

African Linguistics at the Crossroads

Papers from Kwaluseni
1st World Congress of African Linguistics
Swaziland, 18–22. VII. 1994

edited by

Robert K Herbert



RÜDIGER KÖPPE VERLAG KÖLN
1997

**COLOUR CATEGORIES IN AFRICAN LANGUAGES:
A TEST OF THE BERLIN & KAY THEORY OF COLOUR UNIVERSALS**

IAN DAVIES & GREVILLE CORBETT

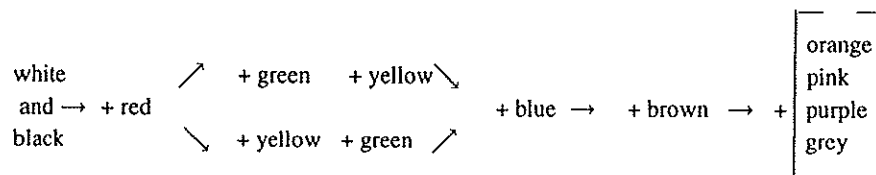
1 INTRODUCTION

Berlin and Kay's (1969) theory of colour universals marks a turning point in ideas about language and colour categorisation. The prevalent view, prior to 1969, was that the variation in colour terminology across languages was prima facie evidence for linguistic relativity and perceptual constructionism. Colour categories originated in cultural needs and there was no biological constraint on how the physical continuum of visible light was segmented (see, for instance, Gleason 1962). Since 1969, however, the prevalent view has shifted towards acceptance of the core of Berlin and Kay's theory - the universality of colour categories (but see, for example, Ratner 1989, Lucy 1992: 177-181 for dissenting views). In this paper we summarise a number of field studies of the colour terms of languages from central and southern Africa carried as further tests of the theory. First, however, we outline the core of the theory.

The original Berlin and Kay theory has been modified a number of times, notably by Kay and McDaniel (1978) and the most recent version is given by Kay, Berlin and Merrifield (1991). In the brief summary of the theory that we give here, we describe the core of the theory which incorporates elements from several stages in its development. The core of the theory is that, rather than being arbitrary, all colour term inventories are drawn from a severely restricted subset of possible colour term inventories. Figure 1 shows the original Berlin

and Kay (1969) hierarchy which specifies sets of "basic" (the most important) colour terms which are possible within the theory. (This is the strictest version of the hierarchy; subsequent developments have successively loosened the constraints of the hierarchy.)

FIGURE 1: *The Berlin and Kay Hierarchy of Basic Colour Terms*



The hierarchy has an implicational structure: if a language has a particular term on the hierarchy, then it should also have all terms to the left of that term. For example, if a language has a term for red, then it should also have terms for white and black. Conversely, if a language does not have a particular term, then it should not have any terms to the right of the given term on the hierarchy. For instance, if there is no basic term for brown, then there should not be terms for orange, pink, purple or grey. Some terms share a common position on the hierarchy in which case there is no implicational relationship between them. Thus, for instance, the presence of a term for pink implies nothing about the presence of orange, purple or grey, whereas the presence of pink does imply that there should be terms for white, black, red, green, blue and brown.

The concept of *basic colour term* is central to the theory. Like the rest of the theory, this concept of BASIC COLOUR TERM has been modified since the original theory was presented (see Crawford 1982 and Moss 1989 for discussions). The essence of the concept is neatly captured by Hardin and Maffi (forthcoming): basic color terms are *general* and *salient*. A term is general if it can be used for a wide range of objects and its meaning cannot be subsumed under that of another basic term. For a term to be salient it must be readily elicitable, it must be found in the idiolects of most speakers, and there must be agreement across speakers as to what a term denotes (particularly over the best exemplar of the term). Hardin and Maffi's characterisation of basic colour terms is based on how terms are currently used; their origin is unimportant. Thus for example, recent loan words, or words that were originally restricted in their range of referents, can be basic, if they currently meet the criteria of generality and salience. In accord with Hardin and Maffi, the measures we use in this paper to establish basic colour terms rely primarily on our informants' behaviour, and we do not consider the etymology of the colour terms directly.

Berlin and Kay's colour universals are universals of the foci (the best examples) of colour categories, whereas the boundaries of colour categories may vary considerably. The variation in category boundaries can be minor for languages with similar numbers of basic terms, but languages with relatively few colour terms can have much larger colour categories than those with more extensive inventories of colour terms. Such large categories may include the foci of two or more of the universal foci as we discuss below. It follows from the emphasis on the universality of foci rather than category boundaries that a necessary criterion for a term to be basic is that there should be good agreement over its focus.

Kay and McDaniel (1978) offer a theory of the nature and development of colour categories based in part on perceptual physiology and formal fuzzy set theory. For our immediate purposes the most pertinent aspect of the theory is that languages with few colour terms should have "composite" categories, that is categories which include two or more of the foci of the eleven universal categories. For example, a language which had say, terms for white, black, red and green, would be expected to subsume most or all of colour space within those four categories. Such an arrangement would probably mean that the universal foci for yellow and orange were subsumed in the red category and that blue would be subsumed in the green category, although other arrangements are possible. Further, as a language with composite categories acquires additional terms, this is accomplished by the splitting of one or more composite categories into its components. To pursue the example given above, an additional term could be acquired by the splitting of the green with blue category into separate green and blue basic terms.

Preliminary studies suggest that the languages we report here have composite categories, most notably a green-with-blue category – "grue" – and one aim of the studies was to see if this expectation was borne out. The operational test for composite categories we apply is to see whether a colour category includes the foci of two or more of the universal colour categories. In addition to looking for evidence for composite categories, we also looked for evidence that composite categories might be in the process of splitting or might have left evidence ("debris") of earlier splitting. For example, increased use of borrowed terms to describe the components of a composite term, in parallel with continued use of the original composite term, might be due to the composite term splitting.

We report data on the languages: Chichewa, Ndebele, Setswana and Xhosa. The languages are all members of the Bantu group,¹ which is a large, relatively homogenous grouping within the Niger-Kordofanian family (see Ruhlen 1987). Within Narrow Bantu, there is a split into Northwest Bantu

(further divided into three zones, A - C) and Central Bantu (divided into thirteen zones, D - S). Chichewa is found in zone N, while Setswana is found in zone S. Ndebele and Xhosa are also found in zone S, both in the same subgroup. The latter two are therefore particularly closely related.

We used two main tasks: a colour list task and a colour naming task, although not all samples did both tasks. The colour list task was suggested by Berlin and Kay as a measure of salience and we used this on all samples; basic terms should occur in most informants' lists. The colour naming task was used to establish the referents of the colour terms, to measure the frequency with which terms are used, and to estimate the levels of consensus as to what terms denote; basic colour terms should be used frequently and there should be good agreement across speakers as to what the terms denote. Although the studies of the four languages used similar procedures, there were some minor variations. In what follows we describe the methods in outline and the full technical details are given in Davies et al. (1992) for Setswana, Davies and Corbett (1994a) for Chichewa, Davies and Corbett (1994b) for Xhosa and Davies, Davies and Corbett (1994) for Ndebele.

2 METHOD

2.1 SUBJECTS

Samples of about 40 people were drawn from mother-tongue speakers of four languages (the country of residence is given in brackets): Chichewa (Malawi), Ndebele (Zimbabwe), Setswana (Botswana) and Xhosa (South Africa). For each language there was one sample of adults from a rural area, with ages ranging from about 20 years to about 80 years. For Chichewa and Ndebele there were further samples that consisted of students (mean age 20 years) and school children (mean age about 12 years), respectively.

2.2 STIMULI

The colour samples in the colour-naming task consisted of 65 "colour-tiles". The tiles were coloured paper, 50mm square, mounted on 3mm thick plywood, sprayed with a light film of transparent varnish to protect them from staining during use. The colours were an evenly spread sample of "colour space" taken from the Color-Aid corporation range. Table 1 gives their Color-Aid codes and their CIE coordinates, and Figure 2 shows their distribution in CIE chromaticity space together with the loci of the 11 "universal" foci, taken from Heider (1971).

Color-Aid code	CIE coordinates		
	BRIGHTNESS Y	REDNESS x	GREENNESS y
Y Hue	64.77	.47	.48
S2	16.99	.41	.44
YOY Hue	47.48	.50	.43
T4	55.63	.45	.41
S2	22.08	.36	.38
YO Hue	39.52	.51	.41
T3	47.02	.48	.41
S3	10.72	.36	.41
OYO Hue	26.51	.54	.37
O Hue	25.00	.54	.37
S1	14.34	.50	.37
S3	9.15	.42	.36
ORO Hue	18.87	.57	.34
T3	36.88	.46	.35
S3	26.51	.33	.32
RO Hue	16.22	.58	.33
T3	32.66	.45	.32
S3	4.19	.37	.34
ROR Hue	15.23	.53	.31
T3	29.82	.42	.30
S3	20.71	.34	.28
R Hue	11.71	.50	.29
T4	24.34	.40	.27
S3	4.81	.33	.30
RVR Hue	9.11	.42	.24
S1	12.79	.35	.25
S3	28.43	.36	.28
RV Hue	6.97	.33	.19
T2	14.51	.31	.19
VRV Hue	6.71	.30	.19
S3	28.42	.36	.28
V Hue	4.67	.26	.17
VBV Hue	4.13	.24	.17
T4	19.05	.25	.20
BV Hue	4.21	.22	.19
S2	7.88	.25	.26
BVB Hue	4.80	.19	.13
S3	26.65	.26	.23
B Hue	9.51	.18	.16
T1	19.02	.20	.19
BGB Hue	9.62	.19	.19
T3	23.08	.20	.23
BG Hue	8.93	.20	.25
T1	16.57	.19	.25
S2	7.42	.21	.26
GBG Hue	10.69	.23	.37
S2	20.79	.20	.25
G Hue	11.99	.24	.42
S3	6.10	.26	.33
GYG Hue	12.89	.25	.44
T4	31.14	.26	.41
S1	15.59	.26	.31
YG Hue	14.66	.28	.48
S3	5.78	.30	.34
YGY Hue	18.92	.30	.51
S3	35.87	.35	.43
Rose Red	17.63	.41	.24
Sienna	13.31	.44	.36
White	81.40	.32	.33
Gray 1	47.55	.32	.33
Gray 2	30.59	.32	.33
Gray 4	18.88	.31	.31
Gray 6	11.20	.31	.31
Gray 8	4.53	.31	.32
Black	3.59	.34	.33

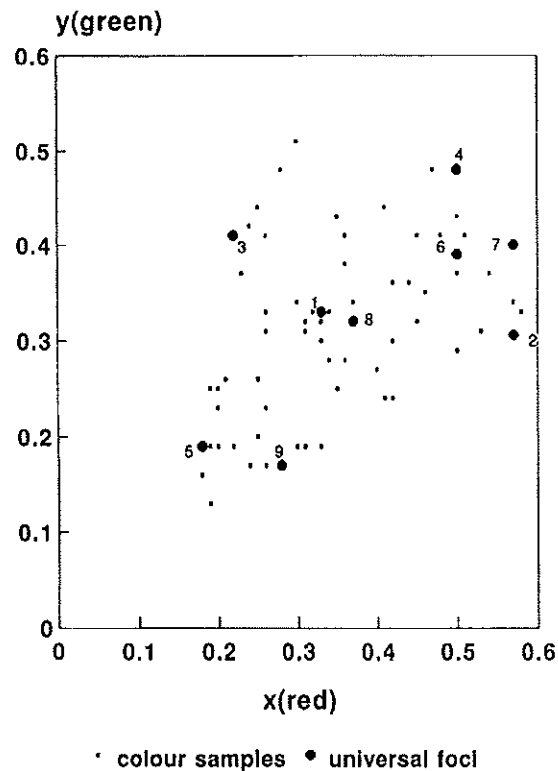


Figure 2: *Foci of the tile colours and the universal foci in CIE coordinates*

1 = achromatic (white, black and grey), 2 = red, 3 = green, 4 = yellow, 5 = blue, 6 = brown, 7 = orange, 8 = pink, 9 = purple

The CIE chromaticity coordinates may be thought of as the proportions of red (x), and green (y), in each colour; a third coordinate, lightness or reflectance (Y) makes up the CIE tri-stimulus values. By implication, the proportion of blue is given by $1-(x+y)$. Every possible colour has a unique locus in 3-dimensional (Y, x, y) space and these tri-stimulus values may be used to convert the Color-Aid stimuli into the more familiar Munsell, or OSA systems, through conversion tables in, for instance, Nickerson (1944). The positions of

the universal foci can be used to interpret the graph: universal green is located towards the top left of the space and has a high y value; universal red lies in the right of the space and has a high x value; and universal blue lies towards the origin and has low values on x and y .

The Color-Aid range is based on the Ostwald colour solid, (see Foss, Nickerson and Granville 1944, for an outline of this system) and is made up from 24 "hues": Y (yellow), O (orange), R (red), V (violet), B (blue) and G (green), plus intermediate values designated by combinations of the previous codes; for instance YOY, YO, and OYO are the intermediate hues between Y and O. In addition to the hues there are seven variants of each hue consisting of four "tints" T1 to T4, and three "shades" S1 to S3; the tints have increasing amounts of white added to the hue as their index number increases, whereas the shades have increasing amounts of black added as their index number increases. In addition, there is also a "grey-scale" and a number of colours of particular significance to painters.

2.3 PROCEDURE

All subjects did the list task and all subjects, except those in the two younger samples, also did the colour naming task. In every case the experiment was administered by a mother-tongue speaker of the appropriate language (all instructions were given in the mother-tongue of the subjects). In the list task, they were asked "please tell me as many colour terms as you know". The experimenter wrote down their responses, encouraged them to continue if they paused, and moved on to the next task when they said they had finished. The task generally took less than two minutes. In the colour naming task, they were shown each of the 65 tiles, one at a time, on a piece of grey cloth, in a different random order for each subject. They were asked "what do you call this colour?" All responses, including "don't know" were recorded. The naming task took usually about 20 minutes.

3 RESULTS

3.1 THE LIST TASK

The mean number of terms offered in the list task varied from 5.1 for the Ndebele adult sample through to 9.4 for the Chichewa student sample. The terms offered by at least three informants are shown in Tables 2 to 5 for Chichewa, Ndebele, Setswana and Xhosa respectively (Tables 2 and 3 give the data for the older and younger samples separately), and Table 6 summarises the data across all the samples.

Table 2: CHICHEWA

Frequencies in the list task

Term	Gloss	Rural (n = 40)	Student (n = 43)
chofiira	red	40	43
chakuda	black	38	41
choyera	white	35	42
chobiriwira	grue	33	41
chikasu	yellow	33	38
chotuwa	grey	24	24
chakhofi	coffee	11	12
chamtambo	sky blue	8	10
chamasamba	leaf green	5	2
chakhaki	khaki	3	9
chabululu	blue	1	15
chanyezinyezi	shiny	1	15
chagirini	green	1	6
chamandalasi	coloured	0	7

Table 3: NDEBELE

Frequencies in the list task

Term	Gloss	Adult (n=39)	Child (n = 34)
kubomvu	red	33	29
kuluhlaza	grue	32	26
kumhlophe	white	31	28
kumnyama	black	25	28
kulithanga	yellow	19	26
ilubende	brown	8	13
luthuli	grey	8	5
okuintwala	grey	6	1
okulithendele	spotted	5	4
okuyisibhakabaka	blue	3	7
bulu	blue	3	15
khaki	khaki	0	17
kyelo	yellow	0	11
igilini	green	0	8

Table 4: SETSWANA

Frequencies in the list task (n=44)

Term	Gloss	Frequency
bontsho	black	33
bosetlha	yellow/brown	28
bosweu	white	27
bohibidu	red	23
botala	grue	21
borokwa	brown	19
bohunohu*	reddish-brown	18
bonala*	red & white	16
bowebu*	roan	14
bonkgwe*	white-backed	12
bophatshwa*	black & white	11
selaole	purple	10
bogweba*	red & white	8
botuba*	dun/grey	7
sethunya	blossom/yellow	6
mathubapula	sand	6
botala jwa tlhaga	grass-grue	6
lephutshe	pumpkin	5
botala jwa legodimo	sky-grue	3

(* < cattle-terms)

Table 5: XHOSA

Frequencies in the list task (n=44)

Term	Gloss	Frequency
mhlophe	white	34
mnyama	black	34
bomvu	red	34
tyheli	yellow	31
hlaza	grue	30
mdaka	brown	21
pinkhi	pink	12
blue	blue	11
purple	purple	5
ngewevu	grey	4
bhelu	pale	4
grey	grey	3

The terms in each table are ordered by the frequency with which they were offered, and the frequency is also given for each term. For each language, the five most frequent terms are terms for white, black, red, grue and yellow². Further, in most cases, the five terms are offered by a majority (often over three-quarters) of the sample. For Ndebele, no other term is offered by the majority of either the adult or the child sample, and for the adults there is a marked fall in the score for the next most frequent term, *ilubende* 'brown'. For each of the other three languages there is a sixth most frequent term whose score stands out from the remaining terms. In Chichewa, that term is *chotuwa* 'grey', which is offered by over half of each sample, and there is then a sharp fall to the frequency of the next term, *khofi* 'coffee'. In Setswana, the next term is *borokwa* 'brown', which is followed in the rank order by five "cattle-terms". These are terms whose use is restricted to describing cattle, and thus fail the generality criterion for basicness (although we shall see that in at least one case this restriction on use is not complete). The next most frequent general term in Setswana is *selaole* 'purple', which is offered by less than a quarter of the sample. Lastly, in Xhosa, the next term – *mdaka* – is also a term for brown; this is offered by almost half of the sample, whereas the next most frequent term – *pink* 'pink' – is offered by just over a quarter of the sample.

TABLE 6: Summary of list task

TERM	Chichewa		Ndebele		Xhosa	Setswana
	R	S	A	C		
black	38	41	25	28	34	33
white	35	42	31	28	34	27
red	40	43	33	29	34	23
grue	33	41	32	26	30	21
blue	8	15	3	15	11	3
green	5	6	0	8	0	6
yellow	33	38	19	26	31	28
brown	0	0	8	13	21	19
purple	0	0	0	0	5	10
pink	0	4	0	4	12	0
orange	0	0	0	0	0	0
grey	24	24	8	7	4	0

R Rural Sample A Adult Sample
S Student Sample C Child Sample

From Table 6 it can be seen that as well as the general patterns described above, there are differences between the samples in the extent to which separate blue and green terms are offered. In Chichewa, about a third of the student

sample offered *chabuluu* 'blue', whereas only one member of the rural sample did. Similarly, in Ndebele, over a third of the children offered *okuyisibhakabaka* 'blue' whereas only three adults did. Both of the younger samples also offered a green term, but in both cases, the frequencies are much lower than those for the corresponding blue term. In Xhosa, a quarter of the sample offered a blue term and none offered a separate green term, whereas in Setswana, six people offered a separate green term and three people offered a separate blue term.

3.2. COLOUR NAMING

Tables 7 to 9 show the most frequently used terms in the colour-tile naming task for Chichewa, Setswana and Xhosa respectively. These data are derived from the full colour-tile by naming-response matrices given in the technical papers referred to earlier. The frequencies (column three) are the total number of times each term was used, summed across informants and colour-tiles.

TABLE 7: Colour naming in Chichewa

TERM	GLOSS	FREQUENCY	DOMINANCE
chobiriwira	grue	382	6
chofiira	red	288	5
chotuwa	grey	188	1
chikasu	yellow	182	5
chakuda	black	118	2
chakhofi	coffee	98	0
chamtambo	blue	72	0
choyera	white	43	1
chamasamba	green	43	0
chonwira	darkish	30	0
chaphulsa	ash-grey	15	0

TABLE 8: Colour naming in Setswana

TERM	GLOSS	FREQUENCY	DOMINANCE
botala	grue	391	9
bohibidu	red	197	5
botuba	dun/grey	135	2
borokwa	brown	124	2
selaole	purple	117	0
bontsho	black	112	2
bowebu	roan	110	0
bosetlha	yellow/brown	109	0
sethunya	blossom (yellow)	106	1
mathubapula	sand	103	0
bosweu	white	69	1

TABLE 9: Colour naming in Xhosa

TERM	GLOSS	FREQUENCY	DOMINANCE
hlaza	grue	144	2
bomvu	red	120	2
blue	blue	113	0
mnyama	black	107	2
tyheli	yellow	104	2
mdaka	brown	82	1
pink	pink	75	0
grey	grey	54	0
mhlophe	white	44	1
purple	purple	43	0
navy	navy	12	0

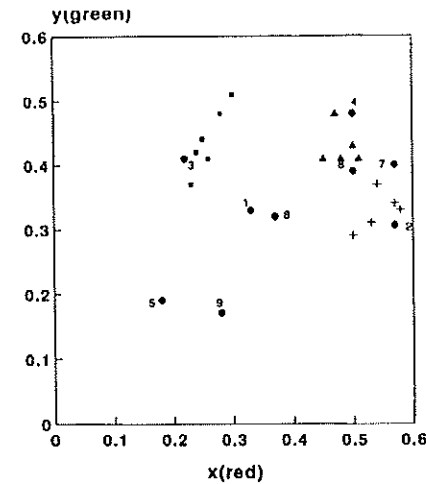
In general, the most frequent terms in the list task are also frequently used in the naming task. However, there are differences in the rank orders of the frequencies across the two tasks. This is particularly clear for terms for white and black, which tend to have lower ranks in the naming task than in the list task. The frequencies in the naming task are a function of both of the salience of colour terms and of the particular colours in the sample. Use of the white terms was mainly restricted to one colour-tile, and use of the black terms was mainly restricted to two colour-tiles. In contrast the opportunity to use several of the other terms was much greater than for the black and white terms, and it is these properties of the colour sample, rather than differences in salience, which produce the discrepancy in rank orders for the white and black terms between the two tasks. There are other discrepancies between the measures based on the two tasks. First, for Setswana, the rank order of *bosetlha* is lower for the naming task than for the list task, and the rank order of *botuba* 'dun' is higher for the naming task than for the list task. Second, for Xhosa, the rank order for *blue* 'blue', is higher in the naming task than in the list task.

One criterion for a term to be basic is that there should be good agreement over what it denotes. Tables 7 to 9 give a measure of the level of agreement across the samples of what the terms denote, the "dominance index". A term is dominant for a given colour-tile if more than half of the sample use that term to name the colour-tile; the dominance index is the total number of colour-tiles for which that term is dominant. With just one exception, all terms which were the five most frequent terms in the list task (terms for white, black, red, grue and yellow) have at least one colour-tile for which each is dominant. The exception is the Setswana term *bosetlha* 'yellow' which although its frequency of use is relatively high (109), is never the dominant term for any colour-tile. The sixth most frequent term in the list task (*chotuwa* 'grey' in Chichewa; *borokwa* 'brown' in Setswana; and *mdaka* 'brown' in Xhosa) are dominant for at least one colour-tile in each case.

There are just two terms other than the ones already given, which have non-zero dominance indices, and they are both Setswana terms. *Botuba* 'dun'

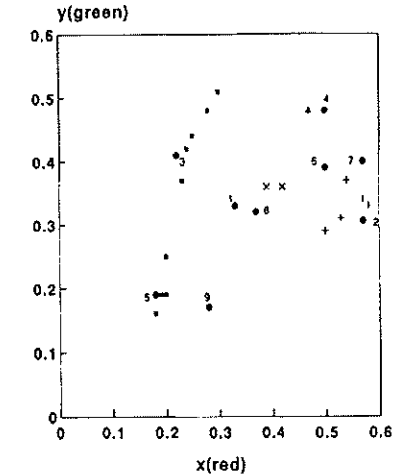
is dominant for two colour-tiles and *sethunya* 'blossom' is dominant for one colour-tile. In table 5, we indicated that *botuba* was a cattle term, but its use to name the colour-tile samples indicates that its use is not restricted to describing cattle.

FIGURE 3: Chichewa



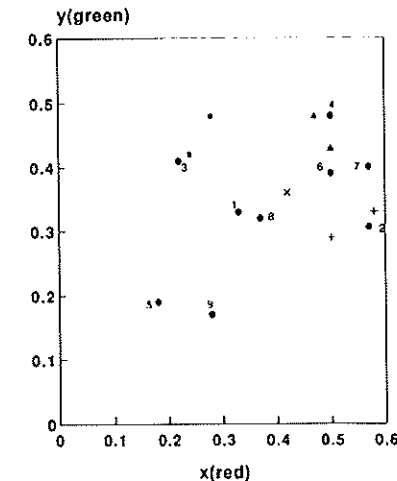
* chobiriwira + chollira ▲ chikasu • universals

FIGURE 4: Setswana



* botala ■ bohlibidu ▲ sethunya × borokwa • universals

FIGURE 5. Xhosa



* hlaza + bomvu ▲ thylli × mdaka • universals

Figures 3 to 5 show the regions of colour space for which each term was dominant for Chichewa, Setswana and Xhosa respectively. The loci of the colour-tiles for which each term was dominant are plotted in the same coordinate system as the universal foci and the full colour-tile sample (figure 2). These figures can be used to check the appropriateness of the glosses that they have been given and to establish whether any of the terms meet the requirements for a composite term (inclusion of the foci of two or more primary terms: white, black, red, green, blue and yellow). It would be expected that the dominant region for a given term would be close to, or would include, the focus of the universal term whose gloss it has been given. Thus, for instance, the Chichewa term *chofiira* 'red' in Figure 3 can be seen to include the focus of the universal red as shown in Figure 2. By and large the data are consistent with the glosses given. However, the support for the gloss 'grue' for *chobiriwira* in Chichewa and *hlaza* in Xhosa is not very strong. *Chobiriwira* (Figure 3) and *hlaza* (Figure 5) are concentrated in the green region of colour space and do not include focal blue. The Chichewa term *chobiriwira* was used to denote blue colours, but in competition with *chamtambo* 'sky'; for such colours typically about 15 people out of 40 used *chobiriwira* and about 5 people used *chamtambo* for colours in the blue region. Xhosa speakers tended to use the borrowed term *blue* but without sufficient consensus for the term to achieve dominance. The uncertainty in both Chichewa and Xhosa over what term to use for blues is reflected in the fact that just under half of the two groups were unable to name the blue colours. The Setswana term *botala* (Figure 4), on the other hand, does include focal green and focal blue; this double inclusion justifies the gloss grue and is consistent with *botala* 'grue' being a composite category. No other terms include the foci of two or more of the primary categories and thus do not meet the requirements of a composite term. On the other hand, some terms have broader ranges than equivalent terms in languages which have the full set of universal categories. For example, the Chichewa term, *chikasu* 'yellow', includes the focus of the universal orange category (orange is not a primary category). Chichewa also has attenuated forms of some colour terms. For instance, *chofirira* 'reddish' and *chobiriwirira* 'grueish' are the attenuated forms of *chofiira* 'red' and *chobiriwira* 'grue'. The regions covered by the combined simple and attenuated forms extend the domain of the terms further into the regions which would be denoted by other universal terms in eleven-term languages (such as English), but even these combined categories do not include the foci of two universal primary terms. Thus, the enlarged *chofiira* 'red' category, includes more of the orange, pink and purple regions of colour space, but it does not extend as far as the focus of yellow or blue.

3.3. COMBINING THE MEASURES

Table 10 summarises the data from the two tasks. For each of the four languages we indicate whether the language shows evidence of having a basic term for each of the eleven terms in the Berlin and Kay hierarchy. If a term is offered by more than half of a particular sample, then that is indicated by + sign in the appropriate column; if less than half the sample offer the term it is indicated by a - sign; and if there are no relevant data, this is indicated by a blank. In condensing the data into a single summary table we have treated green and grue as equivalent terms.

TABLE 10: Combined indicators of basic terms in the four languages

term	CHICHEWA		NDEBELE		SETSWANA		XHOSA	
	list	dom.	list	dom.	list	dom.	list	dom.
white	+	+	+		+	+	+	+
black	+	+	+		+	+	+	+
green	+	+	+		+	+	+	+
yellow	+	+	+		-	+	+	+
blue	-	-	-		-	-	-	-
brown	-	-	-		+	+	+	+
orange	-	-	-		-	-	-	-
pink	-	-	-		-	-	-	-
purple	-	-	-		-	-	-	-
grey	+	+	-		-	+	-	-

All four languages meet both criteria for having basic terms for white, black, red, and green. All the languages except Setswana, meet both criteria for having basic terms for yellow. The problem with Setswana is that one term *bosetlha* was offered frequently on the list task but there was little agreement over how to use it in the naming task; and a second term *sethunya*, was offered infrequently in the list task, but there was good agreement over how to use it in the naming task. Setswana and Xhosa meet both requirements for having a basic term for brown, and Chichewa meets both requirements for having a basic term for grey. Finally, Setswana meets the criterion for having a basic grey term on the naming task but the term *botuba* was not offered very frequently on the list task.

4 DISCUSSION

There is a strong case for arguing that all four languages have basic terms for white, black, red and green (or grue). In every case, the appropriate

terms appeared in the majority of our informants' lists, and there was at least one colour that the majority of informants in a given sample denoted by the given term. Thus the criteria of salience and consensus are met for these terms. Three of the languages – Chichewa, Ndebele, and Xhosa – also appear to have basic terms for yellow, on the same grounds as given above (although, the evidence for Ndebele is based on the list task alone). It is less clear whether Setswana has a basic term for yellow because there appear to be two terms competing for the yellow slot, one (*bosetlha*) which meets the salience criterion but not the consensus criterion, and the other (*sethunya*) which meets the consensus criterion but not the salience criterion.

Even with the uncertainty about the status of the yellow terms in Setswana, the inventories given above are all consistent with the Berlin and Kay hierarchy. However, Chichewa, Setswana and Xhosa have at least one other term which meet the joint criteria of salience and consensus. Chichewa has a basic term for grey, and Setswana and Xhosa have basic terms for brown. The apparently exceptional behaviour of grey is no surprise; instances were known as early as Berlin and Kay (1969: 45); see also Kay and McDaniel (1978: 640). The problem with brown is of greater interest. According to the original Berlin and Kay theory and to its modification by Kay and McDaniel, basic terms for brown should not occur unless there are also basic terms for white, black, red, green, yellow and blue. In Setswana, there is no suggestion of a basic term for blue, and therefore the inventory for Setswana seems to be an exception to the theory. There is more evidence for terms for blue in Chichewa and Xhosa than in Setswana, but there is not sufficient evidence to suggest that there are basic terms for blue in either language (Table 10). These two languages would then also be inconsistent in detail with the theory. However, Kay, Berlin and Merrifield (1991) report that it is not uncommon for terms for brown as well as grey to violate the sequence of the original hierarchy. They accord brown and grey something of a 'wild-card' status, thus somewhat relaxing the strictures of the theory.

According to Kay and McDaniel, languages with less than six basic terms should have at least one composite term. Thus all of the languages described here should show evidence of a composite term. We have concentrated on the possibility that the languages would have a composite green with blue term. The evidence for Setswana is consistent with the term *botala* 'grue' being composite: its range includes the foci of the universal green and blue categories. On the other hand, although the range of the ostensive green with blue categories in Chichewa and Xhosa includes the blue as well as the green regions, the two categories are most clearly concentrated around the universal green focus. That is, although some people use the grue terms to denote blue

and green, there is only strong consensus over its use for green colours. This uncertainty is accompanied by tentative use of a separate blue category in both languages. In Chichewa, a separate blue category was more evident in the student group than in the rural group. It may be that we are observing the decomposition of the original grue categories into separate blue and green categories catalysed by the availability of appropriate loan words. On the other hand, it may be that Kay and McDaniel are wrong; perhaps languages with relatively few terms just have broader categories than languages with the full complement of basic colour terms. The broader categories may include the foci of more than one of the six primary categories, but they do not have to. This would be consistent with the Chichewa categories *chikasu* 'yellow' which includes the focus for orange, and *chofira* 'red' which extends well into pink and purple.

5 SUMMARY

All four languages fit the original Berlin and Kay (1969) theory to a first approximation: the most basic terms in each case form a 'legitimate' set. In addition, the exceptions to the theory – possible basic terms for brown and grey appearing too early – have been accommodated by later versions of the theory by according them wild card status. The evidence for composite categories is relatively weak; there is evidence for relatively broad categories, but these do not usually include the foci of two or more primary terms. The different situations in these relatively closely related languages provides specially favourable conditions for observing diachronic developments in colour term systems.

NOTES

We acknowledge the help and advice of David and Tiny Jerrett in Botswana, Al Mtenje in Malawi, and Lusanda Rataemane in South Africa. This work was supported by the British E.S.R.C. grant numbers R000 23 1958 and R000 23 3978; we are very grateful for this support.

¹ See Blommaert (1985) for wider discussion of primary terms in Bantu languages.

² The glosses we give in the tables and the corresponding assertion that they are terms representing the universal terms white, black, red, grue and yellow, are based on our expert informants' advice and on the definitions given in the available dictionaries. However, when we examine how the terms are used to name colours the aptness of some of the glosses will be questioned.

REFERENCES

- Berlin, B. and Kay, P. 1969. *Basic Color Terms: Their Universality and Evolution*. Berkeley & Los Angeles: University of California Press.
- Blommaert, J.M.E. 1985. The semantics of Bantu colour terms. *Grazer Linguistische Studien* 24:63-75.
- Corbett, G.G. and Davies I.R.L. 1995. Linguistic and behavioural measures for ranking basic colour terms. *Studies in Language* 19:301-357.
- Crawford, T.D. 1982. Defining "basic color term". *Anthropological Linguistics* 24:338-343.
- Davies, I.R.L., MacDermid, C., Corbett, G.G., McGurk, H., Jerrett, D., Jerrett, T. and Sowden, P. (1992). Color terms in Setswana: a linguistic and perceptual approach. *Linguistics* 30:1065-1103.
- Davies, I.R.L. and Corbett, G.G. 1994a. The basic colour terms of Chichewa. *Lingua*.
- Davies, I.R.L. and Corbett, G.G. 1994b. A statistical approach to determining basic color terms: an account of Xhosa. *Journal of Linguistic Anthropology* 4:175-193.
- Davies, I.R.L., Davies, C. F. and Corbett, G.G. 1994. The basic colour terms of Ndebele. *African Languages and Cultures* 7:36-48.
- Foss, C.E., Nickerson, D. and Granville W.C. 1944. Analysis of the Ostwald Colour System. *Journal of the Optical Society of America* 34:361-381.
- Gleason, H.A. 1961. *An Introduction to Descriptive Linguistics*. New York: Holt, Rinehart & Winston.
- Hardin, C. L and Maffi, L. (forthcoming). Editors' Introduction to *Color Categories in Thought and Language*. Cambridge: Cambridge University Press.
- Heider, E.R. 1971. "Focal" color areas and the development of color names. *Developmental Psychology* 4:447-455.
- Kay, P., Berlin, B. and Merrifield, W. 1991. Biocultural implications of systems of color naming. *Journal of Linguistic Anthropology* 1:12-25.
- Kay, P. and McDaniel, C.K. 1978. The linguistic significance of the meanings of basic color terms. *Language* 54:610-646.
- Lucy, J.A. 1992. *Language Diversity and Thought: A Reformulation of The Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
- Moss, A.E. 1989. Basic colour terms: problems and hypotheses. *Lingua* 82:313-332.
- Newhall, S.M., Nickerson, D. and Judd, D.B. 1943. Final report of the OSA Subcommittee on the spacing of the Munsell colors. *Journal of the Optical Society of America* 33:385-418.
- Ratner, C. 1989. A sociohistorical critique of naturalistic theories of color perception. *The Journal of Mind and Behaviour* 10:361-372.
- Ruhlen, M. 1987. *A Guide to the World's Languages: I: Classification*. Stanford: Stanford University Press.

African Linguistics at the Crossroads: Papers from Kwaluseni
RK HERBERT (ed), pp 599-628. Köln: Rüdiger Köppe. 1997.

NOUN CLASSIFICATION IN SWAHILI:

A COGNITIVE-SEMANTIC ANALYSIS USING A COMPUTER DATABASE

ELLEN CONTINI-MORAVA

1 INTRODUCTION

The phenomenon of noun classification has long been of interest to linguists and anthropologists because understanding the basis for grouping nouns together as members of a class hints at a system of cognitive or cultural classification underlying the system of linguistic classification. However, the question what, if any, semantic principles can explain the groupings of nouns into classes in Bantu languages has been controversial. The received wisdom is that although some generalizations can be made, there is a lot of arbitrariness in these systems. In this paper I will suggest that the diagnosis of arbitrariness rests on an overly restrictive definition of what 'semantic coherence' means, and that a cognitive-semantic approach reveals more systematicity than might appear at first.

This paper is a report on research in progress on the semantics and syntax of the noun class system of Swahili. Swahili has a typical Bantu noun class system, but its status as a lingua franca has led to the assimilation of an unusually large number of loanwords from genetically unrelated languages, especially Omani Arabic, Persian, and various Indian languages (and more recently English). The need to accommodate nouns of foreign origin, some of which fit the phonological forms but not the semantic content associated with the various noun classes, has challenged the resources of the system (see Nurse and Hinnebusch 1993, chapter 3; henceforth N + H). Thus Swahili is an interesting case study for looking at continuity and change in noun class systems.