour Terms

The acquisitional model learns using the statistical procedure known as Bayesian inference. This method of learning is based on Bayes' rule, which was proved by Bayes (1763). The rule allows the probability of alternative hypotheses to be calculated given some evidence about which the hypotheses make predictions. In the case of learning colour words, the hypotheses would be specifications of the range of colours which a colour word denoted, and the evidence would consist of examples of which colours a person had observed the colour term being used to identify. The model simplifies the task of learning colour terms by concerning itself only with the dimension of hue, and ignoring variations in colour due to differences in lightness or saturation [1]. A consequence of this restriction is that the model will be concerned only with those colour terms which differ principally on the dimension of hue, which in English are 'red', 'orange', 'yellow', 'green', 'blue' and 'purple'.

The model does not treat the observed examples as being completely reliable, but allows for the possibility that some of them may be erroneous [2]. This allows the model to learn even when it is presented with a small number of incorrect examples. (It would be a poor psychological model if a single incorrect example rendered it completely unable to learn, as empirical evidence shows that people have little difficulty in learning from such unreliable evidence.) The full technical details of the model are somewhat complex, and are not relevant to the issues addressed in this paper. For these reasons they will not be described here, although they are specified in detail in Dowman (2001). What is important for present purposes is that the model is able to calculate for every colour just how likely it is that it comes within a colour word's denotation, and that these probabilities may be used to define the degree of membership of each colour in a colour category. In the learned denotations, each colour will have a different degree of membership, though generally we would expect the degree of membership to be greatest towards the centre of a word's denotation. This is because the model will be most certain that these colours can be named by the colour word, while for colours further away from this point the model will be less certain of their membership of the category.

Figure 3 shows one kind of colour term system which is learnable by the bayesian model. The examples from which it was learned were created by selecting random colours from within the section of the colour space corresponding to each colour word. The axis at the bottom of the graph corresponds to hue, with red at the left, moving on to orange, then yellow, green blue and finally purple. Moving past purple would return us to the left hand side of the graph, as the hue dimension is circular, and so red and purple are adjacent to each other. The vertical axis corresponds to the probability with which the model believes that each particular hue comes within the range of colours identified by a colour word, with the top of the graph corresponding to high degrees of membership, and the bottom of the graph to low degrees of membership.



---

[1] Saturation measures to what extent a colour is diluted by grey, with highly saturated colours being the least grey, and most eye catching.

[2] In the simulations reported in this paper, the model has an *a priori* belief that there is a 0.5 probability of examples being erroneous.

Figure 3 A Learnable Colour Term System of a Type which is Unattested Typologically

We can see that the colour term system shown in Figure 3 contains two overlapping colour terms which both denote hues in the red and yellow part of the colour space. The model has observed ten examples of each of these colour terms, and using these it has been able to determine roughly which colours come within each term's denotation. Where the curves are close to the top of the graph the model is very certain that the colours in that part of the colour space come within the denotation of the colour term. Conversely, where the curves are very near the bottom of the graph the model is very sure that the corresponding colours do not come within the colour term's denotation. We can see that both of these terms display the prototype properties which are characteristic of basic colour terms. Each term rises to a single point which is the best example of the term, but the degree of membership in the colour term category declines gradually moving away from that point. When the probability of membership is an intermediate value, this indicates that the model is unsure whether the colours can be denoted by the colour word or not. Hues in these areas correspond to a colour word's fuzzy boundaries, especially where the probability of membership is close to 0.5, in which case the model thinks that it is almost equally likely that the colour comes within a word's denotation as that it comes outside of it.

In contrast the colour term on the right hand side of the graph has quite different characteristics. This colour word was learned from examples generated in the same way as for the other two colour words, but in this case the model has observed 60 examples of colours named by the colour term. There is a range of hues for which the model has a very high degree of certainty that they are members of the colour category, and in this part of the graph the curve is very flat and very close to the top of the graph. However, for almost all other colours the model is very certain that they cannot be named by the colour term, which is indicated by the curve being very close to the bottom of the graph. There are only a very few colours for which the degree of membership is at an intermediate value, and so the boundaries of the colour category are demarcated by almost vertical lines.

This colour term does not have prototype properties, as it does not have fuzzy boundaries, and the degree of membership of colours in the category is almost completely constant throughout its denotation. (The degree of membership does in fact rise to a single maximum close to the centre of the colour category, but this cannot be seen on the graph because both the colour with the greatest degree of membership and those immediately surrounding it have almost identical degrees of membership). This colour term clearly does not resemble the colour terms seen in real human languages, and so these results show that the Bayesian model is able to learn languages with properties which are unattested typologically.

If we look at the colour term system as a whole, we can identify another property of this system which is not in accord with the colour term systems which have been observed in real languages, and that is that the colour terms do not partition the colour space. Rather than having a single word which can be used to name each range of colours, we have two overlapping colour terms with their foci in almost the same part of the colour space, something which is not usually observed in real languages [3]. There is a further inconsistency between this colour term system and those observed in real languages, and that is that there are large gaps between the overlapping colour terms and the other term, so that many colours are left without any corresponding linguistic label. In contrast, empirical evidence shows that colours terms almost always partition the colour space, so that for every colour there is a corresponding colour word which may be used to name it [4]. If Figure 3 corresponded to a real language, then we would

---

[3] MacLaury (1997) identifies a phenomenon that is seen in some languages, that he names 'co-extension', in which two overlapping colour words denote roughly the same range of colours. However, in such cases each colour term tends to have its prototype in a different part of the colour space, so this phenomenon does not correspond to the case of the overlapping colour terms seen here.

[4] Kay and Maffi (1999) do report the existence of a very small minority of languages in which the colour terms do not appear to partition the colour space. However, such colour term systems are exceptional, so it would seem that we are more in need of an explanation of why almost all languages partition the colour space, rather than an explanation of why a minority do not. Once an explanation of why partition occurs has been developed, we may then be able to explain non-partition as a chance

observe a series of colour terms, with little or no gap between them, and only minimal overlapping of neighbouring terms.

1.2 Simulating Colour Term Evolution

The results of the previous section clearly show that the acquisitional model alone is insufficient to explain the empirical data concerning colour term universals, and so the program was extended so that it could model not only learning, but also the social processes in which language is used, and through which it is passed on to each new generation of speakers. Rather than just using a single model of acquisition and presenting it with random examples, multiple copies of the model were created in order to simulate a whole community of people[5]. These artificial people were then made to talk to each other, and to learn from one-another. This process was simulated over a number of generations, until eventually the simulation was stopped and the properties of the emergent language were examined.

In the initial state of the model, each person had observed a single random colour anywhere in the colour space, together with a colour word which had been used to name it. Initially the colour words known by each person were all different, so that there would be no coherent language in the community. Each person was assigned a random age, varying from zero to the maximum age to which people in the simulation could live. The simulations then proceeded by choosing a speaker and a hearer at random (the only restriction being that these could not both be the same person). A colour for the speaker to name would then be chosen at random, and the speaker would find the word which they thought most likely to be a correct label for the colour. This word, together with the corresponding colour, would then be observed by the hearer and remembered by him[6] as an example. He would then use this example to help determine the best word to choose when it came to be his turn to be the speaker. This procedure was then repeated many times, to simulate people talking to each other and using colour terms. However, one time in every thousand, instead of the speaker choosing the best word based on the observations they had made, they would be creative instead, and make up a completely new word. This occasional creative behaviour is necessary, because otherwise there would be no way for new words to enter the language, or for the overall number of terms known within the community to increase.

A parameter in the model controlled how long each person lived for, measured in terms of how many times a person would speak during their lifetime. The actual life span of each person was varied randomly by an amount of up to 20% either above or below the chosen average life span. Once a person reached the end of their life span they would be replaced by a new person with an age of zero who had not observed any colour term examples. (If a person should ever be chosen as the speaker before they had observed any colour terms, then the program would just go back and choose another person instead.)

Figure 4 shows the result of one experiment, where the simulation was run for a period of time equal to ten average life spans, and where on average each person heard 60 examples during their lifetime. This graph shows the colour words learned by one person in this simulation who was near the end of his life span. It shows that a language has evolved which has six colour terms, each of which is focussed in a different part of the colour space, and each of which has prototype properties. The terms roughly partition the colour space, with the colour terms dividing up the colour space with only small overlaps

---

occurrence, especially if the explanation relies on rules which only make statistical rather than absolute predictions.

[5] In all the simulations reported in this paper ten artificial people were used. Varying the number of people used in the simulations does not appear to have a significant effect on the results, but as the number of people is increased the program tends to run more slowly. Clearly a real language community would contain more than ten people, but increasing the number of people simulated would not seem to be necessary for present purposes.

[6] For convenience I refer to all agents in the simulation as though they were male, although, as no distinction is made in the model concerning the sex of the artificial people, this decision is completely arbitrary.

or gaps between them. All the other people in the simulation who were over a certain age had learned very similar colour term systems, each containing the same six terms. (Although the location of the category foci and boundaries varied slightly between people, as each would have observed a unique set of examples.) The colour term systems of the youngest members of the community were somewhat more variable, as these people would not have observed enough data to determine accurately the correct denotation for all the colour terms, and may not even have observed any examples at all of some terms.

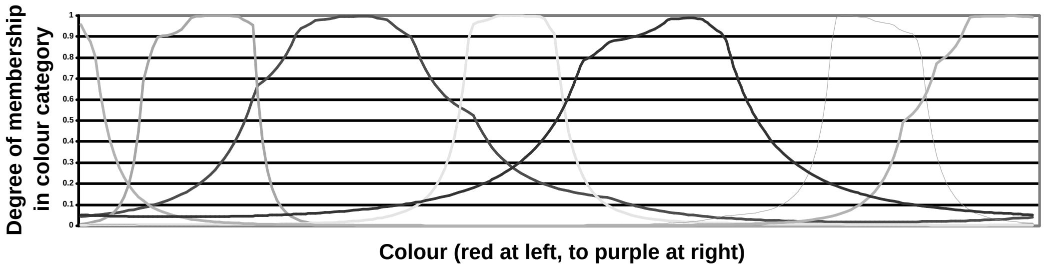


Figure 4 A Colour Term System which Emerged in an Evolutionary Simulation

The simulation has produced a colour term system which appears to have the general properties of colour term systems found in real languages, in that it partitions the colour space, and each term clearly has prototype properties. Repeating the simulation produced similar results, although there was some variation as to the exact number of colour terms which emerged. We might expect that if people observed more colour term examples during their lifetimes, then towards the end of their lives they would learn the denotations of the words with a very high degree of confidence and precision. This would cause words to lose their prototype properties and become like the rightmost term in Figure 3. However, this does not happen in practice during the evolutionary simulations. When the average number of examples which a person observes during their lifetime is increased, the emerging colour term systems tend to have more words, and so the number of examples of each word observed by each speaker remains more or less constant. Conversely, decreasing the number of examples observed by each speaker tends to produce colour term systems with fewer colour terms, but again people will observe a similar number of examples of each term.

3. Discussion

The results of the simulations clearly show that both partition and non-partition colour term systems are learnable by the acquisitional model, as are colour terms with prototype properties and those without. As empirical observations have found that languages do partition the colour space, and that basic colour terms do have prototype properties, it would appear that the acquisitional model fails to sufficiently constrain the range of learnable languages. That at least is the view consistent with Chomsky's (Chomsky 1986) focus on language acquisition and speaker's individual knowledge of language as the primary objects of study in linguistics This is because, if we took as our primary data colour term systems of the type which emerged in the evolutionary simulations, we would reach the conclusion that people are equipped with an innate language acquisition device which forces the learned colour term systems to both partition the colour space and to have prototype properties, as all the systems which emerged in these simulations had those properties. However, in the case of the simulations reported here, we can see that any such conclusion would be completely incorrect, as there is nothing in the acquisitional model which gives any preference to learning colour term systems which conform to the property of partition, and the model is quite capable of learning colour term denotations which do not have prototype properties. The simulations therefore demonstrate that I-language is too narrow a concept to allow us to understand the observed properties of colour term systems.

The extensions which Hurford (1987, 1990) makes to Chomsky's model of language acquisition are uncontroversial, in that it is clear that we learn language from other people, and so the language which provides the input to our language acquisition devices will be determined by other individual's I-languages, and the social context in which language is used. What is controversial about Hurford's model is whether it is necessary to consider the diachronic perspective when understanding central aspects of synchronic language. In the evolutionary simulations of colour term systems we saw new properties emerging which were not predictable from the properties of the acquisitional model, and so this demonstrates that, in this situation, the social processes in which language is used are as important as individual psychology in understanding the properties of colour term systems. This is clearly supportive of de Saussure's (de Saussure, 1959) view that language is simultaneously a social and a psychological phenomenon. It seems that we can only understand the synchronic properties of language through considering the diachronic processes of language evolution, although the nature of diachronic change is determined by synchronic processes.

I believe that this model also exemplifies the value of the computational evolutionary modelling methodology in helping us to gain a better understanding of language. Surprising new properties emerged in the evolutionary simulations, properties which it would have been difficult to predict simply by extrapolating from the properties of the acquisitional model. This raises some interesting questions regarding other computational models of language acquisition. For example, Ellefson and Christiansen (2000) constructed a recurrent neural network which they used to model the acquisition of syntactic rules concerning question formation. They found that the neural network could learn simple artificial languages in which the question formation rules were of the type found in real languages better than it could learn artificial languages in which the question formation rules violated universal constraints on the syntax of question formation. They suggested that this learning bias has caused languages to evolve in such a way that they all conform to what now appears to be a universal rule. However, in the light of the findings of this paper, it would be interesting to investigate whether Ellefson and Christiansen's model would in fact produce languages with the predicted properties if a community of speakers was modelled over several generations, and whether any other unexpected properties would emerge. At present most acquisitional models take too much time to learn to make such simulations practicable, but as computers become more powerful there will be increasing opportunities to make use this kind of evolutionary methodology.

There is one major characteristic of basic colour term systems which the model makes no attempt to explain, and that is the typological patterns of colour terms across languages. Berlin and Kay (1969) showed that there were regularities in the way that languages lexicalise the colour space, so that the properties of colour term systems are partly predictable. For example, if a language has only two colour terms then these will divide up the colour space so that one term denotes white, yellow, red and very light colours, and the other denotes black, blue, green and very dark colours. We never see colour terms which denote both red and blue colours, or terms where yellow or red are grouped with black instead of with white. Berlin and Kay proposed that the development of colour term systems follows an evolutionary sequence, where the number of terms grows from two to eleven over the course of several generations. Biases which made some colour term systems more learnable than other ones could cause languages to evolve in such a way that at each stage of their evolution they conformed to the observed typological patterns. Work is in progress to add such learning biases to the acquisitional model, so that the plausibility of Berlin and Kay's evolutionary hypothesis can be investigated. If the model is then able to account for the typological data then this will clearly support Berlin and Kay's evolutionary hypothesis, but it will also provide further support for the validity of the present model.

In conclusion, this paper has presented evidence which suggests that colour term systems partition the colour space as a result of diachronic processes, and that there is no reason to suppose that any component of an innate language acquisition device, or any aspect of the ontogenetic process, prevents us from learning basic colour terms which either overlap, or which leave large ranges of colour without any corresponding colour term. It was also proposed that the mechanism that we use to learn colour terms may be equally able to learn colour terms which have prototype properties as ones which do not. Colour terms in real languages may have prototype properties only because of the social processes which moderate the number of colour terms which emerge in a language. In general, the synchronic properties of a language may best be understood by placing the language in a diachronic context, and so using existing acquisitional models as part of an evolutionary simulation may increase their explanatory power, thus demonstrating the importance of the computational evolutionary approach in linguistics.

4. Acknowledgements

I would like to thank everyone who has helped me with this paper or given me any feedback on my work on colour terms and Bayesian inference, including Judy Akinbolu, Tony Belpaeme, Tom

Griffiths, Emmanuel Letellier , Rob MacLaury, my Ph.D. supervisor Judy Kay. I am also grateful for financial support from the Australian government and the University of Sydney in the form of IPRS and IPA scholarships respectively.

5. References

Bayes (1763). An Essay Towards Solving a Problem in the Doctrine of Chances. *Philosophical Transactions,* Volume 53, pages 370-418.

Berlin, B. & Kay, P. (1969). *Basic Color Terms.* Berkeley: University of California Press.

de Boer, B. (2001) *The Origins of Vowel Systems*. Oxford: Oxford University Press.

Chomsky, N. (1972). *Language and Mind*. New York, NY: Harcourt Brace Jovanovich Inc.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use.* New York, NY: Praeger.

de Saussure, F. (1959). *Course in General Linguistics*. Eds. C. Bally and A. Sechehaye in Collaboration with A. Reidlinger, trans. Wade Baskin. London: Peter Owen.

Dowman, M. (2001). A Bayesian Approach to Colour Term Semantics. *Lingu@scene,* Volume 1.

Ellefson, M. R. & Christiansen, M. H. (2000). The evolution of subjacency without Universal Grammar: Evidence from artificial language learning. In *Proceedings of the 22nd Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hurford, J. R. (1987). *Language and Number The Emergence of a Cognitive System*. New York, NY: Basil Blackwell.

Hurford, J. R. (1990). Nativist and Functional Explanations in Language Acquisition. In I. M. Roca (ed.), *Logical Issues in Language Acquisition.* Dordrecht, Holland: Foris Publications.

Kay, P. & Maffi, L. (1999). Color Appearance and the Emergence and Evolution of Basic Color Lexicons. *American Anthropologist,* Volume 101, pages 743-760.

Kirby, S. (2000). Syntax without Natural Selection: How Compositionality Emerges from Vocabulary in a Population of Learners. In C. Knight, M. Studdert-Kennedy and J. Hurford (eds.) *The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form.* Cambridge: Cambridge University Press.

MacLaury, R. E. (1997). *Color and Cognition in Mesoamerica: Construing Categories as Vantages.* Austin, Texas: University of Texas Press.

Taylor, J. R. (1989). *Linguistic Categorization: Prototypes in Linguistic Theory.* Oxford: Oxford University Press.