Evolution of Basic Colour Terms

Mike Dowman

Theoretical and Applied Linguistics

School of Philosophy, Psychology & Language Sciences

The University of Edinburgh

mdowman@ling.ed.ac.uk

Abstract

Cross-linguistic typological patterns in basic colour term systems were investigated using computer modelling. An evolutionary expression-induction model was used to simulate the evolution of basic colour terms over several generations. The model used simulated people which were able to infer the denotation of basic colour terms, using a Bayesian inference procedure. The simulated people would repeatedly name colours, so producing data from which other people could learn. When the universal foci were made especially salient, the languages emerging over a number of generations replicated the typological patterns seen in the 110 languages of the world colour survey. It was found that if random noise was added to the data from which the artificial people learned, this had no major effect on the emergent languages. This shows that the Bayesian inference procedure is able to learn effectively even in realistic noisy environments, and thus provides further evidence to support the validity of the model.

1 Introduction

This paper tries to explain well documented typological patterns seen in colour term systems cross-linguistically. In most, if not all, languages a small set of *basic* colour terms, which are highly salient psychologically and which are in frequent use, can be distinguished from other words used to denote colour. The number and denotational ranges of these terms varies greatly between languages. However, Berlin and Kay (1969) were able to show that there are clear typological patterns concerning the denotations of colour terms across languages, and their findings, whilst undergoing several revisions, have been broadly confirmed by subsequent studies. The work reported in this paper used a computer model to simulate the evolution of colour term lexicons, and investigated whether the emergent languages reproduced the empirical typological data. It was found that the typological patterns could be accounted for by the computer model, even when the simulations were made more realistic by adding communicative misunderstandings, and therefore creating a large amount of misleading linguistic data.

Berlin and Kay (1969, p6) listed a number of criteria that they used to distinguish basic colour terms from other words used to denote colour. They considered colour terms to be basic only if they were known by all speakers of the language, they were highly salient psychologically, they did not just name a subset of the colours denoted by another colour word, and their meanings were not predictable from the meanings of their component parts. They also provided some further criteria to deal with any doubtful cases. Application of these criteria seems to distinguish clearly between basic and non-basic colour terms in most languages, although there can still remain some questionable cases. The application of these criteria to English, results in the set of 11 basic colour words, *red, yellow, green, blue, orange, purple, pink, brown, grey, black* and *white*, excluding terms such as *crimson*, *blonde* and *royal blue* (Berlin and Kay, 1969). It is possible that other words should be considered basic for some speakers (for example *turquoise*, *cream* or *beige*), but all of these terms could be excluded on grounds of salience, each of them being much less frequent than any of the words usually considered to be basic colour terms in English[i].

Berlin and Kay (1969) examined a sample of 98 languages, and found that there was very wide variation in the actual ranges of colour denoted by each colour term in different languages. However, they found that this variation was certainly not without limit, as had been presumed by earlier researchers (for example Gleason, 1961). While the number of colour terms varies between languages, which combination of colour terms exist in any given language seems to be partly predictable. The work reported in this paper attempted to address the issue of why such regularities exist.

Berlin and Kay suggested that the regularities were the product of an evolutionary process in which languages gradually evolved from an initial state in which they had only two basic colour terms, and in which more terms were added in predictable orders. One of the goals of this paper is to investigate whether this hypothesis could explain the typological patterns seen in colour term systems. While Berlin and Kay claimed that the patterns were due to an evolutionary process, they left the details of this process unspecified. Cultural evolution of language[ii] is realized through a process in which language is passed from speaker to speaker over a number of generations. The language learned by each subsequent generation will not be exactly the same as that spoken by the previous one, so the language will gradually change over time. Any complete theory of the evolution of colour terms should specify exactly how colour term systems are transmitted between generations, and what properties of either people or their environment are responsible for creating the attested typological patterns.

In order to provide a fully explicit and rigorous theory concerning the evolution of colour terms, a computer model was constructed (Dowman, 2003a). This model aimed to implement the most essential elements of the processes through which colour terms evolve over a number of generations, including specifying how their denotations could be learned. The initial simulations were somewhat artificial in nature, as the learners in those simulations always received completely accurate information about which colours the other speakers were referring to with each colour term. This is somewhat unrealistic, as in reality there are many factors that could interfere with successful language learning, so this paper reports more recent work that has extended the original model by adding 'noise' to the simulations in order to make them more realistic.

2 Colour Terms across Languages

Before going on to consider how colour term typology can be explained, it is first necessary to obtain a clear understanding of the data concerning the properties of basic colour terms, both within individual languages and cross-linguistically. The first property of note is that basic colour terms have *prototype properties*. Colour terms do not simply denote a uniform range of colours, but instead some colours are members of the colour category corresponding to the colour term to a greater extent than other colours are. For each colour term, there will usually be a single colour which speakers of the language consider to be the best example of that colour term, and this colour is called the prototypical colour. The more dissimilar a colour is from the prototype, the less good a member of the colour category it is. At some level of dissimilarity it becomes difficult to determine whether colours come within the colour category or outside of it, and these colours mark the category's fuzzy boundary (Taylor, 1989). There is little agreement between speakers of the same language as to where the colour term boundaries should be placed, but there is usually close agreement on the location of the prototype colours (Berlin and Kay, 1969).

Berlin and Kay (1969) investigated the range of colour term systems in a wide range of languages. Their study gathered data from speakers of twenty different languages, using arrays of Munsell[iii] chips. Firstly the basic colour terms of each language were determined, using the criteria discussed above. This produced the first finding of Berlin and Kay's study, which was that all languages appear to have between 2 and 11 basic colour terms[iv]. Berlin and Kay also used data from published sources such as dictionaries and grammars to bring the total number of languages in their study to 98, and that data appeared to confirm the finding.

The next stage of the research involved asking each person to both map the outer boundary of each of the basic colour terms in their language on an array of Munsell chips, and to identify the best or most typical examples (the prototype) of each term. They discovered that the boundaries of the areas of colour denoted by colour terms varied greatly between languages, which was consistent with earlier findings. However, they found that most of the prototypes were placed in just a few areas of the colour space, either clustering on single Munsell chips, or on a small number of nearby chips, leaving over 70% of the surface of the Munsell array clear of any prototypes at all. They showed that there were 11 clusters of prototypes, corresponding to the locations in which English speakers would place the best examples of each of the basic colour terms in English. This clearly showed that the lexicalization of the colour space in unrelated languages was certainly not random or completely arbitrary, but appeared to conform closely to universal restrictions, and this has recently been supported by rigorous statistical analyses (Kay and Regier, 2003; Regier et al, 2005).

A further finding emerged when Berlin and Kay investigated the combination of colour terms existing in any particular language. They found that when terms were classified based on their prototypes, it was largely predictable which colour terms would exist in a language, if the number of basic colour terms in the language was known. Berlin and Kay expressed these regularities using the implicational hierarchy shown in Figure 1. This hierarchy was constructed partly using evidence derived from participants using Munsell arrays, though evidence for most of the languages used to construct this hierarchy came only from published sources such as dictionaries. All languages appeared to have a term with its prototype at white, and a term with its prototype at black, shown at the left of the hierarchy, but some languages had no other basic colour terms but these. However, if a language had a term for any of the colours further right in the hierarchy, it always had terms for all the colours further left in the hierarchy. For example, if a language had a term with its prototype at yellow, then it would also have terms with their prototypes at white, black and red.

FIGURE 1 ABOUT HERE

It should be noted that while Berlin and Kay proposed that this hierarchy described the general patterns seen in colour term systems cross-linguistically, they did acknowledge the existence of some exceptions and problematic cases. However, given that only six out of the 98 languages in the study appeared to be seriously problematic, Berlin and Kay did not modify their theory, but instead simply decided to treat these languages as exceptions. Since Berlin and Kay published their original study, there has been a great deal of interest in basic colour terms, and much more data has been collected, generally using methods which address the deficiencies of Berlin and Kay's original study. These studies have in large part confirmed Berlin and Kay's original findings, though several modifications have been made to their hierarchy in order to accommodate language types that were not attested in their original study, or that were originally treated as exceptional aberrations (Kay, 1975; Kay et al, 1991).

A very large survey of the colour term systems of 110 minor languages, the World Colour Survey (Kay et al, 1997), has now produced a wealth of high quality data, allowing us to get a much more complete picture of colour term systems worldwide. Using this new data, Kay and Maffi (1999) produced a new classification of colour term systems, which has modified the original hierarchy of Berlin and Kay (1969) considerably, but which still shows that the attested colour term systems are only a small subset of those which are logically possible. Kay et al (1997) noted that it appears that there are six universal focal colours, corresponding to the colours which would typically be the prototypes of red, yellow, green, blue, black and white colour terms, and that the order of appearance of basic colour terms which do not include one of these colours in their denotations, such as the English terms *purple, orange, pink, brown* and *grey*, is less predictable. Their classification of colour term systems was made primarily by considering only terms whose denotation included at least one of the universal foci. These classifications were then simply augmented with a list of which other basic colour terms existed in the language.

However, Kay et al (1991) did note that while purple or brown terms may be seen in languages which do not have separate terms for both green and blue, contrary to Berlin and Kay's hierarchy, it seems that orange or pink terms do not normally appear unless a language has separate terms for both green and blue.

Kay (1975) had already noted that grey terms sometimes appear in languages even when those languages have not developed terms for some of the other colours which Berlin and Kay predicted would normally appear before grey. The general conclusion that we can draw from these findings is that the order in which such terms emerge in a language does not seem to be completely predictable, though there appear to be general trends concerning the order in which these terms emerge. We can note that orange terms tend to be seen only in languages which have already developed purple terms, but that this is not always the case (MacLaury, 1997), and that purple terms tend to emerge once a language has acquired separate terms for green and blue, but that this rule does not apply to all languages. These specific findings are among the data which the evolutionary computer model described here was able to explain.

Another important difference between the analysis of Berlin and Kay (1969) and that of Kay and Maffi (1999) is that, while Berlin and Kay classified terms based simply on the locations of their prototypes, Kay and Maffi paid more attention to terms' full denotational ranges. They classified colour terms depending on which universal foci they contained, rather than just in terms of which universal foci corresponded to the term's prototype, so that, for example, two terms which both had red prototypes would be classified differently if one also named a range which included yellow, but the other did not.

The simplest kind of basic colour term system contains just two colour terms. Systems of this type can be divided into one of two types. The first type of system contains one term denoting all light colours, together with yellow and red hues of medium lightness, and another term denoting all dark colours, together with green and blue hues of medium lightness (Heider and Olivier, 1972). The second type of such systems simply has one term for light colours, and another for dark (MacLaury, 1997). However, in both cases the colour term denotations extend so that they effectively cover the whole of the colour space. This seems to be an almost universal property of basic colour terms systems, although Levinson (2001) reports that there are some exceptional languages in which parts of the colour space have no corresponding basic colour term.

Kay and Maffi (1999) proposed that languages evolve from a state in which they have only two colour terms, and that they then gradually add more terms over time, never losing colour terms once they have gained them. Their classification of languages therefore takes the form of an evolutionary sequence, which begins with two colour terms, and then progressively subdivides the areas of the colour space named by each of these terms until each of the universal foci is named by a separate colour term. They found that 83% of the languages in the World Colour Survey lie somewhere along the trajectory shown in Figure 2, where early stage languages with only two colour terms are at the top of the diagram, in which case one term names light colours together with yellow and red, and the other dark colours together with blue and green[v]. Languages in which each of the universal foci is represented with a separate term are at the bottom of the diagram, and intermediate languages lie somewhere in between. Languages were considered to lie on this trajectory if they appeared to be best classified either as being at one of the five stages, or as being in transition between stages. In a language in transition, there would typically be disagreement between speakers as to how many basic colour terms the language possessed, usually because older speakers would not use a colour term which had entered the language during their lifetimes. Out of 91 languages which Kay and Maffi (1999) placed on this trajectory, 18 were considered to be in transition.

FIGURE 2 ABOUT HERE

Kay and Maffi (1999) acknowledged that not all languages seem to follow the above trajectory, at least not for the whole of their development. They placed seven languages either on a side branch of the main trajectory, or in transition into or out of a side branch. Once a language has reached the second stage of the trajectory shown in Figure 2, it usually gains an extra colour term, so that a black term and a green-blue composite term come to replace the black-green-blue term. However, it seems that a few languages instead split apart the red-yellow composite and replace it with separate red and yellow terms, leaving the black-green-blue term intact. One language in the World Colour Survey seems to be in transition into this state, and two appear to be in transition out of it.

Once a language has reached this stage, it seems that there are two routes it can take. Firstly, the black-green-blue composite can split into a black term and a green-blue composite, in which case the language will return to the main line of the trajectory with a five term system. However, another possibility is that the black-green-blue composite splits to produce a black-blue composite and a separate green term. Three languages in the World Colour Survey were given this classification, while one was analyzed as being in transition into this state, and one as being in transition out of it. Following this stage, the black-blue composite splits, returning the language to a state on the main line of the trajectory.

In order to explain the existence of three other languages, Kay and Maffi (1999) had to postulate another branch to the evolutionary trajectory, although in this case its origin was somewhat unclear, as it could not have developed straightforwardly from the main line of the trajectory. This branch was proposed because two languages in the World Colour Survey had yellow-green-blue composite terms, together with separate black, white and red terms. At the first stage of the main line of the evolutionary trajectory, the yellow and green universal foci appear in separate composites, so languages of this type could not emerge simply through progressive splitting of composite colour terms. There was also one language which contained a yellow-green composite, together with black, white, red and blue terms. This language could be derived from the earlier type if the yellow-green-blue term split into a yellow-green term and a separate blue term, but this does not provide an explanation of how the yellow-green-blue composite arose in the first place.

There were a few languages in the World Colour Survey which did not seem to fit well into Kay and Maffi's (1999) theory of a limited number of fixed evolutionary trajectories. Firstly there were three languages which appeared to be in transition directly from a state in which they had a black-green-blue composite term and separate white, red and yellow terms, to a state in which each of the universal foci was represented by a different colour term. This missed out the expected intermediate stages in which the black-green-blue term would be expected to split first into either black and green-blue or green and black-blue terms. There were also four languages which Kay and Maffi could not place anywhere on the evolutionary trajectories. Each of these languages contained colour terms with their prototypes at black, white and red, but the way in which they divided up the rest of the colour space was inconsistent, as speakers did not agree on the set of basic colour terms, or on the denotations of those terms.

Some linguists have challenged the general findings of Berlin and Kay (1969) and Kay and Maffi (1999), suggesting that colour term systems do not conform to predictable rules to the extent that those researchers have claimed. Saunders (1992) has gone so far as to suggest that some of Berlin and Kay's findings were a product of their methodology, and has suggested that such words can only be understood in relation to other words in the language, and within the context of their secondary connotations and of the belief systems of which they form a part. Levinson (2001) suggested that the idea of colour as a domain for linguistic categorization, in which we would expect to find a set of co-hyponyms each denoting specific ranges of colour might not be applicable to all languages. He noted that some languages conflate other properties, such as texture or variegation, with colour. This contributes to difficulties in determining which colour terms in particular languages are basic, because it makes it more difficult to decide which words are colour terms at all.

However, none of Levinson's (2001) findings seem to be at odds with the basic conclusions of Kay and Maffi (1999). Their data still demonstrate regularities in the way in which informants name colour chips, regardless of how many other aspects of the colour words' meanings or syntactic properties are disregarded. The distinction between basic and non-basic colour terms is clearly not completely arbitrary, and while it seems likely that whether a marginal term was classed as basic or not might in some cases have been determined partly by whether it supported or contradicted Kay and Maffi's trajectories, such an issue cannot be said to falsify Kay and Maffi's theory. We should remember that, ever since Berlin and Kay (1969), exceptions to the hypothesized evolutionary trajectories have been acknowledged, and so, even if classifying a term regarded as non-basic as basic would result in a language no longer fitting on an evolutionary trajectory, this would not be a serious problem for Kay and Maffi's theory of predictable trajectories.

MacLaury (1997) reports some types of basic colour terms that were not included in Kay and Maffi's trajectories. Firstly he notes that some languages have basic colour terms naming broad ranges of desaturated colours, especially terms which include dull brownish, lavender, grey and/or beige colours. Secondly he reports that it is common to see two colour terms that name approximately the same range of colours, in which case one term will be *dominant* and have its prototype near to the centre of the category, and the other *recessive*, with its prototype near the edge. Any explanation of colour term typology should therefore be flexible enough to account for such data, but the particular data against which the performance of the model presented in this paper was measured was that reported by Kay and Maffi (1999).

3 Psycholinguistic and Neurophysiological Findings

Attempts have been made to relate the findings concerning basic colour terms and their distribution across languages to more fundamental properties of the human mind and the human visual system. Hering (1964) noted that red and green appear to be opposite colours, and that the same is true for yellow and blue, because no colour can appear to be reddish-green or yellowish-blue. Hering proposed that the opponency of red and green, and of yellow and blue, was due to the physiology of the human visual system, although he did not have any direct evidence of this. De Valois et al (1966) and De Valois and Jacobs (1968) found evidence that there are opponent cells in the lateral geniculate nuclei of the brains of Macaque monkeys (and therefore also presumably humans) that respond maximally to either red, yellow green or blue light, and minimally to the opposite colour. It was initially assumed that these colours corresponded to the universal foci, and therefore were responsible for universals in colour naming (Kay and McDaniel, 1978). However, it has now become clear that the universal foci do not correspond exactly to what would be expected if they were a direct product of the neural responses of opponent cells (Derrington et al, 1984). Therefore, any neurophysiological basis for the universal foci is at present only hypothetical.

However, regardless of whether or not the location of universal foci is determined by opponent cells, their existence is supported both by the typological literature discussed above, and also by a wide body of psychological research. Firstly, De Valois and De Valois (2001) reported that if a focal red and a focal green light are added together in equal proportions, they will cancel each-other out, so that a neutral grey colour is perceived. Furthermore, after staring at a red surface, a green after image will be seen. Similar effects are observed for yellow-blue opponency. Both of these results indicate a special status for the universal foci.

Further evidence has come from work aimed specifically at explaining properties of colour language. Rosch (who published her earlier work under the name of Heider) found that if children were shown several Munsell colour chips and asked to select one, colours that form the prototypes of English colour terms were picked out much more often than other colours, suggesting that they were especially salient. She also showed that children can pick out the matching colour in an array of colour chips much more easily when shown a prototype colour as compared to a non-prototype one (Heider, 1971). Heider (1972) showed that monolingual speakers of Dani, a language with only two basic colour terms, could remember the colour of colour chips corresponding to the prototypes of English colour terms better than the colour of other chips. This was taken as evidence confirming the hypothesis that these colours are more memorable for all people, regardless of what language they speak. Heider (1972) also showed that Dani speakers found it easier to learn words for prototype rather than for non-prototype colours, and Rosch (1973) demonstrated that Dani speakers found artificial colour categories easier to learn when they had a prototype colour at their centre, compared to cases in which the prototype was peripheral, or the category did not include a prototype colour. Rosch suggested that this showed that these colours were more easily kept in long term memory, which is presumably why they tend to form the prototypes of colour terms. In the studies, Rosch looked at the prototypes of the English words *red*, *yellow*, *green*, *blue*, *purple*, *orange*, *brown* and *pink*, but she suggested that the red, yellow, green and blue foci, were more memorable than the other ones.

Roberson, Davies and Davidoff (2000) sought to replicate some of Rosch's work, but this time using speakers of Berinmo, a language with just five basic colour terms, as well as British English speakers, as participants. They showed that some of Rosch's results could have been influenced by the languages spoken by the participants, and that others may simply have been due to participants' tendency to always point at prototype colours in the Munsell array, regardless of whether they were shown a prototype or non-prototype colour chip. They were also unable to show that Berinmo speakers found it easier to learn names for prototype colours than for other colours. Lucy and Shweder (1979) showed that, by using a different Munsell array, which had been corrected so that all chips were equally discriminable, then in most cases prototype colours were not remembered better than other colours. Hence it would seem that we can produce or remove effects concerning the special status of the prototype colours simply by changing the colour array used in experiments. As it is difficult to argue that any particular array of colour chips is the correct one to use, this would seem to invalidate results that could be explained in terms of ease of discrimination.

Roberson et al and Lucy and Shweder's studies have cast doubt on much of the evidence concerning the universal special status of prototype colours. However, the cross linguistic evidence clearly demonstrates the special status of the prototype colours, because, cross-linguistically, most people place the prototypes of basic colour terms on the universal prototype colour chips, or else chips immediately adjacent to them. The evolutionary computer model described in this paper aims to give an explanation of colour term typology, suggesting that it is the product of evolutionary processes occurring under the influence of universal biases. The model rests on the assumption that the red, yellow, green and blue universal foci are especially salient, which is supported by both psychological and linguistic evidence, even though not all of the studies are in complete agreement.

4 Evolutionary Modelling

Given the typological data reviewed above, we need to consider just what kind of explanation is appropriate to account for it. Chomsky (1965, 1972, 1986) has emphasized that languages can be seen primarily as psychological phenomena, in that the ability to use and understand language is an ability that we have as individual people. Given this perspective, regularities across languages are generally seen as products of an innate Language Acquisition Device. The possible human languages are those which the Language Acquisition Device is able to learn, and so, if a particular linguistic structure is not attested, then this is presumably because the Language Acquisition Device is not capable of acquiring it.

Kay and McDaniel (1978) came close to Chomsky's position in proposing a limited set of universal colour categories, which were determined by the neural response functions of cells in the retina of the eye. They proposed that some of the universal properties of colour term systems were due to all the colour categories in the world's languages being chosen from this finite inventory. This was clearly an attempt to explain properties of the colour term system in terms of innate structure, and so was in this way similar to Chomsky's Universal Grammar. However it appears that on the one hand Kay and McDaniel's proposal was too restrictive, because there is considerable variation in the exact denotations of similar colour terms in different languages, but also that it was not sufficiently constraining, in that it predicted the existence of types of colour categories that have never been observed.

However, explaining language as a primarily psychological phenomenon does not necessitate a strongly nativist position. For example, linguists such as that of Rumelhart and McClelland (1986), who modelled the acquisition of the past tense of English verbs, have retained Chomsky's focus on a psychological, ontogenetic account of linguistic phenomena. Their neural network, together with the algorithm used to train it, can be seen as a Language Acquisition Device. At the end of the learning process, the final knowledge of language will correspond to the neural network and the learned connection strengths. Rumelhart and McClelland's model reproduced many of the patterns observed when children learn English, and so it was argued that these patterns could be explained as resulting from properties of the children's Language Acquisition Devices, which they have suggested may learn in a similar way to the neural network (Rumelhart and McClelland, 1986, p267).

This paper explains language in terms of a process that is both psychological and social. Hurford (1987, 1990) put concepts of language involving social dimensions on a more concrete footing, by placing Chomsky's (1972) concept of a Language Acquisition Device within a social context. Chomsky's conceptualization of language acquisition neglects to specify how the data from which people learn, the primary linguistic data, is produced. Hurford noted that this data is produced by other speakers of the language, and so Chomsky's acquisition-based explanation of language can be extended to include the processes in which language is produced and therefore passed on to subsequent generations of speakers, as shown in Figure 3 (adapted from Hurford, 1987, p22).

FIGURE 3 ABOUT HERE

The new component in Hurford's concept of language, which is missing from Chomsky's, is the arena of language use, through which language passes from one generation of speakers to the next. The arena of language use is partly psychological, but it also includes all those factors relating to the context in which we communicate, and the aspects of the world that determine what we talk about, and so what utterances (or written language) are produced to form the primary linguistic data for the next generation of speakers. While Hurford's spiral retains Chomsky's acquisitional paradigm, it allows for a wider range of possible explanations for linguistic phenomena. Explanations of linguistic phenomena under Chomsky's model are limited to a single generation, but Hurford's spiral allows for the possibility of explaining language universals and language typology in terms of diachronic processes.

Most computer models of language, such as that of Rumelhart and McClelland (1986) have focussed on the issue of acquisition, and on psychological representations. However, there are a number of computer models, called *expression-induction models* (Hurford, 2002) or *iterated learning models* (Brighton and Kirby, 2001), that make use of Hurford's diachronic model of language, and therefore explain linguistic phenomena as part of an evolutionary process. The expression-induction methodology involves creating a number of artificial people (often referred to as *agents*), who are able to infer a simple language based on example utterances, and then to express themselves using that language. The expressions of the agents will form the input from which the next generation of learners will induce their grammars (although in some models all artificial people both express language and learn from other people at the same time, so there need not be any clear cut generational gap). Generally, after a period of time, all speakers will come to share a common language, although the internalised languages of each individual person may in some cases be slightly different. Because language is transmitted from person to person indirectly, via the arena of language use, the transmission of language between generations may be imperfect, resulting in language drift, which can be viewed as a cultural evolutionary process. The expression-induction methodology originated with Hurford (1987), and there now exist a wide range of expression-induction models, explaining phenomena as diverse as syntactic compositionality (Kirby, 1999), and vowel system typology (de Boers, 1999).

Dowman (2003b) used a model of colour term evolution to show that the kinds of colour word systems that emerged in an expression-induction model were only a subset of those that could be learned by the acquisitional part of the model, demonstrating the importance of the evolutionary aspect of such models. Belpaeme (2002) also constructed an expression induction model of colour term evolution, although the details of his model were somewhat different from the model presented here, as were the aspects of colour term systems for which it was able to account. Belpaeme's simulations typically contained ten artificial people, each of whom was able to represent colour categories using adaptive networks, a kind of neural network. Colour in the model was represented in terms of the CIE L*a*b* space, which represents colour in terms of three dimensions, one of which corresponds to its degree of redness or greenness, one to the degree of yellowness or blueness, and the third to the lightness or darkness of the colour[vi]. The networks acted as fuzzy membership functions, allowing colour categories corresponding to a volume of the three dimensional CIE L*a*b* space of almost any size or shape to be represented. Each artificial person could also remember a number of word forms, each of which could be paired with a colour category.

In the initial state of the simulation, the artificial people did not know any colour categories or colour words, so, the first time one of them spoke, they would have to create a new category and corresponding word. In general, communication proceeded by first choosing one colour to be a topic, and several others to form a context, and then choosing one person to be a speaker, and another to be a hearer. The speaker would then try to communicate to the hearer which of the colours was the topic, by choosing a word which included the topic, but not the context, in its denotation. If the word that the speaker used was known by the hearer, and the colour category which the hearer had associated with that word included only one of the colours, then the hearer would understand that that colour was the topic. If this was correct, then communication would have been successful, and so the association between the topic colour and the colour word would be strengthened. When communication was not successful, the correct topic would be shown to the hearer, and the hearer would then adapt the word's colour category to reflect this new information. The artificial people would eventually forget categories and words that were not used.

The simulations varied with respect to whether the same artificial people existed for the whole simulation, or whether people were periodically replaced with new people who did not know any colour words, so that the transmission of colour word systems between different generations of people could be simulated. However, in both of these conditions, after a period of time, coherent colour lexicons shared by all the artificial people emerge, as the colour space was divided up into a number of regions, each of which was named by a different colour word. While there was some variation in the colour term systems learned by each individual person, their languages were consistent enough for them to be able to communicate successfully in more than 85% of all interactions. There was, however, little correlation between the colour term systems emerging in Belpaeme's model and those observed empirically, as Belpaeme's model did not reproduce the typological patterns seen in colour naming[vii].

The model of colour term evolution presented in this paper builds on the work of Belpaeme (2002). It further investigates the properties of a computer model first reported in Dowman (2003a), which was also an expression-induction model of colour term evolution. However, the learning mechanism used is Bayesian inference, not adaptive networks, and the artificial people receive no feedback concerning whether communication has been successful. Instead they simply try to mimic each other's language, though they receive no reward for successful imitation. Some of the findings of the psychological studies of colour perception and colour naming reviewed above were incorporated into the model, and this allowed it to account for much of the typological data concerning colour words, as will be shown below. In addition, it is shown that the model presented in this paper is extremely robust to noise, and so colour term evolution is not greatly affected when large quantities of random data is added to the data from which colour terms are learned.

5 The Evolutionary Model

The expression-induction model of colour term evolution combined a Bayesian model of colour term acquisition, and an evolutionary model which contained artificial people who learned colour term denotations using a Bayesian model that was originally described in Dowman (2004). They then communicated about colour using the learned denotations. The Bayesian model allowed the denotations of colour words to be inferred based on examples of colours which those words had been used to identify. The model was used to simulate the evolution of languages over several generations of speakers, so that the properties of the emerging languages could be investigated.

The acquisitional model used a circular hue space to represent the range of possible colours. No attempt was made to model the dimensions of lightness or saturation, and hence the model is unable to account for colour terms that are distinguished primarily on these dimensions, such as English *white*, *black*, *brown*, *pink* and *grey*. This simplification made the creation of the computer model much simpler, and it remains possible to investigate highly saturated colour terms, such as English *red, orange, yellow, green, blue* and *purple*. There is no reason in principle why the model could not be extended so that it made use of the full three-dimensional colour space, and this remains an aim for future research. It should be noted, however, that a previous model of basic colour terms, that of Kay and McDaniel (1978), was also essentially restricted to one dimension, as the fuzzy set operations which Kay and McDaniel proposed were only specified for single dimensions.

The acquisitional part of the model used Bayes optimal classification (Bayes, 1763; Barnett, 1982) to determine the denotation of colour terms from a number of example colours believed to be correct denotata of the colour term. Bayes optimal classification is a statistical technique, and, if the possibility of erroneous data is considered when a Bayesian inference procedure is formulated, then it can be made robust to noise. The Bayesian model of colour term acquisition assumed that only half of the colour examples it received were accurate, and that the other half were completely random, and hence gave no indication of a colour terms' correct denotation. However, the model did not receive information about which, if any, examples were erroneous, so it would have to rely on the Bayesian procedure to determine this.

The data from which denotations were learned was a set of example hues which had been used to name the colour term whose denotation was being learned. Each colour term was learned independently, so example denotata for other colour terms would not be considered in this process. The model assumes that colour terms will denote a contiguous part of the colour space, but has no inbuilt restrictions or biases to assume that the denotation will be of any particular size, or that it is more likely to correspond to one part of the colour space as opposed to any other. Therefore a learned denotation could cover almost the entire colour space, just a small range of colours, or anything in between. The Bayesian mechanism allows the model to assign a probability to each hue in the colour space, corresponding to how likely it is that that hue can correctly be denoted by the colour term. These probabilities can be equated with degree of membership in the colour term category, and hence the model effectively produces fuzzy sets describing each colour term's denotation, with each hue's membership varying from 1 (full membership) to 0 (definitely not a member).

As described up to the present point, the colour space in the model is completely uniform, with all parts of the colour space having identical properties. Such a model could clearly not explain the typological patterns seen in basic colour term systems, so the model was modified in accordance with the linguistic and psycholinguistic evidence discussed above. The four universal foci, red, yellow, green and blue, were made especially salient. Heider (1972) provided psycholinguistic evidence to suggest that people find these colours especially salient, and therefore remember them more easily. Hence the model was modified so that examples of possible denotata of colour words were more likely to be remembered by the artificial people when they corresponded to these specific colours. More precisely, the artificial people would always remember colour examples that fell at exactly these points, but would only remember one in twenty of other examples.

It was also proposed that the universal foci are not equally spaced in the conceptual colour space that is used in colour classification. The foci of green and blue were placed the closest together of any neighbouring universal foci, whilst blue and red were hypothesized to be furthest apart. The red and yellow and yellow and green foci were placed at intermediate distances, with the yellow and green foci somewhat closer together than the red and yellow foci. Independent verification for these parameter settings is somewhat weak, although MacLaury (1997) does suggest that there is some evidence that the green and blue foci are more similar than the other universal foci. Hence the primary justification for these parameters is provided through the results of simulations using the model. As these distances are in terms of a conceptual, psychological space, rather than any physical measurement, it is not clear that we will ever be able to get any more direct evidence of the true conceptual distances.

The model, as described up to this point, is able to learn colour term denotations when presented with example colours, and therefore can be used as the induction part of an expression-induction model. In order to create a full expression-induction model, Dowman (2003a) incorporated it into a simple evolutionary simulation, in which there were ten artificial people, each of which would remember their own set of colour terms and corresponding colour examples, and generalize to new colours using the Bayesian model.

Simulations were begun with each person knowing a different colour term, and a single example colour that it could be used to name. Each stage of the simulation would then involve choosing one person to be the speaker, and another to be the hearer. A colour would then be chosen for the speaker to name, and they would determine to what extent that colour was a member of the colour category corresponding to each colour term that they knew. They would then say the colour term in whose colour category the colour had the highest degree of membership, and the hearer would observe this term, together with the colour. The colour would then serve as an example of the colour term for the hearer the next that they spoke. In order to allow new colour terms to enter the language, one time in a thousand, instead of the speaker saying a term they already knew, as described above, they would make up a completely new term, and say that instead. Periodically one of the older speakers would die and be replaced by a new person who did not know any colour words at all, and so the transfer of colour terms between several generations of people could be simulated. All simulations were performed over a number of iterations equal to twenty average lifespans for the simulated people, and results were based on the colour term systems present in each community of ten artificial people at the end of the simulations.

6 Evolutionary Simulations

Dowman (2003a) performed evolutionary simulations using the expression-induction model, to investigate whether it could account for the typological data. The simulation was run with the average lifespans of the simulated people set so that they would remember, on average, 18, 20, 22, 24, 25, 27, 30, 35, 40, 50, 60, 70, 80, 90, 100, 110 or 120 colour examples by the end of their lifetimes. The simulation was run 25 times in each of these conditions. Analysis of the resultant languages showed that, over a number of generations, coherent colour lexicons emerged shared by all the adult speakers in the community. These systems usually partitioned the colour space, although some hues were not clear members of any colour category, while other hues had strong membership in more than one colour term category. There was also some variation between the exact denotations of the colour terms known by different speakers in each community, and sometimes some speakers would know more or fewer colour terms than other ones. Each colour word also had prototype properties, in that there was a single colour that had the highest degree of membership in the colour term category, and, moving away from this prototype, colours became progressively less good examples of the colour category. The results presented in this paper were not derived from individual artificial people, but are instead based on an analysis of those colour terms that were known by most individuals in each community. Those artificial people who had observed only a small number of colour term examples were excluded, as were any terms for which individuals had observed less than four example colours.

It was found that the colour term systems which evolved tended to conform to the types contained in Kay and Maffi's evolutionary trajectories (Kay and Maffi, 1999). There were a small number of systems which did not fit these patterns, which is consistent with the empirical data, and a small proportion of colour words were of types not attested empirically as basic colour terms. (Because it is sometimes difficult to determine objectively which colour terms are basic, when such terms exist in real languages they might well simply have been classed as non-basic, and hence would not be considered to constitute exceptions to the evolutionary trajectories.) It was clear that the majority of colour term systems and individual colour terms were of types attested typologically.

However, one aspect of the simulations reported in Dowman (2003a) appeared to be unrealistic, in that the data from which the artificial people learned was completely free from noise. The data from which the artificial people learned consisted of colour words paired with specific colours that those words could name. The justification for using this kind of data is that this is presumably the same type of data from which real people learn colour words. Children appear to learn the meanings of words largely by observing the speech of other people (Bloom, 2000). In order to learn word meanings, children must infer not only the words which people use, but also their intended meaning. In the case of colour words, the intended meaning corresponds to a particular colour that the speaker intends to identify.

However, inferring the intended referent of a word used by another person would seem to be a somewhat difficult task, and so it seems unlikely that this could be accomplished without ever pairing a colour with a word which cannot denote that colour. Furthermore, there are added complications, because the speaker could use an incorrect word, or other errors could occur, such as the learner mishearing a word. For all these reasons it would seem that not all the data from which children learn colour words is likely to be accurate. The Bayesian acquisitional model was designed to be able to cope with erroneous data, but in the initial simulations no such erroneous data was present. The exact colour that the speaker had intended to name was always passed to the hearer along with the corresponding colour term. Therefore it could not be said that the simulations were realistic in that respect.

Therefore it was decided to investigate whether coherent colour term vocabularies would emerge in the presence of large quantities of random noise, and whether the colour term systems would still reflect the typological patterns. Further simulations were conducted using the same computer model, but now 50% of the colour examples passed to hearers were chosen at random. This simulated situations in which hearers mistake the intended referent of a colour word, and so believe that it was used to identify an unrelated colour. A level of random noise of 50% is very high, and is probably higher than that encountered by people when they learn languages. However, if simulations performed in both this condition, and in the condition where there was no random noise reproduce the empirical typological data, it would seem reasonable to presume that the model is likely to be valid for any intermediate level of random noise.

The model was run 170 times, with artificial people remembering the same number of *accurate* colour examples as before. The model was run ten times for each lifespan simulated. Figure 4 shows results from these simulations, compared to those in which there was no random noise. The average number of basic colour terms emerging in each condition was measured and plotted on the graph. A term was considered basic if a person had seen at least 4 examples of it. Terms were included only if they were known by at least half of the artificial people in the community whose age was over half the average lifespan. (This final restriction was added because younger people would be likely to know fewer terms, as they would not have had sufficient opportunity to learn all of the terms which were widely used in their community.)

FIGURE 4 ABOUT HERE

We can see that the number of colour words emerging in the languages was, on average, roughly proportional to the number of colour examples observed by the artificial people during their lifetimes. This result might seem to be unsurprising, because if people use colour words more often, then they might find it useful to have more such words in their languages. However, in these simulations there were no truly functional pressures, because the artificial people receive no benefit or reward for achieving successful communication. Hence this model suggests that when we use terminology within a particular domain frequently, we might gain more words making more fine grained distinctions within that domain, regardless of whether such distinctions have any functional advantage.

Perhaps surprisingly, the number of words emerging seems to be dependent solely on the number of *accurate* examples of colour words which people observe during their lifetimes. Even though in the condition with 50% noise, twice as many examples were observed by each artificial person as when there was no random noise, the extra random examples seem to have little or no effect on the number of colour terms emerging. (The small differences between the no noise and 50% noise conditions can be attributed simply to random variation.) This result seems somewhat counter intuitive, because no parameter was changed between the simulations which would have given the model any indication that there were varying amounts of random noise, and there was no indication given to any of the artificial people which would allow them to distinguish accurate from random examples.

The most important consideration, however, was whether the simulations would still mirror the typological patterns despite the presence of so much random noise. Figure 5 compares the proportions of basic colour terms which were classified as red, yellow, green, or blue, or as composites of these terms, in each condition of having no noise, or 50% noise, to the proportions of terms which were placed in each of these classes in Kay and Maffi's (1999) analysis of the world colour survey (WCS) data. We can see that the typological patterns in the relative frequencies of each type of colour term are roughly reproduced in each condition. The only major differences between the condition with no noise and that with noise is that there are fewer green terms and blue terms when there is a high level of noise, and a greater proportion of yellow-green-blue terms. In some ways adding noise has resulted in the simulations more closely reflecting the empirical data, especially in that the proportion of blue terms is now almost the same as that found in the world colour survey. In other ways, adding noise has caused the simulation results to diverge somewhat from the empirical data. However, we cannot be sure that such small discrepancies are not simply due to random variation. We might think that by performing more simulations we could resolve this issue. However, as the World Colour Survey is simply a snapshot of a fixed number of languages, we cannot be sure of how representative it is of the range of possible colour term systems in human languages. Small discrepancies between the simulations and the empirical data are therefore not necessarily due to inaccuracies in the model, but could be due to deficiencies in the data instead.

FIGURE 5 ABOUT HERE

Kay and Maffi (1999) did not provide data on the occurrence of purple and orange terms, but it is acknowledged that purple terms occur more frequently than orange terms (MacLaury, 1997). In the noiseless condition 76.9% of terms without a universal focus were purple, while 19.2% of such terms were orange, while with 50% noise these figures were 60.6% for purple and 26.8% for orange. Hence in both conditions the empirical finding that orange is less common than purple was supported by the simulations. The corresponding figures for lime and turquoise terms were 3.8% and 0% with no noise, and 9.9% and 0.3% with 50% noise. These results are consistent with the empirical data, in that in general basic lime and turquoise terms are not found in natural languages. (Neither of these terms is generally considered basic in English.) Possibly the occurrence of basic lime terms is somewhat more frequent than should be expected, although it is not clear how often such terms are simply ignored in linguistic analyses as there is a theoretically motivated expectation that they will not be basic. Overall the results with and without random noise were very similar.

7 Conclusion

The simulations reported in this paper investigated whether an existing explanation of colour term typology was robust to the addition of random noise, or whether adding noise would prevent it from replicating the typological data. This issue is important, because linguistic theories, including those with computer implementations, typically pay little attention to the possibility of language learners encountering erroneous language data. However, in the real world, there are numerous sources of erroneous data, such as speakers who themselves have not learned a word or construction correctly, and so use it incorrectly, and hearers who misunderstand or misinterpret the communicative intent of a speaker, and so believe that a word was used to denote something other than its intended referent. There are therefore many reasons why we should expect a large proportion of the linguistic data that people use when learning languages to be erroneous, so any linguistic theory should be compatible with such a situation.

When a very high level of random noise, consisting of colours chosen entirely at random and presented to the model as additional examples, was added to the evolutionary model, the model was still able to account for the empirical data. The results in this condition were very similar to those produced when no random noise was generated, demonstrating that the model is extremely robust to noise. This would seem a very desirable property for a model of language evolution, as it would be a very poor model of language which was unable to account for empirical data when attempts were made to reproduce conditions similar to those in which real language evolution takes place.

8 Acknowledgments

This work was supported by IPA and IPRS scholarships from the University of Sydney and the Australian Government respectively, and an Economic and Social Research Council post-doctoral fellowship. I would like to thank my Ph.D. supervisor Judy Kay, and everyone else who has assisted me with the research reported here.

9 References

Barnett, V. (1982). *Comparative Statistical Inference*, Second Edition. Chichester: John Wiley & Sons Ltd.

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

Belpaeme, Tony (2002). *Factors influencing the origins of color categories*. PhD Thesis, Artificial Intelligence Lab, Vrije Universiteit Brussel.

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

Bloom, P. (2000). *How Children Learn the Meanings of Words*. Cambridge, MA: MIT Press.

Brighton, H. & Kirby, S. (2001). The Survival of the Smallest: Stability Conditions for the Cultural Evolution of Compositional Language. In J. Kelemen & P. Sosík (Eds.), *Advances in Artificial Life.* Berlin: Springer.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT 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: Praeger.

de Boer, B. (1999). Evolution and self-organisation in vowel systems. *Evolution of Communication* 3(1): 79–103.

Derrington, A. M., Krauskopf, J. and Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357:241-265.

De Valois, R. L., Abramov, I. & Jacobs, G. H. (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America*, 56: 966-977.

De Valois, K. K. & De Valois, R. L. (2001). Color Vision. In N. J. Smelser and P. B. Baltes (Eds. in Chief), *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Elsevier.

De Valois, R. L. & Jacobs, G. H. (1968). Primate color vision. *Science,* 162, 533-540.

Dowman, M. (2003a). Explaining Colour Term Typology as the Product of Cultural Evolution using a Bayesian Multi-agent Model. In R. Alterman and D. Kirsh (Eds.) *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Dowman, M. (2003b). Modelling Language as a Product of Learning and Social Interactions. *Cognitive Systems*, 6(1), 41-54.

Dowman, M. (2004). *Colour Terms, Syntax and Bayes: Modelling Acquisition and Evolution*. Ph.D. Thesis, University of Sydney.

Gleason, H. A. (1961). *An Introduction to Descriptive Linguistics*. Revised Edition. New York, NY: Holt, Rinehart and Winston.

Heider, E. R. (1971). “Focal” Color Areas and the Development of Color Names. *Developmental Psychology,* 4(3):447-445.

Heider, E . R. (1972). Universals of Color Naming and Memory. *Journal of Experimental Psychology,* 93:10-20.

Hering, E. (1964). *Outlines of a Theory of the Light Sense*. Trans. L. M. Hurvich and D. Jameson. Cambridge, MA: Harvard University Press.

Heider, E. R. & Olivier, D. C. (1972). The structure of the color space in naming and memory for two languages. *Cognitive Psychology*, 3:337-354.

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. Roca (ed.) *Logical Issues in Language Acquisition*. Dordecht, Holland: Foris Publications.

Hurford, J. R. (2002). Expression/induction models of language evolution: dimensions and issues. In T. Briscoe (ed.) *Linguistic Evolution through Language Acquisition: Formal and Computational Models*. Cambridge: Cambridge University Press.

Kay, P. (1975). Synchronic variability and diachronic change in basic color terms. *Journal of Language in Society* 4:257-70.

Kay, P., Berlin, B., Maffi, L. & Merrifield, W. (1997). Color Naming Across Languages. In C. L. Hardin & L. Maffi (eds.) *Color Categories in Thought and Language*. Cambridge: Cambridge University Press.

Kay, P., Berlin B., and Merrifield, W. R. (1991). Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology,* 1: 12-25.

Kay, P. & McDaniel, K. (1978). The Linguistic Significance of the Meanings of Basic Color Terms. *Language*, 54 (3): 610-646.

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

Kay, P. & Regier, T. (2003) Resolving the question of color naming universals. *Proceedings of the National Academy of Sciences of the United States of America,* volume, 100, issue 3:9085-9089.

Kirby, S. (1999). *Function Selection and Innateness: The Emergence of Language Universals*. Oxford: Oxford University Press.

Lammens, J. M. G. (1994). *A Computational Model of Color Perception and Color Naming*. Doctor of Philosophy dissertation, State University of New York at Buffalo.

Levinson, S. C. (2001). Yélî Dnye and the Theory of Basic Color Terms. *Journal of Linguistic Anthropology*, 10(1):3-55.

Lucy, J. A. and Shweder, R. A. (1979). Whorf and his critics: linguistic and nonlinguistic influences on color memory. *American Anthropologist,* 81:581-615.

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

Regier, T., Kay, P. and Cook, R. S. (2005). Focal Colors are Universal after all. *Proceedings of the National Academy of Sciences* 102(23):8386-8391.

Roberson, D., Davies, I. & Davidoff, J. (2000). Color Categories are Not Universal: Replications and New Evidence from a Stone-Age Culture. *Journal of Experimental Psychology: General*, 129(3): 369-398.

Rosch, E. H. (1973). Natural Categories. *Cognitive Psychology*, 4:328-350.

Rumelhart, D. and McClelland, J. (1986). On learning the past tenses of English verbs. In McClelland, J., Rumelhart, D., and the PDP research group (eds.). *Parallel Distributed Processing: Explorations in the microstructure of cognition. Volume 2: Psychological and Biological Models*. Cambridge, MA: MIT Press.

Saunders, B. A. C. (1992). *The Invention of Basic Color Terms*. Utrecht: ISOR.

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

---

[i] This frequency data was obtained from the 100 million word British National Corpus.

[ii] Throughout this paper, *evolution* always refers to cultural evolution, rather than real phylogenetic (biological) evolution.

[iii] Munsell chips are small pieces of cardboard which are painted in carefully controlled pigments, so that the colours of the chips are systematically spaced over the range of all possible colours.

[iv] Although note that Russian and Hungarian both have a twelfth term that should probably be considered basic, at least for some speakers MacLaury (1997).

[v] This trajectory appears to ignore languages with two basic colour terms, which simply make a split into dark and light colours, rather than grouping red and yellow with white and green and blue with black.

[vi] This colour space was chosen because Lammens (1994) found that his computer model of colour naming worked best in it.

[vii] Belpaeme (2002) did suggest that his model might be able to explain the split into light and dark colours seen in languages with only two colour terms, because this might be the easiest way to divide up the colour space. However, in its present form, the model could not explain the typological restrictions seen in more complex types of colour term system.

$$\begin{bmatrix}\text{white}\\ \text{black}\end{bmatrix} < \begin{bmatrix}\text{red}\end{bmatrix} < \begin{bmatrix}\text{green}\\ \text{yellow}\end{bmatrix} < \begin{bmatrix}\text{blue}\end{bmatrix} < \begin{bmatrix}\text{brown}\end{bmatrix} < \begin{bmatrix}\text{purple}\\ \text{pink}\\ \text{orange}\\ \text{grey}\end{bmatrix}$$

Figure 1. Berlin and Kay's (1969) Implicational Hierarchy.

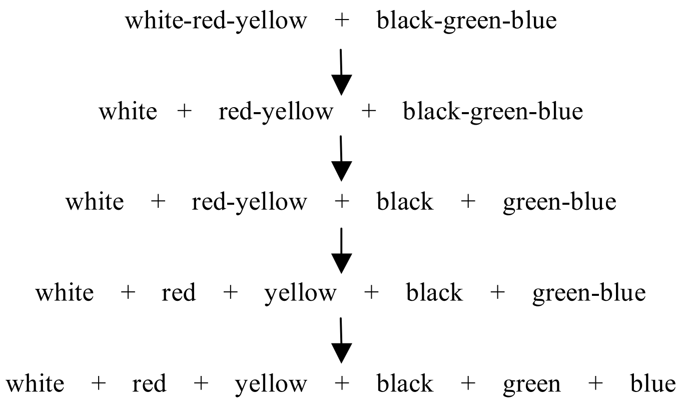

Figure 2. The Main Line of Kay and Maffi's Evolutionary Trajectory

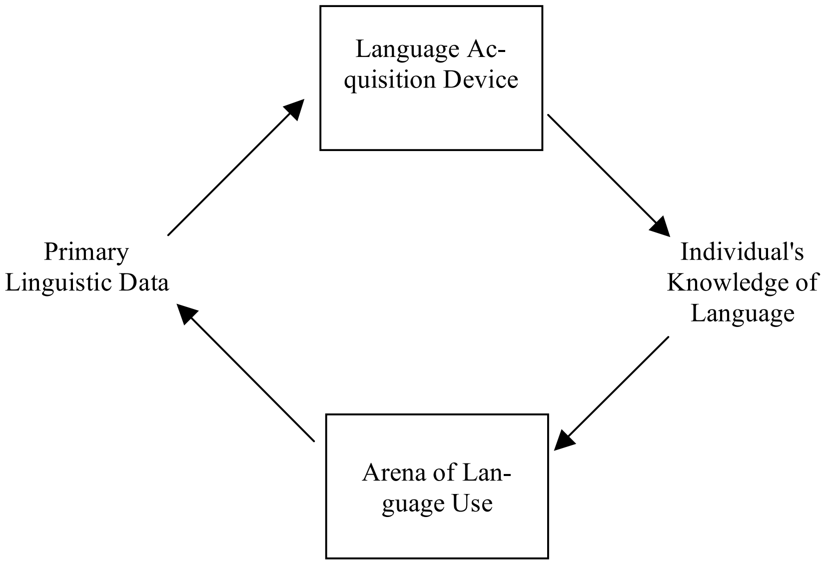

Figure 3. Hurford's Diachronic Spiral

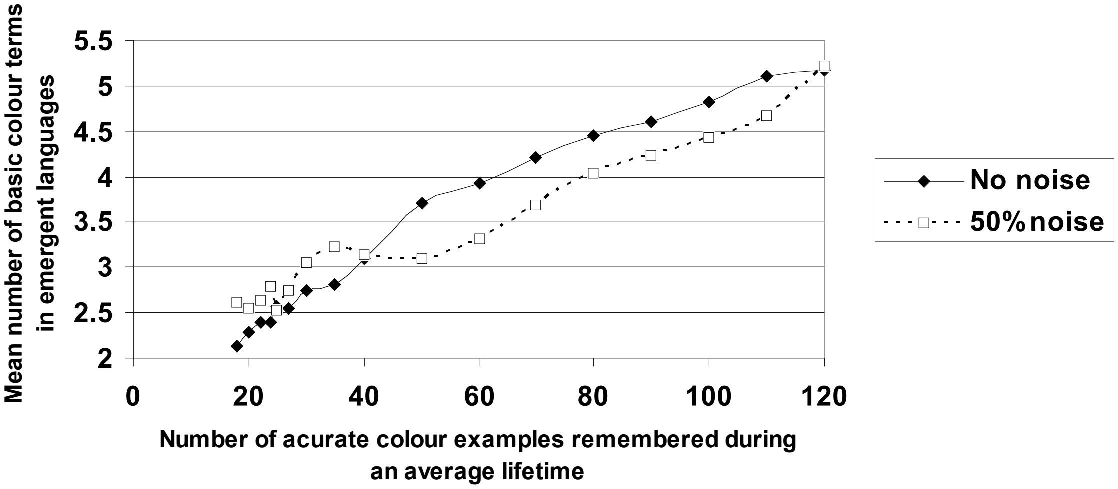

Figure 4. The Number of Basic Colour Terms in Emergent Languages

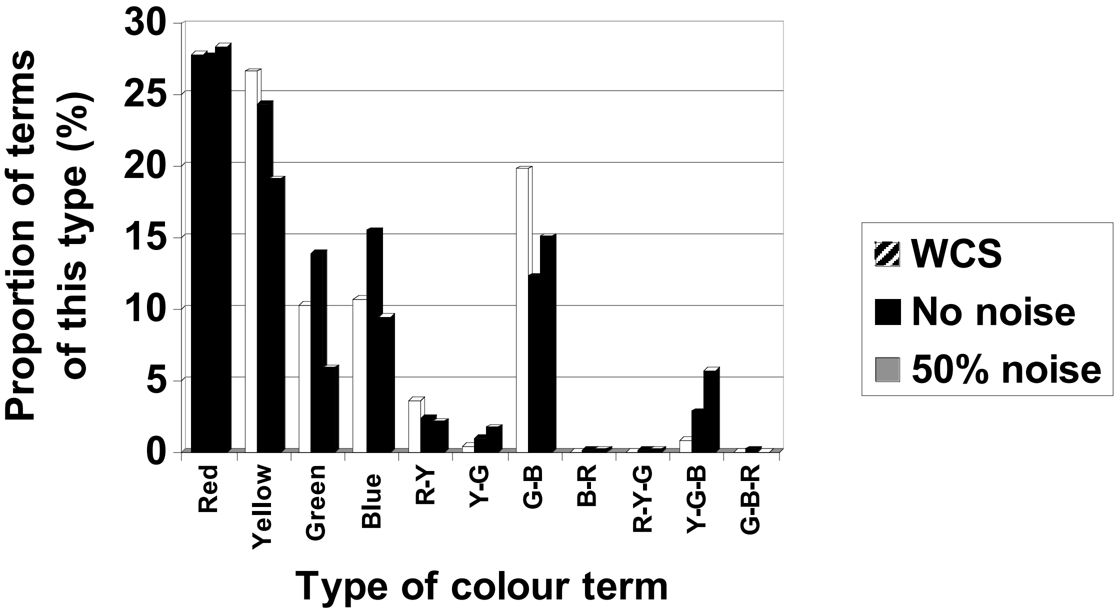

Figure 5. The Frequencies of Each Type of Color Term