Effects of language on color discriminability

Nathan Witthoft (witthoft@mit.edu)
Jonathan Winawer (winawer@mit.edu)
Lisa Wu (skuld@mit.edu)
NE20-419, MIT, 77 Mass Ave
Cambridge, MA 02139 USA

Michael Frank (mcfrank@stanford.edu)
Alex Wade (wade@white.stanford.edu)
Department of Psychology
Stanford, CA 94305

Lera Boroditsky (lera@mit.edu)
NE20-456, MIT, 77 Mass Ave
Cambridge, MA 02139 USA

Abstract

Different languages divide the color spectrum in different ways. Can such linguistic codes affect color discrimination? Results of three experiments suggest that color language can influence people's color judgments even in conditions when all color stimuli are present at the same time and need not be stored in memory. Two experiments showed that color-discrimination performance within a language group is affected by verbal interference (and not spatial interference). The third experiment showed that color discrimination performance across a boundary that exists in one language but not another can be altered by linguistic interference only for the language group that codes that linguistic distinction.

General Introduction

Does the language you speak affect the way you perceive the world? The strong Linguistic Determinism view – the idea that all aspects of thought (even low-level perceptual abilities) are determined by language – is most closely associated with the writings of Benjamin lee Whorf (1956). Whorf's ideas have generated interest and controversy, with much of the empirical work focusing on color perception.

Different languages divide the color spectrum differently; does this lead speakers of different languages to actually perceive colors differently? Early studies claimed no differences in color perception (Heider, 1972), but recent cross-linguistic studies (Davidoff et al 1999, Kay et al 1984) as well as studies of categorical learning (Goldstone, 1994) have claimed that linguistically learned categories can indeed affect people's perceptions of colors and shapes.

In another set of recent studies, Roberson & Davidoff (2000) have shown that language is involved directly online in a number of perceptual decision tasks. For example, Roberson & Davidoff showed subjects a target color chip picked from a blue-green continuum. Then, the chip was removed and replaced by two new chips. The subjects' task was to indicate which of the two new chips was identical in color to the target chip they had seen before. The two new chips were always the same perceptual distance apart, but could either be from the same color category (both blue or both green) or from different color categories (one blue and one green). Subjects were considerably more accurate on trials when the two test chips were from two different categories (a between-category trial) than when they were form the same category (a within-category trial). This was true despite the fact that the colors were normed such that between-category and within-category pairs were equidistant in Munsell color space.

To test whether this categorical affect was due to the English linguistic categories of "blue" and "green," Roberson and Davidoff introduced linguistic interference into the delay between the presentation of the target chip and the two test chips. After seeing the target chip, subjects read words for five seconds, and then were tested on the two test chips. Roberson and Davidoff found that this linguistic interference condition completely erased the advantage of between-category trials (but the same was not true for a non-linguistic visual interference task). These results suggest that language does indeed play a role online in color judgment tasks.

So what are the limits of linguistic effects on our perceptions of color? For example, is the influence of language limited to acting as a dual code in memory across a delay? Can language affect color judgment even in tasks with no memory component? What if all colors are shown at the same time? Could linguistic boundaries influence color decisions even when all colors are simultaneously visible? This paper aims to address these questions.

In Experiment 1, we replaced the memory paradigm used by Roberson & Davidoff with a simultaneous presentation paradigm (the target chip and the two test chips are all presented on the screen together at the same time). Just as before, we found that between-category decisions were easier than within-category decisions (as measured in reaction time), and that this effect could be reduced with verbal interference. In Experiment 2 we found that verbal interference had a selective effect on this between-category advantage (a spatial interference condition did not produce the same effect). In Experiment 3, we compared English and Russian speakers on a categorical boundary in the blue range that exists in Russian but not in English. These results again suggested that even when all colors are visible at the same time, the categories present in one's language can play a role in one's color judgments.

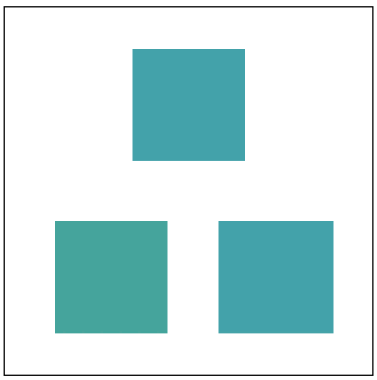

Figure 1: An example stimulus triad used in Experiments 1 & 2. Subjects were asked to report which of the two bottom squares was identical in color to the square on top.

Experiment 1

In Experiment 1, subjects were shown three color patches on the screen at the same time (as shown in Figure 1). Subjects were instructed to indicate which of the bottom two patches was identical to the top patch as quickly as possible. For half of the trials subjects were asked to perform a verbal interference task (rehearsing a string of numbers) while making their color judgments.

Methods

Participants 22 native English speakers participated in the study in exchange for payment. Subjects were recruited and tested at MIT.

Materials & Procedure The stimuli were seven blue, green, and blue-green squares, 3.81 cm, identical to those used by Roberson and Davidoff. They were presented on a computer screen with a white background. All subjects were tested in a quiet, darkened room. Subjects completed 96 trials with verbal interference and 96 trials without. The two conditions were blocked and the order of the blocks was randomized across subjects.

On no- interference trials, subjects were presented with 3 colored squares as shown in Figure 1. On each of these color matching trials, the subjects indicated the match by pushing buttons marked right or left. Stimuli remained on the screen until the subject responded. For the verbal interference block, subjects were presented with a 7-digit number for 3 seconds and were instructed to rehearse it for later test. They were then presented with 8 color matching trials identical to those used in the no-interference condition, and were then asked to type in the 7-digit number. There were a total of 12 numbers to recall for each subject within the verbal interference block.

Results

Results are summarized in Figure 2. Overall, subjects were faster on between-category trials than on within-category trials, but this advantage was reduced with verbal interference. Details of the analyses are described below.

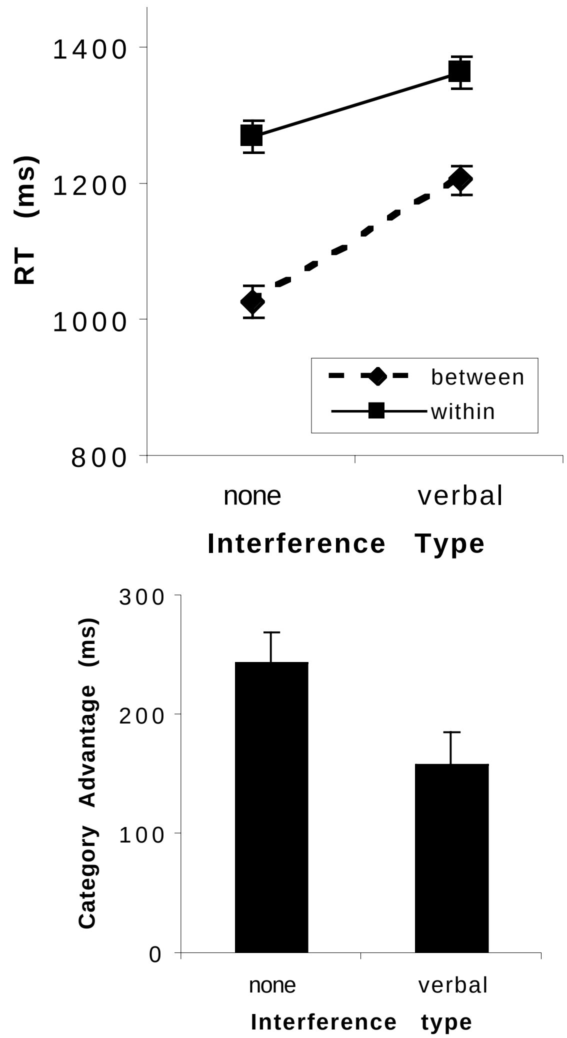

Figure 2: Results of Experiment 1. The top panel shows mean RTs for between-category and within-category trials with and without verbal interference. The bottom panel shows the between-category advantage (RT on within-category trials minus RT on between-category trials). As shown in the bottom panel, the between-category advantage was reduced with verbal interference.

Reaction time data was trimmed by eliminating any trial with either an RT greater than 3500 ms or an incorrect color match, as well as any trial that took place during maintenance of an interference sequence if it was subsequently recalled incorrectly. Additionally, subjects were eliminated entirely if they failed to get 80% of the matching trials correct, or 80% of their responses in under the 3500 ms deadline, or if they answered fewer than 50% of the verbal interference tests correctly. This eliminated 3 subjects, leaving 19 subjects for the data analysis.

Reaction times for the matching trials were analyzed using a 2 (interference type: none vs. verbal) x 2 (category: between vs. within) repeated measures ANOVA. There was a main effect of interference type, (F {1,18} = 8.498, p = 0.009), and category, (F {1,18} = 25.056, p < 0.001). Critically, as in Experiment 1, there was also a significant interaction between interference type and category, (F (1,19) = 4.998, p < 0,05), such that the between-category advantage became smaller with verbal interference (Fig. 2).

Discussion

In Experiment 1 the verbal interference task produced a decrease in the between-category advantage even when all colors are presented simultaneously. These results suggest that language is involved in color judgments even in tasks that have very limited if any memory requirements. However, one worry might be that the reduction in this between-category advantage is not specific to *verbal* interference per se, but rather could be a function any kind of secondary task. To address this question, Experiment 2 introduced a new spatial interference condition in addition to the verbal interference and no interference conditions.

Experiment 2

Methods

Participants 25 native English speakers participated in the study in exchange for payment.

Materials & Design This experiment used the same stimuli, equipment, and testing rooms as Experiment 1. Trials were blocked by interference type, (spatial, verbal, or none). The color matching trials were identical to those in Experiment 1. For the spatial interference block, subjects were presented with a square 4x4 black and white grid. Within the grid pattern, 4 of the 16 squares had been blacked out. Subjects were presented with the grid for 3 seconds and asked to remember it by maintaining a visual image of the pattern. Subjects then performed 8 color matching trials. Subjects were then shown the original grid and a foil and were asked to indicate which one they saw before. Foils were created by moving a single black square from the sample grid pattern one space in any direction. On verbal interference trials, subjects were presented a 7-digit number and asked to remember the number by rehearsing it. Just as with the spatial interference, after 8 matching trials, subjects were shown the original number string and a foil and were asked to indicate which one they had seen before. Verbal foils were created by randomly changing one of the 7 digits. Pilot experiments showed that the verbal and spatial interference tasks were matched in difficulty. Again trials were balanced across interference conditions by category (between-category or within-category). There were 96 matching trials for each of the 3 interference types (none, spatial, and verbal) for a total of 288 trials. For the verbal and spatial interference blocks there were 12 7-digit numbers and 12 grids to be recalled respectively.

Results

Results are summarized in Figure 3. As before, subjects were faster on between-category trials than on within-category trials, but this advantage was reduced with verbal interference. Importantly, spatial interference did not have the same effect despite being matched in difficulty with the verbal interference task. This suggests that linguistic information per se is involved in the color judgment tasks in this study. Details of the analyses are described below.

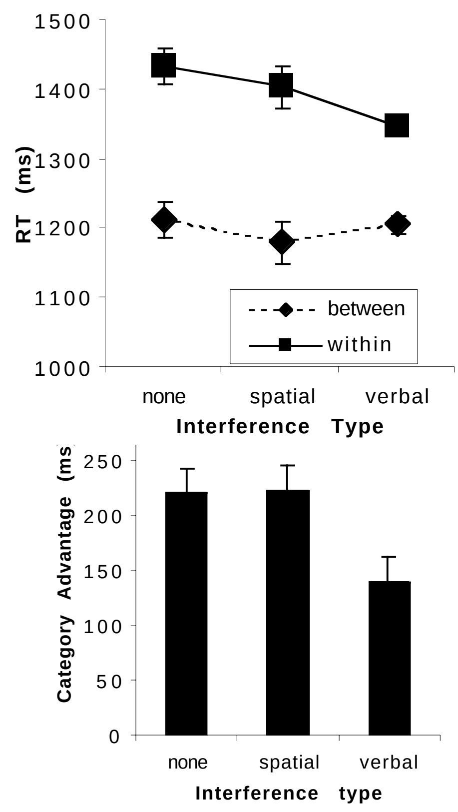

Figure 3: Results of Experiment 2. The top panel shows mean RTs for between-category and within-category trials with no, verbal, and spatial interference. The bottom panel shows the between-category advantage (RT on within-category trials minus RT on between-category trials). As shown in the bottom panel, the between-category advantage was reduced with verbal interference and not with spatial interference.

Reaction time was trimmed in the identical fashion as for Experiment 1, except that the threshold for correct responses on interference trials was set to 75%, not 50%, because of the shift from free response to 2AFC. Ten subjects were eliminated for poor performance, leaving 15 in the analysis.

Results were analyzed using a 3 (interference type: none, spatial, verbal) x 2 (category: between, within) repeated measures ANOVA. There was a main effect of category ($F\{1,14\} = 67.73$, $p < 0.001$). There was also an interaction of interference type by category, ($F\{1,14\} = 6.579$, $p = 0.022$; Fig. 3). Further analysis showed that the size of the category advantage was not different between the spatial interference (M=221 ms) and no-interference conditions (M=224 ms; $F\{1,14\} = 0.003$, $p = .954$). However, there was a significant difference in reaction times between the no-interference trials (M= 221 ms), and the verbal interference trials (M=140 ms; $F(1,14) = 6.579$, $p = 0.022$). There was also a trend in the difference between the categorical advantage on verbal interference (M=224 ms) and spatial interference trials, (M=140 ms, $F\{1,14\} = 3.958$, $p = 0.067$).

Discussion

These results show that the reduction in the between-category advantage during verbal interference observed in Experiments 1 and 2 is specific to *verbal* interference and not simply due to an overload of general processing resources by a secondary task. An equally difficult spatial interference task did not cause a reduction in the between-category advantage. These results parallel those of Roberson and Davidoff on the memory task, in which they compared verbal-, visual-, and no-interference conditions.

Experiment 3

While the results of Experiments 1 and 2 provide evidence for an on-line role of language in color discrimination, they leave open the question of whether this influence applies to linguistic color boundaries other than blue/green, and whether languages with different color coding schemes will show a pattern of results specific to the language.

Therefore, Experiment 3 examined whether the pattern of results found in Experiments 1 and 2 is specific to the coding system of the language. To this end, we sought to replicate the previous results with native Russian speakers, using a color boundary that exists in Russian but not in English. Instead of using the blue/green boundary (as in Experiments 1 and 2), we used the siniy/goluboy color boundary in Russian (roughly dark blue/light blue). Russian does not have a single word for blue. Russian speakers use 'goluboy' to describe light blue shades and 'siniy' to describe dark blues. The task was the same as used in Experiments 1 and 2. Three color squares were shown, and subjects had to say which of the bottom two squares was identical to the top square. Subjects performed the task either with no interference, verbal interference, or spatial interference. We predicted that for Russian subjects, verbal (but not spatial) interference would reduce between-category advantage on the color task. Further, we predicted that interference would have no effect on English speakers.

Methods

Participants 26 native Russian speakers and 25 age-matched native English speakers participated in this experiment (Russians, 28.9±10.2 yr; English 26.3±9.2 yrs; M±SD). All subjects were recruited and tested at MIT. The age at which Russian subjects started learning English ranged from 7 to 21 (11.9±3.8 yr).

Materials and Design A series of twenty blue square patches were made, ranging from light blue to dark blue, separated by equal intervals in CIE coordinate space. The overall design and procedure were the same as in Experiment 2. Test pairs were created by making each possible pair of colors that were 2 steps apart in the 20-step sequence of blues created for this task. As in Experiments 2, category (within or between) and interference (none, spatial, or verbal) were treated as within-subject variables in a repeated measure ANOVA. Additionally, native language (English or Russian) was used as a between subject variable.

Procedure All subjects were instructed in English. Russian and English subjects were tested on the same computers in the same quiet rooms. The experiment was blocked, with 142 color matching trials in each of three blocks (block order randomized across subjects). As in Experiment 2, on interference blocks subjects remembered either a grid pattern (spatial interference), or a number sequence (verbal interference) over 8 color matching trials and then responded to a 2AFC query. At the end of the experiment, all subjects were further presented with each color patch twice, one at a time (in random order), and were asked to classify them as "siniy" or "goluboy" for Russian speakers, and "light blue" or "dark blue" for English speakers. This was done to establish each subject's individual boundary. For the purposes of analysis, 3 between category test pairs were defined for each subject straddling the individual boundary, as were the 3 within category pairs immediately on either side of the boundary.

Results

Results are summarized in Figure 5. Overall, the results suggest a cross-linguistic difference in color-judgments. Russian speakers' color judgments across the goluboy/siniy boundary were affected by verbal interference, and English speakers' judgments were not. Neither group was affected by the spatial interference task. Details of the analyses are described below.

Data were trimmed according to the criteria used in Experiment 2, which resulted in the loss of 3 subjects in each language, leaving 23 Russian and 22 English subjects for analysis.

On average, both the goluboy/siniy boundary and the light blue/dark blue border fell on color-chip 8.6 (on a scale from 1 to 20, lightest blue to darkest blue), based both on reaction time and on classification (see Figure 4).

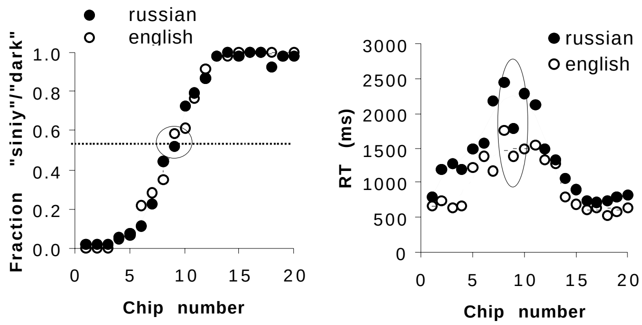

Figure 4: Siniy/Goluboy and light/dark blue borders for Experiment 3.

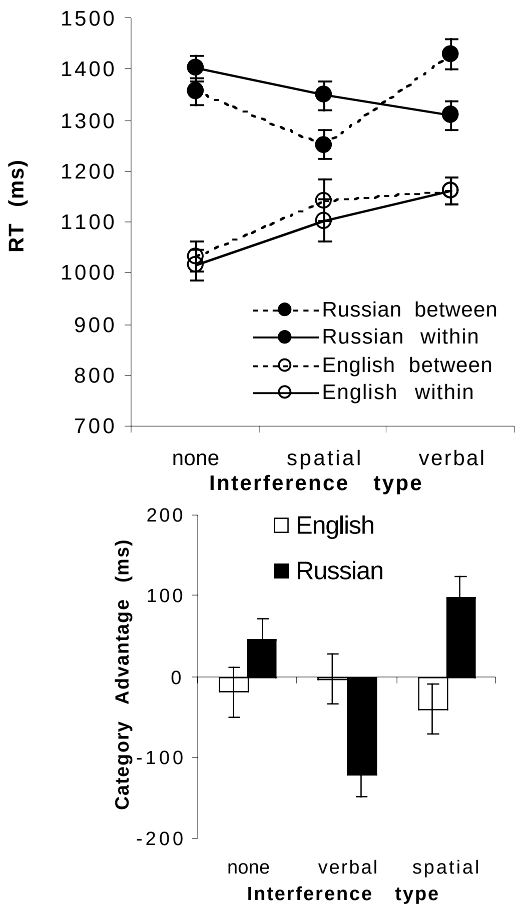

Figure 5: Results of Experiment 3. The top panel shows mean RTs for between-category and within-category trials with no, verbal, and spatial interference for Russian and English speakers. The bottom panel shows the between-category advantage (RT on within-category trials minus RT on between-category trials). Russian speakers' color judgments across the goluboy/siniy boundary were affected by verbal interference, and English speakers' judgments were not.

As predicted, verbal interference for Russian speakers diminished the between-category advantage compared to no interference (no interference, M=46 ms; verbal, M=-121 ms; F {1, 22} = 8.672, p = .007; Fig. 5). The effect of interference was specific to language, as there was no difference in the between-category advantage between the no-interference blocks (M=46 ms) and the spatial interference blocks (M=98 ms, F {1, 22} = .633, p = .435; Fig. 5). Furthermore, verbal interference reduced the between-category advantage relative to spatial interference (-121 ms vs 98 ms; F {1, 22} = 11.760, p = .002; Fig. 5). This is the same pattern of results that we found for English speakers for blue/green in Experiment 3 (compare to Fig. 3).

A different pattern was found for English speakers in the siniy/goluboy task. English speakers showed no between-category advantage in any interference condition (none, -18 ms; spatial -40 ms; verbal, -2 ms; Fig. 5). Furthermore, the between category "advantage" (actually a slight disadvantage for English speakers) did not differ across any interference condition (p > .6, all pairwise comparisons).

The most compelling evidence that the color discrimination task is affected specifically by language, and only in those subjects whose native language codes the border, comes from a direct comparison between English speakers and Russian speakers. In comparing no interference to verbal interference trials, the cross category advantage reduced more for Russians than for English speakers (F {1, 1, 44} = 4.772, p = .035, 3-way interaction of category, interference type, and language). Likewise, in comparing spatial interference to verbal interference, the between-category advantage reduced more for Russian speakers than for English speakers (F {1, 1, 44} = 7.143, p = .011, 3-way interaction).

Discussion

The results of Experiment 3 provide strong evidence that the on-line availability of language can strongly affect color discrimination. The between-category advantage for Russians was reduced and reversed by verbal, but not spatial, interference. Despite the norming done on the spatial and verbal interference tasks for Experiment 2, one might be tempted to argue that the verbal interference task was more difficult according to some critical criterion that we did not measure. This argument cannot apply to the results of Experiment 3, however, as both the English speakers and Russian speakers had exactly the same verbal and spatial interference tasks, yet the differential effect of interference was only seen in Russian speakers.

Interestingly, the boundary between light blue and dark blue according to the English speakers was on average the same as the siniy/goluboy boundary in Russian. Possibly, this may reflect a naturally salient division. However, light and dark blue are not basic color categories in English, and conversely, Russian does not have a single basic category to refer to both siniy and goluboy. We suggest that it is the habitual and mandatory division of the blue color space into two categories in Russian, but not in English, that accounts for the difference in the results.

Summary

Results of three experiments suggest that color language can influence people's color judgments even in conditions when all color stimuli are present at the same time and need not be stored in memory. Two experiments showed that color-discrimination performance within a language group is affected by verbal interference (and not spatial interference). The third experiment showed that color discrimination performance across a boundary that exists in one language but not another can be altered by linguistic interference only for the language group that codes that linguistic distinction.

Let us now return to the question of whether color language can shape color perception. While color language has been shown here to affect performance on a rather basic color discrimination task, is this enough to conclude that different language groups actually perceive colors differently? This conclusion does not seem consistent with the results of these studies. Because the effect of language can be altered by linguistic interference, it seems that language is acting a secondary process in the color discrimination tasks used in these studies. This secondary process can alter the results or speed of a color judgment, but it seems more likely that this interference happens at a decision stage rather late in the processing stream. This conclusion is in agreement with earlier findings of Kay and Kempton (1984), who showed that the availability of a naming strategy caused the perceptual distance to stretch between color s crossing a linguistic boundary, whereas without the naming strategy the perceptual distance returned to the expected values based on the number of just noticeable differences separating the test colors. Further experiments are planned to test whether speeding color judgments, or using even simpler perceptual tasks could completely erase the effects of language by not allowing the linguistic processes enough time to interfere with the perceptual decisions.

Acknowledgments

This research was funded by a grant from the Searle Scholars Foundation to Lera Boroditsky. The authors would like to thank the citizens of Cognation for many valuable discussions of this work.

References

Davidoff, J., Davies, I. & Roberson, D. (1999) Colour categories of a stone-age tribe. *Nature* 398, 203-204.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General*, *123*, 178-200.

Heider, E.R. (1972) Universals in color naming and memory. *Journal of Experimental Psychology*, 93, 10-20.

Kay, P., and W. Kempton (1984) What is the Sapir-Whorf hypothesis? *American Anthropologist* 86:65-79.

Roberson, D. & Davidoff, J. (2000) The categorical perception of colours and facial expressions: The effect of verbal interference. *Memory & Cognition*, 28, 977-986.