

A cross-cultural study of English and Setswana speakers on a colour triads task: A test of the Sapir–Whorf hypothesis

Ian R. L. Davies*, Paul T. Sowden, David T. Jerrett, Tiny Jerrett
and Greville G. Corbett

Department of Psychology, University of Surrey, Guildford, Surrey GU2 5XH, UK

We report a cross-cultural study of speakers of Setswana and of English carried out as a test of the linguistic relativity hypothesis (the Sapir–Whorf hypothesis). These languages differ in their number of ‘basic’ colour terms—English has eleven and Setswana has five—and in the position of some colour category boundaries. Speakers of the two languages did a ‘triads’ task in which they chose which of three colours was least like the other two. There were two types of triad: ‘controls’, for which any linguistic influences should lead to the same choices, and ‘experimental’, for which any linguistic influences should lead to different choices by the two groups. Thus the universalist position predicts that the choices of the two samples should be essentially the same for all triads, whereas the relativist position predicts that choice should be the same for the control triads, but differ for the experimental triads. The most striking feature of the results was that the choices made by the two samples were very similar for both kinds of triads, thus supporting universalism. But, there were also small but reliable differences associated with the linguistic differences, thus supporting Whorfianism. Overall, it appears that there is a strong universal influence on colour choice but this universal influence can be moderated by cultural influences such as language, a position consistent with ‘weak Whorfianism’.

The Sapir–Whorf hypothesis (Sapir, 1921; Whorf, 1956) is perhaps the most well known and the most provocative conjecture about the relationship between language and thought. According to the theory, speakers of different languages necessarily construe the world differently, and are locked into the world view given to them by their language: ‘The “real world” is to a large extent unconsciously built on the language habits of the group’ (Sapir, 1951 [1929]). The status of the Whorfian hypothesis within psychology has changed from one of broad acceptance before about 1970, to one of general rejection since 1970 (see Brown, 1976). However, we believe that there are still grounds for believing that colour language can affect colour cognition, and we report here a cross-cultural study that tested this conjecture (see Davies & Corbett, 1997 for a review of Whorfianism and colour cognition).

The shift from belief in linguistic relativity to belief in colour universals was fuelled more by a general conceptual shift in the *Zeitgeist* than by any substantial body

* Requests for reprints.

of empirical evidence. Further, this shift was more marked in psychology than in its allied disciplines such as anthropology or sociology (see Saunders & van Brakel, 1997). Two key studies fuelled the shift. First, Berlin & Kay (1969), working within linguistics and anthropology, claimed to have shown that underlying the apparent diversity in colour lexicons, there was a simple universal structure. Their key insight was that the foci (the best exemplars) of colour categories were relatively invariant, whereas category boundaries varied considerably across languages. Further, Kay & McDaniel (1978) proposed that the constraints on possible colour lexicons originated in the underlying perceptual physiology. Second, Rosch-Heider (e.g. Heider, 1972; Heider & Olivier, 1972), in a crucial series of studies on the Dani of Indonesia, appeared to have shown that these 'stone-age' people had cognitive representations of colour much like Americans, despite having only two basic colour terms available to describe colour. She interpreted her data as supporting universal perceptual physiological structures underlying colour cognition (but see Lucy, 1992; Lucy & Shweder, 1979; Ratner, 1989; and Simpson, 1991 for alternative interpretations).

Even if there are colour universals based on common physiology, this does not exclude the possibility that colour cognition might be modulated by language. Perception can be modified by experience, and experience can be shaped by language. For instance, discrimination on basic sensory dimensions can be improved with practice (e.g. Fahle & Edelman, 1993). Further, learning to categorize a perceptual dimension, such as brightness or saturation, improves discrimination along that dimension (Goldstone, 1994). In learning to speak a language, it is likely that attention would be directed at the foci and to the boundaries of the basic colour categories of the language. Thus, if colour perception is 'plastic', the process of learning colour language should result in enhanced discrimination to these critical regions of colour space.¹ If such perceptual learning occurs, then speakers of languages with different colour categories would have different distributions of sensitivity across colour space due to the differing positions of category boundaries and foci encoded in the languages.

Davis & Corbett (1997) found evidence consistent with the above conjecture. They compared speakers of Setswana (spoken in Botswana in southern Africa) with English speakers on a colour grouping task. Setswana has a single term for BLUE or GREEN, whereas English has two basic terms for the same region. Participants were asked to sort a representative set of 65 colours into groups so that 'members of a group looked similar to each other'. The basic pattern of grouping was strikingly similar for the two samples but, nevertheless, the Setswana sample was more likely to group BLUE colours with GREEN colours than the English sample was. This result is consistent with the presence of the *blue-green* category boundary in English, amplifying the perceptual differences of stimuli on differing sides of the boundary (learned distinctiveness), or with shared category inclusion attenuating perceptual differences for the Setswana sample (learned equivalence).

Kay & Kempton (1984) compared the Tarahumara (from Mexico), whose language also has a BLUE or GREEN term (*siyóname*) with speakers of English on a

¹ Bornstein & Korda (1984, 1987) and Boynton, Fargo, Olson & Smallman (1989) have found categorical effects on hue judgment that are consistent with this speculation. However, these studies are restricted to a single language, English.

'triads' task. Participants were asked which of three colours from the BLUE-GREEN border region was the least similar to the other two. When two colours both fell in the same English category, say *blue*, and the third colour fell in the English category *green*, English speakers were more likely than the Tarahumara speakers to select the third colour. Kay & Kempton interpret their results as showing language learning 'stretching' physical differences across category boundaries, and thus supporting 'modest Whorfianism'. However, the number in the sample was very small (there were four Tarahumara speakers and five English speakers) and ideally there should have been a 'control' task in which the expected choices of the two language groups were the same. For instance, three colours that all fall in the English category *green* should produce the same choices in both language groups as there is no category boundary separating any of the colours in either language. Similar choice by the two groups would reinforce the inference that the differences in the experimental task were due to linguistic effects, rather than reflecting some general influence on choice. As it stands, the results are consistent with the Sapir-Whorf hypothesis, but it would be reassuring to see such an important result replicated on a larger sample, using other language groups, and incorporating the control task mentioned above.

We report here a comparison of English and Setswana using a version of the triads task. The two languages differ in the number of 'basic' colour terms in the language (Berlin & Kay, 1969), and in the position of several colour category boundaries. English has 11 basic colour terms: *white, black, red, green, yellow, blue, brown, purple, pink, orange* and *grey* (Davies & Corbett, 1995). Setswana has five basic terms: *boswen* 'white', *bontsho* 'black', *bobibidu* 'red', *botala* 'grue' and *borokwa* 'brown' (Davies *et al.*, 1992). The most obvious difference between the languages is the relatively small number of basic terms in Setswana compared to English. One manifestation of the differing number of basic colour terms is the particularly clear difference in how the languages categorize the BLUE-GREEN region described earlier; we exploit this difference in our selection of colours for the triad task. While the difference in categorization of the BLUE-GREEN region is the clearest difference in the positions of colour category boundaries, there is also a range of interesting, but more graded differences. First, the range of the Setswana term *bobibidu* 'red' includes much of ORANGE and PINK. Second, although Setswana does not have a clear basic term for YELLOW, there are two competing terms for that region. *Sethunya* 'yellow-blossom' denotes focal YELLOW, but is not widely used to denote other parts of the YELLOW region. Rather, *bosetlha* 'light brown' is the most prevalent term used for non-focal YELLOWS. Third, there also are several non-basic terms used to denote the GREY-BROWN region such as *bowebu* 'grey' and *botuba* 'dun'. In brief, there are graded differences in the position of category boundaries that we exploit in the design of the triads.

There were two kinds of triads: 'control' and 'experimental'. For the controls, linguistic influences should lead to the same choice in each language, whereas for the experimental triads linguistic influences should lead to different choices by the two language samples. We scaled the potential linguistic influences using a 'nominal-overlap index' (NOI): the NOI measures the extent to which two colours were designated by common colour terms, and the NOI thus varies in accord with the differing positions of category boundaries across languages. Even for the control

triads, where the tile colour with the lowest NOI was the same for both languages, the extent of this 'nominal isolation' varied. Thus, if there is a linguistic effect on colour choice, there should be an ordinal effect, and a quantitative effect detectable in participants' choices. First, the most frequently chosen colour in the triads task should be the one with the least nominal overlap with the other two triad members; and second, the frequency with which a colour is chosen should be a function of the degree of nominal isolation.

Kay & Kempton (1984) used triads composed of relatively similar colours whereas we used some triads with relatively large perceptual differences between the colours. We measure 'perceptual distances' as distances in CIE (1976)² L^* , u^* , v^* colour space. Kay & Kempton's intra-triad perceptual distances range from about 20 to 40, whereas ours range from about 25 to 120. To the extent that the measure of perceptual distance is valid, if perceptual isolation determines choice, participants should choose the most isolated tile colour in colour space as the odd one out.

We treat the correlation between the NOI and the observed choices as one test for linguistic influence on performance. If there are linguistic influences on colour choice, then the intra-language correlation between the NOI and observed choice should be high. Conversely, if the primary influence on choice is some universal perceptual/cognitive process, then first, the correlation of the observed choices between languages should be high; and second, this correlation should be high, even for the experimental triads, despite the two languages differing in which colour is the most nominally isolated. Besides the correlational approach just outlined, we also compare performance on the control triads with that on the experimental triads for the two languages using the classical experimental approach. To the extent that there are linguistic influences on colour choice, scores should be essentially the same for both languages for the control triads, but differ for the experimental triads.

² Within the CIE system each colour is made up from three notional primaries— X , Y and Z —and the proportions of these three in any colour— x , y and z —must sum to 1. The x and y values, plus a measure of lightness (Y), form the coordinates of the CIE chromaticity diagram (z is given implicitly by $z = 1 - (x + y)$). Every possible colour has a unique locus in three-dimensional (Y , x , y) space, and these tri-stimulus values may be used to convert the Color-Aid stimuli into Munsell, or OSA notations, through conversion tables in, for instance, Newall, Nickerson & Judd (1943).

Although each colour has a unique location in tri-stimulus space, the space does not map onto colour perception very usefully. In particular, equal distances in the space do not correspond to equal perceptual steps. The CIE 1976 system (L^* , u' , v') represents colours in a transformed space which is approximately perceptually equal: L^* is a transform of Y and u' and v' are transforms of x and y respectively. BLUE colours have relatively low u' and v' coordinates and lie towards the origin; GREEN colours have relatively low u' but high v' scores; RED colours have high u' but low v' values; and YELLOW colours have relatively high u' and v' values; the achromatic colours have u' values of around 0.20 and v' values of around 0.47, but they vary on L^* .

In order to represent perceived distance between colours (total colour distance) the CIE recommend using a further colour space: (L^* , u^* , v^*). The coordinates u^* and v^* are transforms of u' and v' respectively. This transformation takes into account that the perceptual consequences of variations in (u' , v') depend in part on the value of L^* . The total colour difference in (L^* , u^* , v^*) is given by the Euclidean distance:

$$[(L^*_1 - L^*_2)^2 + (u^*_1 - u^*_2)^2 + (v^*_1 - v^*_2)^2]^{1/2}.$$

See Hunt (1987) for further information on the CIE system.

Method

Participants

There were two groups of participants that we designate by the first language spoken by each group: English and Setswana. There were 40 people in each sample, half were men and half were women, and each group had a similar age distribution: mean = 39 years; range = 23–68 years. The English sample was from Guildford, in Britain; and the Setswana sample was from villages around the small town of Kanye in southern Botswana; they were members of the Bangwaketse tribe. All were mother-tongue speakers of their respective languages, and none of the Setswana sample knew more than a few words of English.

Stimuli

There were 15 triads each consisting of three 'tile colours' that were 50 mm squares of Colour-aid coloured paper mounted on thin plywood. Triads C1 to C9 are the control triads—the nominal predictions are the same for each language—and triads E1 to E6 are the experimental triads—the nominal predictions are different for each language. The tile colours used in the triads are shown in Fig. 1. This shows the loci of the tile colours in each triad in the (u^* , v^*) plane of the CIE 1976 (L^* , u^* , v^*) colour space; the two-dimensional distances between loci give some indication of the perceptual distances between stimuli, but true perceptual distances are distances in the full three-dimensional space.³

The triads were displayed on a tray covered in uniform grey material. The experiment was carried out under natural light at the same time of year for each sample; this was the summer for the English sample and winter for the Setswana sample. The experimenter avoided presenting the stimuli in deep or dappled shade or in direct sunlight, but no formal measure of the illuminant was taken.

Table 1 shows the most frequent term used to describe each tile colour, and the percentage of the sample that used that term to name the tile colour (the data are taken from Davies & Corbett, 1995 for English; and Davies *et al.*, 1992 for Setswana and none of the two samples reported here were involved in the naming tasks).

Table 2 gives a measure of nominal overlap (NOI), and a measure of 'perceptual isolation' within a triad, for each tile colour and for each language. The NOI is the sum of a tile colour's nominal overlap with each of the other tiles in the triad. The nominal overlap between two tiles is the sum of the overlap scores for each common term used to name a pair of tile colours. For each common term the smaller of the two measures of frequency of use across the two tiles contributes to the overlap score for the two tile colours. For example, for triad C1 for the English sample, the terms used by more than one person and the percentage of the sample that used each term for each tile are as follows: gray-2 = *grey* (95.7 per cent); YOY-T4 = *yellow* (74.5 per cent), *orange* (4.3 per cent), *cream* (6.4 per cent); Y-hue = *yellow* 89.4 per cent (none of the terms used by just one person were in common across two tile colours). The overlap scores between gray-2 and the other two tile colours is thus zero as there were no common terms in either case. On the other hand, the overlap score between YOY-T4 and Y-hue is the smaller of 74.5 and 89.4, and this score is shown in Table 1, column 4. If there had been a further term in common (e.g. if Y-hue had been called *cream* with a score of 4.3 per cent) then the smaller of the two scores for the second common term would be added to the scores for the first common term.

The perceptual distance measure is the proportion of the total intra-triad perceptual distance due to each tile. Specifically, the perceptual distances are the distances between tile colours in CIE (L^* , u^* , v^*) coordinates; the index for each tile colour is the sum of its distances from the other two tile colours, divided by the sum of the three inter-tile colour distances. Table 2 also shows in bold which tile colour in each triad is the most nominally isolated (the smallest NOI), and which tile colour is the most perceptually isolated (the largest perceptual distance).

³ The full CIE specification of the stimuli is available from the authors.

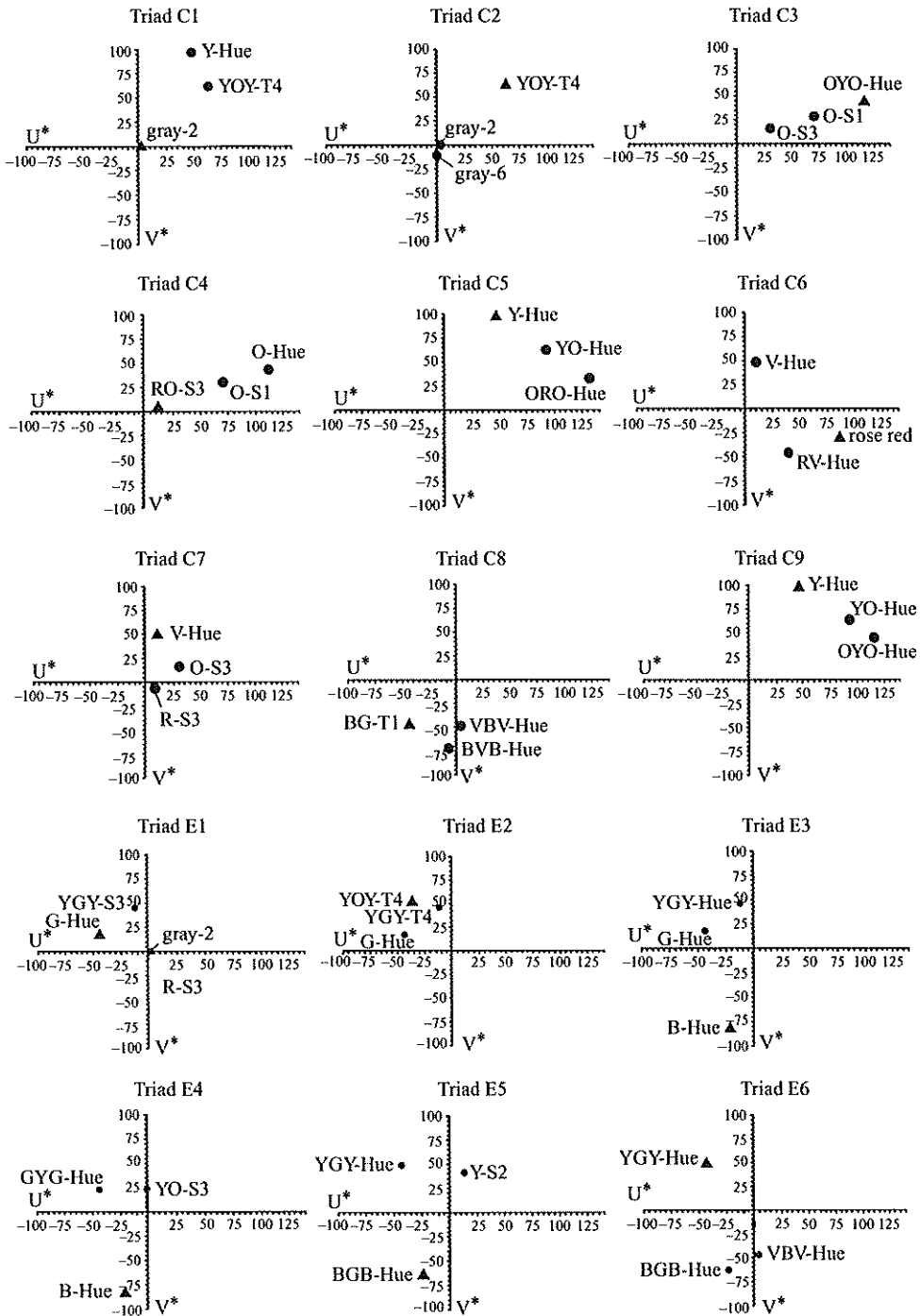


Figure 1. Perceptual distances between triad members for the control triads (C) and the experimental triads (E). The most perceptually isolated tile is shown as a ▲.

Table 1. The most frequent term and the percentage of the sample that used that term (%F) for each language for each tile colour in the control (C) and experimental (E) triads

Triad	Colour-aid code	English		Setswana	
		Term	%F	Term	%F
C1	Gray-2	grey	95.7	botuba	67.5
	YOY-T4	yellow	87.5	bosetlha	45.0
	Y-hue	yellow	89.4	sethunya	92.5
C2	YOY-T4	yellow	89.4	bosetlha	45.0
	Gray-2	grey	95.79	botuba	67.5
	Gray-6	grey	85.1	boebu	42.5
C3	OYO-hue	orange	80.9	bohobidu	20.0
	O-S1	tan	42.6	borokwa	35.0
	O-S3	brown	91.5	borokwa	77.5
C4	O-hue	orange	83.0	bohobidu	60.0
	O-S1	tan	42.6	borokwa	35.0
	RO-S3	brown	80.9	borokwa	65.0
C5	ORO-hue	orange	42.6	bohobidu	70.0
	YO-hue	yellow	61.7	mathubapula	45.0
	Y-hue	yellow	89.4	sethunya	92.5
C6	Rose red	pink	55.3	bohobidu	22.5
	RV-hue	purple	78.7	selaole	30.0
	V-hue	purple	78.7	selaole	40.0
C7	V-hue	purple	78.7	selaole	40.0
	R-S3	brown	25.5	borokwa	29.8
	O-S3	brown	91.5	borokwa	82.5
C8	VBV-hue	purple	59.6	selaole	40.0
	BVB-hue	blue	74.5	botala	37.5
	BG-T1	blue	63.8	botala	47.5
C9	OYO-hue	orange	95.7	bohobidu	20.0
	YO-hue	orange	63.8	mathubapula	45.0
	Y-hue	yellow	89.4	sethunya	92.5
E1	Gray-2	grey	95.7	boebu	47.5
	YGY-S3	green	85.1	bosetlha	7.5
	G-hue	green	91.5	botala	82.5
E2	YOY-T4	yellow	74.5	bosetlha	45.0
	YGY-S3	green	85.1	bosetlha	7.5
	G-hue	green	91.5	botala	82.5
E3	B-hue	blue	83.0	botala	62.5
	G-hue	green	91.5	botala	82.5
	YGY-S3	green	85.1	bosetlha	7.5
E4	B-hue	blue	83.0	botala	62.5
	GYG-hue	green	89.4	botala	70.2
	YO-S3	green	51.1	botuba	10.0
E5	BGB-hue	blue	85.1	botala	60.0
	YGY-hue	green	87.2	botala	52.5
	Y-S2	green	61.7	botuba	10.0
E6	YGY-hue	green	87.2	botala	52.5
	BGB-hue	blue	85.1	botala	60.0
	VBV-hue	purple	59.6	selaole	35.0

Key. *botuba* 'pale'; *bosetlha* 'light brown'; *sethunya* 'yellow-blossom'; *boebu* 'dun'; *bohobidu* 'red'; *borokwa* 'brown'; *mathubapula* 'sand'; *selaole* 'purple'; *botala* 'grue'.

Table 2. Perceptual distance and nominal overlap scores for each tile colour for each language. The most isolated tile colour in each triad on the perceptual distance measure and on the nominal overlap measures are indicated in bold (C = control triad; E = experimental triad)

Triad	Colour-aid	Perceptual distance	Nominal overlap	
			English	Setswana
C1	Gray-2	41.9	0.0	10.0
	YOY-T4	26.8	74.5	77.5
	Y-hue	31.4	74.5	72.5
C2	YOY-T4	43.8	0.0	5.0
	Gray-2	26.2	85.1	87.5
	Gray-6	30.1	85.1	82.5
C3	OYO-hue	38.4	4.3	5.0
	O-S1	25.0	34.0	35.0
	O-S3	36.6	34.0	35.0
C4	O-hue	35.2	36.2	0.0
	O-S1	25.0	78.7	72.5
	RO-S3	39.8	42.6	72.5
C5	ORO-hue	36.8	0.0	0.0
	YO-hue	25.0	68.1	30.0
	Y-hue	38.2	68.1	30.0
C6	Rose red	41.2	0.0	22.5
	RV-hue	25.1	34.0	42.5
	V-hue	33.7	34.0	35.0
C7	V-hue	38.6	12.8	0.0
	R-S3	26.3	38.3	35.0
	O-S3	35.0	25.5	35.0
C8	VBV-hue	30.9	38.3	22.5
	BVB-hue	29.5	61.7	52.5
	BG-T1	39.6	83.0	45.0
C9	OYO-hue	33.6	36.2	20.0
	YO-hue	25.0	78.7	82.5
	Y-hue	41.4	42.6	82.5
E1	Gray-2	33.3	0.0	7.5
	YGY-S3	28.5	85.1	7.5
	G-hue	38.2	85.1	0.0
E2	YOY-T4	29.1	0.0	7.5
	YGY-S3	31.3	85.1	7.5
	G-hue	39.6	85.1	0.0
E3	B-hue	41.2	0.0	62.5
	G-hue	26.7	85.1	62.5
	YGY-S3	32.1	85.1	25.0
E4	B-hue	41.6	0.0	62.5
	GYG-hue	29.2	51.1	62.5
	YO-S3	29.2	51.1	0.0
E5	BGB-hue	40.1	0.0	60.0
	YGY-hue	30.2	61.7	60.0
	Y-S2	29.7	61.7	0.0
E6	YGY-hue	43.3	0.0	60.0
	BGB-hue	28.9	19.2	60.0
	VBV-hue	27.8	19.2	15.0

Procedure

The experimenters were first-language speakers of the appropriate language and the trials were conducted in the appropriate language. The instruction ('which colour looks least like the other two colours') was translated into Setswana and back-translated into English. The back-translation matched the original English at the first attempt. Pilot work had shown that the task was understood by samples from both language groups. Nevertheless, a practice triad consisting of two RED tile colours and a GREEN tile colour was presented first to check that the instructions were understood. In each case participants selected the GREEN tile colour as the odd one out, and then the 15 test triads were presented, one at a time, on the grey cloth. The experimenter placed the three tile colours on the grey cloth, at approximately equal distances from each other, in a triangular layout, and asked the participant which was the least like the other two. A different random order was used for each of the 40 participants within a language sample, but the same 40 sequences were used for each language sample. The position of the tile colours within a triad relative to the participant was also randomized. After the completion of the 15 triads, participants were asked how they did the task.

Results

Table 3 shows the frequency with which each language group selected each tile colour within a triad as the odd one out. It can be seen that the pattern of choice across the 15 triads is similar for the two languages: for 12 out of 15 triads the most frequent choice is the same for both languages. The three exceptions are triads C3, E5 and E6. In triads C3 and E5 two tile colours have joint equal top scores for Setswana, and one of these tile colours is the same as the most frequent tile colour for English. Thus it is only for triad E6 that the most frequently chosen tile colour is different for the two languages: YGY-hue is the most frequent choice for English with a score of 31, whereas VBV-hue is the most frequent choice for Setswana with a score of 23. The broad similarity across the two languages is reflected in the large correlation between the scores for the two samples for the 45 tile colours (15 triads \times 3 tiles): $r = .87$; $p < .001$.

For the control triads (same nominal predictions) the most frequent choices for the two languages are the same for eight out of the nine triads. For four of the control triads (triads C4, C5, C8 and C9) the perceptual distance prediction and the nominal prediction conflict (see Table 3). For triads C4, C5 and C8, the most frequent choice is consistent with the perceptual distance prediction, whereas for triad C5 the most frequent choice is consistent with the nominal prediction.

Despite the three exceptions to the nominal predictions just described, the correlation between the NOI score and the frequency of choice across the 45 tiles is reasonably large for English ($r = -.74$, $p < .01$), but the correlation is smaller for Setswana ($r = -.35$, $p < .05$). (In fact the correlation between the Setswana frequencies and the English NOIs is larger than the correlation between the Setswana frequencies and the Setswana NOIs: $r = -.64$, $p < .01$.) For the experimental triads, the correlation for Setswana is just $-.10$ (n.s.) compared to $.85$ ($p < .001$) for English. The correlation between perceptual distance and frequency of choice ($r = .58$) is lower than the equivalent NOI-frequency correlation ($r = -.74$) for English, but for Setswana the perceptual distance-frequency correlation ($r = .61$) is greater than the equivalent NOI-frequency correlation ($r = -.35$).

Table 3. Frequencies with which tiles were selected for each triad for each language (C = control; E = experimental)

Triad	Colour-aid code	Frequency tile chosen	
		English	Setswana
C1	Gray-2	36	35
	YOY-T4	0	3
	Y-hue	4	2
C2	YOY-T4	38	32
	Gray-2	0	0
	Gray-6	2	8
C3	OYO-hue	29	19
	O-S1	3	19
	O-S3	8	2
C4	O-hue	11	10
	O-S1	1	0
	RO-S3	28	30
C5	ORO-hue	21	27
	YO-hue	4	0
	Y-hue	15	11
C6	Rose red	21	25
	RV-hue	11	2
	V-hue	8	13
C7	V-hue	15	9
	R-S3	2	3
	O-S3	23	28
C8	VBV-hue	16	16
	BVB-hue	1	0
	BG-T1	23	24
C9	OYO-hue	5	12
	YO-hue	2	0
	Y-hue	33	28
E1	Gray-2	36	27
	YGY-S3	1	1
	G-hue	3	12
E2	YOY-T4	37	32
	YGY-S3	1	1
	G-hue	2	7
E3	B-hue	36	29
	G-hue	0	1
	YGY-S3	4	10
E4	B-hue	34	23
	GYG-hue	1	5
	YO-S3	5	12
E5	BGB-hue	28	19
	YGY-hue	1	2
	Y-S2	11	19
E6	YGY-hue	31	15
	BGB-hue	2	2
	VBV-hue	7	23

The degree of similarity between the two samples' scores seems to be lower for the experimental triads than for the control triads. This can be seen in two ways. First, comparing the size of the largest frequency for each triad across the two languages shows that for the control triads, Setswana has the highest score for five triads (C4, C5, C6, C7 and C8) and English has the highest score for four triads (C1, C2, C3 and C9). However, for the six experimental triads, English has the highest score in each case. Second, the distribution of scores within a triad is about the same for the two languages for the control triads, but the distribution is flatter for Setswana than for English for the experimental triads. This difference reflects the fact that the scores for Setswana for the tile colour predicted by the Setswana NOI are greater than those for the English sample.

To test the impression that the degree of inter-language similarity differs for the control and experimental triads, we derived a score for consistency with the NOI prediction for each participant for the two types of triad. This score was the number of triads for which the participant followed the nominal prediction, expressed as a proportion of the number of triads in the set. There are of course two such NOI predictions for the experimental triads (one for each language) and hence more than one consistency score could be derived. Here we report consistency scores with the English NOI, but, as we shall see, the results are essentially the same if we use the Setswana NOI. If participants follow their respective nominal predictions, then scores should be similar across languages for the control triads, but less similar for the experimental triads. Fig. 2 shows the mean scores across participants for the

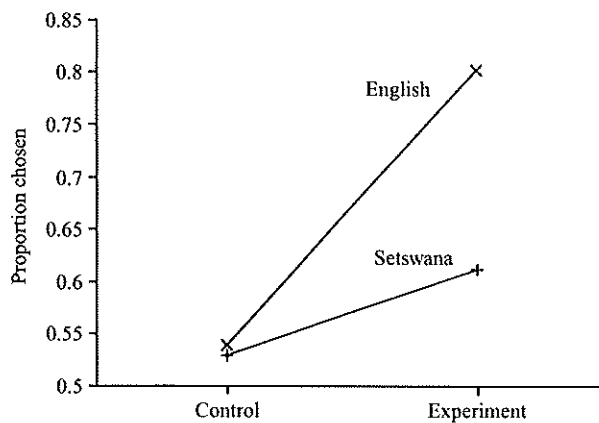


Figure 2. Proportion of tiles chosen predicted by the English NOI, for both language groups.

control and experimental triads for both language groups. Two-way ANOVAs on language by triad type (control-experimental) with repeated measures on the last factor found that both main effects and the interaction were significant. The mean scores collapsed across triad type were lower for Setswana than for English ($M_s = 0.57$ and 0.67 , respectively) ($F(1, 78) = 7.74$, $p < .01$). The mean scores collapsed across language were lower for the control triads than for the experimental

triads ($M_s = 0.53$ and 0.70 , respectively) ($F(1, 78) = 54.71, p < .001$). Finally, as can be seen in Fig. 2, although the English scores were higher than the Setswana scores for both the control (0.54 compared to 0.53) and the experimental triads (0.80 compared to 0.61), the size of this difference was greater for the experimental triads than for the control triads ($F(1, 78) = 14.94, p < .001$). This interaction is consistent with the Setswana sample being less likely to select the English NOI prediction for the experimental triads than the English sample, even allowing for the somewhat lower mean for the Setswana sample compared to the English sample on the control triads. The critical result is the interaction between language and triad type. The equivalent interaction is still significant if we compare the likelihood that participants select the colour predicted by the Setswana NOI, rather than the English NOI.

Discussion

The most prominent pattern in the results is the similarity in the choices of the two language groups. For most triads the most frequently chosen tile colour was the same for both languages, and this similarity is reflected in the large positive inter-language correlation between the scores for the 45 tile colours ($r = .87$). The tile colours chosen as the odd ones out tend to be the same for both languages even for those triads such as triad C6 and C8 where consensus in Setswana over the tile colour names is relatively low (see Table 3). Further, the similarity between the two language groups is also high for the experimental triads, although the most nominally isolated tile colours are not the same for the two language groups. These results suggest that the association between the tile colour chosen and the degree of nominal isolation in Setswana must be relatively low. The correlation (r) between the NOI and the observed choices is $-.35$ for Setswana, compared to $-.74$ for English. However, if we only consider the experimental triads, the correlation for Setswana is $-.10$ compared to $-.85$ for English. The low correlation for Setswana for the experimental triads reflects the fact that the Setswana sample tend to select the same tile as the English sample, rather than the tile colour with the greatest nominal isolation in Setswana.

On the other hand, there are differences in choice that are reliably associated with linguistic differences. The Setswana sample was more likely to select the tile colour predicted by the Setswana nominal prediction for the experimental triads than was the English sample. This difference was small, but significant. For four out of the six experimental triads the most frequent choices were the same for both samples, and were consistent with the English nominal prediction. However, the Setswana scores for the tile colour predicted by the English NOI were smaller than the English scores in every case. Conversely, the Setswana scores for the tile colour predicted by the Setswana NOI were higher than the English scores in every case. Further, for triad E5 for the Setswana sample, the tile colour predicted by the English NOI had the same score as the tile colour predicted by the Setswana NOI, while for triad E6 the tile colour predicted by the Setswana NOI was the most frequent choice for the Setswana sample.

Remember that for five out of six of the experimental triads the most perceptually isolated tile colour defined in terms of CIE distance (see later discussion) was the

same as the most nominally isolated for English, but not for Setswana. Thus, linguistic and perceptual influences were opposed for the Setswana sample. Further, the size of the perceptual influence that had to be overcome in order for linguistic influences to be expressed was relatively large (see Fig. 1 and Table 2). However, the likely degree of linguistic influence also varied across triads. The NOI does not reflect linguistic distance perfectly; it sometimes fails to distinguish between nominal isolation due to a tile colour being anominate (not included in any linguistic category) and nominal isolation due to a tile colour being included in a different category to the other two tile colours in a triad. For five of the experiment triads (E1 to E5) one tile colour is virtually anominate in Setswana: the most frequent name for these tile colours is used by just 10 per cent of the sample or less (see Table 1). For the remaining experimental triad (E6), however, two of the tile colours fall within a category boundary in Setswana (*botala* 'grue') and the third tile colour has a reasonable degree of inclusion in a different category (*selaole* 'purple') used by 35 per cent of the sample. It is perhaps notable that the largest deviation from the English NOI prediction for the Setswana sample occurred for triad E6. It appears that inclusion in a different category has more influence on choice than only non-inclusion in the category to which the other two tile colours belong.

We have so far relied on the universalist premise that distance in CIE 1976 (L^* , u^* , v^*) space is a valid measure of perceptual distance for both language groups (although it does not represent perceptual distances perfectly even for English speakers). At the same time we have used the method of triads to establish perceptual distances within our triads.⁴ To the extent that our data reflect perceptual judgments, they imply that the perceptual spaces of the two samples differ. Thus our assertion that for the Setswana there were large perceptual distances to be overcome to allow nominal effects to emerge only makes sense within a universalist framework. Within a relativist framework, however, it is more useful to think of perceptual space being shaped or perhaps 'warped' by language (and possibly other influences). Our data then suggest that the colour spaces of English and Setswana speakers differ. We cannot of course on the basis of our data offer a full revision of CIE perceptual space for Setswana speakers. The data do suggest that, although the topology of the CIE space may be universal, the local metrics may differ across cultures. In accord with this space-warp framework, Kay & Kempton (1984) interpret their triads' data as reflecting the presence of a category boundary 'stretching perceptual differences' for their English-speaking sample, while the Tahuramura, with no equivalent category boundary, choose on the basis of 'undistorted' perceptual distances.

So far we have assumed that the differences in our data reflect perceptual differences. It is also possible that the differences between the two language samples may be due to some participants using a labelling strategy, rather than a perceptual one as they were instructed to do. That is, perhaps some participants chose the tile colour with the different name as the odd one out. We cannot rule this possibility out; however, all participants asserted during their debriefing that they used a perceptual strategy.

⁴ Compare other studies using the method of triads such as Indow (1988).

We set out to test a question derived from the Whorfian hypothesis: can language affect perception? We have found differences in the behaviour of speakers of different languages associated with differences between the languages. However, in trying to decide whether the behavioural differences reflect perceptual differences, we are hindered by the vagueness of the concept of perception, and by conceptual disagreement over the notion of perception (e.g. Fodor & Pylyshyn, 1981; Turvey, 1977). In exploring performance differences across languages further, we try to avoid the debate over direct versus indirect perception. Rather, we are trying to map out the nature and locus of linguistic influence using a battery of tasks that range from traditional 'low-level' perceptual tasks (measurement of difference thresholds) through to tasks traditionally regarded as having a significant cognitive component, recognition memory tasks. As the cognitive component increases, so does the incentive to use a labelling strategy, and we thus expect a corresponding increase in the differences between language groups. However, if there are also detectable effects on the 'low-level' tasks, this would suggest that the effect was truly perceptual according to most approaches to perception.

In summary, our data show a striking similarity between language groups in their choice of similarities and differences amongst colours. This pattern supports the universalist position rather than the relativist position. But, in addition, there are small, but reliable differences between the two samples associated with linguistic differences. Thus overall, we may conclude that the universal constraints on colour perception may be modulated by small cultural influences, including language.

Acknowledgements

This work was supported by the ESRC (grant R000 23 1958), which we gratefully acknowledge. We would like to thank three reviewers for their constructive comments.

References

- Berlin, B. & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley and Los Angeles: University of California Press.
- Bornstein, M. H. & Korda, N. O. (1984). Discrimination and matching within and between hues measured by reaction times: Some implications for categorical perception and levels of information processing. *Psychological Research*, *46*, 207–222.
- Bornstein, M. H. & Korda, N. O. (1987). Some psychological parallels between categorization processes in vision and audition. In S. Harnad (Ed.), *Categorical Perception*. New York: Cambridge University Press.
- Boynton, R. M., Fargo, L., Olson, C. X. & Smallman, H. S. (1989). Category effects in color memory. *Color Research and Application*, *14*, 229–234.
- Brown, R. (1976). In memorial tribute to Eric Lenneberg. *Cognition*, *4*, 125–153.
- Davies, I. R. L. & Corbett, G. G. (1995). A practical field method for identifying probable basic colour terms. *Languages of the World*, *9*, 25–36.
- Davies, I. R. L. & Corbett, G. G. (1997). A cross-cultural study of colour-grouping: Evidence for weak linguistic relativity. *British Journal of Psychology*, *88*, 493–517.
- Davies, I. R. L., MacDermid, C., Corbett, G. G., McGurk, H., Jerrett, D., Jerrett, T. & Sowden, P. (1992). Color terms in Setswana: A linguistic and perceptual approach. *Linguistics*, *30*, 1065–1103.
- Fahle, M. & Edelman, S. (1993). Long-term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research*, *33*, 397–412.

- Fodor, J. A. & Pylyshyn, I. (1981). How direct is visual perception?: Some reflections on Gibson's 'Ecological Approach'. *Cognition*, **9**, 139-196.
- Goldstone, R. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, **123**, 178-200.
- Heider, E. R. (1972). Universals in color naming and memory. *Journal of Experimental Psychology*, **93**, 10-20.
- Heider, E. R. & Olivier, C. C. (1972). The structure of color space in naming and memory for two languages. *Cognitive Psychology*, **3**, 337-354.
- Hunt, R. W. G. (1987). *Measuring Colour*. Chichester: Ellis Horwood.
- Indow, T. (1988). Multidimensional studies of Munsell color solid. *Psychological Review*, **95**, 456-470.
- Kay, P. & Kempton, W. (1984). What is the Sapir-Whorf hypothesis? *American Anthropologist*, **86**, 65-79.
- Kay, P. & McDaniel, 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 Reformation of the Linguistic Relativity Hypothesis*. Cambridge: Cambridge University Press.
- Lucy, J. A. & Shweder, R. A. (1979). Whorf and his critics: Linguistic and nonlinguistic influences on color memory. *American Anthropologist*, **81**, 581-615.
- Newhall, S. M., Nickerson, D. & 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. *Journal of Mind and Behaviour*, **10**, 361-372.
- Sapir, E. (1921). *Language*. New York: Harcourt Brace.
- Sapir, E. (1951 [1929]). The status of linguistics as a science. In D. Mandelbraum (Ed.), Berkeley: University of California Press. Originally published in *Language*, **5**, 207-214.
- Saunders, B. A. C. & van Brakel, J. (1997). Are there nontrivial constraints on colour categorization? *Behavioral and Brain Sciences*, **20**(2).
- Simpson, C. (1991). Colour perception: Cross-cultural linguistic translation and relativism. *Journal for the Theory of Social Behaviour*, **21**, 409-429.
- Turvey, M. T. (1977). Contrasting orientations to the theory of visual information processing. *Psychological Review*, **84**, 67-88.
- Whorf, B. (1956). *Language, Thought and Reality*. Cambridge, MA: MIT Press.

Received 9 August 1996; revised version received 21 November 1996

