

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Asymmetry in Stroop interference arising from naming practice: shape beats color

Permalink

<https://escholarship.org/uc/item/1ms751vm>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 33(33)

ISSN

1069-7977

Authors

Protopapas, Athanassios
Markatou, Artemis

Publication Date

2011

Peer reviewed

Asymmetry in Stroop Interference Arising from Naming Practice: Shape Beats Color

Athanassios Protopapas (aprotopapas@phs.uoa.gr)

Department of Philosophy and History of Science, University of Athens
GR-157 71 Athens, Greece

Artemis Markatou (art_markatou@hotmail.com)

Graduate Program in Basic and Applied Cognitive Science, University of Athens
GR-157 71 Athens, Greece

Abstract

Stroop interference is characterized by strong asymmetry between word and color naming such that the former is faster and interferes with the latter but not vice versa. This asymmetry is attributed to differential experience with naming in the two dimensions, i.e., words and colors. Here we show that training on a visual-verbal paired associate task equivalent to color and shape naming leads to strongly asymmetric interference patterns. 28 adults practiced naming colors and novel shapes, one dimension more extensively (10 days) than the other (2 days), using nonsense syllables. Despite equal training, color naming was strongly affected by shape even after extensive practice, whereas shape naming was more resistant to interference. To reconcile these findings with theoretical accounts of interference, reading may be conceptualized as involving visual-verbal associations akin to “shape naming.” An inherent advantage for naming shapes may provide an evolutionary substrate for the invention and development of reading.

Keywords: Naming; Automaticity; Training; Interference.

Asymmetry in Stroop Interference

Stroop interference (Stroop, 1935) is commonly considered to be among the most familiar, most cited, and most investigated phenomena in all of cognitive psychology. It is well established that it takes longer to name the color in which a word is printed when the word means a different color (e.g. the word “red” printed in green ink), whereas it makes no difference in reading the word what color it is printed in. A complete explanation of this basic asymmetry remains elusive. MacLeod (1991) surveyed the landscape two decades ago and charted a list of challenges for future theorists. A number of comprehensive accounts have approached the topic from different angles, including automaticity (Cohen, Dunbar, & McClelland, 1990), attentional filtering (Phaf, Van der Heuden, & Hudson, 1990) or conflict monitoring (Botvinick, Braver, Barch, Carter, & Cohen, 2001) in connectionist networks; information theoretical considerations (Melara & Algom, 2003); and verification (Roelofs, 2003) or utility learning (Lovett, 2005) in goal-directed production systems.

All of these accounts, in one way or another, are concerned with the strong asymmetry observed between color naming and word reading, and between their relative interference and facilitation effects. Crucially, the *source* of the asymmetry is entirely experience-dependent in these approaches. That is, reading interferes with color naming rather than vice versa due to the extensive history of reading compared to color naming, that is, greater practice in the word naming dimen-

sion (MacLeod, 1991, p. 182). Lovett (2005, p. 496) suggested that “utility” mediates the effects of practice, a distinction of consequence only when competing processes differ in task efficiency. In the “tectonic theory,” dimensional imbalance arises from access efficiency gradients in long-term memory due to developmental history, namely processing experiences, such as perceptual experiences with words and associated oral responses (Melara & Algom, 2003, p. 430).

When a structural asymmetry is posited, it is specific to the nature of reading having to do with direct associations between written and spoken word forms through verbal processes. In the model of Phaf et al. (1990), the asymmetry was introduced ad hoc, to account for the “privileged status” of inherent compatibility between written and spoken words (p. 310). Likewise, Roelofs (2003) posits an inherent privilege for written words, accessing their stored lemmata and spoken word forms directly. In contrast, shape and color naming is conceptually mediated and initially symmetric. Extensive practice may support the formation of direct links between shapes or colors and the corresponding naming responses, thus becoming “similar to reading aloud” (p. 117). Thus, in every current model of interference between two processing dimensions, practice is the crucial factor behind the dimensional imbalance that determines the interference.

Specific Effects of Training

A surprisingly small number of studies have examined the development or malleability of interference through practice. MacLeod (1998) found reduction of interference but no reverse interference (from incongruent color to word naming), despite 5 or 10 days of training on color naming, attesting to the robustness of the asymmetry. MacLeod and Dunbar (1988) trained participants to respond to familiar shapes with color names, in a visual-verbal paired associate learning task using color words. The resulting “shape naming” was vulnerable to interference from incongruent colors in the early stages of training but the asymmetry was eventually reversed: After 20 days of practice, color names for the shapes interfered with regular color naming, consistent with a practice-based account of interference, in which the novel shape-word pairings became sufficiently automatic to cause interference.

Pritchatt (1968) trained participants to respond with nonsense syllables to colors in a paired-associate learning task, and then tested for interference in naming color patches and

the same nonsense syllables printed in color, replicating the interference effect with the newly formed color “names” (Exp. 3, p. 356). Asymmetric interference effects arose in comparison to a condition in which participants learned the reverse associations, that is, responding with color names to nonsense syllables; however the relative magnitude of the effects was not examined. Glaser and Dolt (1977, Exp. 2) trained participants to respond with nonsense syllables to colors and to color words, introducing an additional step of presumed association. This caused substantial and equal interference in both directions, eliminating any asymmetry. Unfortunately it is not known precisely how much each dimension was practiced, because participants were simply instructed to overlearn the dimensions within a week at their leisure.

Rationale of the Present Study

The aforementioned findings indicate that novel paired-associate learning may lend itself to investigation of the development of Stroop interference. In particular we are interested in the origin of the fundamental asymmetry between word and color naming. If this is related specifically to reading, or to the relative amount of practice in naming each dimension, as assumed by current theories, then no asymmetry should arise in shape vs. color naming after equal practice.

In the present study, we trained participants to respond with nonsense syllables to color patches and to unfamiliar visual forms (Chinese characters). One group trained more on the colors and another trained more on the shapes, in precisely counterbalanced training schedules. According to current theories of interference, equal amounts of cross-dimension interference should be expected with incongruent stimuli.

Method

Participants

Twenty eight Greek volunteers (8 male) 19–36 years old participated in the experiment, assigned randomly into two groups of 14 (after removing data from a 29th participant with excessive error rates). Most were graduate students at the University of Athens. None of them had any knowledge of or experience with the Chinese language and ideograms.

Materials

Materials included the simplified Chinese characters for “red,” (红 U+7EA2) “blue,” (蓝 U+84DD) and “green” (绿 U+7EFF) presented in red (RGB: 0xFF0000), blue (0x0000FF), green (0x00FF00), and white (0xFFFFFFFF) color on a black background, as well as plain red, green, and blue rectangular patches of the same colors and dimensions. Characters and patches occupied blocks of 55×55 pixels. For initial practice, we used single recordings of the corresponding words in standard Mandarin (pinyin: *hóng*, *lán*, and *lǜ*, respectively), pronounced clearly by a male native speaker.

Procedure

There were four training steps and two testing steps in the experimental procedure, spread over two weeks. In Step 1,

participants familiarized themselves with one dimension, by looking at three stimuli presented simultaneously and permanently on the screen, clicking at will on each to hear the corresponding syllable. They were instructed to memorize the image-sound associations. Group A was exposed to the three white characters (“shapes”) and Group B to the three patches (“colors”), associated with the corresponding syllables. This step was completed in a few minutes, as soon as each participant was confident to remember the verbal responses.

In Step 2, participants practiced naming the stimuli they were familiarized with. Each stimulus appeared on the screen for 1 s, with a 1 s interstimulus interval. Participants were required to name aloud each stimulus as quickly as possible without making mistakes. They completed 5 blocks per day, of 60 repetitions each (including an equal number of each character/color), for 2 consecutive days. Thus, Group A accumulated 600 trials of practice naming white characters and Group B the same number of trials naming color patches.

In Step 3, on the following day, participants were familiarized with the other dimension, in the same manner as in Step 1. Group A now saw the 3 color patches while Group B saw the characters, associated with the same three syllables, for a few minutes. These first 3 steps all took place at home, unsupervised, based on detailed written instructions and prepared materials installed on the participants’ computers.

In Step 4, the first interference measurement was made by the experimenters, denoted below as Time 1 (T1). All participants were tested in color naming first, followed by shape naming. Prior to the main procedure they were briefly familiarized with a few two-dimensional stimuli, that is, characters drawn in a color that was associated with the same syllable (congruent) or a different syllable (incongruent). They were then administered 144 naming trials, including 36 trials in each character/color combination condition (resulting in 36 congruent, 72 incongruent, and 36 neutral trials), counterbalanced in color, character, and (for the incongruent condition) distracting element. The neutral stimuli were color patches for the color naming test and white characters for the shape naming test. Trials in all conditions were presented mixed, in a pseudorandom order (determined individually), within fully counterbalanced 24-trial blocks, under the control of a DMDX script (Forster & Forster, 2003). Each stimulus appeared on the screen for 3350 ms, with a 167 ms interstimulus interval. Responses were recorded into wav files by DMDX. The entire testing session lasted about 25 minutes.

In Step 5, participants practiced naming the other dimension, using the exact same procedure and materials as in Step 2, interchanged among groups. So Group A now practiced naming color patches while Group B practiced naming white characters. They were required to complete 5 blocks of 60 trials per day, for 10 days, thus accumulating 3000 trials of practice on the previously unpracticed dimension.

Finally, in Step 6 the second interference measurement was taken, denoted Time 2 (T2), using the exact same procedures and materials as in Step 4 (T1).

Analysis

Vocal naming responses were processed with CheckVocal (Protopapas, 2007) to rate accuracy and correct mistriggered response times. Log-transformed response times (logRT; for correct responses only) and binomial accuracy data were analyzed using linear and generalized linear mixed-effects models, respectively, fitted with restricted maximum likelihood estimation using package lme4 (Bates & Sarkar, 2007) in R (R development core team, 2007). LogRTs were fitted linearly (Baayen, Davidson, & Bates, 2008); errors were fitted with binomial distributions (Dixon, 2008) via a logit transformation (Jaeger, 2008). Effect size estimates (β coefficients) are reported anti-logged ($\hat{\beta}' = \exp(\hat{\beta})$), to be interpretable as time ratios and odds ratios, respectively.

Results

Figure 1 plots mean response times and error rates for each group at each measurement time and condition. There were 16,128 error observations and 15,746 logRT observations (due to 381 incorrect responses and 1 excluded outlier RT < 50 ms). An omnibus model was fitted first to each data set. For example, the logRT model was specified as:

```
logRT ~ group * time * task * cond + (1|sID),
```

with 2 levels of group (A and B), measurement time (T1 and T2), and task (color and shape), and 3 levels of cond(ition) (congruent, incongruent, and neutral), as well as a random effect of subjects. Specific effects were tested with restricted versions of this model on appropriately selected data.

Learning

We first examined direct learning effects of training, considering the neutral condition only. In the group \times time \times task models, naming at T2 was faster and more accurate than at T1 (main effect, logRT: $\hat{\beta}' = .856, t = -10.14, p < .00005$; error: $\hat{\beta}' = .399, z = -1.97, p = .048$), while shape naming was overall faster and much more accurate than color naming (main effect, logRT: $\hat{\beta}' = .899, t = -6.99, p < .00005$; error: $\hat{\beta}' = .056, z = -2.72, p = .006$). Due to several significant interactions, more specific tests are reported next.

In the most trained dimension, color naming time improved during the 10 days between T1 and T2 for Group A (logRT: $\hat{\beta}' = .841, t = -11.73, p < .00005$; error: $\hat{\beta}' = .398, z = -1.95, p = .051$), as did shape naming time for Group B (logRT: $\hat{\beta}' = .870, t = -11.47, p < .00005$; error: $\hat{\beta}' = .193, z = -1.34, p = .179$). The marginal interaction of group/task \times time for these data indicates that Group B may have improved somewhat less in shape naming time than Group A in color naming time ($\hat{\beta}' = 1.035, t = 1.79, p = .073$). There was no significant difference in accuracy improvement during this period between the two groups ($\hat{\beta}' = .495, z = -.54, p = .591$). In the previously trained dimension, Group A also improved in shape naming time between T1 and T2 (logRT: $\hat{\beta}' = .966, t = -2.90, p = .004$; error: $\hat{\beta}' = 8.707, z = 1.86, p = .062$), whereas Group B deteriorated in color naming time during this period (logRT:

$\hat{\beta}' = 1.060, t = 3.43, p = .0006$; error: $\hat{\beta}' = 1.379, z = 0.87, p = .383$). The interaction group/task \times time in this dimension was significant for response time only (logRT: $\hat{\beta}' = 1.097, t = 4.49, p < .00005$; error: $\hat{\beta}' = .163, z = -1.50, p = .135$).

To examine whether practice was equally effective in the two dimensions (tasks) we compared color naming performance to shape naming performance in 3 cases: at T1 after minimal exposure and familiarization only (color for Group A, shape for Group B), at T1 after 2 days' practice (shape for Group A, color for Group B), and at T2 after 10 days' practice (color for Group A, shape for Group B). None of the time differences approached statistical significance ($p > .5$). The error differences were significant only after 2 days' practice ($\hat{\beta}' = 14.365, z = 2.33, p = .020$).

In sum, neutral shape naming became faster with either type of extensive practice (shape or color), while color naming became slower after extensive shape naming practice. There was evidence for shape naming being more accurate than color naming after two days' practice. However, the lack of significant differences in response times between the two groups/tasks after comparable practice suggests that any differential effects of learning may not be simply indicative of one naming dimension being easier than the other.

Interference

Interference was examined as a condition effect, ignoring the congruent condition and considering only the neutral and incongruent condition. At T1, there was significant interference in accuracy for Group A in color naming ($\hat{\beta}' = .527, z = -2.24, p = .025$) and Group B in shape naming ($\hat{\beta}' = .250, z = -2.81, p = .005$) but not vice versa ($p > .2$). There was also significant interference in response time only for Group B in shape naming ($\hat{\beta}' = .961, t = -3.37, p = .0008$) and no other group/task combination ($p > .3$). At T2, there was significant interference in response times in every case, that is, for Group A in color ($\hat{\beta}' = .931, t = -5.69, p < .00005$) and shape ($\hat{\beta}' = .949, t = -4.72, p < .00005$), and for Group B in color ($\hat{\beta}' = .907, t = -6.46, p < .00005$) and shape ($\hat{\beta}' = .969, t = -2.96, p = .003$). There was only marginal interference in error rates, for Group B in color ($\hat{\beta}' = .635, z = -1.65, p = .098$) and shape ($\hat{\beta}' = .127, z = -1.89, p = .059$); and no error interference for Group A ($p > .3$). There was no difference at T2 in either color or shape naming time interference between the two groups ($p > .15$), and no difference between color and shape naming time interference for Group A ($p > .3$). There was, however, more time interference in color naming than in shape naming for Group B ($\hat{\beta}' = 1.070, t = 3.46, p = .0005$). There was no difference in error interference between the two groups.

Change in interference as a result of practice was tested as an interaction of condition (neutral vs. incongruent) by time (T2 vs. T1). Response time interference increased for Group A in both color ($\hat{\beta}' = .936, t = -3.38, p = .0007$) and shape naming ($\hat{\beta}' = .955, t = -2.91, p = .0037$), whereas for Group B it increased only in color ($\hat{\beta}' = .893, t = -5.20, p < .00005$) and remained unchanged in shape ($\hat{\beta}' = 1.008,$

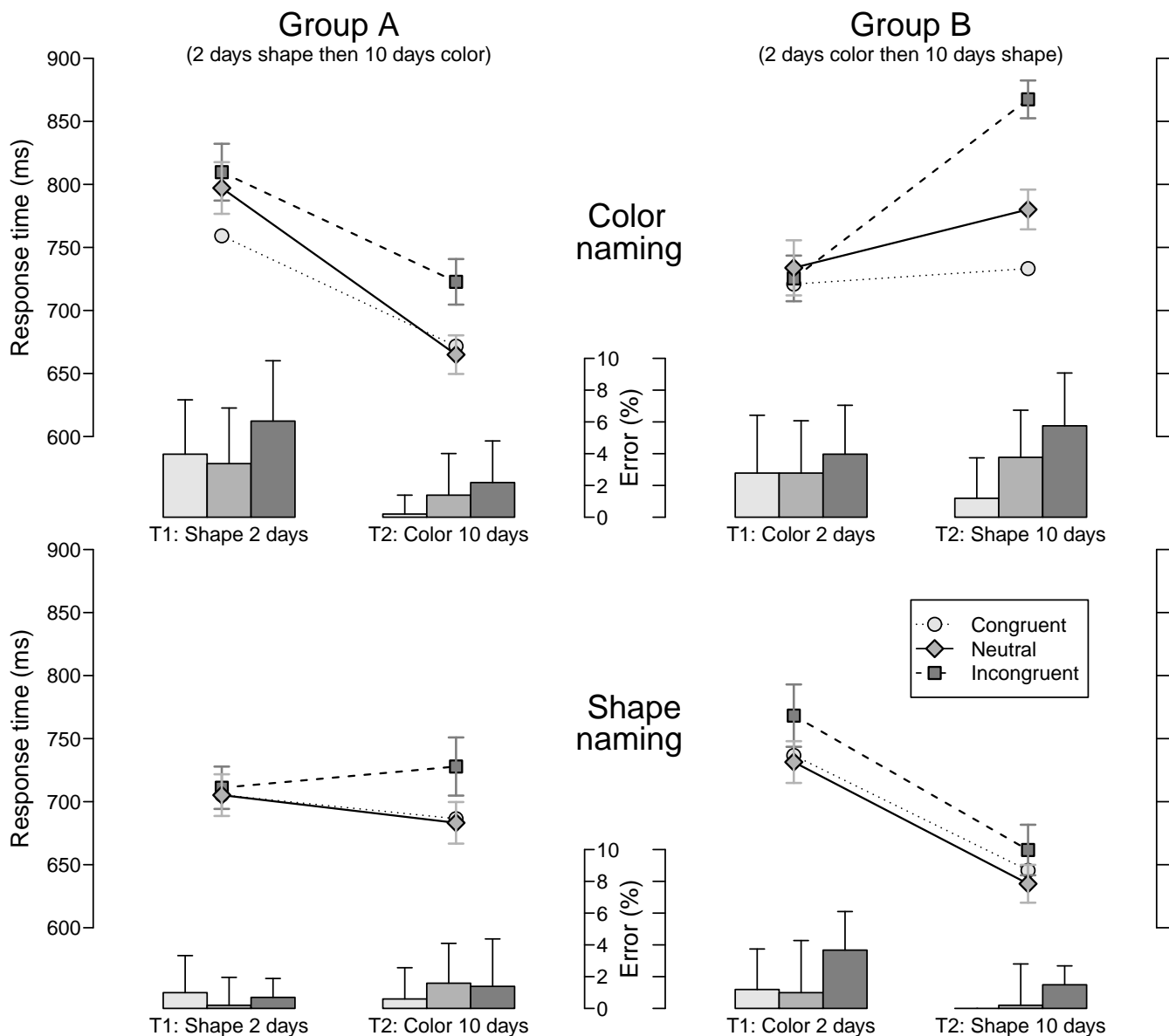


Figure 1: Performance of each group in each naming task at each time point. Response times, shown by markers and connecting lines, refer to the higher vertical axis; error proportions, shown by filled vertical bars, refer to the lower vertical axis. Gray shades are matched between markers and bars for the same condition. Error bars show between-subjects standard error.

$t = .50, p = .615$). There was no change over time in interference as measured by error rates ($p > .2$).

Testing the three-way interaction to compare interference change between the two groups or the two tasks, we found no difference between the groups in color naming interference change (logRT: $\hat{\beta}' = .954, t = -1.60, p = .109$; error: $\hat{\beta}' = .784, z = -.36, p = .718$) and no difference between color and shape for Group A (logRT: $\hat{\beta}' = 1.018, t = .71, p = .479$; error: $\hat{\beta}' = 3.445, z = .97, p = .333$). There was, however, a significant difference between groups in shape naming time interference change (logRT: $\hat{\beta}' = 1.056, t = 2.38, p = .017$; error: $\hat{\beta}' = .127, z = -1.24, p = .213$) as well as between color and shape for Group B (logRT: $\hat{\beta}' = 1.128, t = 4.37, p < .00005$; error: $\hat{\beta}' = .549, z = -.49, p = .624$).

If practice effects on interference were symmetrical, we would expect no three-way interaction between condition (neutral vs. incongruent), time (T2 vs. T1), and group, either for the dimension practiced during this interval (color for Group A and shape for Group B) or for the dimension practiced previously (shape for Group A and color for Group B). Yet this interaction was significant in logRT analyses both for the practiced ($\hat{\beta}' = 1.077, t = 2.92, p = .004$) and the unpracticed dimension ($\hat{\beta}' = .935, t = -2.50, p = .012$); there was no interaction in error rates (both $p > .2$). Evidently, Group B showed greater interference increase in the unpracticed dimension and less in the practiced dimension than Group A.

In sum, extensive practice in either color or shape naming produced increased interference in color naming, whereas in-

interference in shape naming was produced only as a result of practice in color naming and was less than the corresponding interference in color naming after comparable practice.

Facilitation

Facilitation was also examined as a condition effect, comparing the congruent and neutral condition (ignoring the incongruent condition). At T1, there was time facilitation only for Group A in color naming ($\hat{\beta}' = 1.050$, $t = 3.24$, $p = .0012$) and no accuracy facilitation in any condition ($p > .1$). At T2, there was time facilitation only for Group B in color naming ($\hat{\beta}' = 1.062$, $t = 3.82$, $p = .0001$) and at least marginal accuracy facilitation for both Group A ($\hat{\beta}' = 7.085$, $z = 1.83$, $p = .068$) and Group B ($\hat{\beta}' = 3.322$, $z = 2.45$, $p = .015$) in color naming; there was no shape naming facilitation ($p > .15$).

Discussion

The results show strong asymmetries in interference between color and shape “naming,” that is, responding to colors and to Chinese characters with nonsense syllables. The asymmetries emerged following a paired-associate learning procedure with the same responses used for stimuli in both dimensions. The difference is obvious in the top right and bottom left panels of Figure 1, which plot responses to the least practiced dimension: With equal amounts of cross-dimension training, color naming interference for Group B grew much greater than shape naming interference for Group A. Increased color naming interference for Group B appears largely due to slowing down of responses to color, apparently as a result of shape naming, much more so for the incongruent stimuli. In contrast, shape naming was hardly affected. In fact, response times to neutral shapes *decreased* even after extensive color naming practice, suggesting that shape naming is overall a more robust visual-verbal association than color naming.

Our findings cannot be explained simply by shape naming being somehow easier than color naming, because there were no significant differences in response times or errors in the neutral condition after comparable amounts of training in most cases. So, the explanation of the asymmetry cannot rely on an overall faster or more accurate response to shapes compared to colors. Only interference from shapes to color naming was strong and increasing after training on either shape or color naming. In some sense, shape naming can be said to dominate color naming in our experimental paradigm.

Our findings are not necessarily inconsistent with any theoretical mechanism about the interference itself, as all theoretical explanations must somehow account for an asymmetry. However, our findings present a challenge for theoretical accounts to more specifically address the differential efficiency of practice in inducing whatever causes the asymmetry in the corresponding framework. It seems that the asymmetry may be to some extent specific to aspects of naming rather than simply due to different amounts of practice. At the same time, the asymmetry cannot be too specifically tied to reading because it was shown to occur with novel shapes and nonsense syllables. Expert reading also involves mapping from visual

patterns to phonological forms, but neither the visual nor the phonological component are arbitrary and isolated. Rather, visual patterns in reading are composed of familiar (letter) sequences forming coherent units at multiple grain sizes, and map to meaningful lexical forms within a language system. Our findings show that interference asymmetries are caused outside of reading proper, in the absence of differential experience. We thus narrow down the domain in which to search for the causes of the asymmetry, and perhaps for the interference itself, pointing to simple visual-verbal associations.

A possible objection might be raised that our design contained an initial asymmetry and was not fully counterbalanced, in that colors were already associated with words, and the words might have interfered with learning of the novel color-syllable associations. Although logically conceivable, this argument seems unlikely for two reasons: First, the exact opposite, that is, the *lack* of automaticity in color naming is typically cited as explanation of the *lack* of interference from colors to words in the standard Stroop paradigm throughout the literature. And second, if learning of the novel associations was hindered by some existing factor, we should have observed differences in the efficacy of paired-associate training, which was not the case. Moreover, our response set did not consist of words, and could arguably not compete with lexical candidates, as it was crafted to be as plain association and as dissimilar from reading as possible. It should also be noted that color naming was always tested first at both testing sessions, so as to be free from immediate effects of shape naming. Therefore, the observed heightened interference in color naming can only be attributed to the training schedule.

Our findings are not inconsistent with previous demonstrations that training in one dimension affects interference. In particular, they are fully consistent with the rising interference from shapes to color naming found by MacLeod and Dunbar (1988), because they only tested the dominant (shape) dimension taking over, and not the other way around. We predict that in a complementary experiment, namely, learning to respond to colors with shape names, it should prove much more difficult to establish a reversal of interference.

How should our findings be interpreted? One possible route relates to studies arguing against an innate capacity to categorize and name colors (Zhou et al., 2010), consistent with the lack of universal color terms (Regier & Kay, 2009). In contrast, shapes are inherently nameable in that visual forms define object categories, and are conceptually and lexically associated with distinct representational entities (cf. review in Prevor & Diamond, 2005, in the context of object-color interference). It is obvious that if we paint a banana red we get a red banana and not an elongated apple. This trite asymmetry may be related to why line drawings work so well, how reading is at all possible, why orthographic systems with thousands of ideograms emerged and survived, and what is a locus of risk and potential failure for children who cannot become fluent readers despite much effort and support.

As we accept that rats readily develop taste aversion yet

quail are more predisposed to color aversion, that is, there is species-specific associative bias (Anderson, 2000), it may be that humans naturally associate visual forms with verbal labels. Thus reading is made possible by the parasitic development of specialized visual form processors on an evolved substrate of shape-word associations supporting object recognition, categorization, and lexicalization (naming). Since our brains could not have evolved for reading, it remains to be determined what was the pre-existing structure that was so efficiently hijacked by the invention of written language.

This line of thought is compatible with findings that visual-verbal paired associate learning, but not intramodal (visual-visual or verbal-verbal) learning, is strongly correlated with reading performance, especially in irregular words, even after phonological processing skills are statistically controlled (Hulme, Goetz, Gooch, Adams, & Snowling, 2007).

Finally, it should be noted that to say that shapes are inherently namable whereas colors are not is not an explanation. It is merely an observation, prerequisite to explanation. If it proves robust in further experimentation then a theoretical explanation will need to describe the nature of visual-verbal associations, their interaction with other processes to produce naming, facilitation, and interference, and their contribution to the development and expression of reading skills.

Acknowledgments

We thank Ying Liu for the Chinese characters and words, Eleni Vlahou for testing participants and checking responses, and Fotis Fotiadis for comments on the manuscript.

References

- Anderson, J. R. (2000). *Learning and memory: An integrated approach* (2nd ed.). New York: Wiley.
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.
- Bates, D., & Sarkar, D. (2007). *lme4: Linear mixed-effects models using S4 classes* (R package ver. 0.99875-6).
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.
- Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.
- Dixon, P. (2008). Models of accuracy in repeated-measures designs. *Journal of Memory and Language, 59*, 447–456.
- Forster, K. I., & Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers, 35*, 116–124.
- Glaser, W. R., & Dolt, M. O. (1977). A functional model to localize the conflict underlying the Stroop phenomenon. *Psychological Research, 39*, 287–310.
- Hulme, C., Goetz, K., Gooch, D., Adams, J., & Snowling, M. J. (2007). Paired-associate learning, phoneme awareness, and learning to read. *Journal of Experimental Child Psychology, 96*, 150–166.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*, 434–446.
- Lovett, M. C. (2005). A strategy-based interpretation of Stroop. *Cognitive Science, 29*, 493–524.
- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.
- MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. *Memory & Cognition, 26*, 201–211.
- MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 126–135.
- Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review, 110*, 422–471.
- Phaf, R. H., Van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology, 22*, 273–341.
- Prevor, M. B., & Diamond, A. (2005). Color-object interference in young children: A Stroop effect in children 3½–6½ years old. *Cognitive Development, 20*, 256–278.
- Pritchatt, D. (1968). An investigation into some of the underlying associative verbal processes of the Stroop colour effect. *Quarterly Journal of Experimental Psychology, 20*, 351–359.
- Protopapas, A. (2007). CheckVocal: A program to facilitate checking the accuracy and response time of vocal responses from DMDX. *Behavior Research Methods, 39*, 859–862.
- R development core team. (2007). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing (www.R-project.org).
- Regier, T., & Kay, P. (2009). Language, thought and color: Whorf was half right. *Trends in Cognitive Sciences, 13*, 439–446.
- Roelofs, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review, 110*, 88–125.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.
- Zhou, K., Mo, L., Kay, P., Kwok, V. P. Y., Ip, T. N. M., & Tan, L. H. (2010). Newly trained lexical categories produce lateralized categorical perception of color. *Proceedings of the National Academy, 107*, 9974–9978.