J. Child Lang. **26** (1999), 23–47. Printed in the United Kingdom
© 1999 Cambridge University Press

Why is colour naming difficult?*

NICK BRAISBY
Department of Psychology, London Guildhall University

JULIE DOCKRELL
Division of Psychology, South Bank University

(*Received* 30 *May* 1995. *Revised* 18 *May* 1998)

ABSTRACT

A suggestion exists in the child language literature that the meanings of natural kind terms are acquired before the meanings of colour terms. Explanations have typically claimed that object terms are more salient than property terms. Such explanations, however, tend to ignore the fact that natural kind terms refer to categories with sharp, clear boundaries while colour terms refer to categories with unclear or variable boundaries. Nonetheless, there has been little evidence to show that the delay in the acquisition of colour terms arises from these semantic properties. This study compares natural kind and colour naming (and corresponding comprehension) by 48 children, ranging in age from 3;0 to 5;5. The results suggest that, *contra* the salience view, the apparent delay in colour naming may be explained on solely semantic grounds.

INTRODUCTION

Why is colour naming difficult?

The semantic relation between colours and colour terms appears to be particularly difficult for children to gauge, even though the basic cognitive prerequisites to succeed are present. Young children can sort correctly by

[*] The order of authorship is alphabetical. We are grateful to Robin Campbell, Edward Chan, John Clibbens, Robin Cooper, Bradley Franks, James Hampton, Helen Moss, Philip Smith and two anonymous referees for helpful comments and discussions; and to Jonathan Patrizio for help with data entry. Nick Braisby also acknowledges the British Academy under the partial support of whose postdoctoral research fellowship this research was conducted. Address for correspondence: Nick Braisby, Department of Psychology, London Guildhall University, Calcutta House, Old Castle Street, London E1 7NT, UK. tel: +44 171 320 1086; fax: +44 171 320 1236; e-mail: braisby@lgu.ac.uk.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

colour, and colour terms are introduced to them in the first few years of life often with an obvious link to a particular object such as 'red ball' or 'blue ribbon'. Moreover, children can make relatively accurate discriminations along the dimensions of hue, saturation and brightness (Cook, 1931) and could, at least in principle, use colour terms with their conventional reference. Yet purported delays in the development of children's colour lexicons were commented on in the late nineteenth century (Darwin, 1877, cited in Bornstein, 1985*a*). More recently difficulties have been noted both in acquiring an initial colour lexicon (Rice, 1980) and in the subsequent course of development (Cruse, 1977; Campbell, MacDonald & Dockrell, 1982; Andrick & Tager-Flusberg, 1986; Davidoff & Mitchell, 1993). Bornstein (1985*b*) suggests that children only start to use colour terms with their true referential meaning between the ages of four and seven. Indeed, Davidoff (191:149) claims that 'children are like the blind in knowing which words are colour names but being unable to link them with visual stimulation.'

Data supporting such views are based mainly on informal observation and normative assessment of the colour lexicon. Implicit contrasts are drawn between the protracted period of development for accurate use of colour terms and the ease with which children acquire other lexical items to which they are frequently exposed. The difficulty in initially acquiring colour terms was highlighted in a training study by Rice (1980). She reports taking an average of 1080 trials to teach children who had no colour terms the terms 'red,' 'green' and 'yellow' – an extremely surprising phenomenon given the speed at which children acquire lexical items in fast mapping studies (see Dockrell & Campbell, 1986). Data collected by Johnson (1977) indicates how small the productive colour lexicons of children actually are. Children at the age of 2;6 correctly named a mean of three colours. No direct comparisons were made with other lexical items but, once again, the suggestion from the literature is that there is something special about mapping colour terms to their correct denotations. Moreover, in a comparative experimental study, Bornstein (1985*a*) has demonstrated that children have greater difficulty acquiring colour-to-name mappings than shape-to-name mappings. The claims of delayed acquisition in development are based on alleged difficulties with the colour lexicon *per se*, and not with specific colour terms or with more general cognitive factors.

The purported delay in colour naming is surprising given that the average three-year-old also has all the necessary linguistic tools to use colour terms correctly. They know the type of attribute that these terms denote, i.e. colour, and they have a range of entries in their lexicons which refer to colours. Young children have little difficulty in generating colour terms when asked to do so. Thus it appears that a specific difficulty exists in establishing the correct mappings between colour terms and their referents, that is colours (see also Soja, 1994, in support of this argument).

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

Although empirical evidence is equivocal, since colour is often not specifically contrasted with other semantic domains or with other attributes of varying salience, there are a number of possible explanations as to why this difficulty might occur for the young child. One possibility is that the mapping may be difficult to establish because it is not salient to the child. From this perspective colour is viewed as a peripheral attribute in the developing lexicon. Arguably, if a child's attention could be drawn to the relevant dimension then mappings should be encouraged (Campbell & Stewart, 1992). A second possible explanation is that the child's difficulties with colour terms arise because of the conceptual structure of the colour domain. Keil (1987) has argued that different types of concept may be represented differently and that these different representations might have major consequences for the acquisition of word meanings. One possibility is that the child could have difficulties establishing the basis for the BOUNDARIES of colour terms. In this case, the difficulties would be primarily determined by the semantic properties of colour terms: two properties in particular offer credence to this view. First, colour terms are vague. That is, colour categories tend to merge into one another, rather in the manner that 'cup' and 'bowl' may merge into one another (cf. Labov, 1973). Secondly, although following from the first, as Miller & Johnson-Laird (1976) indicate, colour categories are not hierarchically organized. For instance, to claim that burgundy is a kind of red is not to claim that 'burgundy' and 'red' lie in the same hierarchical relation as do 'chicken' and 'bird.' While it is true that all chickens are birds, not all burgundies are acceptable instances of 'red' – some may be regarded as more purple than red. Indeed, evidence for such a possibility can be found in no less a source than a dictionary. Burgundies range from 'blackish-purple to purplish-red,' olive is 'yellow-green,' aquamarine is 'greenish-blue' and cream is a 'yellowish-white' (William Collins & Sons Co. Ltd, 1982). Thus colours may have no single superordinate, a fact that is perhaps unsurprising given that colours can be produced via the mixing of other colours. However, note the contrast with domains which are hierarchically organized, such as the animal kingdom. A chicken is unequivocally a bird, and a bait is unequivocally a mammal: despite the fact that bats have characteristics of birds, they are NOT bird-mammals. Domains which are genuinely hierarchical in this way are characterized by categories having a unique immediate superordinate, quite unlike the case of colour.

In support of the salience argument, Lewandowsky (1908), arguing for an ONTOLOGICAL explanation, claimed that colour naming is difficult for the child because colour is an inherent part of an object, and hence that colour naming must necessarily be secondary to object naming. A prerequisite for colour naming would then appear to be appropriate knowledge of the colours of objects. However, Davidoff & Mitchell (1993) have shown that the converse occurs: children can have more difficulty with object-colour

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

knowledge than with colour naming. Equally troubling for such an ontological explanation is the fact that it suggests that all inherent parts of objects (including, for example, shape and size) will be subject to a naming delay – hence, one would not expect colour terms SPECIFICALLY to be delayed.

A variation on the salience theme is presented by Macnamara (1972). He adopts a LINGUISTIC explanation of the apparent delay in colour naming, whereby learning the names of objects is simply more important for the child than learning the names of colours. As a consequence, properties which are more pertinent to object naming, such as shape, assume a greater salience for the child than does colour, being less pertinent to object naming. Baldwin (1989) reports data which arguably support this explanation: when objects are not labelled children tend to sort them by colour, but when objects are labelled they tend to sort them by non-colour category. Thus, the salience of colour and, hence, the delay in naming could be seen as a product of linguistic context. This perspective is described as the linguistic hypothesis in subsequent sections.

The evidence in support of a conceptual or SEMANTIC distinction is less direct. Markman (1989) suggests that, for adults, nouns and adjectives betray a general psychological difference, where this difference relates to their differing semantics. This view is reinforced by Soja (1994) who reports that two-year-olds have specific difficulties in establishing the mappings between colour terms and their referents during acquisition. Equally suggestive is the adult literature on concepts. While prototype theory is thought to provide an effective account of colour (adjective) concepts (Rosch, 1975; Rosch & Mervis, 1975; Rosch, 1977), recent research suggests that natural kind (noun) concepts, contra prototype theory (Osherson & Smith, 1981; Cohen & Murphy, 1984; Smith, Osherson, Rips & Keane, 1988; Hampton, 1993), may be best understood in terms of underlying theories (Murphy & Medin, 1985) where these pertain to putative essences (Medin & Ortony, 1989). We will return to the semantics of colour terms towards the end of this paper. However, it is important to note how different are the salience and semantic explanations of the difficulties in colour naming. Although it is unclear as to precisely what is meant by salience, accounts often advert to notions such as frequency or familiarity. A semantic account, however, adverts principally to the properties of the category referred to by a term, and any ramifications these have for the corresponding concept. Thus, in principle, the two accounts will be independent of one another. Categories with similar semantic properties may nonetheless differ in salience, and categories of similar salience may nonetheless have very different semantic properties.

Despite the possibility that the developmental delay in colour naming may be offered a semantic explanation, as Markman points out, there is, as yet, little evidence in the developmental literature which bears directly on the issue. Indeed, we believe Markman's point is quite apposite. In spite of

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

consistent reports of delayed colour naming, and of reasonably well specified models of the colour naming process, we believe that there is, as yet, insufficient evidence for clear conclusions to be drawn. Studies of children's colour naming often do not directly compare colour with non-colour naming. Further, according to both ontological and linguistic explanations of the delay in colour naming, all colour terms should, in principle, be equally subject to a delay relative to non-colour control terms. However, no *a priori* grounds have been advanced for such a homogeneity of colour terms. In fact there are clear indications that colour terms do not form a homogeneous set. Berlin & Kay (1969) surveyed the basic colour terms for nearly a hundred languages as described in the ethnological literature, and tested informants speaking 20 of them. They concluded that there are certain focal colours which are more codable and more memorable than non-focal colours. Similarly, Heider (1971) demonstrated that certain colours are focal within colour categories. Given this non-homogeneity of colour categories it would be unsurprising if colour terms were not homogenous, contra both ontological and linguistic explanations. Lastly, relatively little is known concerning the full range of errors that children make in attempting to name colours, nor of their relation to errors made in non-colour naming. Yet an error analysis might allow the other factors underlying naming strategies to be identified, and may help to decide between the competing hypotheses concerning delayed colour naming.

The experiment reported here is an attempt to investigate systematically the production of colour terms in comparison with the production of non-colour terms. The non-colour terms chosen were terms for animals since these are natural kind terms for which there is already a considerable developmental and adult literature. Children of two age groups (3;0 to 4;2, and 4;9 to 5;5) were assessed for naming (twice) and comprehension (once) over a period of six weeks. Our aim was to fill the gaps in our understanding of the apparent development delay in colour naming, as discussed above. If colour terms are subject to a naming delay because colours are, in general, less salient than properties such as shape, then this should be true for the majority of colour terms. A separate analysis of the errors presented by the children is used as a means of tapping the children's representations of the two domains.

METHOD

Subjects

Forty-eight children from a London school were selected to participate in the experiment. Children were seen on four separate occasions over a period of six weeks. Ten children were excluded from the analysis because they missed

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

one or more testing sessions, leaving a total of 38 children who participated in all phases of the experiment, and for whom a complete data set is available. The class teachers were asked to exclude from the sample children with speech and language problems, hearing difficulties or any visual problems. The children were divided into two age groups: group 1 (Gp1), the younger children, mean age 3;5, range 3;0 to 4;2; and group 2 (Gp2), the older children, mean age 5;3, range 4;9 to 5;5. Within the constraints of the subject pool, groups were balanced for sex: Gp1 comprised 8 boys and 11 girls and Gp2 10 boys and 9 girls.

Materials

The lexical items were chosen to reflect a range of word frequency within the domains of colour terms and animal terms. The terms were further categorized as being either high or low frequency for each domain. Since there are no available measures of spoken word frequency the word frequency measures were taken from Francis & Kucera (1982) as a best approximation. Thus, the high frequency category contained terms such as 'cow' and 'bear,' 'red' and 'blue,' while the low frequency category contained terms such as 'platypus' and 'scorpion,' 'beige' and 'violet.' Appendix A lists the items used and their word frequency.

The animal stimuli comprised 14 monochrome line drawings chosen from the BPVS. Each picture was mounted on card. For the comprehension task four pictures were mounted on card. The location of the correct picture was randomly varied across trial items. The colour stimuli comprised 14 colour patches from a series of commercially available colour charts. Patches were chosen to be the best examplars of the corresponding colour. All stimuli were piloted on 10 adults, randomly selected from a subject pool, to ensure that the appropriate target stimuli were consistently selected in response to the animal and colour names. Each patch was mounted on card. In the comprehension task four colour patches were mounted on card and the location of the correct colour was similarly randomly varied across trial items. For both animal and colour stimuli, the choice items for the comprehension tasks were specifically chosen to reflect variations in similarity to the target item: thus there was one highly similar alternative, one similar alternative and one dissimilar alternative.

Procedure

Children were tested over a four week period. In the first week (week 1), a pretest was carried out to establish that children's lexical production and comprehension was within 1 S.D. from the mean. To establish the children's comprehension age, the British Ability Scales (BAS) was administered and

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

their level of vocabulary skills was assessed by the British Picture Vocabulary Scale. Children's naming (lexical production) of the animals and colours (test stimuli) was assessed in week 2 and the task was repeated again in week 4. Lexical comprehension of the test stimuli was assessed in the intervening week, week 3. The naming task was repeated to establish consistency of naming and to investigate acquisition of two trained items. The procedure was identical at both times.

For the production task, 14 animals and 14 colours were presented in blocks to the children. On the assumption that animal naming would be a familiar activity for the children, they always received the block of animal items followed by the block of colour items. This order helped ensure that children were clear about the task requirements for the colour naming. Within each block the 14 items were randomly presented. In the comprehension task, the block of animal items was presented first followed by the colour items. Within each block the items were randomly presented.

Children were individually tested in a room adjoining their class. The tasks were referred to as games and children were asked to come and play some games with the experimenter. For the production and comprehension tasks the children were asked to help a poor puppet who didn't know any words. Children were told 'This is my friend Sophie. She doesn't know many words. Can you help Sophie learn some new words?' In the production task the children were asked to 'Tell Sophie the names of the animals/colours' and in the comprehension task children were asked to 'Show Sophie the colour/animal.' Children were happy to participate in the study. Testing lasted approximately 15 minutes.

RESULTS

All children successfully completed the pretest. Table 1 contains the overall pattern of responding in the production task for the two domains (colour and animals) – responses were coded as correct, incorrect, don't know or forgotten. Each child contributed 28 responses for the comprehension task (total responses for the sample = 1064) and 56 responses for the production tasks (total responses for the sample = 2128). Children succeeded in the production task on approximately 50% of all trials. Table 2 gives patterns of responding in the production task according to word frequency (high or low), domain (natural kind or colour) and age group (Gp1 or Gp2). The incorrect responses were analysed using a specially designed coding frame and these are discussed following the principal analyses.

First, a hierarchical log–linear analysis was performed on the data from the production tasks, treating response type (correct, incorrect or don't know), age group (Gp1 or Gp2), domain (natural kind or colour), word frequency (high or low), and time (time 1 or time 2) as variables. Additionally, subjects

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

TABLE 1. *Overall pattern indicating numbers and kinds of response in the production task*

	Animal terms		Colour terms	
	Number correct	Correct (%)	Number correct	Correct (%)
Correct	532	50	489	46
Incorrect	309	29	458	43
Don't know	213	20	117	11
Forgotten	0	0	0	0

Note: Due to rounding, columns may not sum to 100.

TABLE 2. *Frequency of responses by semantic domain, word frequency and age group*

	High frequency				Low frequency			
	Animal		Colour		Animal		Colour	
	Gp1	Gp2	Gp1	Gp2	Gp1	Gp2	Gp1	Gp2
Correct	63	78	67	89	23	37	11	18
Incorrect	25	16	25	6	43	32	73	67
DK	12	6	8	5	33	31	17	15
For.	0	0	0	0	0	1	0	0

Note: The figures represent the percentage of responses of each type according to condition. Due to rounding, columns may not sum to 100.

TABLE 3. *Significant results in a hierarchical log-linear analysis of the frequency of all production responses*

Effect	χ^2	d.f.	Significance
Response type	373·44	2	0·0001
Response type × age	68·13	2	0·0001
Response type × time	14·01	2	0·001
Response type × domain	64·38	2	0·0001
Response type × word frequency	653·12	2	0·0001
Response type × domain × word frequency	62·16	2	0·0001

was entered as a factor, the standard approach to analysing the results of within-subject designs using log-linear models (D. B. Wright, personal communication, March 15, 1996; see also D. B. Wright, 1996, and Mills & D. B. Wright, in preparation). The approach treats the effect of subjects as a fixed effect (for a similar treatment see Brewer, 1988). The few responses

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

of 'forgotten' were excluded from the analysis to ensure that no cell had an expected frequency of less than five. All significant results are reported in Table 3.

The two-way interactions suggest that the pattern of responding in the production task differs significantly between the younger and older children, between natural kind and colour terms, between high and low frequency terms, and between the two times of testing. None of the two-way interactions challenge the explanatory power of the linguistic or salience hypotheses. However, the three-way interaction between response type, domain and word frequency is NOT PREDICTED by these hypotheses. This interaction effectively means that the difference in the pattern of production between high and low frequency terms itself differs significantly between natural kind and colour terms. Table 4 indicates the locus of this effect. While the pattern

TABLE 4. *Frequency of responses in the production task by semantic domain and word frequency*

	High frequency		Low frequency	
	Animal	Colour	Animal	Colour
Correct	70	78	30	14
Incorrect	21	16	38	70
DK	9	6	32	16

Note : The figures represent the percentage of responses of each type according to condition. Due to rounding, columns may not sum to 100.

of children's naming is virtually identical for high frequency colour and natural kind terms, for low frequency terms naming of colours is significantly worse than naming of natural kinds. Table 4 suggests that with low frequency animal terms children reply that they DO NOT KNOW the name whereas with low frequency colour terms children use other colour names to label the colours. Again, *contra* the linguistic and salience hypotheses, the analysis suggests that children's performance on low frequency colour terms is not to be explained in terms of a generally worse performance with colour terms, nor with a generally worse performance with low frequency terms. Indeed, *contra* these hypotheses, for high frequency instances, children's colour naming rivals, if not betters, their naming of natural kinds.

A further analysis was performed in order to counter any lack of robustness in the hierarchical log-linear analysis due to its assumption that the effect of subjects is fixed. Since each child was tested on seven terms from the four different kinds of term (low frequency colour, high frequency colour, low frequency animal, high frequency animal), a MANOVA was performed on the number of each kind that were correctly named. Each child was therefore

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

given four scores of 0–7. Within-subject factors were word frequency (2), domain (2), and age (2) was a between-subject factor. The results of the analysis are shown in Table 5.

TABLE 5. *Significant results in a MANOVA analysis of the number of correct production responses*

Effect	F	d.f.	Significance
Word frequency	576·95	1, 36	0·001
Domain × word frequency	49·53	1, 36	0·001
Time	41·44	1, 36	0·001
Age	19·68	1, 36	0.001
Time × domain	14·58	1, 36	0·001
Age × domain × word frequency	4·20	1, 36	0·048

All the effects which bear essentially on the competing linguistic, salience and semantic hypotheses were reproduced. Most importantly, a two-way interaction between domain and word frequency is observed as shown in Table 6. This confirms the picture, provided by the log-linear analysis, that low frequency colour terms prove particularly difficult.

However, there are two further effects which were not revealed by the earlier log–linear analysis. The first, a two-way interaction between time and domain, suggests that, over the two times of testing, naming for animal terms improves more markedly (mean week 1 = 3·2 out of 7; mean week 2 = 3·9) than for colour terms (mean week 1 = 3·1; mean week 2 = 3·3). The difference between colour and animal terms (overall) at week 1 is not significant ($F(1, 36) = 0{\cdot}05$, $p = 0{\cdot}82$), but a difference emerges at week 2 ($F(1, 36) = 6{\cdot}54$, $p < 0{\cdot}05$). We suggest that, for animal terms, the presence of highly similar nonmembers may facilitate their acquisition, since they help provide an informative contrast, whereas there can be no such cases for colour terms because of their vagueness. The second, weak three-way interaction between age, domain and word frequency might be expected on the view that the developmental delay in colour naming manifests itself in the interaction between domain and word frequency. If this is the case, we might expect the two-way interaction to be sensitive to age. It appears that, for all kinds of term, the older group names more items successfully than the younger group (low frequency colour: older group mean = 1·2, younger group mean = 0·7; high frequency colour: older group = 6·3, younger group = 4·7; low frequency animal: older group = 2·5, younger group = 1·6; and high frequency animal: older group = 5·4, younger group = 4·4). However, while the naming of low frequency colour terms is worse than low frequency animal terms ($t(18) = 3{\cdot}83$, $p < 0{\cdot}05$, for the older group; $t(18) = 4{\cdot}92$, $p < 0{\cdot}01$, for the younger group), only the older group show this difference

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

TABLE 6. *Mean number (maximum 7) of correct naming responses in the production task by semantic domain and word frequency*

	Animal terms	Colour terms
High frequency	4·9	5·5
Low frequency	2·1	1·0

for high frequency terms ($t(18) = -3{\cdot}07$, $p < 0{\cdot}01$, for the older group; $t(18) = -0{\cdot}76$, $p = 0{\cdot}46$, for the younger group).[1] This picture is consistent with the notion that the developmental delay in colour naming results from low frequency colour terms. Indeed, for high frequency terms, the naming of colour appears to improve more for colour terms than it does for animal terms; for low frequency colour terms, naming in the older group is better than in the younger group, but this improvement is smaller than that seen in low frequency animal terms.

Despite the convergence of our log–linear and MANOVA analyses, it could be argued that the patterns of response described above simply indicate that children have no knowledge of the denotation of the colour terms and that when this occurs for colours they opt to produce a high frequency term. Alternatively, their increased 'don't know' responses in the animal condition may reflect the fact that they know the term's denotation but cannot recall the accurate name. Such explanations would then support the view that children's colour lexicons are delayed compared to other domains – the interaction described above could then be argued to be the result of an artifact, and the linguistic and salience hypotheses maintained.

To evaluate this possibility similar analyses were performed but, in this case, were restricted so as to include a child's production responses ONLY for those items where the child succeeded in the corresponding comprehension task. Hence, the analysis only concerns production responses for those items where a child has some knowledge of its denotation. By this means we attempted to ensure that the poor production of low frequency colour terms was not due to extraneous factors (such as lack of knowledge of the term's denotation or lack of clarity as to what the experimental situation required) that might lead to poor performance in the task.

Restricting attention to trials for which an item was successfully comprehended means the analysis concerns 77·7% of all production trials. A hierarchical log-linear analysis was performed on these data, treating response type, age group, domain, word frequency and time as variables. Again, subjects was entered as a factor. This analysis yielded an identical

[1] A Bonferroni-type adjustment has been made to the significance levels for these post-hoc comparisons, and those reported later.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

TABLE 7. *Significant results in a hierarchical log-linear analysis of the frequency of production responses, where the corresponding comprehension trial yielded a correct response*

Effect	χ^2	d.f.	Significance
Response type	543·32	2	0·0001
Response type × age	26·66	2	0·0001
Response type × time	14·36	2	0·001
Response type × domain	27·35	2	0·0001
Response type × word frequency	394·07	2	0·0001
Response type × domain × word frequency	43·61	2	0·0001

TABLE 8. *Frequency of responses in the production task by semantic domain and word frequency, where the corresponding comprehension was successful*

	High frequency		Low frequency	
	Animal	Colour	Animal	Colour
Correct	74	85	39	24
Incorrect	18	11	33	60
DK	8	4	28	16

Note : The figures represent the percentage of responses of each type according to condition. Due to rounding, columns may not sum to 100.

TABLE 9. *Significant results in a MANOVA analysis of the number of correct production responses, where the corresponding comprehension was correct*

Effect	F	d.f.	Significance
Word frequency	530·50	1, 36	0·001
Domain × word frequency	52·63	1, 36	0·001
Time	34·58	1, 36	0·001
Age	23·77	1, 36	0·001
Time × domain	18·50	1, 36	0·001
Age × word frequency	6·25	1, 36	0·017

pattern of results to the analysis conducted on all responses. Table 7 contains all significant results.

These results are identical to those given previously (which concern all productions) and so we offer an identical interpretation. The locus of the three-way interaction of response type × domain × word frequency is illustrated in Table 8. Again, this shows that the poor performance with colour terms is entirely due to poor naming associated with low frequency

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

colour terms. However, because the analysis is restricted to instances of production where the corresponding term has been comprehended, this interaction cannot be explained by assuming that children have poorer knowledge of the denotations of low frequency colour terms. The restriction to instances of production where comprehension is correct, means that only cases where there is good evidence that children do have knowledge of denotations are considered. Once more a MANOVA analysis was performed on the number of correct productions in order to counter any lack of robustness in the log-linear analysis due to its assumption that the effect of subjects is fixed. The results of this analysis are shown in Table 9.

Again, all of the effects which bear essentially on the competing linguistic, salience and semantic hypotheses were reproduced including, most importantly, the two-way interaction between domain and word frequency. Most importantly, a two-way interaction between domain and word frequency is observed as shown in Table 10. This confirms the picture,

TABLE 10. *Mean number* (*maximum* 7) *of correct naming responses in the production tasks by semantic domain and word frequency*, *where the corresponding comprehension was correct*

	Animal terms	Colour terms
High frequency	4·7	5·2
Low frequency	2·1	0·9

provided by the log–linear analysis, that low frequency colour terms prove particularly difficult.

However, there are two further effects which were not revealed by the log–linear analysis. The first, a two-way interaction between time and domain, suggests that, over the two times of testing, naming for animal terms improves more markedly (mean week 1 = 3·2 out of 7; mean week 2 = 3·9) than for colour terms (mean week 1 = 3·1; mean week 2 = 3·3). Once again, the difference between colour and animal terms (overall) at week 1 is not significant ($F(1, 36) = 0{\cdot}63$, $p = 0{\cdot}43$), but a difference emerges at week 2 ($F(1, 36) = 8{\cdot}17$, $p < 0{\cdot}01$). As before, we suggest that the contrast provided by the presence of similar non-members may facilitate the acquisition of animal terms, whereas this is problematic for colour terms due to their vagueness. The second, two-way interaction between age and word frequency, tends to confirm the picture that the developmental delay in colour naming manifests itself as an effect of word frequency. Although, for low and high frequency terms, the older group names more items successfully than the younger group (low frequency: older group mean = 1·9, younger group mean = 1·1, $t(36) = -2{\cdot}99$, $p < 0{\cdot}05$; high frequency: older group = 5·7,

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

younger group = 4·2, $t(36) = -4{\cdot}93$, $p < 0{\cdot}01$), this improvement is more marked for high frequency terms than for low frequency terms. Again, this is consistent with the notion that the difficulty with colour naming may have its roots in a difficulty with low frequency terms.

Overall, these results appear not to support the linguistic and salience explanations of colour naming: colour terms do not, IN GENERAL, elicit worse naming than their non-colour matches. Colour naming is only worse in the case of low frequency terms. That is, there appears to be an intrinsic naming difficulty with low frequency colour terms and any explanation of the apparent delay in colour naming must surely advert to the particular properties of low frequency colour terms.

One source of evidence rests with the type of naming errors produced by the children. Error analysis can be used to develop a theoretical model of the children's naming processes (see Vitkovitch, Humphreys & Lloyd-Jones, 1993, for an analogous approach with adults' naming errors). All the children's errors were classified. The classification system was constructed on the basis of errors previously reported in the adult literature, developmental data and the responses provided by the children.

The error classification system was devised to capture general errors in production, as well as errors that rely upon the different properties of the animal and colour domains. General errors were (not mutually exclusively) classified as PHONETIC (sharing either an initial or final phoneme) NONSENSE (e.g. 'a bumpy'), LIKE A (e.g. 'like a dinosaur'), CIRCUMLOCUTIONS (i.e. a relevant comment about the item, though not including the item, e.g. 'big, large thing'), or as COMBINATIONS (e.g. 'fish horse', 'blue green'). Errors that seem to depend upon the semantic properties of animal terms (natural kind terms) were classified according to the properties of the animal referred to by the incorrect animal term. Attributes included CLOSE PHYSICAL RESEMBLANCE (e.g. 'mouse' for 'rat'), SAME LOCOMOTION (e.g. 'swimming', 'flying'), SIMILAR HABITAT (e.g. water for 'whale' and 'seal'), and SHARED IMMEDIATE SITUATION (e.g. seaside for 'seahorse' and 'fish'). Incorrect names were also classified according to whether they were SUPERORDINATES or SUBORDINATES of the term required, or if they were contrastive terms at the SAME LEVEL.

Error that depend upon the semantic properties of colour terms were classified according to the similarity of the colours referred to by the incorrect term and by the cord term: HIGHLY PLAUSIBLE (i.e. a variant of the same shade, e.g. 'pink' for 'red'), PLAUSIBLE (i.e. adjacent in the spectrum, e.g. 'orange' for 'red' or 'yellow'), or IMPLAUSIBLE (i.e. neither of the above). Errors also included cases where a colour was referred to RELATIVE to another (i.e. as a qualified version of another, e.g. 'dark red' for 'burgundy' or 'light green' for 'aquamarine').

A series of hierarchical log-linear analyses were performed on the classified error data, one for each error type. In addition to error type, age group (Gp1

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

or Gp2), domain (colour or natural kind), word frequency (high or low), and time (week 1 or week 2) were entered as variables. Subjects was also entered as a factor.

General errors

On average 10% of the errors were phonologically based. The percentage of these errors did not vary across word class or age group. However, both phonetic initial (change in likelihood $\chi^2_2 = 7{\cdot}796$, $p < 0{\cdot}01$) and phonetic end (change in likelihood $\chi^2_2 = 3{\cdot}790$, $p = 0{\cdot}05$) responses were more common for high frequency terms. It is worth noting that this pattern of errors might suggest the phonological influence occurs following semantic retrieval. That is the child has identified the appropriate target but has difficulty with the output phonological representation.

Nonsense responses, 'like a' responses and circumlocutions were infrequent (an average of 3% of all errors). There was no overall difference in 'like a' responses between colour and animal terms. However, nonsense responses were significantly more common for animal terms (change in likelihood $\chi^2_2 = 19{\cdot}863$, $p < 0{\cdot}0001$) as were circumlocutions (change in likelihood $\chi^2_2 = 4{\cdot}092$, $p < 0{\cdot}05$). In contrast, responses which were combinations of terms (such as 'blue green' or 'fish horse') occurred for both colours and animals terms. There was no difference in the production of combinations across the two age groups for the animal terms (8·6% and 5·3%). However, for the colour terms older children produced many more combinations (19·2% versus 2·4%). Overall these combinations were significantly more common for low frequency terms (change in likelihood $\chi^2_2 = 8{\cdot}438$, $p < 0{\cdot}01$).

Specific errors

Selection errors, where a lexical item from the same semantic domain was chosen present an interesting and contrasting picture. The majority of animal choices were of animals that bore a similarity to the target item on one or more dimension. Items which reflected a structural similarity to the target were more common for low frequency items (for physical similarity, change in likelihood $\chi^2_2 = 30{\cdot}905$, $p < 0{\cdot}0001$; for similar mode of locomotion, change in likelihood $\chi^2_2 = 5{\cdot}917$, $p < 0{\cdot}05$). In contrast 43% of the colour choices bore no relation to the target colour whatsoever. Moreover, if we only include in the analysis those children who were correct in comprehension a similar proportion (37·5%) of the colour errors were totally at odds with the target. For example, a child correctly identified brown in the comprehension task but called brown 'blue' in the production task. Some of the younger children's responses can be explained in terms of a crude strategy where

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

children use one high frequency colour term to label a range of colours. Six of the younger children label five or more discrete colours with the same term. The terms used include 'blue,' 'green,' 'pink,' 'grey' and 'white.'

There is no single explanations of these response patterns. However, our interpretation of the colour selection errors can be enhanced if we consider the difference between low frequency and high frequency terms. Significantly more low frequency terms were labelled in a relative fashion (change in likelihood $\chi^2_2 = 9{\cdot}015$, $p < 0{\cdot}01$). Moreover 50% of the low frequency errors were highly plausible alternatives indicating that the children were over-extending a known term. In contrast only 15% of the high frequency errors were highly plausible.

Further, we analysed the gross pattern of errors in terms of whether a response of 'don't know' was made, or an incorrect term produced. A four-way MANOVA was performed on the number of incorrect responses made minus the number of don't know responses made (thus each child could score between +7 and −7 in each condition): time (2) × domain (2) × word frequency (2) × age (2). Only three effects were significant: word frequency ($F(1, 36) = 17{\cdot}60$, $p < 0{\cdot}001$), domain ($F(1, 36) = 25{\cdot}32$, $p < 0{\cdot}001$) and the interaction domain × word frequency ($F(1, 36) = 89{\cdot}42$, $p < 0{\cdot}001$). Table 11

TABLE 11. *Mean number of incorrect responses minus the number of 'don't know' responses as a function of word frequency and domain*

	High frequency	Low frequency	(Row means)
Animal terms	0·83	0·39	0·61
Colour terms	0·66	3·76	2·21
(Column means)	0·74	2·08	1·41

contains the relevant means. Although there is a general preponderance to produce more incorrect than 'don't know' responses, note how striking is the cell for low frequency colour terms: compared with low frequency animal terms, in naming colours with low frequency colour terms children tend to produce many more cases of incorrect responses than 'don't know'.

These errors support the view that children treat natural kinds and colour terms in quite different ways. Children's colour errors suggest that they are sensitive to the vagueness of colour boundaries and use this fact when they offer alternative labels for colours. In contrast, children much more readily admit to not knowing an animal term and when they produce an alternative do so on the basis of a much more complex similarity matrix (see results of structural similarity reported above), in contrast to the high proportion of implausible errors generated for colour terms.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

Data concerning how adults judge the different animal categories in relation to each other, and the different colour categories in relation to each other provides important illuminating evidence to help explain the word frequency effect. As is shown in Table 12, adults typically judge certain pairs

TABLE 12. *Adults'* ($N = 14$) *judgements that the given categories overlap with or belong in other categories* (*either high or low in word frequency*)

	Animal terms	Colour terms
High frequency items reported as a kind of low frequency item	0	17 (17%)
Low frequency items reported as a kind of high frequency item	0	29 (30%)

of low and high frequency colours to be at least equivocal members of the same colour category. That is, colours are seen to form a continuum and adults do not unequivocally draw boundaries between them. Thus both burgundy and red are seen as instances of both the category 'red' and the category 'burgundy' (confirmation that colour terms do not form a hierarchy in the way that animal terms do). In contrast, adults do not consider any of the animal categories to be instances of any other of the animal categories: cows are not thought to be pigs, and pigs are not thought to be cows, etc. Thus, these results appear to suggest that the difference between colour and object naming may be a consequence of the domains being structured in very different ways.

DISCUSSION

According to linguistic or salience explanations, the delay in colour naming is due to colour being less salient than properties pertinent to object naming. Object naming is deemed to be more important to the developing child. However, our evidence casts doubt on this as the sole explanation since, for high frequency colour and natural kind terms, there appears to be no contrast in children's production (or, in fact, in comprehension, although this was not reported for sake of brevity). For low frequency terms, however, there is a marked contrast between colour and natural kind terms in children's production (and also comprehension). Indeed, our results suggest that previous reports of a delay in colour naming may have paid insufficient attention to the structure of the colour domain and also to factors such as word frequency.

Consequently, we are not faced with explaining a general developmental delay in colour naming but, rather a differential deficit for low frequency colour terms. This deficit cannot be explained in terms of a gross deficit (such

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

as salience) for colour terms or a gross deficit for low frequency terms. Rather, problems with colour naming may be entirely due to special factors which only (or largely) affect low frequency colour terms.

There are two important aspects of our results which we believe any account of children's colour naming must explain. First, an explanation is required of the fact that low frequency colour terms tend to elicit much poorer naming than their natural kind counterparts. Secondly, an explanation is required of the different patterns of responding for natural kind and colour terms (natural kinds tend to elicit more responses of 'don't know,' while colour terms tend to elicit more production of incorrect colour names). Explanations of these findings will lead to a better understanding of colour naming in particular, and also of naming in general. Our argument is that an explanation of these findings can be couched SEMANTICALLY, i.e. in terms of the broad semantic properties of natural kind and colour terms. According to this view, the first finding reflects the fact that colour and natural kind terms possess differently structured boundaries; the second reflects the fact that colour and natural kind terms possess different kinds of (semantic) criteria of application.

What evidence is there that the boundaries of colour and natural kind terms differ?

First, an explanation of the frequency effect is suggested by the judgements of our adult subjects. Generally, they considered instances of colour terms also to be instances of other colour terms, and they responded in this way more for low frequency colour terms than for high frequency colour terms. These results suggest that colour terms tend to overlap such that their boundaries merge and are relatively indistinct, and that this overlapping may be more consequential for low frequency terms. This explanation is supported by evidence which indicates that natural kind and colour terms differ markedly in the kinds of boundaries they possess.

The psychological literature on concepts contains evidence to the effect that adult concepts for natural kind and colour terms differ. Prototype theories (e.g. Rosch, 1975; Rosch & Mervis, 1975) are generally thought to provide accurate accounts of colour concepts. Indeed, Rosch (1977) argued that 'the domain which has most readily lent itself to the demonstration of a prototype and surround structure is that of colour'. While she attempted to extend this analysis to other semantic categories (see Lakoff, 1986), the literature has since re-evaluated the suggestion that natural kind concepts are organized around a prototype in the way that colour concepts are thought to be. The work of Keil (1987), Rips (1989), Medin & Ortony (1989), Gelman & Coley (1991), Murphy (1991) and others, has adopted what Medin &

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

Ortony label 'psychological essentialism'. Natural kind concepts according to this view specify possibly incomplete theories of their corresponding kind. For example, the concept 'water', rather than comprising a prototype of water, may consist of theories concerning the putative essential properties of water (cf. Kripke, 1972; Putnam, 1975).

This psychological difference in the concepts for natural kind and colour terms corresponds directly to a difference in their semantics. Psychological essentialism, for instance, is consistent with, and respects many intuitions of, the rigid designation view of meaning proposed by Kripke (1972) and Putnam (1975). According to this view, natural kind terms apply to objects in virtue of their possessing certain (almost exclusively non-observable) essential properties.[2] Whether or not an entity is a *cat* depends upon that entity possessing the essential properties of the underlying kind *cat* and possession of these properties is thought to be a strictly binary matter: either an entity possesses them or it does not, thus ensuring that the boundaries of natural kind terms are clear and sharp (cf. Kamp & Partee, 1995). And further evidence for the sharpness of natural kind concepts is provided by Armstrong, Gleitman & Gleitman (1983) who demonstrated that subjects who give graded goodness-of-example ratings to categories such as *fruit* and *vegetable* nonetheless regard their membership as being all-or-none.

In contrast and, according to C. Wright (1975) because they are observational (see later), colour terms are necessarily vague (Dummett, 1975). The significance of this point is that vague terms are thought to have unclear boundaries (Zadeh, 1965; Goguen, 1969; Fine, 1975), variable boundaries (Bosch, 1983; Goldstein, 1988) or, according to Sainsbury (1991), no boundaries at all. That is, unlike natural kind terms for which it appears for every instance there is a determinate answer as to whether or not it falls under the term, for colour terms there may be instances for which there is NO determinate answer as to whether they fall under a colour term.

Thus there is a convergence of views on the position that natural kind terms have clear, sharp and non-overlapping boundaries, a consequence of which is that each instance of a natural kind falls under one and only one natural kind term (excepting superordinate terms). There is also convergence concerning colour terms. They do not have clear, sharp boundaries – their boundaries are at best indistinct, leading to the consequence that particular colour shades, lying at the putative boundaries between different colour terms, may be considered to (partially) fall under a number of different colour terms. A particular shade, for example, may equally be considered to

[2] What is meant here is non-observable to most language users. Whether such properties are non-observable to all language users presumably depends upon one's views on the philosophy of science.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

be an instance of 'orange' as an instance of 'red' (just as dictionaries consider 'olive' to be 'yellow-green' – William Collins Sons & Co. Ltd, 1982). These differing consequences represent the first central feature of our semantic explanation detailed later. The second feature of our explanation concerns the different patterns of naming error for natural kind and colour terms.

What evidence is there that colour and natural kind terms differ in what counts as a semantically appropriate response?

Our error analysis shows that colour naming tends to elicit more productions of incorrect colour names than do natural kinds, and natural kinds tend to elicit more productions of 'don't know' than do colours. The strength of the semantic explanation we offer thus depends on explaining this differential pattern of responding. However, there are reasons to believe that this pattern is to be expected given the semantic properties of the two types of term.

First, natural kind terms are deemed to obey Putnam's hypothesis of the Division of Linguistic Labour (Putnam, 1975), whereby the arbiters of reference, in the case of apparent borderlines, are the appropriate scientific subset of the language community. Uncertainty over whether a natural kind term applies to borderlines can be settled, at least in principle, via recourse to this scientific community. This ensures that, even in cases of uncertainty, the boundaries of natural kinds can nonetheless be thought to be sharp and non-overlapping. There is also convergence that though there are clear criteria for the application of a natural kind term, these are determined by theories that make reference to (possibly unknown) essential properties. From either proposition it follows that individual language users may be unaware of the precise criteria for the application of a natural kind term (indeed this is an essential feature of the argumentation of both Kripke and Putnam). As a consequence, faced with naming a putative instance of a kind a response of 'don't know' is semantically appropriate – it is consistent with there being a determinate answer as to the name of the instance, but simply reflects a lack of information on the part of the language user.

The semantics of colour terms, on the other hand, indicate that colour terms differ substantially in this regard. First, colour terms are OBSERVATIONAL: according to C. Wright (1975:338), 'if one can tell at all what colour something is, one can tell just by looking at it'. He continues, 'the look of an object decides its colour, as the feel of an object decides its texture or the sound of a note its pitch. The information of one or more senses is decisive of the applicability of an observational concept' (p. 338). Unlike natural kind terms, the reference of colour terms is determined by observation, with the caveat that this be under normal conditions. This caveat is important for understanding the claim that vague terms are observational. The claim does not contrast with the observation that colour judgements, for instance, are relative to certain normal observing conditions. The claim is,

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

rather, that when those normal conditions are satisfied, the application of the term is then determined by observation and by observation alone.[3]

However, because colour terms are observational, the criteria for applying colour terms are NECESSARILY available to individual language users (assuming they are able to make the requisite observations). Thus, it appears that each individual language user should, in principle, be able to determine whether a given colour term applies. Indeed, for colour terms there is no 'colorific' subset of the language community which can arbitrate on questions of reference. Thus, assuming that a language user possesses the relevant colour terms and can make the required observations, a response of 'don't know' would not indicate a lack of information, but, in effect, would reflect an admission of a genuine indeterminacy. Hence, and in contrast to natural kind terms, responding 'don't know' in colour naming may be less semantically appropriate than generating an incorrect colour name.

The semantic explanation

In sum, the substantial semantic differences between natural kind and colour terms arises at their putative boundaries. While central or focal cases may betray no naming difference, borderline cases for natural kinds and colours do. Further, the pattern of naming errors ('don't know' being more predominant for natural kinds than for colours) reflects the different semantic properties of these two types of term. This type of explanation is strengthened by the fact that our adult subjects tended to regard low frequency colour terms as borderline instances of high frequency colour terms. In broad terms, the semantic explanation seems consistent with our findings concerning children's naming. In more detail, this semantic explanation proceeds as follows.

We assume that the frequency effect arises because of two factors. First, animal terms are non-overlapping, while colour terms overlap and merge at their boundaries. Thus, for a given colour, as Table 8 suggests, there are often (at least) two plausible names for that colour such that either could, in principle, be produced in colour naming. We assume that, in general, there will be asymmetries between these candidate names (e.g. perhaps they may differ in terms of word frequency, or along some other pertinent dimension). If there is more than one candidate name, this conflicts with the pragmatics of the naming task, where it is presupposed that there is only one appropriate term which should be produced. Thus, lastly, we assume that the asymmetries between the candidate names are exploited such that one name is preferred (e.g. perhaps the one higher in frequency; or the one that is more

[3] This is not to say that colour identification is not relative to observer and observing conditions. The claim that colour terms are observational is simply that their applicability is determined by observation provided those conditions are met.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

basic, cf. Berlin & Kay, 1969). In contrast, the fact that animal terms do not overlap means that there is less equivocation as to which animal name should be produced.

Secondly, the semantic difference between colour and natural kind terms offers an explanation of the gross pattern of naming errors we obtained. Colours tended to yield more productions of incorrect names, and fewer 'don't know' responses than animals, consistent, as suggested above, with Putnam's view of the semantic properties of natural kind terms. In contrast, the fact that colour terms are observational suggests that people should not in general defer in naming colours, that for such terms a response of 'don't know' is less semantically appropriate, and that any errors that arise should be cases where an incorrect colour name is produced. This is precisely what our findings reveal. Our claim, then, is that the two key aspects of our overall findings directly reflect the semantic properties of natural kind and colour terms.

Finally, this semantic explanation offers an interesting interaction with explanations of difficulties with colour terms that focus on meta-linguistic strategies. Soja (1994) suggests that the difficulty that two year olds experience in learning the mapping between colour terms and their referents may result from the contravention of the principles of contrast (Clark, 1987) and mutual exclusivity (Markman & Wachtel, 1988). For example, an instance of purple might be referred to as *purple*, *bluey red*, *pinky blue*, and so on, contravening mutual exclusivity. Soja argues that the contravention of these principles may lead to the mappings between colour terms and their referents being particularly difficult to infer. However, the semantic explanation which we have offered suggests that these principles are necessarily contravened in the case of colour terms BECAUSE they are vague and have ill-defined boundaries. Almost by definition, colours which are borderline between several colour terms necessarily may be referred to with different colour terms, thus contravening mutual exclusivity and contrast. So this meta-linguistic explanation of the difficulties in acquiring colour terms may itself be explained in terms of the semantic properties of colour terms.

CONCLUSIONS

In this paper, we have investigated the apparent delay in colour naming and offered an alternative, semantic explanation of this. Object terms, such as natural kind terms, refer to categories with clear or sharp boundaries. Colour terms, on the other hand, are known to have vague semantics, referring to categories with no clear boundaries. It follows that our results should also be obtained in comparing other natural kind terms (say, vegetable and mineral terms) with other vague terms (say, size and quantity adjectives). While evidence exists that children have difficulties in acquiring some adjectives it

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

remains a possibility that these difficulties rest in their vagueness, rather than solely in differential salience or in the application of meta-linguistic strategies. The semantic explanation advocated here would then predict that difficulties should arise primarily with borderline instances of the adjective category, and not with central instances.

Our semantic explanation also presupposes that older children attach more refined boundaries to their colour terms. However, divining boundaries for vague terms is an extremely difficult problem to solve, given that colour terms may have no clear boundaries or, at best, ones which vary according to context. In fact, as adults we know that borderline instances of vague terms are, almost by definition, very hard to classify. However, given what we know about the semantics of vague and natural kind terms, one way of investigating this problem is to contrast the ways in which colour and natural kind terms are conveyed and the patterns of deference attached to each.

Clearly, there is a need to extend our understanding of the child's developing lexicon. Our results suggest that a successful account will only be forthcoming once facts concerning the semantics of the relevant lexical items are incorporated.

REFERENCES

Andrick, G. R. & Tager-Flusberg, H. (1986). The acquisition of colour terms. *Journal of Child Language* **13**, 119–34.

Armstrong, S. L., Gleitman, L. R. & Gleitman, H. (1983). What some concepts might not be *Cognition* **13**, 263–308.

Baldwin, D. A. (1989). Priorities in children's expectations about object label reference: form over color. *Child Development* **60**, 1291–1306.

Berlin, B. & Kay, P. (1969). *Basic colour terms: their university and evolution*. Berkeley, CA.: University of California Press.

Bosch, P. (1983). 'Vagueness' is context-dependent: a solution to the Sorites paradox. In T. T. Ballmer & M. Pinkal, (eds), *Approach vagueness*. Amsterdam: Elsevier.

Bornstein, M. H. (1985*a*). Colour-name versus shape-name learning in young children. *Journal of Child Language* **12**, 387–93.

Bornstein, M. H. (1985*b*). On the development of colour naming in young children: data and theory. *Brain and Language* **26**, 72–93.

Brewer, W. F. (1988). Memory for randomly sampled autobiographical events. In U. Neisser & E. Winograd (eds), *Remembering reconsidered: ecological and traditional approaches to the study of memory*. Cambridge: C.U.P.

Campbell, R. N., MacDonald, T. & Dockrell, J. E. (1982). The relationship between comprehension and production and its ontogenesis. In F. Lowenthal, J. Vandamme and J. Cordier, (eds), *Language and language acquisition*. Plenum.

Campbell, R. N. & Stewart, S. (1992). Contextual salience and the elicitation of colour vocabulary. *Child Language Seminar*, *University of Glasgow*.

Clark, E. V. (1987). The principle of contrast: a constant on language acquisition. In B. MacWhinney, (ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Erlbaum.

Cohen, B. & Murphy, G. L. (1984). Models of concepts. *Cognitive Science* **8**, 27–58.

William Collins Sons & Co. Ltd (1982). The new Collins concise dictionary of the English language London: Author.

Cook, W. M. (1931). Ability of children in color discrimination. *Child Development* **2**, 303–20.

Cruse, D. A. (1977). A note on the learning of colour names. *Journal of Child Language* **4**, 305–11.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

Davidoff, J. B. (1991). *Cognition through colour*. Cambridge, MA: MIT/Bradford Books.
Davidoff, J. B. & Mitchell, P. (1993). The colour cognition of children *Cognition* **48**, 121–38.
Dockrell, J. E. & Campbell, R. N. (1986). Lexical acquisition strategies. In S. Kuczaj & M. Barrett, (eds), *Semantic development*. New York: Springer Verlag.
Dummett, M. (1975). Wang's paradox. *Synthese* **30**, 301–324.
Fine, K. (1975). Vagueness, truth and logic. *Synthese* **30**, 265–300.
Francis, W. N. & Kucera, H. (1982). *Frequency analysis of English usage : lexicon and grammar*. Boston: Houghton Mifflin.
Gelman, S. A. & Coley, J. D. (1991). Language and categorisation: the acquisition of natural kind terms. In S. A. Gelman & J. A. Byrnes, (eds), *Perspectives on language and thought*. Cambridge: C.U.P.
Goguen, J. A. (1969). The logic of inexact concepts. *Synthese* **19**, 325–73.
Goldstein, L. (1988). The Sorites as a lesson in semantics. *Mind* **97**, No. 387, 447–55.
Hampton, J. A. (1993). Prototype models of concept representations. In J. van Mechelen, J. A. Hampton, R. Michalski & P. Theuns, (eds), *Categories and concepts : theoretical views and inductive data analysis*. London: Academic Press.
Heider, E. R. (1971). 'Focal' colour areas and the development of colour names. *Developmental Psychology* **4**, 447–55.
Johnson, E. G. (1977). The development of color knowledge in preschool children. *Child Development* **48**, 308–11.
Kamp, H. & Partee, B. (1995). Prototype theory and compositionality. *Cognition* **57** (2), 129–91.
Keil, F. (1987). Conceptual development and category structure. In U. Neisser, (ed.), *Concepts and conceptual development*. Cambridge: C.U.P.
Kripke, S. A. (1972). Naming and necessity. In D. Davison & G. Harman, (eds), *Semantics of natural languages*. Dordrecht: Reidel.
Labov, W. (1973). The boundaries of words and their meanings. In C. J. N. Bailey & R. Shuy (eds). *New ways of analysing variation in English*. Georgetown: Georgetown University Press.
Lakoff, G. (1986). *Women, fire and dangerous things : what categories tell us about the nature of thought*. Chicago: Chicago University Press.
Lewandowsky, M. (1908). Ueber abspaltung des farbensinnes. *Monatsschrift fuer psychiatrie und neurologie*, **23**, 488–510.
Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review* **79**, 1–13.
Markman, E. M. (1989). *Categorisation and naming in children : problems of induction*. Cambridge, MA: MIT Press.
Markman, E. M. & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology* **20**, 121–57.
Medin, D. & Ortony, A. (1989). Psychological essentialism. In S. Vosniadou & A. Ortony, (eds), *Similarity and analogical reasoning*. Cambridge: C.U.P.
Miller, G. A. & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge: C.U.P.
Mills, C. & Wright, D. B. (in preparation). *Introducing log-linear modelling*. London: Sage Publications.
Murphy, G. L. (1991). Meaning and concepts. In P. J. Schwanenflugel (ed.), *The Psychology of word meanings*. Hillsdale, NJ: Erlbaum.
Murphy, G. L. & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review* **92**, 289–316.
Osherson, D. N. & Smith, E. E. (1981). On the adequacy of prototype theory as a theory of concepts. *Cognition* **9**, 35–58.
Putnam, H. (1975). The meaning of 'meaning'. In *Mind, language and reality, volume 2 : Philosophical papers*. Cambridge: C.U.P.
Rice, M. (1980). *Cognition to language : categories, word meanings and training*. Baltimore: University Park Press.
Rips, L. (1989). Similarity, typicality and categorisation. In S. Vosniadou & A. Ortony, (eds), *Similarity and analogica reasoning*. Cambridge: C.U.P.

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press

Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology : General* **104**, 192–233.
Rosch, E. (1977). Classification of real-world objects: origins and representations. In P. N. Johnson-Laird & P. C. Watson, (eds), *Thinking : readings in cognitive science*. Cambridge: C.U.P.
Rosch, E. & Mervis, C. B. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology* **8**, 573–605.
Sainsbury, R. M. (1991). *Concepts without boundaries*. London: King's College Department of Philosophy. (ISBN: 095172990X).
Smith, E. E., Osherson, D. N., Rips, L. J. & Keane, M. (1988). Combining prototypes: a selective modification model. *Cognitive Science* **12**, 485–527.
Soja, N. N. (1994). Young children's concept of colour and its relation to the acquisition of colour words. *Child Development* **65**, 918–37.
Vitkovitch, M., Humphreys, G. W. & Lloyd-Jones, T. (1993). On naming a giraffe as a zebra: picture naming errors across different categories. *Journal of Experimental Psychology* **19**, 243–59.
Wright, C. (1975). On the coherence of vague predicates. *Synthese* **30**, 325–65.
Wright, D. B. (1996). *Understanding statistics : an introduction for the social and behavioral sciences*. London: Sage Publications.
Zadeh, L. A. (1965). Fuzzy sets. *Information and Control* **8**, 338–53.

APPENDIX

WORDS USED AND THEIR WORD FREQUENCY

Animal	Frequency	Colour	Frequency
High			
cow	46	red	169
bear	24	blue	126
fly	23	green	85
mouse	20	grey	74
seal	17	brown	66
pig	14	yellow	52
peacock	6	purple	11
Low			
seahorse	2	pink	6
crocodile	1	olive	2
raccoon	1	cream	2
woodpecker	1	beige	1
porcupine	1	burgundy	0
scorpion	0	aquamarine	0
platypus	0	violet	0

https://doi.org/10.1017/S0305000998003638 Published online by Cambridge University Press