Learning in Context: Linguistic and attentional constraints on children's color term learning

Catherine G. O'Hanlon and Debi Roberson

Department of Psychology, University of Essex

Key words: COLOR TERMS, COLOR CATEGORIZATION,

LINGUISTIC CONTRAST, PERCEPTUAL SALIENCE, ATTENTION, MUTUAL

EXCLUSIVITY

Abstract

Three experiments investigated whether linguistic and / or attentional constraints might account for preschoolers' difficulties when learning color terms. Task structure and demands were equated across experiments, and both speed and degree of learning were compared. In Experiment 1, three-year-olds who were matched on vocabulary score were taught new secondary color terms by either corrective, semantic, or referential linguistic contrast. Corrective contrast produced more rapid and more extensive learning than either semantic or referential contrast, supporting the hypothesis that targeted linguistic feedback facilitates learning. Experiment 2 replicated and extended the first experiment with Italian children and found cross-cultural differences in the amount learned about colors named differently in the two languages. In Experiment 3, some of the children were introduced to the new terms within a context of enhanced perceptual salience. These children learned as fast and performed as accurately as those given corrective linguistic feedback in Experiment 1.

In contrast to the ease with which count nouns are learnt, often from a single exposure (Carey, 1978; Markman, 1989), there is substantial evidence (Clark, 1973a; Keil & Carroll, 1980; Macnamara, 1982) that young children find it particularly difficult to learn dimensional adjectives for relative size (e.g. *wide, tall)* and color terms (e.g. *red, blue)*. Even though, by eighteen months, children differentiate nouns from adjectives (Golinkoff, Mervis & Hirsch-Pasek, 1994), it takes several years of exposure before children master all eleven basic color terms of English (e.g. Bornstein, 1985b; e.g. Pitchford & Mullen, 2002; Shatz, Behrend, Gelman & Ebeling, 1996). Some recent studies (Franklin, Clifford, Williamson & Davies, 2005; Pitchford & Mullen, 2002; 2005) have found that many children can name the best examples of eleven color terms by four years of age, and Pitchford and Mullen (2001) found no selective difficulty in matching or comprehension, compared to other dimensional terms, when only two contrasting colors (red and green) were tested in isolation. However, with a larger set, that includes intermediate exemplars, children's color term use continues to be error prone up to six years of age (Roberson, Davidoff, Davies & Shapiro, 2004; Sandhofer & Smith, 1999)

At the same time, young children appear to understand that color terms form an independent lexical semantic category. Two-year-olds regularly respond with a color term to the question "what color is this?" although often incorrectly, but do not always respond appropriately to questions referring to other dimensional terms such as "how many?" or to the question "what is it?" (Backsheider & Shatz, 1993; Bartlett, 1977). However, knowledge of color terms seems unrelated to the ability to use them, at this stage. Roberson et al. (2004) tested 3-year-olds in two very different cultures, in the UK and in South-West Africa. When asked, "tell me all the colors that you know", both groups of children were as likely to list terms that they could not use correctly as those that they could. Sandhofer and Smith (1999)

also found that two-year-olds could classify red as a color word before they applied it

correctly. They suggested that learning color terms promotes selective attention to color, so

children only achieve a comprehensive conceptual representation of the color domain after

acquiring a sizeable color vocabulary.

A number of possible explanations have been offered as to why color terms are

relatively hard to master (Andrick & Tager-Flusberg, 1986; Baldwin, 1989; Braisby &

Dockrell, 1999; Macario, 1991; Rice, 1980; Sandhofer & Smith, 1999; Soja, 1994). The

difficulty cannot arise from perceptual limitations, since infants under 6 months of age make

fine grained perceptual distinctions both of form (Quinn & Eimas, 1996) and color (Franklin

& Davies, 2004). Moreover, with intensive training, over a sustained period, children can

learn to apply several color terms correctly before their second birthday (Cruse, 1977;

Mervis, Bertrand & Pani, 1995). Soja (1994) suggested that the problem may primarily be

attentional, since 2-year-olds can select, classify and match objects, as well as draw simple

inferences, on the basis of color, prior to learning color terms, providing nothing else varies

(shape, texture, size, etc.) (Soja, 1994; Tomikawa & Dodd, 1980). There is a well

documented 'shape bias' in novel word interpretation which might make young children less

likely to attend to color when matching objects that vary along competing perceptual

dimensions (Baldwin, 1989; Smith, Jones & Landau, 1992; Soja, Carey & Spelke, 1985).

Shape is highly diagnostic of object category (for natural kinds) and function (for artifacts),

so a bias towards attending to shape over other dimensions is useful for categorization

(Carey, 1991). Two-year-olds match objects on the basis of color in the absence of a label,

but on the basis of shape when a [unfamiliar] label is given (Baldwin, 1989), and are more

likely to give novel words a shape interpretation rather than a color interpretation when both

dimensions vary (Bornstein, 1985a). If children's difficulties in learning color terms arise

from attentional biases then, once their attention is successfully drawn to the color dimension,

learning to apply the set of color terms they can produce should be rapid. If, however, their difficulties arise at some other level, attentional manipulations alone may not suffice.

When learning to apply novel words children use mutual exclusivity (the expectation that novel words only apply either to novel objects, or to a novel property or feature of a known object; Markman & Wachtel, 1988). A child who already knows some color words, such as green and purple, but overextends these to inappropriate referents, might resist a color interpretation for a new word, such as *mauve*, because they rule out color as a potential referent on the basis of mutual exclusivity. In this case, explicit linguistic contrast might facilitate color term learning.

A number of studies have attempted to teach novel color words using linguistic contrast ("It's not X, it's Y"), and varying the non-linguistic contextual information to draw attention to the color dimension (Au, 1990; Au & Laframboise, 1990; Au & Markman, 1987; Carey & Bartlett, 1978; Dockrell, 1981; Gottfried & Tonks, 1996). Results have shown that this kind of feedback significantly improves 3-5 year olds' ability to learn novel color words even when provided on a single occasion (Au & Laframboise, 1990; Carey & Bartlett, 1978). Three kinds of contrast have been investigated within the domain of color term learning: semantic, corrective and referential.

Semantic linguistic contrast introduces a new color term by contrasting it with random familiar color terms. For example, a child might hear, "See this, it is mauve. It is not green and it is not yellow". The linguistic feedback should direct the child's attention to the color domain but the contrast does not generally provide useful information about which color it is. Au and Markman (1987) contrasted color and texture to investigate the effects of semantic contrast on color term acquisition in 3 and 4-year-olds. Children in the color condition were told, "Pass me the mauve square. See this, it is not green and it is not red. It is mauve". Children in the texture condition heard, "Pass me the rattan square. See this, it is not cotton

and it is not wood. It is rattan*". However, children at both age groups and in both conditions were more likely to interpret the new word as a texture term. Semantic linguistic contrast did not facilitate the learning of a new color word when the competing perceptual dimension was texture.

One possible reason for the poor performance with semantic contrast is that it presents the child with an 'unmotivated denial' of the random color terms used by the experimenter (Au & Laframboise, 1990). This might divert the child's attention and encoding processes to the negated items at the expense of the target word/referent.

Corrective linguistic contrast attempts to overcome this difficulty by targeting the child's own overextension errors (e.g. "It's not red, it's mauve"). Corrective contrast has been shown to be more effective than semantic contrast (Au & Laframboise, 1990), ostensive definition ("This is X") (Gottfried & Tonks, 1996), or simple negation ("It's X" – implying, within the context, "It's not Y") (Merriman, 1986) at teaching three-to-five year olds novel color terms. This kind of feedback explicitly informs the child of a lexical error, overcoming assumptions about mutual exclusivity. A third type of contrast has been termed referential linguistic contrast as it makes reference to known objects or dimensions that are in view at the time a new word is given, and uses a naturalistic, rather than an instructive context for word learning. Thus it more closely resembles everyday adult-child interactions. Carey and Bartlett (1978) used referential contrast in a kindergarten playroom setting by asking 3-year-old children to "bring me the chromium tray, not the red one" in the presence of two trays of identical shape. The children were free to interpret the novel word *chromium* in any way. However, hearing a known word (red) and an unknown one (chromium) being applied to a familiar object (a tray) should have directed their attention to the color domain. When assessed after one week on a comprehension task ("Which is the chromium/red one?"), a hyponym task ("Is chromium/red a color?") and a production task ("What color is this?"),

about half of the children had learned something about the word 'chromium'. However, children in a control group, who heard "pass me the chromium tray" without referential contrast, performed similarly to those in the experimental group. They might have picked up the correct tray by simply excluding the known, red one and ignoring the word *chromium* (Au, 1990). Heibeck and Markman (1987) found a similar level of success simply asking children to "bring me the chromium tray, not the other one". The benefits of referential linguistic contrast are controversial, and have not been compared to those of corrective or semantic contrast.

Across studies that have taught novel terms using different types of linguistic contrast, memory and attentional demands have also varied considerably. The poor results associated with linguistic contrast in some studies (Au & Markman, 1987; Dockrell, 1981) might be the result of increased task demands, especially if children in the experimental groups were directing attention to other perceptual dimensions at the time that the novel words were introduced.

The first experiment reported here attempted to reconcile previous findings on the effectiveness of linguistic contrast in children's acquisition of color terms. Learning contexts using corrective, semantic and referential linguistic contrast were presented within short cartoon-like video clips on a laptop computer. Task structure and demands were kept constant across all groups so that the effects of three different learning conditions could be meaningfully compared. If corrective contrast facilitated learning more than the other two kinds, this would further support the hypothesis that mutual exclusivity assumptions render color term learning harder. In addition, referential contrast might produce more learning than semantic contrast because the contrastive objects are in view when the linguistic information is given and because the contrastive input explicitly instructs a child to direct attention to the target referent by excluding known colors. The benefits of hearing new words contrasted with

the child's own label (corrective linguistic contrast), and those of having the contrastive objects in view at the time the contrastive input is given (referential linguistic contrast), might therefore produce similar results, but for different reasons.

The second experiment partially replicated and extended the results of Experiment 1 by exploring the effects of corrective linguistic contrast among a group of native Italian three-year-olds. Italian children might make less rigid assumptions of mutual exclusivity for one of the three colors tested (*teal*), because Italian adults use three terms to describe the range of colors in the English basic category blue ('blu' for dark blues, 'celeste' for light blues and 'azzuro' for blues towards the green boundary).[1] Italian children thus have a less consistent linguistic input for colors in this region than English children (because different terms are used) and should have a weaker expectation that terms are mutually exclusive in the blue area. If novel term learning is indeed hindered by assumptions of mutual exclusivity, they should learn more about 'teal' than English speaking children.

The third study addressed the possibility that difficulties in learning color words arise because children are biased to attend to other perceptual dimensions of stimuli (such as shape or texture) during word learning (Soja, 1994). This study examined the possibility that three-year-olds might find it easier to learn if the color dimension was made particularly salient at the time the new words were heard. If learning was facilitated under these conditions, then late color-word acquisition might be explained in terms of attentional rather than linguistic biases.

These studies assess the learning of secondary color terms in young children who have already mastered some basic terms, rather than the learning of initial basic terms because the

---

[1] Other languages bordering the Mediterranean also use more than one basic term in this area, (Greek: Androulaki, Pestana, Lillo & Davies, 2001; Russian:Davies & Corbett, 1997; Turkish: Ozgen & Davies, 1998)

use of linguistic contrast (its not X, its Y) requires that children already use some color terms, even if they fail to apply them correctly.

In sum, the present studies investigated the hypothesis that delayed color competence in preschoolers might be accounted for either by linguistic constraints in early word learning, or attentional biases which might prevent children from directing attention to the dimension of color within early word learning contexts. Finally, a direct comparison was undertaken of the of the corrective linguistic contrast employed in Experiments 1 and 2 with the perceptual salience manipulation used in Experiment 3

Experimental studies

Three-year-olds matched on chronological and vocabulary age (BPVS II score range 90-110) were taught three very low frequency color words (beige, crimson and teal; mean frequency=5, 10, 15 respectively) (Carroll, Davies & Richman, 1971) over five weeks. Computer-generated stimuli were used to keep the learning environment constant across groups and over time. The computer games and the linguistic content were designed to match conditions in previous similar studies but test objects varied only in color on any one training session. The objects used and their layout on the screen were kept constant within each training session, but varied from one session to another to avoid the possibility that children might simply learn an object-color pairing, or a target location. Children's color knowledge and their own labels for the targets were assessed through pre-test naming and comprehension tasks. Both comprehension and production measures were used to establish color knowledge (see Roberson et al., 2004 for a detailed description). A pre-test listing task was administered to investigate children's understanding that color terms form a discrete lexical set. The assessment tasks used at pre-test were also used on all subsequent assessment trials. Comprehension, rather than production, was used as the most sensitive measure of

learning[2]. Children were followed-up two weeks after the last training trial to assess learning retention.

The speed at which children learned information about target words, and the degree of learning achieved overall, as well as at different stages of the experimental timeframe, was calculated. Appendix I gives a detailed description of the method for calculating speed and degree of learning scores.

General Method

Participants

Native English speaking children were recruited from local nursery schools. All were tested for normal color vision with the City Color Vision Test (Fletcher, 1980), and had normal language development as assessed by the BPVS II measure of receptive vocabulary (Dunn, Dunn, Whetton & Burley, 1997). Groups were matched on vocabulary age and all children had vocabulary age between 3;0 -3;11 years (BPVS II score range 90-110).

Apparatus and Materials

An Apple Macintosh PowerBook G4 laptop computer was used to generate, measure and present test stimuli. Color stimuli were presented in cartoon-like video clips (approximately 32 x 20 cm) using SuperCard. Sound recordings were added using Sound Studio and presented via a Sennheiser HD495 headset. The fourteen colors used were typical examples of the eleven basic colors of English (white, black, red, blue, green, yellow, pink, purple, orange, brown, grey), and three very low frequency colors, beige, crimson and teal. An independent sample of 100 native English-speaking adults categorized the eleven basic colors selected; 98% overall naming agreement was obtained. Of the three low frequency targets beige falls within the range of the brown category for English speakers, crimson falls

---

[2] Asymmetry of success on productive and receptive tasks in young children may be accounted for by a difference in task difficulty, since, for productive naming the number of available responses is potentially infinite (Naigles, 2002; Zelazo, Frye & Rapus, 1996; Zelazo & Reznick, 1991).

within red and teal falls within blue, although all three are outside the range of best examples given by Berlin and Kay (1969). These targets were named inconsistently by adults, using a variety of low frequency terms (63%) or modified basic terms (37%) (e.g. light brown for beige). Only 6% used the terms beige, crimson and teal. Test stimuli were circles, stars, flowers or cars. Circles were used for naming and comprehension tasks. The other three shapes were used in the experimental training trials. A Minolta TV-Color Analyzer II was used to measure the physical properties of the fourteen color stimuli. CIE tristimulus co-ordinates, computed CIELuv $\Delta\varepsilon'$ values, and corresponding RGB values, are given in Appendix II.[3] Computed $\Delta\varepsilon'$ differences between targets and best examples (focals, Rosch, 1972) of the categories within which they fall were as follows: beige–brown = 18.32; crimson–red = 55.18; teal-blue = 31.05; teal-green = 36.47.

Procedure

Each child was tested individually in a quiet corner of the nursery. For computerized tasks children wore headphones and sat facing the screen at a distance that enabled them to touch it (approximately 40 cm). The study spread over a period of six weeks for each child. During the first week five pre-tests were administered, in the following order: (i) a listing task, (ii) a color vision screening test (iii) an assessment of vocabulary age (BPVS II), (iv) a naming task and (v) a comprehension task. During the successive five weeks, four experimental training sessions were run, each followed by an assessment of color term comprehension. There was a delay of five days between training and assessment, and a delay of at least one day between assessment and the next training session. The naming, comprehension, and training tasks used computer-generated video-clips in which a (colorless)

---

[3] While the brightness of the three target stimuli was not identical, distracters at both training and test were organised so that the target was at an intermediate brightness level within the set.

cartoon character interacted with the child. All instructions, elicitations of responses and appropriate linguistic feedback were delivered by the cartoon character.

Listing task. Children were asked to tell the experimenter all the color words they knew. Those who failed to list any words were prompted with, "what is your favorite color?" and then, "are there any other colors you can think of?".

Naming task. The cartoon character introduced the game and, on the appearance of the first circle asked, "what color is this one?". The experimenter recorded the child's response. The first circle was then replaced by another colored circle in the same location. The sequence was repeated until all of the fourteen colors were displayed once. If a child failed to respond to the computerized voice, the experimenter prompted him/her saying, "do you remember this one?". Colored circles were presented in random order.

Comprehension task. Fourteen colored circles were shown and the child was asked, "can you show me where my X color is?". The experimenter recorded the child's response, and the child was then asked, "are there any more X ones?". The sequence was repeated for the eleven basic colors at pre-test, and for all the fourteen colors on subsequent assessments.

Experimental training trials. Each training session consisted of three trials, one for each target color. Three colored objects, identical in shape, were displayed on each trial, one target and two distracters chosen from red, blue, green and yellow; within- or near-category sets were avoided (e.g. the crimson target never appeared with the red distracter; the blue and green distracters never appeared within the same set). Color sets were kept constant within each training session. The first two sessions were played with stars, the third with flowers, and the fourth with cars. The centre of the screen was blank on all three kinds of displays. Stars were aligned in the centre-top half of the screen, and the cartoon character appeared in the centre of the bottom half. Flowers appeared to the right, with the character to the left.

Cars were displayed to the left, with the character to the right. The objects and color sets used are summarized in Appendix III.

Two-week follow-up. The naming, comprehension and listing tasks were administered again two weeks after the final training session.

<div align="center">EXPERIMENT 1</div>

Experiment 1 investigated the extent to which linguistic feedback of the corrective, semantic and referential linguistic contrast kind facilitates learning of novel color terms. Two control conditions were run. One matched the structure of the corrective and semantic linguistic contrast conditions, and introduced target words by ostensive definition. The other matched the referential linguistic contrast condition, introducing the novel terms in a naturalistic setting, without explicit instruction. The three kinds of linguistic feedback were expected to promote faster and more overall learning compared to the two control conditions. Among the experimental groups, corrective contrast was expected to promote more rapid and more overall learning than semantic contrast, whereas similar patterns of learning were expected for the corrective and referential contrast groups.

<div align="center">*Method*</div>

Participants

Sixty children (25 boys, 35 girls) aged between 3 and 3;11 (mean age = 3;4) participated in the study. Twelve children were randomly assigned to one of five groups. The distribution of children across groups is given in Table 1.

Procedure

Each training trial initially showed three pale grey circles on the screen. The game was explained by the character and then the circles were replaced by the test objects.

In the corrective and semantic linguistic contrast conditions, and the first control condition, the child was asked, "can you show me where my crimson star is?". Once the child

pointed at the screen, the game proceeded with the character saying, "there it is! I am going to get it!" and moved across the screen to the target object, whether the child's response was correct or not. The two distracters were replaced by pale grey circles, and the appropriate linguistic input was given. Children in the [first] control group heard, "see this, this is my crimson star"; children in the corrective linguistic contrast group heard, "see this, this is my crimson star, it is not my [child's own label] star", and in the semantic linguistic contrast group they heard (for example), "see this, this is my crimson star, it is not my yellow star, and it is not my green star".

In the referential linguistic contrast condition the child was asked, "can you show me where my crimson star is, not my red star, and not my green star?". In the [second] control condition the child was asked, "can you show me where my crimson star is?". The three color-objects remained on display throughout the trial. If the child gave a correct response, the character jumped up, moved across the screen to the target object, pointed to it, and said, "thank you! I found it! I found my crimson star!". If the child gave an incorrect response, the character looked left and right (stars game), or up and down (flowers and cars games), jumped up and down remaining in the same location, and said, "I've looked for my crimson star, shall we look for another one?"

In all conditions, the next trial was 'played' with the same objects but a different target color and a different pair of distracters. At the end of the third trial, the child was thanked for their participation. Targets were presented in random order. The number of times that target words were spoken was equated across trials and groups. Training sessions lasted about 10 minutes.

Table 1 about here

Results

Pre-test listing task. Children listed a mean of three color words at pre-test, and four at the two-week follow up (Time 5). The difference was significant by a paired samples t-test, $t(59) = 3.36$, p < .01.

Color knowledge at pre-test and Time 4. Table 2 summarizes the percentage of children who knew each of the eleven basic colors at pre-test and at the end of the experimental period (Time 4). Significantly more color words were known at the end of the study (9) than at pre-test (8) by a paired samples t-test, $t(23) = 2.25$, p < .05. Of the three basic terms associated with the targets, 92% of children knew red, 93% blue. However, only 55% of children knew brown.

Table 2 about here

Overall learning. Two overall comprehension scores were computed for each child, one for speed of learning and one for degree of learning over the training period. These scores pooled data across the three targets. All scores ranged 0-1. Speed of learning scores quantified how early on during the experimental timeframe children succeeded at assessment in response to at least one target; degree of learning scores quantified the proportion of successful responses across the four assessments (Appendix I). These scores are summarized in Table 3.

Independent samples t-tests showed that there were no differences between the two control groups on either speed of learning, $t(22) = .03$, p > .05, or overall degree of learning, $t(22) = .34$, p > .05, scores; data for these two groups was collapsed for subsequent analyses. Two (speed and degree of learning) one-factor (Condition) between subjects ANOVAs with four levels (control, corrective, semantic, referential) were conducted. A significant main effect of condition was found on speed of learning scores, $F(3, 56) = 4.90$, MSE = .07, p < .01, and on overall degree of learning scores, $F(3, 56) = 6.16$, MSE = .02, p < .01. Post hoc analysis (Tukey HSD) showed that the corrective group learned faster than controls (p < .01)

and than children in the semantic group (p < .05), as well as learning more overall than both

these groups (p < .01).

<div align="center">Table 3 about here</div>

Individual target learning. If assumptions of mutual exclusivity prevent children

learning novel color terms then these assumptions should be strongest for novel terms

subsumed within a known, rather than an unknown basic category. Thus more should be

learnt about beige than either crimson or teal, because only 55% of children initially passed

the knowledge criteria for brown. Performance with the three targets is summarized in Table

4.

These data were analyzed in a 4 (Condition: control, corrective, semantic, referential) x

3 (Target: beige, crimson, teal) x 2 (Time: degree of learning at Time 1, Time 4) mixed

design ANOVA with repeated measures over the second and third factors. Significant main

effects were found of condition, $F(3, 56) = 4.89$, MSE = .12, p < .01, target type, $F(2, 112) =$

24.84, MSE = .11, p < .01, and time of assessment, $F(2, 56) = 53.36$, MSE = .12, p < .01.

There was no interaction between target and condition, $F(6, 112) = 1.80$, MSE = .11, p > .05,

whereas there was a significant interaction between target and time, $F(2, 112) = 8.29$, MSE =

.09, p < .01, and the interaction between time and condition approached significance, $F(3, 56)$

= 2.54, MSE = .11, p = .07. There was no three-way interaction, $F(6, 112) < 1$.

Post-hoc investigation of the main effect of condition showed that the corrective group

(but not the referential group) learned more than the control (p < .05) and the semantic (p <

.01) group. The difference between the corrective and referential groups was not significant.

Post-hoc pairwise comparisons of the interaction between target and time (using the

Bonferroni adjustment) showed that that there was more learning of the beige target

compared to both the crimson and teal targets at Time 4 (p < .01).

<div align="center">Table 4 about here</div>

Two-week follow up (Time 5). Performance at the two-week follow up (Table 3) was compared to performance at first assessment (Time 1) using a 4 (Condition: control, corrective, semantic, referential) x 3 (Target: beige, crimson, teal) x 2 (Time: degree of learning at Time 1, Time 5) mixed design ANOVA with repeated measures over the second and third factors. There was a significant main effect of condition, $F(3, 56) = 2.86$, MSE = .09, p < .05, a significant main effect of target type, $F(3, 56) = 24.20$, MSE = .09, p < .01, a significant main effect of time of assessment, $F(1, 56) = 20.10$, MSE = .06, p < .01, and a significant interaction between target and time, $F(2, 112) = 6.69$, MSE = .06, p < .01. The interaction between target and condition approached significance, $F(6, 112) = 1.99$, MSE = .09, p > .05, whereas there was no significant interaction between time and condition, $F(3, 56) < 1$, and no three-way interaction, $F(6, 56) < 1$. Post hoc pairwise comparisons (using the Bonferroni adjustment) showed that the corrective linguistic contrast group recalled more information at Time 5 than controls (p < .05), whereas there were no significant differences between other groups (p > .05), and that more was recalled at Time 5 about the beige target than either the crimson or teal targets (p < .01), whereas there was no difference in the amount recalled about the latter two (p > .05).

To investigate whether learning color terms promotes selective attention to color, so that children only achieve a comprehensive conceptual representation of the color domain after acquiring a sizeable color vocabulary (Sandhofer & Smith, 1999), the number of color terms known over the study period was investigated. A significant correlation was found between the number of color terms known and the degree of learning of the novel color terms, $r(240) = .29$, p < .01. Sandhofer & Smith (1999) found that children knowing 5 terms or more significantly out-performed children who knew fewer than five, suggesting a critical threshold for new term learning. It was not possible to compare these groups statistically

within the present data set, since 52 / 60 children knew 5 or more terms, however the overall

degree of learning for children knowing 5+ or < 5 terms was .18 and .08, respectively.

Discussion

Experiment 1 investigated the extent to which three different kinds of linguistic

feedback facilitate new color term acquisition in three-year-olds. The effects of corrective,

semantic and referential linguistic contrast were compared under conditions matched for

structure and task demands. There was a significant improvement in all children's

performance over the five-week training period, demonstrating that this kind of learning

environment successfully captured the attention of the children and facilitated learning.

Results demonstrated a superior effect of corrective over semantic linguistic contrast, and the

control conditions, replicating previous findings (Au, 1990; Au & Laframboise, 1990; Au &

Markman, 1987; Gottfried & Tonks, 1996).

Children in the semantic contrast group might have been disadvantaged because they

were required to elaborate more information at training. They heard new words contrasted

with two random known color terms, whereas children in the corrective group heard these

contrasted with one, and controls with none. However, Au and Laframboise (1990) found

that performance of children in a 'One Random Label Contrast Condition' and a 'Two

Random Labels Contrast Condition' did not differ. Poor learning among children given

semantic feedback may instead result from confusion, because the random terms used in the

contrastive information refer to colors that are not present in the display. The performance of

children in the referential contrast condition was intermediary between that observed for

children in the corrective condition and controls (and the semantic group).

There are two reasons why corrective contrast might be superior to referential contrast.

Firstly, within the corrective context, linguistic and non-linguistic information converge, so

that a child's attentional resources can be focused on the intended object. Children in the

referential contrast condition were explicitly instructed to direct attention to the target and to ignore the two familiar colors 'in view' (e.g. "show me the crimson star, not the blue one and not the yellow one"); however, they would have had to assess the distracters in order to exclude them. Thus, attention was divided between three objects, making the task more difficult, and the learning process slower.

Alternatively, corrective contrastive input may help because it explicitly contradicts assumptions of mutual exclusivity, by contrasting the novel word with the child's own label (Au, 1990). Our data support this hypothesis, since children learned more about beige than either crimson or teal, despite the fact that beige is much closer in overall perceptual distance from brown than either crimson from red or teal from blue. Fewer children had already acquired brown (the most likely basic color associate for beige) than red, blue or green (the associates for crimson and teal). If brown is acquired later, a finding reported by a number of other authors (see Pitchford & Mullen, 2002, 2003), then learning the new term beige involved less violation of the mutual exclusivity assumption, and corrective feedback would have been particularly useful under these circumstances. However, this effect may be attenuated by overall number of colors known since children who know more than 5 terms learn more about the novel terms (see Sandhofer & Smith, 1999).

To further investigate the possibility that assumptions of mutual exclusivity might impact negatively on color term acquisition, Experiment 2 replicated the corrective contrast condition and respective control condition of Experiment 1, with a group of native Italian speaking three-year-olds. Italian-speaking preschoolers encounter a more numerous set of linguistic color categories in the blue-green region compared to English-speaking preschoolers. This is not the case with 'red' and 'brown', for which, as for English, there is a single basic term [*rosso* (red) and *marrrone* (brown)]. Having already learned that colors in the blue-green range can be named with more than one term, Italian children might be less

constrained by assumptions of mutual exclusivity, and more willing to accept another term to describe a shade of bluish-green (teal). If children in Experiment 1 found beige easier to learn than crimson and teal because the former was classified within a less well-established (and less rigid) linguistic category, Italian children should find teal easier to learn than English children.

<div align="center">EXPERIMENT 2</div>

Experiment 2 replicated the control and the corrective linguistic contrast conditions of Experiment 1, with a sample of Italian speaking three-year-olds, matched on vocabulary age (and schooling) with the English sample. Italian adults use four linguistic categories in the blue-green region, "blu", "azzuro", "celeste" and "verde" (green), instead of two – "blue" and "green". These terms are present in the idiolect of all adult observers (one condition of a 'basic' term), but there is some individual variation in the ranges of color to which they are applied. Italian children thus already hear adults separate this region into multiple categories and may hear very similar shades referred to by different terms on different occasions, in contrast with the consistency with which they hear the term red used. This should weaken their assumptions of mutual exclusivity for color names in this region and make them more willing to learn a new term under conditions of explicit linguistic instruction. Thus of particular interest was the degree to which Italian preschoolers might learn about the teal target. If they learned more about teal than English children, this would provide converging evidence that assumptions of mutual exclusivity impede young children's learning of novel color words.

<div align="center">*Method*</div>

Participants

Twenty-four native Italian-speaking children (11 boys, 13 girls) aged between 3 and 3;11 (mean age = 3;8) participated in the study. Children were recruited from a nursery

school in Naples. Vocabulary age was assessed using the Italian adaptation and standardization (Stella, Pizzoli & Tressoldi, 2000) of the [English] Peabody Test for receptive vocabulary (Dunn & Dunn, 1981). The distribution of children across groups is given in Table 1.

Apparatus, Materials and Procedure

The same Apparatus, Materials and Procedure were used as in Experiment 1. Beige and crimson translate into the corresponding "beige" and "cremisi" in Italian, whereas teal was translated into "avion" – a secondary term for a particular shade of blue. Frequency counts for the three Italian words used were as low as those for the English words. All sound recordings were translated from English into Italian by a native bilingual speaker. The content was kept exactly the same as that used in Experiment 1.

Results

Data were analyzed in the same order and manner as those described in Experiment 1, however, it was not possible to conduct the two-week follow up assessment with the children who participated in Experiment 2.

Pre-test listing task. As found with English children, Italian children listed a mean of three color words at pre-test.

Color knowledge and target classification at pre-test and Time 4. Knowledge of the eleven basic colors at pre-test and at the end of the experimental period (Time 4), and pre-test knowledge of the two further two terms used by Italians in the blue-green, *azzurro* and *celeste*, is summarized in Table 2. A paired samples t-test showed that the mean number of color words known differed significantly between pre-test (8) and Time 4 (9), $t(23) = 3.24$, p < .01. Similarly to the English children, there was ceiling knowledge of red (100%). Knowledge of *blu* among the Italian sample (88%) did not differ from that of blue among the English group (93%); however, 67% of children also knew *azzurro* (67%) and 17% also

knew *celeste*. Fewer Italian children knew *marrone (33%),* (brown), compared to the English

sample (55%).

Table 5 summarizes the way in which the three targets were classified at pre-test and at

Time 4 by English and Italian children. There were marked differences between the English

and Italian groups in the way that the beige and teal targets were categorized at pre-test. 46%

of Italian children classified the teal target as *verde* (green), 17% classified it *blu, 29%*

classified it as *azzuro* and 8% classified it as *celeste*. Additionally, among Italian children,

beige was more likely to be classified as grey (50%), than brown (25%), whereas English

children predominantly classified it brown (60%).

Table 5 about here

Overall learning. Speed and overall degree of learning scores obtained across groups of

Experiment 2 are summarized in Table 6. Two independent samples t-tests showed that the

corrective group out-performed controls on both speed of learning, $t(22) = 2.45$, p < .05, and

on degree of learning, $t(22) = 2.24$, p < .05.

Table 6 about here

Individual target learning. Performance with the three targets is summarized in Table

7. A 2 (Condition: control, corrective) x 3 (Target: beige, crimson, teal) x 2 (Time: degree of

learning at Time 1, Time 4) mixed design ANOVA with repeated measures over the second

and third factors was used to analyze these data. The effect of condition approached

significance, $F(1, 22) = 3.61$, MSE = .16, p = .071, but there were significant main effects of

target type, $F(2, 44) = 7.70$, MSE = .10, p < .01, and time of assessment, $F(1, 22) = 72.84$,

MSE = .08, p < .01. The interaction between target and condition was not significant, $F(2,$

44), p = 1, the interaction between target and time approached significance, $F(2, 44) = 2.86$,

MSE = .18, p = .068, and there was a significant interaction between time and condition, $F(1,$

$22) = 7.02$, MSE = .08, p < .05. There was no three-way interaction, $F(12, 44) < 1$. Post hoc

analysis of the time by condition interaction (using the Bonferroni adjustment) showed that children given corrective feedback performed better than controls at Time 4 (p < .05). Planned post hoc analysis of the interaction between target and time (Bonferroni) showed that more was learned at Time 4 about beige than crimson (p < .01) but not more than teal (p > .05). The difference between teal and crimson was not significant.

Table 7 about here

There was a significant correlation between the number of color terms known at pre-test and the degree of learning of the novel color terms, $r(96) = .26$, p < .01. No children in this sample knew less than 5 color terms.

Comparison between Experiments 1 & 2

To examine whether Italian children differed from English children in their success at learning the three novel terms, data from experiments 1 and 2 were compared. Independent samples t-tests showed that there were no differences between the two control groups on either speed of learning, $t(22) = .22$, p > .05, or overall degree of learning scores, $t(22) = .99$, p > .05; data for these two groups were collapsed for subsequent analyses. A 3 (Condition: control, corrective contrast Exp 1, corrective contrast Exp 2) x 3 (Target: beige, crimson, teal) x 2 (Time: degree of learning at Time 1, Time 4) mixed ANOVA with repeated measures over the second and third factors was computed. There were significant main effects of condition, $F(2, 45) = 6.79$, MSE = .11, p < .01, target, $F(2, 90) = 23.31$, MSE = .11, p < .01, and time of assessment, $F(1, 45) = 103.98$, MSE = .09, p < .01. There was also a significant interaction between condition and target type, $F(4, 90) = 2.48$, MSE = .11, p = .05, a significant interaction between condition and time, $F(2, 45) = 4.23$, MSE = .09, p < .05, and a significant interaction between target and time, $F(2, 90) = 6.61$, MSE = .15, p < .01, but no three-way interaction, $F(4, 90) < 1$. Pairwise comparisons of the interaction between target and condition (using the Bonferroni adjustment) found that children in the control group and

English children in the corrective group learned more about beige than crimson or teal (both

p < .01). However, Italian children in the corrective group learned more about beige than

crimson (p < .05) but there was no difference in their learning rate for beige and teal (p >

.05). Thus, while showing similar learning rates for beige and crimson, Italian children in the

corrective condition learned relatively more about teal.

Discussion

Experiment 2 showed that, just as for the English children, Italian three-year-olds given

corrective feedback learned faster and more about the three novel colors than controls. Thus,

Experiment 2 replicated the results of Experiment 1, showing that linguistic feedback of the

corrective kind facilitates learning of the color domain, across languages. In addition, the

degree of learning was similar between the two groups. Control groups performed similarly

and, in both cases, learned less than children given corrective linguistic contrast.

Differences between the two language groups emerged in their classification of teal,

prior to the training period, and in the amount of learning of this target. Whilst English

children learned significantly more about beige than the other targets, for Italian children, a

similar amount was learned about teal and beige. This supports the hypothesis that the

availability of a larger number of terms for colors in this region, combined with less certainty

about which category to place the target in (*blu, azzuro, celeste* or *verde*), leads to weaker

assumptions of mutual exclusivity[4].

Explanations based on mutual exclusivity alone, however, cannot account for the fact

that children who know a greater number of color terms are more likely to learn the novel

targets. If knowing a larger number of terms means that more of the distracters can be

discounted, reducing the number of alternative referents for the novel term, then paucity of

---

[4] Italian children learned as much about beige as English children, although they tended to classify it as grey, rather than brown at pre-test (Table 7). Since even fewer children knew grey at pre-test, learning a novel term in this area should not be more difficult.

term knowledge might make the task more difficult or, as in the case of semantic linguistic contrast, divert attention away from the targets to be learned. Attentional accounts of children's difficulty in learning color terms have emphasized that children show a bias towards attending to other aspects of novel stimuli (e.g. shape). These biases, combined with a large number of distracters to consider, might also impact on children's learning of novel color terms.

<div align="center">EXPERIMENT 3</div>

To examine the alternative explanation that attentional biases make it difficult for children to learn novel color terms (Soja, 1994), Experiment 3 assessed whether boosting the perceptual salience of the color dimension within a word learning context might facilitate learning the meaning of color words. Two groups of three-year-olds, experimental and controls, played computerized games similar in structure to those used in Experiment 1. The task required children to match two objects (among a choice of three different pairs) on the basis of a novel color word; experimental children obtained the expected outcome only when the pairing was based on the correct color match; control children obtained the expected outcome regardless of their choice.

Faster and/or more learning by the experimental group would support an attentional account for children's comprehension difficulties in the color domain. Since children in the control group always obtained the expected outcome, they might fail to learn any new word-referent mappings if attention at training were directed to a dimension other than color.

<div align="center">*Method*</div>

Participants

Twenty-four children (10 boys, 14 girls) aged between 3 and 3;7 (mean age = 3;2) participated in the study. The distribution of children across groups is given in Table 1.

Apparatus and Materials

The same Apparatus was used as those described in Experiment 1. Telephones, cars and stars were used as target stimuli. Stars and cars were of the same dimensions as those used in Experiment 1. Telephones measures 3 x 3.5 cm. The distracters used on training trials (Appendix III) were not selected on the basis of prior term knowledge, but instead were selected among the eleven basic colors used in the study matched to the target only on approximate brightness level (Appendix II). Within-category sets were avoided.

Procedure

The first two sessions were played with telephones, the third with cars, and the fourth with stars. Test objects were always displayed in the middle of the screen, either horizontally or vertically. On either side of the test objects was a cartoon character. These were identical in shape but one was colorless (as described in Experiment 1), and the other was colored. The colored cartoon character was either beige, crimson, or teal, and the task involved matching the target-color test object with the colored character.

On the telephones game, once the test objects were revealed, the child was asked, "can you show me which telephone I should use, to phone my friend *beige Tippy*?". Upon the child's response, the character proceeded, "ok ready, I am going to get the [chosen color] telephone!*"* and moved across the screen to touch the chosen object; this begun to ring, and then children in both groups saw the colored character respond and communicate with the other if a correct response was given. If an incorrect response was given, children in the control group saw the same sequence of events, whereas children in the experimental group did not get a response from the colored character, despite seeing the chosen telephone ring. Successful trials ended with "shall we call another friend?". Unsuccessful trials ended with (for example), "my friend teal Tippy did not answer, shall we call another friend?*".* On trials using cars and stars children were asked to "…*find beige Tippy's car"*, otherwise the context remained unchanged – the colorless character touched the object chosen by the child, and if

the answer was correct this moved (cars) or bounced (stars) to the colored character who then

thanked the child "…*for finding my car*".

<div align="center">*Results*</div>

Pre-test listing task. Children listed a mean of three color words at pre-test, and four at

the two-week follow up (Time 5). The difference was significant by paired samples t-test,

$t(23) = 3.60$, p < .01.

Color knowledge at pre-test and Time 4. Knowledge of the eleven basic colors at pre-

test and at Time 4 is summarized in Table 2. Children knew a mean number of seven colors

at pre-test, and a mean of eight at the end of the study. The difference was found to be

significant by a paired samples t-test, $t(23) = 2.25$, p < .05. Knowledge of the three color

terms associated with the targets, red (crimson), blue (teal) and brown (beige), was similar to

that found in Experiment 1. 92% of children knew red and blue, and 58% knew brown.

Overall learning. Table 8 summarizes overall speed and degree of learning scores

across groups. The experimental group out-performed controls on both speed and overall

degree of learning by two independent samples t-tests, speed of learning, $t(22) = 2.37$, p <

.05, degree of learning, $t(22) = 2.16$, p < .05.

Individual target learning. Table 9 summarizes performance with the three targets at

first and last assessment. The data were analyzed in a 2 (Condition: control, corrective) x 3

(Target: beige, crimson, teal) x 2 (Time: degree of learning at Time 1, Time 4) mixed design

ANOVA with repeated measures over the second and third factors. The main effect of

condition was not significant, $F(1, 22) = 3.14$, MSE = .14, p = .09, but there were significant

main effects of target type, $F(2, 44) = 9.34$, MSE = .12, p < .01, and time of assessment, $F(1,$

$22) = 38.50$, MSE = .14, p < .01. There was no interaction between target and condition, $F(2,$

$44)$, p < 1, but a significant interaction between target and time, $F(2, 44) = 5.39$, MSE = .09,

p < .01, and no interaction between time and condition, $F(1, 22)$, p < 1. The three-way

interaction was not significant, $F(2, 44) < 1$. Post hoc analysis (using the Bonferroni adjustment) showed that more was learned about the beige target at fourth assessment than was about the crimson or teal targets (p < .01).

Tables 8 and 9 about here

Two-week follow up (Time 5). Performance at Time 5 (Table 8) was analyzed in a 2 (Condition: experimental, control) x 3 (Target: beige, crimson, teal) x 2 (Time: degree of learning at Time 1, Time 5) mixed ANOVA with repeated measures over the last two factors. There was no significant main effect of condition, $F(1, 22) = 2.57$, MSE = .07, p > .05, but a significant effect of target type, $F(2, 44) = 29.99$, MSE = .06, p < .01, and time of assessment, $F(1, 22) = 27.20$, MSE = .06, p < .01. There was no significant interaction between condition and target type, $F(2, 44) < 1$, and no significant interaction between condition and time, $F(1, 22) < 1$, but a significant interaction between target and time, $F(2, 44) = 12.73$, MSE = .05, p < .01, and no three-way interaction, $F(2, 44) < 1$. Pairwise comparisons (Bonferroni) of the interaction between target and time showed that children in both groups recalled more about beige than crimson or teal at Time 5 (both p < .01).

There was again a significant correlation between the number of color terms known at pre-test and the degree of learning of the novel color terms, $r(96) = .36$, p < .01. Twenty / twenty-four children in this sample knew 5 or more color terms. There was also a significant correlation between number of terms known and the likelihood of making the correct match on the first training trial (before any feedback was given), $r(24) = .511$, p < .05. However, on all subsequent training trials there was no relationship between the number of color terms known and the likelihood of giving color-based responses on training trials (at Time 2: $r(24) = .274$, p > .05). Three of the five children who knew less than five color terms prior to the study matched objects by color by the fourth / final trial.

Comparison of Linguistic and Non Linguistic training (Experiments 1 & 3): Speed and degree of learning performance

To compare speed and degree of learning between children trained with linguistic and non-linguistic feedback, overall speed and degree of learning scores from the corrective group of Experiment 1 and the experimental group of Experiment 3 were compared by independent samples t-tests (the two control groups did not differ on either measure). For both speed and degree of learning there was no significant difference between the two experimental groups. Thus corrective linguistic feedback and boosting the perceptual salience of the color dimension facilitated novel color term learning to the same degree.

<center>*Discussion*</center>

Experiment 3 demonstrated that manipulating the salience of the color dimension did facilitate target learning. Moreover, it appeared to increase performance over controls to a similar extent as did linguistic contrast. Once again, knowledge of a number of basic color terms at pre-test was related both to successful learning outcomes and to the probability of adopting a color matching strategy from the outset. So possession of a number of terms may be an indication that the child has successfully learned to attend to color.

At the same time, whilst term knowledge predicted choosing the correctly colored target on the first training trial, all children made more correct choices by the fourth trial, regardless of the number of color terms known. This paradigm promotes attention to color even for children with poor term knowledge.

Overall, the evidence suggests that attentional constraints (such as the shape bias) are at least as important as linguistic factors (mutual exclusivity assumptions) in children's color term learning. However, as in Experiment 1, children learned more about beige than either of the other two targets, and attentional factors alone cannot account for this difference. We return to this issue in the general discussion.

GENERAL DISCUSSION

The three learning studies reported here replicated and extended previous findings that corrective linguistic contrast is more successful than either semantic or referential linguistic contrast in teaching children new color terms (Au & Laframboise, 1990; Gottfried & Tonks, 1996), under conditions where linguistic ability, prior color term knowledge and experimental contexts were matched. Learning was facilitated over five weeks of training (rather than a single learning event) even when three novel targets were introduced (rather than just one), and when assessment required selection of the targets from fourteen alternatives (rather than two).

The overall results of Experiment 1 (for combined targets) are consistent with either a linguistic or an attentional explanation of children's difficulties in color term learning, because attentional and linguistic aspects converge in the corrective context (where only one color was in view when the linguistic contrast was given and the contrasted alternative was the term the children themselves had used). However, the pattern of results for individual targets supports the hypothesis that assumptions of mutual exclusivity, combined with the over-extension of known terms, prevents children interpreting novel words as color terms, because more is learned about beige, (associated to brown), than crimson or teal, and fewer children know brown at the outset than know red or blue. This hypothesis is further supported by the results of Experiment 2, where Italian children, for whom the blue space is already divided into more than one category, learned as much about the teal (*avion*) target as they did about the beige target. In other respects their novel term learning was at a similar level to English children.

An explanation based on perceptual distinctiveness, or the distance of the targets from the best examples of the basic categories with which they are associated is unlikely for two reasons. Even though beige is perceptually closer to brown than either crimson is to red or

teal is to blue, both groups learned more about beige than crimson. It might have been the case that children attend more to differences in brightness than to differences in hue, since beige differs from brown principally in lightness, whereas both crimson and teal differ more from red and blue, respectively, along the dimension of hue. However, Italian children learned as much about teal as they did about beige, which contradicts such a purely attentional explanation. English children tended to classify the beige target as brown, whereas Italian children classified it as grey. Few children in either group knew brown or grey at pre-test, in line with previous findings (Pitchford & Mullen, 2003; Roberson et al., 2004). Thus, while the structure of the linguistic input facilitates learning for both groups, cultural differences are already apparent in the terms that children are learning for the same set of stimuli.

Over and above the structure of the linguistic input, for both language groups, those children who knew more basic terms also learned more about the novel terms, despite being matched for overall vocabulary age. Possession of a sizeable color vocabulary may be a useful indicator of children's readiness to attend to (and abstract) the color dimension (Sandhofer & Smith, 1999). Assumptions of mutual exclusivity within a domain such as color cannot operate until the child has learnt to attend to (and abstract) that dimension of stimuli. However, without linguistic feedback, children in control conditions learned less about the targets, regardless of the number of terms known. So the structure of the linguistic input contributes over and above the child's readiness to attend to color.

Experiment 3 therefore manipulated attentional variables without linguistic feedback and showed that boosting the perceptual salience of color promotes learning to the same extent as linguistic feedback. This experiment presented children with an overt perceptual match between the target color and the color of the cartoon character indicated on screen. For this group of children, as for those in Experiments 1 and 2, there was a significant correlation

between overall learning and number of color terms known at the outset. An existing color vocabulary indicates that the child has already learnt to abstract color as a relevant dimension of stimuli and override initial attentional biases that mediate against attending to the color dimension (Soja, 1994). This increases the likelihood of a color matching strategy being adopted without instruction. However, increasing the salience of color by linking it to an action enables children to re-direct attention to this dimension without knowledge of numerous terms.

Thus it appears that both attentional and linguistic factors contribute to the process of color term learning. In the present studies, more was learned about the novel color terms when both attentional and linguistic aspects of the learning context converged in the corrective linguistic contrast condition and this was the case for both English and Italian children. Whether this combination of constraints is uniquely instrumental in the learning of color terms, or whether these principles apply more broadly to other dimensional terms remains to be addressed.

References

Andrick, G.R., & Tager-Flusberg, H. (1986). The acquisition of color terms. *Journal of Child Language, 13*, 119-134.

Androulaki, A., Pestana, N.G., Lillo, J., & Davies, I.R.L. (2001). *Greek colour terms: Evidence for a natural category boundary within blue*. Paper presented at the European Conference on Visual Perception, Kusadasi, Turkey.

Au, T.K. (1990). Children's use of information in word learning. *Journal of Child Language, 17*, 393-416.

Au, T.K., & Laframboise, D.E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development, 61*, 1808-1823.

Au, T.K., & Markman, E.M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development, 2*, 217-236.

Backsheider, A.G., & Shatz, M. (1993). *Children's acquisition of lexical domains*. Twenty-ninth regional meeting of the Chicago Linguistic Society, Chicago.

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

Bartlett, E.J. (1977). *Semantic organization and reference: Acquisition of two aspects of the meaning of color terms*. Paper presented at the biennial meeting of the Society for Research on Child Development, New Orleans.

Bornstein, M.H. (1985a). Colour-name versus shape-name learning in young children. *Journal of Child Language, 12*, 387-393.

Bornstein, M.H. (1985b). On the development of color naming in young chidren: Data and theory. *Brain and Language, 26*, 72-93.

Braisby, N., & Dockrell, J. (1999). Why is colour naming difficult? *Journal of Child Language, 26*, 23-47.

Carey, S. (1978). The child as a word learner. In M. Halle, J. Bresnan & G. A. Miller. (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: MIT Press.

Carey, S. (1991). Whorf versus continuity theorists: Bridging data to bear on the debate. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development*. Cambridge, England: Cambridge University Press.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports in Child Language Development, 15*, 17-29.

Carroll, J.B., Davies, P., & Richman, B. (1971). *Word frequency book*. American Heritage Publishing Co., Inc. New York.

Clark, E.V. (1973a). What's in a word? On the child's acquisition of semantics in his first language. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language*. (pp. 65-110). New York: Academic Press.

Cruse, D.A. (1977). A note on the learning of colour names. *Journal of Child Language, 4*, 305-311.

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.

Dockrell, J. (1981). *The child's acquisition of unfamiliar words: An experimental study*. Unpublished doctoral dissertation. University of Stirling, Stirling, Scotland.

Dunn, L.M., & Dunn, L.M. (1981). Peabody picture vocabulary test (ppvt). American Guidance Service (AGS), Circle Pines, Minnesota.

Dunn, L.M., Dunn, L.M., Whetton, C., & Burley, J. (1997). The british picture vocabulary scale (second ed.) (bpvs ii). The NFER-NELSON Publishing Company Ltd, Berkshire, UK.

Fletcher, R. (1980). The city university colour vision test: Windsor, Berks: Keeler Ltd.

Franklin, A., Clifford, A., Williamson, E., & Davies, I.R.L. (2005). Colour term knowledge

　　does not affect categorical perception of colour in toddles. *Journal of Experimental*

　　Child Psychology, 90, 114-141.

Franklin, A., & Davies, I.R.L. (2004). New evidence for infant colour categories. *British*

　　Journal of Developmental Psychology, 22(3), 349-377.

Golinkoff, R.M., Mervis, C.B., & Hirsch-Pasek, K. (1994). Early object labels: The case for a

　　developmental lexical principles framework. *Journal of Child Language, 21*, 125-156.

Gottfried, G.M., & Tonks, J.M. (1996). Specifying the relation between novel and known:

　　Input affects the acquisition of novel color terms. *Child Development, 67*, 850-866.

Heibeck, T.H., & Markman, E.M. (1987). Word learning in children: An examination of fast

　　mapping. *Child Development, 58*, 1021-1034.

Keil, F., & Carroll, J. (1980). The child's acquisition of "tall": Implications for an alternative

　　view of semantic development. *Papers and Reports in Child Language Development,*

　　19, 21-28.

Macario, J.F. (1991). Young children's use of color in classification: Foods and canonically

　　colored objects. *Cognitive Development, 6*, 17-46.

Macnamara, J. (1982). *Names for things: A study of human learning*. Cambridge, MA: MIT

　　Press.

Markman, E.M. (1989). Categorization and naming in children: Problems of induction.

Markman, E.M., & Wachtel, G.A. (1988). Children's use of mutual exclusivity to constrain

　　the meanings of words. *Cognitive Psychology, 20*, 121-157.

Merriman, W.E. (1986). Some reason for the occurrance and eventual correction of children's

　　naming errors. *Child Development, 57*, 942-952.

Mervis, C.B., Bertrand, J., & Pani, J.R. (1995). Transaction of cognitive-linguistic abilities and adult input: A case of the acquisition of color terms and color-based subordinate object categories. *British Journal of Developmental Psychology, 13*, 285-302.

Naigles, L.R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition, 86*, 157-199.

Ozgen, E., & Davies, I.R.L. (1998). Turkish color terms: Tests of berlin & kay's theory of color universals and linguistic relativity. *Linguistics, 36*, 919-956.

Pitchford, N.J., & Mullen, K. (2005). The role of perception, language, and preference inthe developmental acquisition of basic color terms. *Journal of Experimental Child Psychology, 90*, 275-302.

Pitchford, N.J., & Mullen, K.T. (2001). Conceptualization of perceptual attributes: A special case for color? *Journal of Experimental Child Psychology, 80*, 289-314.

Pitchford, N.J., & Mullen, K.T. (2002). Is the acquisition of basic colour terms in young children constrained? *Perception, 31*, 1349-1370.

Pitchford, N.J., & Mullen, K.T. (2003). The development of conceptual colour categories in preschool children: Influence of perceptual categorisation. *Visual Cognition, 10*, 51-77.

Quinn, P.C., & Eimas, P.D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology, 63*, 189-211.

Rice, M. (1980). *Cognition to language categories: Word, meaning and training*. Baltimore: University Press.

Roberson, D., Davidoff, J., Davies, I.R.L., & Shapiro, L.R. (2004). The development of color categories in two languages: A longitudinal study. *Journal of Experimental Psychology: General, 133*, 554-571.

Sandhofer, C.M., & Smith, L.B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*, 668-679.

Shatz, M., Behrend, D., Gelman, S.A., & Ebeling, K.S. (1996). Color term knowledge in two-year-olds: Evidence for early competence. *Child Language, 23*, 177-199.

Smith, L.B., Jones, S.S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Child Development, 55*, 363-380.

Soja, N. (1994). Young children's concepts of color and its relation to the acquisition of color words. *Child Development, 65*, 918-937.

Soja, N., Carey, S., & Spelke, E. (1985). Constraints on word learning., *Paper presented at the 1985 Biennial convention of the Society for Research in Child Development.* Toronto, Canada.

Stella, G., Pizzoli, C., & Tressoldi, P.E. (2000). Peabody test di vocabolario recettivo. Italian adaptation and standardization of the english peabody picture vocabulary test (ppvt) by l.M. Dunn and l.M. Dunn.: Omega Edizioni.

Tomikawa, S.A., & Dodd, D.H. (1980). Early word meanings: Perceptually or functionally based? *Child Development, 51*, 1103-1109.

Zelazo, P.D., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development, 11*, 37-63.

Zelazo, P.D., & Reznick, J.S. (1991). Age-related asynchorony of knowledge and action. *Child Development, 62*, 719-735.

APPENDIX I *Method for calculating Speed and Degree of Learning scores.*

Speed of Learning scores. These were computed using a linear equation accounting for the fact that (i) four assessments session took place (one following each training session), and (ii) on each occasion the within-session score could vary from 0 to 3. Each child had a single speed of learning score; the less training required to learn, the higher the score obtained. The equation derived was thus: *Speed of Learning = (Assessment Trial + (Correct targets/Targets))/k*, where *Assessment Trial* corresponds to the 'time' at which one or more correct responses were given; values for this item varied from 3-0 (one score for each of the four assessments). A score of 3 was given when one or more correct responses were given at the first assessment (i.e. following the first training session); a score of 0 was given when one or more correct responses were given on the fourth, and final assessment (following the fourth/final training session). *Correct targets* quantifies the number of targets labeled / pointed to correctly within-session (0-3), and *Targets* refers to the number of available correct responses (targets) at each assessment (3). The constant *k* was set to a value of 4, so that scores varied 0-1. A child who labeled *one* target correctly at the *first* assessment obtained a speed of learning score of $3+(1/3))/4 = 0.833$. A child who correctly labeled *two* targets at *second* assessment obtained a SL score of $2+(2/3))/4 = 0.667$. A child who correctly labeled *three* targets at *fourth* assessment obtained a SL score of $0+(3/3))/4 = 0.250$. The score for failing to give correct responses on all assessments was given by $(0+(0/3))/4 = 0$.

Degree of Learning scores. Overall degree of learning scores were computed for responses given to the three targets overall (i.e. not separately for each target type). They were obtained by summing the total number of correct responses given during the study period, and dividing the outcome by the total number of available correct responses (i.e. 12; three for each of the four assessments). Overall degree of learning scores were stringent. For

example, ceiling performance with one of the targets (e.g. correctly labeling the teal target on

assessments 1 through 4) produced an overall degree of learning score of 4/12 = .333.

Degree of learning was also assessed separately in response to each target at the first

assessment (Time 1) and at the fourth/final assessment (Time 4), as well as at the two-week

follow up (Time 5). These graded measures of learning were less stringent as at each

assessment stage scores were independent of performance on previous trials. Children scored

1 for giving a correct response and 0 for giving an incorrect response. By comparing

accuracy in response to the targets at the beginning and at the end of the training period, it

was possible to capture the changes that took place from Time 1 (first assessment) to Time 4

(fourth/final assessment) and then to Time 5 (two-week follow up), and compare them across

groups. As these analyses were computed separately for each of the three targets, and

children within each group scored 1 for giving a correct response and 0 for giving an

incorrect response, effectively, the within-group scores for graded measures of learning

correspond to the proportion of children who responded correctly at each assessment stage

(e.g. if seven [out of twelve] children in the control group responded correctly to the teal

target at Time 4, the degree of learning at Time 4 for the control group would be 7/12 = .58).

APPENDIX II

RGB values, L*u*v* co-ordinates, and Cieluv Δε' values, for the fourteen color stimuli used in Experiments 1, 2 and 3.

	RGB values			u' v' L*			Cieluv Δε'
Color	R	G	B	u'	v'	L*	
White	252	252	252	0.203	0.481	100	0
Black	0	0	0	0.203	0.444	28.555	71.445
Red	254	0	0	0.361	0.502	58.665	131.693
Blue	18	19	127	0.171	0.357	42.257	66.596
Green	65	197	36	0.169	0.526	80.924	45.837
Yellow	255	255	1	0.220	0.534	95.858	26.769
Pink	233	97	145	0.246	0.461	75.658	50.026
Purple	83	0	93	0.229	0.394	42.763	61.472
Orange	240	96	0	0.279	0.518	71.555	83.344
Brown	109	74	32	0.232	0.493	56.088	48.820
Grey	128	128	128	0.120	0.466	73.546	26.618
Beige	167	127	78	0.224	0.493	74.403	33.175
Crimson	152	1	10	0.325	0.483	47.837	92.111
Teal	0	110	94	0.160	0.460	60.486	52.670

APPENDIX III

(i) Objects and color sets used on training-trials in Experiments 1 & 2

	Color Sets		
Objects	Target: Beige	Target: Crimson	Target: Teal
Stars	*Beige, Red, Green*	*Crimson, Green, Blue*	*Teal, Red, Yellow*
Flowers or			
Cars	*Beige, Blue, Yellow*	*Crimson, Green, Yellow*	*Teal, Red, Yellow*

(ii) Objects and color sets used on training-trials in Experiment 3

	Color Sets		
Objects	Target: Beige	Target: Crimson	Target: Teal
Telephones,			
Stars, Cars	*Beige, Pink, Orange*	*Crimson, Purple, Blue*	*Teal*, Red, *Brown*

Table 1

Distribution of children according to chronological age, vocabulary age and number of colour terms known at pre-test, across the five groups of Experiment 1 (Controls, Corrective / Semantic / Referential linguistic contrast) and the two groups of Experiment 2 (Control, Corrective) and Experiment 3 (Control, Experimental). Ages are expressed in years;months.

	Chronological Age	Vocabulary Age	N Color Terms
EXPERIMENT 1	Mean *(SE)*	Mean *(SE)*	Mean *(SE)*
Control 1	3;4 (*.07*)	3;4 (*.06*)	8 (*.71*)
Control 2	3;3 (*.08*)	3;4 (.05)	7 (*.78*)
Corrective	3;5 (*.08*)	3;5 (*.07*)	8 (.67)
Semantic	3;6 (*.09*)	3;6 (*.08*)	8 (*.66*)
Referential	3;5 (*.06*)	3;6 (*.06*)	8 (*.74*)
EXPERIMENT 2			
Control	3;9 (.07)	3;8 (.50)	8 (.46)
Corrective	3;8 (.08)	3;9 (.95)	8 (.46)
EXPERIMENT 3			
Control	3;2 (.05)	3;4 (.06)	8 (.87)
Experimental	3;2 (.19)	3;3 (.21)	7 (.78)

Table 2

Percentage of children who passed the criterion for knowing the eleven basic colors (naming and comprehension) at pre-test and at fourth assessment (Time 4), in Experiments 1, 2 and 3.

Color	Experiment 1 (N=60)		Experiment 2 (N=24)		Experiment 3 (N=24)	
	Pre-test	Time 4	Pre-test	Time 4	Pre-test	Time 4
White	63	65	71	88	54	67
Black	73	77	63	63	63	67
Red	92	92	100	100	92	92
Blue	93	95	88	96	92	92
Green	93	90	96	96	83	83
Yellow	83	83	96	92	79	83
Pink	67	62	79	83	58	67
Purple	67	77	54	75	71	75
Orange	78	82	83	92	67	79
Brown	55	68	33	38	58	54
Grey	18	18	13	17	24	24
Azzurro	-	-	67	-	-	-
Celeste	-	-	17	-	-	-

Table 3

Mean (*standard error*) comprehension scores for speed of learning, overall degree of

learning, and degree of learning at first assessment (Time 1) and the two-week follow up

(Time 5), across groups of Experiment 1. Data are pooled across targets. Scores range 0-1.

	Control (combined)	Corrective	Semantic	Referential
Speed of learning	.24 (*.04*)	.57 (*.09*)	.29 (*.09*)	.36 (*.08*)
Degree of learning (overall)	.13 (*.02*)	.29 (*.04*)	.10 (*.03*)	.21 (*.05*)
Degree of learning (Time 1)	0	.14 (*.05*)	.05 (*.04*)	.03 (*.03*)
Degree of learning (Time 5)	.14 (*.04*)	.25 (*.06*)	.11 (*.06*)	.19 (*.05*)

Table 4

Degree of Learning (*standard error*) of the beige, crimson and teal targets at the first and fourth assessment (Time 1, Time 4) and at the two-week follow up (Time 5), across groups of Experiment 1. Scores range 0-1.

		Control (combined)	Corrective	Semantic	Referential
BEIGE	Degree of learning, Time 1	0	.33 (*.14*)	0	.08 (*.08*)
	Degree of learning, Time 4	.54 (*.10*)	.92 (*.08*)	.33 (*.14*)	.58 (*.15*)
	Degree of learning, Time 5	.29 (*.09*)	.58 (*.15*)	.25 (*.13*)	.50 (*.15*)
CRIMSON	Degree of learning, Time 1	0	.08 (*.08*)	0	0
	Degree of learning, Time 4	.25 (*.09*)	.33 (*.14*)	0	.33 (*.14*)
	Degree of learning, Time 5	.08 (*.06*)	.17 (*.11*)	0	.08 (*.08*)
TEAL	Degree of learning, Time 1	0	0	0	0
	Degree of learning, Time 4	.13 (*.07*)	.17 (*.11*)	.08 (*.08*)	.17 (*.11*)
	Degree of learning, Time 5	.04 (*.04*)	0	.08 (*.08*)	0

Table 5

Proportion of English (N=60) and Italian (N=24) children who pointed to the crimson referent when asked for "red", to the teal referent when asked for either "blue" (*blu*, *azzurro*, *celeste* for Italians) or "green", and to the beige referent when asked for "brown" or "grey", at pre-test.

	Crimson	Teal		Beige	
	Red	Blue	Green	Brown	Grey
English	.77	.58	.25	.57	.22
Italian	.88	.13	.13	.13	.54

Table 6

Mean (*standard error*) comprehension scores for speed of learning and overall degree of learning, across groups of Experiment 2. Scores range 0-1.

	Control	Corrective
Speed of learning	.26 (*.08*)	.53 (*.08*)
Degree of learning (overall)	.18 (*.06*)	.35 (*.04*)

Table 7

Degree of Learning (*standard error*) of the beige, crimson and teal targets at the first and fourth assessment (Time 1, Time 4) and at the two-week follow up (Time 5), across groups of Experiment 2. Scores range 0-1.

		Control	Corrective
BEIGE	Degree of learning, Time 1	.08 (*.08*)	.08 (*.08*)
	Degree of learning, Time 4	.58 (*.15*)	.83 (*.11*)
CRIMSON	Degree of learning, Time 1	.08 (*.08*)	0
	Degree of learning, Time 4	.08 (*.08*)	.42 (*.15*)
TEAL	Degree of learning, Time 1	0	.08 (*.08*)
	Degree of learning, Time 4	.33 (*.14*)	.50 (*.15*)

Table 8

Mean (*standard error*) comprehension scores for speed of learning, overall degree of

learning, and degree of learning at first assessment (T1) and the two-week follow up (T5),

across groups of Experiment 3. Scores range 0-1.

	Control	Experimental
Speed of Learning	.27 (*.07*)	.49 (*.06*)
Degree of Learning (overall)	.16 (*.04*)	.29 (*.04*)
Degree of Learning (Time 1)	0	.06 (*.04*)
Degree of Learning (Time 5)	.19 (*.06*)	.28 (*.06*)

Table 9

Degree of learning (*standard error*) of the beige, crimson and teal targets at the first and fourth assessment (Time 1, Time 4) and at the two-week follow up (T5), across groups of Experiment 3. Scores range 0-1.

		Control	Experimental
BEIGE	Degree of learning, Time 1	0	.17 (*.11*)
	Degree of learning, Time 4	.67 (*.14*)	.75 (*.13*)
	Degree of learning, Time 5	.67 (*.14*)	.58 (*.15*)
CRIMSON	Degree of learning, Time 1	0	0
	Degree of learning, Time 4	.17 (*.11*)	.33 (*.14*)
	Degree of learning, Time 5	.17 (*.11*)	0
TEAL	Degree of learning, Time 1	0	.08 (*.08*)
	Degree of learning, Time 4	.17 (*.11*)	.42 (*.15*)
	Degree of learning, Time 5	0	0