

The visual–auditory color–word Stroop asymmetry and its time course

ARDI ROELOFS

Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Four experiments examined crossmodal versions of the Stroop task in order (1) to look for Stroop asymmetries in color naming, spoken-word naming, and written-word naming and to evaluate the time course of these asymmetries, and (2) to compare these findings to current models of the Stroop effect. Participants named color patches while ignoring spoken color words presented with an onset varying from 300 msec before to 300 msec after the onset of the color (Experiment 1), or they named the spoken words and ignored the colors (Experiment 2). A secondary visual detection task assured that the participants looked at the colors in both tasks. Spoken color words yielded Stroop effects in color naming, but colors did not yield an effect in spoken-word naming at any stimulus onset asynchrony. This asymmetry in effects was obtained with equivalent color- and spoken-word-naming latencies. Written color words yielded a Stroop effect in naming spoken words (Experiment 3), and spoken color words yielded an effect in naming written words (Experiment 4). These results were interpreted as most consistent with an architectural account of the color–word Stroop asymmetry, in contrast with discriminability and pathway strength accounts.

Some of the most widely used tasks in academic and applied psychology are the color–word Stroop task (Stroop, 1935) and its analogues, which have become the “gold standard” of attentional measures (MacLeod, 1992). Between 1965 and 2005, some 2,000 articles have appeared in which such tasks were examined (for reviews, see MacLeod, 1991; MacLeod & MacDonald, 2000). In the color–word Stroop task, participants have to name the presentation color of written color words (one basic task variant) or name the written words themselves while ignoring the presentation color (the other basic variant). Typically, participants are much slower and make more errors in naming the presentation color of an incongruent color word (e.g., saying “red” to the written word BLUE presented in red) than in naming the color of a colored series of Xs in the control condition; this effect is called *Stroop interference*. Furthermore, participants are faster in the congruent condition, when color and color word agree, than in a control condition; this effect is called *Stroop facilitation*. In general, the difference in performance on trials containing incongruent and congruent

stimuli is termed the *Stroop effect*. When the task is to name a word and to ignore its presented color, there is no interference from incongruent colors or facilitation from congruent colors relative to controls, and hence no Stroop effect. This is known as the absence of a *reverse Stroop effect*. The difference in effect between written-word distractors in color naming (i.e., interference and facilitation) and color distractors in word naming (i.e., no effect) is usually referred to as the *color–word Stroop asymmetry*.

The asymmetry in Stroop effects between color naming and word naming does not seem to be due to the relative speed of processing of the colors and words per se (but see Melara & Algom, 2003). As was already observed by Cattell (1886a, 1886b) in the early days of experimental psychology, with the font sizes that are normally used, color words are typically named faster than the corresponding colors—namely, some 100–200 msec faster (MacLeod, 1991). However, when one compensates for the slower processing of colors by presenting color patches in advance of the words to be named (e.g., by 300 or 400 msec), a *stimulus onset asynchrony (SOA) manipulation*, still no interference of color with word naming is observed (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989). In this study, presenting the distractor before the target will henceforth be called a *distractor-first SOA* and indicated by a minus sign (e.g., “–400 msec”), and presenting the distractor after the target will be called a *distractor-second SOA*. Presenting words either before or after color patches reduces interference in color naming. That is, interference from incongruent words peaks when color and word are presented simultaneously. Facilitation from congruent words, on the other hand, is constant at distractor-first SOAs (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989).

The preparation of this article was supported by a VICI Grant from the Netherlands Organization for Scientific Research (NWO). The research was supported by the Max Planck Institute for Psycholinguistics, F. C. Donders Centre for Cognitive Neuroimaging, and the Nijmegen Institute for Cognition and Information. I am indebted to Bicoor Bolla-Bong and Sascha Oberrecht for their help in preparing and running the experiments, and to Daniel Algom, Wilhelm Glaser, James McQueen, Antje Meyer, Rebecca Gross, and the members of the Utterance Encoding Group at the Max Planck Institute for Psycholinguistics for helpful comments. Correspondence concerning this article should be addressed to A. Roelofs, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands (e-mail: ardi@mpi.nl).

According to the dominant Stroop account in the literature, henceforth the *relative pathway strength* account, the color–word Stroop asymmetry is due to a difference in strength between color- and word-naming pathways (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Cohen, Dunbar, & McClelland, 1990; Cohen & Huston, 1994). The idea is that the inadvertent activation of a stronger pathway interferes with the use of a weaker pathway, but not the other way around. Pathway strength is assumed to be reflected in the speed of responding. The fact that word naming is typically faster than color naming suggests that the pathway for word naming is stronger than the pathway for color naming. This difference in pathway strength between word naming and color naming would explain why distractor words interfere with color naming, but not vice versa. The difference in pathway strength between color naming and word naming remains with preexposure of the color or word, which explains why a reversal of the difference in processing time of colors and words, as achieved by manipulating the SOA, does not cause a reversal of the asymmetry in Stroop effects (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989).

The relative pathway strength account does less well in explaining other aspects of the time course findings, however (see Roelofs, 2003, for an extensive discussion). In particular, in the various computer implementations by Botvinick et al. (2001), Cohen et al. (1990), and Cohen and Huston (1994), the account predicts that the largest Stroop effect of distractor words on color naming should be obtained with preexposure of the word (e.g., the effect is predicted to be much larger at an SOA of -400 msec than of 0 msec). Increasing the distractor preexposure time increases the amount of activation that will have been built up along the distractor pathway (i.e., the word pathway in color naming) when the target is presented. However, empirically, the greatest impact of words on color naming occurs when word and color appear within 100 msec of each other (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989). Moreover, the account predicts a small Stroop effect of colors on word naming at long distractor-first SOAs, whereas the empirical data show that there is no effect at all (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989).

According to a related account advanced by Melara and Algom (2003), henceforth the *dimensional discriminability* account, the color–word Stroop asymmetry arises because of a mismatch in dimensional discriminability between color and word. *Dimensional discriminability* refers to the ease with which stimulus values can be separated along a dimension (e.g., red, green, and blue along the color dimension). It is reflected in the speed and accuracy of identifying the values as they vary randomly from trial to trial. Dimensional discriminability is assessed by measuring the times of responding to the values of one dimension while keeping the other dimension constant (the control condition in color–word Stroop experiments). Dimensional discriminability is typically manipulated by increasing or decreasing the physical size of stimuli,

such as the font size of written color words. According to this account, when two dimensions are pitted against each other in an experiment, the more discriminable dimension interferes with the less discriminable dimension, but not the other way around. According to the dimensional discriminability account, the more discriminable dimension has been the word rather than the color in most color–word Stroop experiments published to date. Consequently, words (the more discriminable dimension) have interfered with the naming of colors (the less discriminable dimension), but not the other way around, causing the color–word Stroop asymmetry. It is unclear, however, how dimensional discriminability explains why a reversal of the difference in processing time of colors and words, as achieved by manipulating the SOA, does not cause a reversal of the asymmetry in Stroop effects.

According to a third account, recently developed in much detail and implemented in a computer model called WEAVER++ by Roelofs (1992, 1997, 2003; Roelofs & Hagoort, 2002), henceforth the *word production architecture* account, the color–word Stroop asymmetry arises from the structure of the word production architecture. This account follows Smith and Magee's (1980) suggestion that pictures have priority access to meaning before pronunciation, whereas the reverse holds true for words (see Roelofs, 1992). The same principles are assumed to be true for colors and words. In particular, according to the account, color naming requires conceptual preparation and planning of the spoken color name, including selection of the "lemma" of the name and encoding of the word form (see Levelt, Roelofs, & Meyer, 1999). The *lemma* of a word is a representation of the word as a syntactic entity, mediating between word meaning and form. Word naming may be achieved by a "shallow" form-to-form mapping of print onto the word form of the spoken color name, requiring no lemma selection. Moreover, it is assumed that a lemma only significantly activates the corresponding word form when a speaker actually wants to name the color (see Levelt et al., 1999). According to the word production architecture account, the shallow form-to-form mapping of print onto the color name and the controlled mapping of lemmas onto word forms shields word naming from interference by colors. Thus, the asymmetry in effects between colors and words arises from asymmetry in the functional architecture for color naming and word naming. The difference in architecture between color and word naming remains with preexposure of the color or word, which explains why a reversal of the difference in processing time of colors and words, as achieved by manipulating the SOA, does not cause a reversal of the asymmetry in Stroop effects. As discussed in Roelofs (2003), Stroop interference occurs in WEAVER++ when the activations of the target and distractor temporally overlap, which happens when target and distractor are presented close together in time. Consequently, maximal interference in the model occurs around $SOA = 0$ msec. Furthermore, interference increases with decreasing distractor preexposure and decreases with increasing distract-

tor postexposure. That is, the model predicts that the SOA curve of interference has an inverted-U shape around $SOA = 0$ msec, as empirically observed. Facilitation is constant at distractor-first SOAs in the model, because of a floor effect in the speeding up of responses.

It was shown that *WEAVER++* not only successfully simulates the color–word Stroop asymmetry and its time course, but it also simulates several other classic Stroop data sets, mostly taken from the review by MacLeod (1991), which included response set, semantic-gradient, stimulus, spatial, multiple-task, manual, bilingual, training, age, and pathological effects (Roelofs, 2003). Using only three free parameters with two values each to accommodate task differences (i.e., color naming, picture naming, word reading, and manual responding), the model accounted for 96% of the variance in 16 classic studies (250 data points). Moreover, *WEAVER++* successfully simulated the human brain's blood flow response during Stroop task performance in neuroimaging studies, particularly the fMRI BOLD response in the anterior cingulate cortex, one of the classic brain areas involved with Stroop task performance (Roelofs & Hagoort, 2002).

Whereas an extensive literature has documented the Stroop asymmetry for naming pictures and colors, on the one hand, and naming written words, on the other, it is surprising that no test exists of whether the Stroop asymmetry also occurs between color naming and spoken-word naming, the auditory counterpart of word reading. (MacLeod, 1991, remarked that it is also surprising that only a few studies have investigated word reading with color–word Stroop stimuli, as in Stroop's, 1935, Experiment 1.) Although studies have investigated Stroop interference from spoken color word distractors on color naming and aspects of the time course of this interference, the existence of a crossmodal Stroop asymmetry has not been verified experimentally. Cowan and Barron (1987) were the first to investigate the crossmodal color–word Stroop interference effect. Their participants named colors from a card containing 100 stimuli (as in Stroop's, 1935, original study) while listening to an audiotape containing a random sequence of spoken words. Color naming in this study was slowed more when the spoken words were the names of colors on the stimulus card than when they were noncolor words or when there was nothing on the tape. However, Miles, Madden, and Jones (1989) failed to replicate this crossmodal interference effect, and this failure was followed by an extensive, heated debate over the exact timing of the spoken words on the auditory tape that produced the interference (Cowan, 1989a, 1989b; Miles & Jones, 1989).

Later studies improved on the experimental methodology by measuring color-naming latencies for individually presented items rather than for 100 items on a card. The existence of crossmodal color–word Stroop interference is now well established (Elliott & Cowan, 2001; Elliott, Cowan, & Valle-Inclan, 1998; Hanauer & Brooks, 2003; Shimada, 1990), and the same holds for crossmodal picture–word interference effects (Damian & Martin,

1999; Meyer & Schriefers, 1991; Meyer & van der Meulen, 2000; Schriefers, Meyer, & Levelt, 1990; Stuart & Carrasco, 1993). Moreover, evidence has been obtained about the time course of the crossmodal Stroop effect in color naming.

Shimada (1990) tested the effect of spoken color word distractors on color naming in a crossmodal Stroop experiment conducted in Japanese. The colors to be named were presented one at a time, and each color was accompanied by an auditory stimulus with an onset varying from 200 msec before to 200 msec after the onset of the color, in steps of 100 msec. Color-naming latencies were longer when the spoken color words were incongruent than when a buzzer sound was presented instead. This interference effect was obtained at SOAs of 0 and 100 msec (i.e., the color led the auditory stimulus by 100 msec), but not at other SOAs. Moreover, color-naming latencies were shorter when the spoken color words were congruent than when the buzzer sound was presented, at all SOAs except 200 msec. Elliott et al. (1998) compared the effects of spoken color and noncolor words and silence on color naming. The auditory stimuli were presented 500 msec before or simultaneously with the color. Color-naming latencies were longer in the spoken-color-word condition than in the spoken-noncolor-word condition, and both were longer than those in the silence condition at $SOA = 0$ msec. However, at $SOA = -500$ msec, no difference among conditions was obtained.

The studies of Shimada (1990) and Elliott et al. (1998) suggest that the time course of the Stroop effect obtained with spoken color word distractors in color naming is similar to the time course of effects in similar experiments with written distractors, where interference also peaks when color and word are presented simultaneously (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989; Long & Lyman, 1987; Roelofs, 2003). The time course of the visual–auditory color–word Stroop effect challenges the relative pathway strength account (which predicts that most interference should occur at the longest distractor-first SOAs) and supports the word production architecture account (which predicts that most interference should occur around $SOA = 0$ msec).

However, an unusual aspect of Shimada's (1990) study was that the absolute color-naming latencies with the spoken distractors were shortest around $SOA = 0$ msec in all Stroop conditions, including the incongruent condition, whereas color-naming latencies with visual distractors have been longest around $SOA = 0$ msec in all Stroop conditions, including the congruent condition (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989). Thus, although the time course of Stroop effects is similar with spoken and written distractors, the pattern of absolute color-naming latencies is very different. Although the time course of the crossmodal Stroop effect supports the word production architecture account, the pattern of absolute latencies poses a problem, since according to that account absolute latencies should be longest around $SOA = 0$ msec. The absolute latencies across SOAs for

the incongruent condition in Shimada's study are in line with the predictions of the relative pathway strength account, which holds that the longest color-naming latencies should occur at the longest distractor-first SOA, although the shortest color-naming latencies in the congruent condition are predicted to occur at the longest distractor-first SOA, contrary to Shimada's data. However, in contrast to what Shimada observed, absolute color-naming latencies in Elliott et al.'s (1998) study were longer at SOA = 0 msec than at SOA = -500 msec, corresponding to the pattern that is normally obtained with visual distractor words (see, e.g., M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989). Moreover, the picture-naming latencies in extant auditory picture-word interference studies were also longest around SOA = 0 msec (Damian & Martin, 1999; Meyer & Schriefers, 1991; Schriefers et al., 1990). Thus, it may be that the pattern of absolute naming latencies across SOAs obtained by Shimada is not characteristic for spoken distractors. To resolve this issue, it is important to examine whether the pattern obtained by Shimada can be replicated in a new study.

Moreover, pitting color naming against spoken-color-word naming in the visual-auditory color-word Stroop paradigm provides a new test ground for the different hypotheses about the cause of the Stroop asymmetry. Because of the unfolding of the speech signal over time, spoken-word recognition is typically a few hundred milliseconds slower than written-word recognition. This trend is reflected, for example, in the finding that the onset of semantic effects of distractors in the picture-naming task requires longer distractor-first SOAs with spoken-word distractors (i.e., SOAs of -200 msec or -150 msec) than with written-word distractors (i.e., SOAs of 0 msec or 100 msec), as was observed by Damian and Martin (1999), W. R. Glaser and Dünghoff (1984), and Schriefers et al. (1990), among others. Given that the time required for spoken-word recognition is some 100-200 msec longer than that for written-word recognition, color-naming and spoken-word-naming latencies are likely to be similar. This similarity allows for a test of the dimensional discriminability account of the asymmetry without having to manipulate the physical size of the word stimuli, such as by using an artificially small font in the case of written words (recall that according to the dimensional discriminability account, the Stroop asymmetry results from a mismatch in color- and word-naming times). Shimada (1990) mentioned that color-naming latencies (without spoken distractors) and spoken-word-naming latencies were measured before the reported crossmodal Stroop color-naming experiment (Shimada performed no crossmodal Stroop spoken-word-naming experiment) and that the latencies did not differ. Thus, Shimada obtained a Stroop effect of spoken-word distractors on color naming with equivalent response latencies for color naming and spoken-word naming, challenging the dimensional discriminability account of the Stroop asymmetry. However, as mentioned, the pattern of absolute color-naming latencies across SOAs was unusual in this study, so it is important to see whether Shimada's results can be replicated.

The aim of the present study was twofold. First, as indicated, there is a remarkable lack of research testing the visual-auditory color-word Stroop asymmetry and its time course. The present study was intended to fill this gap by examining performance on the crossmodal Stroop task, testing both color naming and spoken-word naming. Could the findings of Shimada (1990) be replicated, including the unusual pattern of absolute color-naming latencies across SOAs, or would the pattern typically obtained with written distractors be observed (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989)? Moreover, would a color-word Stroop asymmetry be obtained, with the total absence of a Stroop effect of colors on spoken-word naming, as is the case with word reading? Would a color-word Stroop asymmetry also be obtained with equivalent latencies for color naming and spoken-word naming? Finally, would asymmetries be obtained between the Stroop effects of written distractors in spoken-word naming and spoken distractors in written-word naming? The second aim of the present study was to test the relative pathway strength, dimensional discriminability, and word production architecture accounts of the color-word Stroop asymmetry. If a visual-auditory color-word asymmetry were obtained with similar latencies for color and spoken-word naming, and with incongruent and congruent colors having no differential effect at any SOA, this evidence would favor the architectural over the discriminability and pathway strength accounts.

Overview of the Experiments

Four experiments are reported below that examined performance on crossmodal versions of the Stroop task. In Experiment 1, participants named color patches while ignoring spoken color words presented at various SOAs. The onset of the spoken word could be 300, 200, or 100 msec earlier than the onset of the presentation of the color patch; the onsets of the spoken word and the color could coincide; or the onset of the spoken word could be 100, 200, or 300 msec later than that of the color patch. In Experiment 2, participants named spoken words while ignoring color patches presented with the same SOAs as in Experiment 1. A secondary visual detection task assured that participants looked at the colors in both tasks (i.e., the secondary task prevented them from simply closing their eyes in Experiment 2 to prevent interference). The secondary task consisted of the detection of a cross randomly presented instead of a color patch on some of the trials. This task was used in all four experiments. Note that at the distractor-second SOAs in Experiment 2 (i.e., presentation of the color patch or cross after the onset of the spoken word), a participant could not start naming the spoken color word immediately but had to wait until either a color (meaning "go") or the cross (meaning "no-go, press the button instead") appeared on the screen. In Experiment 3, participants named spoken color words while ignoring written color words. If Stroop effects were to be obtained from written words, this would show that spoken-word naming is not immune to interference from visual stimuli. In Experiment 4, participants named writ-

ten color words while ignoring spoken color words. In order to make these experiments comparable to previous studies, the three tasks (color naming, spoken-word naming, and written-word naming) were tested between participants (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989) and the trials were blocked by SOA (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989; Shimada, 1990).

EXPERIMENT 1

The intent of Experiment 1 was to test for a crossmodal color-word Stroop effect and its time course, in order to serve as a benchmark for Experiment 2. Participants named color patches while ignoring spoken color words presented at various SOAs. To make the experiment comparable to Experiment 2, which tested for effects of color patches on spoken-word naming, participants performed the same secondary task as in Experiment 2, which required the detection of a cross appearing on some of the trials.

On every trial in all experiments, participants either named a Stroop stimulus or pressed a key in response to a cross. An alternative would have been to ask for both a naming response and a manual response on all trials (see Posner & Boies, 1971). For example, participants could have been asked to indicate by a left/right buttonpress whether or not a cross was presented. However, in order to make the present experiments comparable to earlier Stroop studies (M. O. Glaser & Glaser, 1982; Shimada, 1990), it was decided to ask for only a naming response on the Stroop trials rather than a naming and a manual response in parallel.

Method

Participants. The experiment was conducted with 14 paid participants from the pool of the Max Planck Institute. All participants were young adult native speakers of Dutch.

Materials and Design. The color-word stimuli consisted of red, green, and blue color patches and the corresponding spoken Dutch color words "rood," "groen," and "blauw." The colors were presented as colored rectangles 1.5 cm high and 4.5 cm wide. The auditory word stimuli were spoken by a native Dutch female speaker and recorded using a DAT recorder. They were digitized and stored on the hard disk of the experimental computer as WAV files. The durations of the spoken words "rood," "groen," and "blauw" were 701, 683, and 697 msec, respectively. The cross (+) was made up of 3-point lines fitted within a 1.5×4.5 cm virtual frame.

There were two independent variables, which varied within participants. The first independent variable was the distractor. In order to have an equal number of different color-word pairings in each distractor condition, only three of the six possible incongruent pairings were used (cf. Roelofs, 2003). There were three congruent pairings ("rood"-red, "groen"-green, "blauw"-blue), three incongruent pairings ("rood"-blue, "groen"-red, "blauw"-green), and three control stimuli (the color patches without spoken distractors). Moreover, each spoken color word was also paired with a cross (i.e., "rood"-+, "groen"-+, "blauw"-+). The second independent variable was SOA, with seven levels: -300, -200, -100, 0, 100, 200, and 300 msec. The cross and the spoken words were presented with the same SOAs as the colors and the spoken words. SOA was varied within participants but between trial blocks. The order of test-

ing the SOAs was counterbalanced across participants. Each of the congruent, incongruent, control, and cross pairings occurred three times within an SOA block. Thus, for each SOA, the participants received a congruent stimulus nine times, an incongruent stimulus nine times, a control stimulus nine times, and a cross nine times. Thus, there were 36 trials per SOA. The stimuli were presented in random order.

Procedure and Apparatus. The participants were tested individually. They were seated in front of a computer monitor (NEC Multisync), a Sennheiser microphone connected to an electronic voice key, and a response button. The distance between the participant and the screen was approximately 50 cm. The participants were asked to name the color patches and to ignore the spoken color words, which were presented over closed Sennheiser headphones. When the cross appeared, they simply had to press the response button instead.

After each participant had read the instructions, a block of 12 practice trials with SOA = 0 msec was administered, which was followed by the seven experimental SOA blocks. The structure of a trial was as follows: A trial started with the presentation of the visual stimulus (color or cross) and the spoken stimulus with the appropriate SOA. The visual stimuli were presented for a maximum of 1.5 sec. Before the start of the next trial, a blank screen appeared for 1.0 sec. Thus, the total duration of a trial was 2.5 sec. A Harmac computer controlled the stimulus presentation and the data collection, including the voice key.

Analysis. After each trial, the experimenter coded the response for errors. Five types of incorrect responses were distinguished: a wrong response word, wrong pronunciation of the word, a disfluency, triggering of the voice key by a nonspeech sound, and failure to respond within 1.5 sec after target presentation. Moreover, buttonpress errors (e.g., misses of the cross) were coded. Incorrect responses were excluded from the statistical analyses of the response latencies. The production latencies and errors were submitted to analyses of variance, with the crossed variables distractor and SOA. All variables were tested within participants. A main effect of distractor was further explored by paired comparisons between the incongruent, congruent, and control conditions by means of paired *t* tests. Interactions between distractor and SOA were also statistically explored through paired *t* tests. In particular, pairwise comparisons tested for Stroop interference (i.e., differences between the incongruent and control conditions), Stroop facilitation (i.e., differences between the congruent and control conditions), and Stroop effects (i.e., differences between the incongruent and congruent conditions) at each SOA.

Results and Discussion

Figure 1 gives the mean color-naming latencies and the mean error percentages for each distractor condition by SOA. There were virtually no technical errors, and most errors involved disfluencies. Therefore, the types of errors are not separated in Figure 1. Table 1 gives the mean latencies and error percentages for the manual responses to the cross.

The statistical analysis of the color-naming latencies yielded a main effect of distractor [$F(2,26) = 12.04$, $MS_e = 2,436$, $p < .001$] but not of SOA [$F(6,78) = 1.14$, $MS_e = 2,814$, $p = .35$]. Distractor also interacted with SOA [$F(12,156) = 3.72$, $MS_e = 1,188$, $p < .001$]. Pairwise comparisons revealed that the incongruent condition was slower than the control condition at all SOAs except 200 msec, and interference in this condition was larger at SOA = 0 msec than at the distractor-first and distractor-second SOAs. The congruent condition was faster than

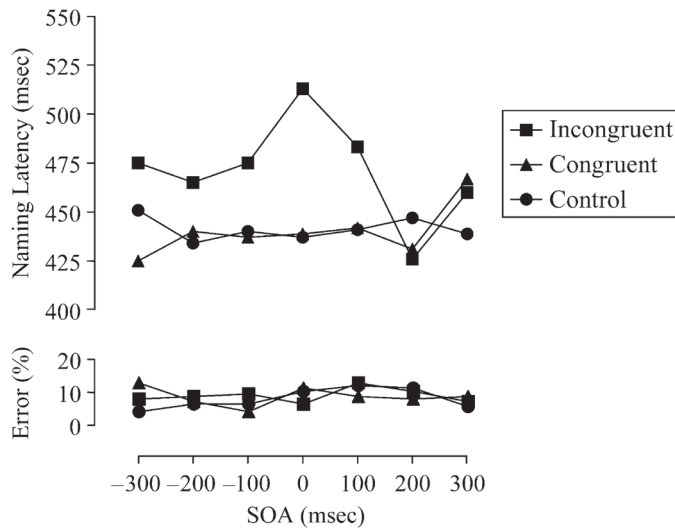


Figure 1. Mean latencies and error percentages for color naming while hearing spoken color word distractors in the incongruent, congruent, and control conditions in Experiment 1. The within-participants 95% confidence interval for the latencies is ± 18.1 msec.

the control condition at SOA = -300 msec. A difference between the incongruent and congruent conditions was obtained at all SOAs except 200 and 300 msec.

The statistical analysis of production errors yielded no main effect of distractor [$F(2,26) < 1, p = .73$] or of SOA [$F(6,78) < 1, p = .63$]. Moreover, there was no interaction of distractor and SOA [$F(12,156) = 1.15, p = .33$]. The error rates were high compared with those usually obtained (around 5%) in similar studies. Moreover, the error rates are quite consistent across conditions, whereas one typically finds more errors in the incongruent than in the other conditions. It seems that the high error rates are due to the combination of a naming response and a go/no-go decision on the critical trials, which the participants found demanding.

The statistical analysis of the manual-response latencies yielded no effect of SOA [$F(6,78) = 1.93, MS_e = 1,341, p = .09$], and neither did the statistical analysis of the errors [$F(6,78) = 1.0, p = .43$]. Moreover, the overall error rate did not differ from zero [$F(1,13) = 1.0, p = .34$].

To summarize, Stroop interference and facilitation effects were obtained with the crossmodal color-word Stroop task, thus replicating Elliott et al. (1998) and Shimada (1990). As in those earlier studies, most interference was obtained around SOA = 0 msec, which also replicates the classic findings that M. O. Glaser and W. R. Glaser (1982) obtained with written words. However, in contrast with Shimada's study, the absolute color-naming latencies in the incongruent condition were longest around SOA = 0 msec, replicating the results with spoken distractor words (Elliott et al., 1998), those with auditory picture-word interference (Damian & Martin, 1999; Meyer & Schriefers, 1991; Schriefers et al., 1990), and color-naming results with written distractor words (M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989).

This suggests that Shimada's results are not characteristic of the effect of spoken distractor words in color naming.

The finding of Stroop facilitation only at an SOA of -300 msec was a surprise. Shimada (1990) obtained Stroop facilitation of congruent distractors relative to a buzzer sound from SOA = -200 msec to SOA = +100 msec. This corresponds to the results obtained with written distractors (see, e.g., M. O. Glaser & Glaser, 1982). It may be that presenting a buzzer sound (Shimada, 1990) is more interfering than presenting no auditory distractor (as in the present experiment). Consequently, congruent distractors may help relative to a buzzer sound, even when they do not help relative to silence, and the difference in the facilitation effect between the studies may be due to a difference in control stimuli.

EXPERIMENT 2

The intent of Experiment 2 was to examine whether the classic color-word Stroop asymmetry obtained with written words and colors also exists for the crossmodal Stroop task. Participants in this experiment named spoken color words while ignoring color patches. The secondary visual detection task (used in Experiment 1) was included to guarantee that participants looked at the colors.

Method

Participants. The experiment was conducted with 14 paid participants from the pool of the Max Planck Institute. All participants were young adult native speakers of Dutch. None of the participants had taken part in Experiment 1.

Materials, Design, Procedure, Apparatus, and Analyses. These were all similar to their counterparts in Experiment 1. The color-word stimuli consisted of the red, green, and blue color patches and the corresponding spoken Dutch color words "rood," "groen," and "blauw" from Experiment 1. In addition, an "empty"

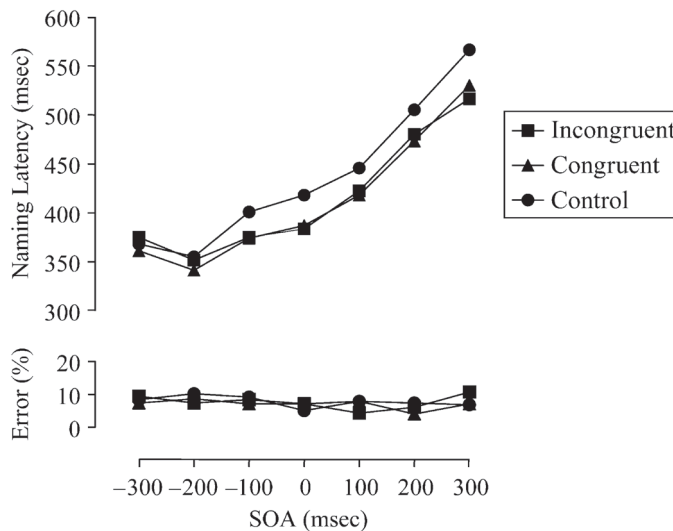


Figure 2. Mean latencies and error percentages for spoken-color-word naming while seeing color patch distractors in the incongruent, congruent, and control conditions in Experiment 2. The within-participants 95% confidence interval for the latencies is ± 16.5 msec.

rectangle (made up of white lines on a black background) 1.5 cm high and 4.5 cm wide served as stimulus in the control condition. There were again three congruent pairings (“rood”–red, “groen”–green, “blauw”–blue), three incongruent pairings (“rood”–blue, “groen”–red, “blauw”–green), and three control pairings (“rood”–rectangle, “groen”–rectangle, “blauw”–rectangle). Moreover, each spoken color word was also paired with a cross (i.e., “rood”–+, “groen”–+, “blauw”–+). The cross and the spoken words were presented with the same SOAs as the colors and the spoken words. The participants were instructed to name the spoken words as quickly as possible while trying to make no mistakes. They were also told to start the naming of the spoken word as soon as possible and not to wait until the end of the auditory signal. When the cross appeared, they simply had to press the button instead.

Results and Discussion

Figure 2 gives the mean spoken-word-naming latencies and the mean error percentages for each distractor condition by SOA. Table 1 gives the means of the latencies and the error percentages for the manual responses to the cross.

The statistical analysis of the naming latencies yielded main effects of distractor [$F(2,26) = 14.54, MS_e = 1,054,$

$p < .001$] and SOA [$F(6,78) = 34.11, MS_e = 5,022, p < .001$]. There was no interaction of distractor and SOA [$F(12,156) < 1, MS_e = 995, p = .51$]. Pairwise comparisons between distractor conditions showed that both the incongruent [$F(1,13) = 13.96, MS_e = 1,488, p < .002$] and the congruent [$F(1,13) = 21.12, MS_e = 1,184, p < .001$] conditions were faster than the control condition and, importantly, that they did not differ from each other [$F(1,13) < 1, MS_e = 491, p = .54$]. The statistical analysis of naming errors yielded no main effect of distractor [$F(2,26) < 1, p = .67$] or of SOA [$F(6,78) = 1.20, p = .32$], and no interaction of the two factors [$F(12,156) = 1.40, p = .17$].

Neither the statistical analysis of the manual-response latencies [$F(6,78) < 1, MS_e = 6,029, p = .93$] nor the statistical analysis of the errors [$F(6,78) = 1.23, p = .30$] yielded an effect of SOA. Moreover, the overall error rate did not differ from zero [$F(1,13) = 2.54, p = .14$].

To summarize, incongruent and congruent colors did not differentially affect spoken-color-word naming at any SOA. Only a few errors were made in the cross detection

Table 1
Mean Manual-Response Latencies (*M*, in Milliseconds), Standard Deviations (*SD*), and Error Percentages (*E%*) per SOA for Experiments 1–4

| SOA | Experiment 1 | | | Experiment 2 | | | Experiment 3 | | | Experiment 4 | | |
|-------|--------------|-----------|-----------|--------------|-----------|-----------|--------------|-----------|-----------|--------------|-----------|-----------|
| | <i>M</i> | <i>SD</i> | <i>E%</i> | <i>M</i> | <i>SD</i> | <i>E%</i> | <i>M</i> | <i>SD</i> | <i>E%</i> | <i>M</i> | <i>SD</i> | <i>E%</i> |
| -300 | 443 | 88 | 0.0 | 493 | 134 | 1.6 | 533 | 144 | 3.2 | 443 | 95 | 0.0 |
| -200 | 437 | 111 | 0.0 | 503 | 146 | 1.6 | 522 | 184 | 0.8 | 428 | 108 | 0.8 |
| -100 | 452 | 92 | 0.8 | 488 | 119 | 1.6 | 543 | 190 | 0.8 | 456 | 119 | 0.8 |
| 0 | 443 | 101 | 0.8 | 482 | 130 | 0.8 | 526 | 146 | 0.0 | 446 | 123 | 0.0 |
| 100 | 476 | 149 | 0.0 | 515 | 160 | 0.0 | 558 | 174 | 0.8 | 440 | 85 | 1.6 |
| 200 | 437 | 93 | 0.0 | 499 | 145 | 0.8 | 521 | 140 | 0.8 | 452 | 125 | 0.0 |
| 300 | 443 | 100 | 0.8 | 503 | 142 | 0.0 | 564 | 173 | 1.6 | 459 | 140 | 0.0 |
| Total | 447 | 107 | 0.3 | 497 | 140 | 0.9 | 538 | 166 | 1.1 | 446 | 115 | 0.5 |

task, indicating that participants did look at the computer screen to see whether a cross or color was presented. Moreover, the spoken-word-naming latencies increased with increasing distractor-second SOAs, suggesting that participants postponed naming the spoken word until they saw on-screen a color (in the incongruent and congruent conditions) or an empty rectangle (in the control condition), and thus that they perceived the distractor colors. Spoken-word-naming latencies were shorter with colors than with the rectangle, suggesting that the colors were identified faster than was the rectangle.

Whereas incongruent and congruent spoken color words differentially affected color naming in Experiment 1, there was no such effect from colors on spoken-word naming in the present experiment; in short, there was a visual–auditory color–word Stroop asymmetry. Comparison of the latencies of color naming (Experiment 1) and spoken-word naming (Experiment 2) in the control conditions revealed that they were very similar. Averaged across SOAs, the means were 441 msec for color naming (Experiment 1) and 439 msec for spoken-word naming (Experiment 2). The mean latency for spoken-word naming includes the time that participants waited for the empty rectangle at distractor-second SOAs. Thus, it would seem better to compare the means at $SOA = 0$ msec, where the timing of the stimulus presentation was the same for both tasks. Statistical analysis showed that the latencies did not differ between color naming and spoken-word naming [$F(1,26) < 1$, $MS_e = 120,567$, $p = .39$]. Shimada (1990) also obtained equivalent latencies for color naming and spoken-word naming in Japanese (but Shimada did not test for a color–word Stroop asymmetry). To be completely sure that the latencies for color naming and spoken-word naming were equivalent, for my study a new group of 14 participants named the colors and spoken words in isolation (i.e., without distractors and without the secondary task). Each color and word was repeated 10 times in a block of trials, and the order of presenting the stimuli was random. Trials were blocked by task, with the order of tasks counterbalanced across participants. The means for color naming and spoken-word naming were 476 and 473 msec, respectively, which did not differ [$F(1,13) < 1$, $MS_e = 2,908$, $p = .70$]. The error rates for color naming and spoken-word naming were 2.6 and 5.2, respectively, which also did not differ [$F(1,13) = 3.39$, $p = .09$].

It may be argued that a mismatch in discriminability between color naming (Experiment 1) and spoken-word naming (Experiment 2) resulted because the mean production latency across congruent and incongruent conditions at $SOA = 0$ msec was longer for color naming (476 msec) than for spoken-word naming (386 msec), since the more salient dimension, spoken word, would intrude on the less salient dimension, color. The problem with this argument is that the congruent and incongruent Stroop conditions are not good estimates of baseline performance, as is evident from a comparison between these distractor conditions and the control condition (color naming without a

spoken distractor) in Experiment 1. Therefore, baseline performance is usually measured by a control condition (MacLeod, 1991) rather than by estimating it from the incongruent and congruent conditions. Melara and Algom (2003) also used measures of baseline performance based on color and word naming without distractors.

To conclude, the visual–auditory color–word Stroop asymmetry was obtained with equivalent latencies for color naming and word naming, with spoken words producing a Stroop effect in color naming (Experiment 1) and colors producing no Stroop effect in spoken-word naming (Experiment 2). These results support the architectural account rather than the discriminability or the pathway strength account of the color–word asymmetry.

EXPERIMENT 3

The aim of Experiment 3 was to examine whether a Stroop effect of distractors in spoken-word naming is present when written color words rather than color patches are used as distractors. Earlier research has shown that written words yield a Stroop effect in written-word naming (W. R. Glaser & Glaser, 1989). According to the word production architecture account advanced in Roelofs (2003), written color words should affect spoken-word naming even when colors have no effect (as observed in Experiment 2). According to this account, both spoken and written words have priority access to pronunciation over meaning, so both spoken and written words directly activate their names. Therefore, spoken and written words should interfere when pitted against each other experimentally, and the traditional word–word Stroop effect (W. R. Glaser & Glaser, 1989) should be replicated with written distractor words in spoken-word naming.

Method

Participants. The experiment was conducted with 14 paid participants from the pool of the Max Planck Institute. All participants were young adult native speakers of Dutch, and none took part in Experiments 1 or 2.

Materials, Design, Procedure, Apparatus, and Analyses. These were similar to their counterparts in Experiment 2. The stimuli consisted of the spoken Dutch color words “rood,” “groen,” and “blauw” of Experiment 1 and their written counterparts. The written words were presented in 36-point lowercase Arial font. In addition, a row of 5 Xs served as the stimulus in the control condition. There were again three congruent pairings (“rood”–ROOD, “groen”–GROEN, “blauw”–BLAUW), three incongruent pairings (“rood”–BLAUW, “groen”–ROOD, “blauw”–GROEN), and three control pairings (“rood”–XXXXX, “groen”–XXXXX, and “blauw”–XXXXX). Moreover, each spoken color word was also paired with a cross (i.e., “rood”–+, “groen”–+, “blauw”–+). The participants were instructed to name the spoken words as quickly as possible while trying to make no mistakes. They were also told to start naming the spoken word as soon as possible and not to wait until the end of the spoken signal. When the cross appeared, the participants simply had to press the button instead.

Results and Discussion

Figure 3 gives the mean spoken-word-naming latencies and the mean error percentages for each distractor

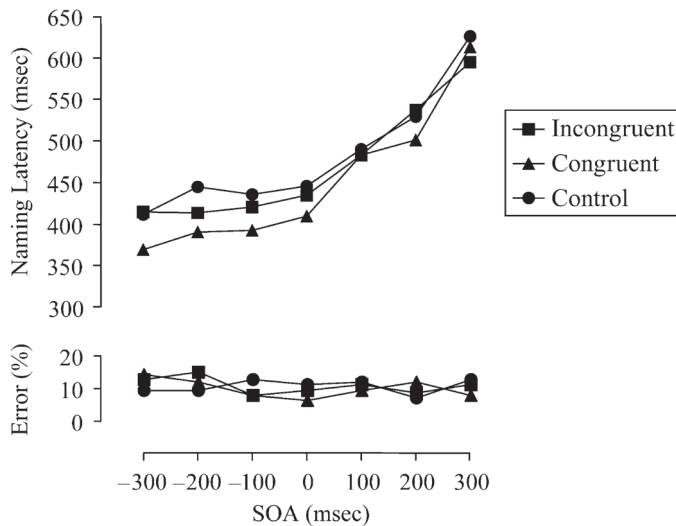


Figure 3. Mean latencies and error percentages for spoken-color-word naming while seeing written color word distractors in the incongruent, congruent, and control conditions in Experiment 3. The within-participants 95% confidence interval for the latencies is ± 21.7 msec.

condition by SOA. Table 1 gives the means of the latencies and the error percentages for the manual responses to the cross.

The statistical analysis of the spoken-word-naming latencies yielded main effects of distractor [$F(2,26) = 9.51$, $MS_e = 2,453$, $p < .001$] and SOA [$F(6,78) = 34.74$, $MS_e = 7,193$, $p < .001$]. There was no interaction of distractor and SOA [$F(12,156) = 1.55$, $MS_e = 1,713$, $p = .11$]. Pairwise comparisons between distractor conditions showed that the incongruent and control conditions did not differ [$F(1,13) = 1.93$, $MS_e = 3,655$, $p = .19$], but that the congruent condition was faster than the control condition [$F(1,13) = 15.00$, $MS_e = 3,063$, $p < .002$]. Moreover, the incongruent and congruent conditions differed from each other [$F(1,13) = 26.56$, $MS_e = 639$, $p < .001$]. The statistical analysis of production errors yielded no main effect of distractor [$F(2,26) < 1$, $p = .86$] or of SOA [$F(6,78) < 1$, $p = .76$] and no interaction of the two factors [$F(12,156) < 1$, $p = .65$].

The statistical analysis of the manual-response latencies yielded no effect of SOA [$F(6,78) < 1$, $MS_e = 9,427$, $p = .80$], and the statistical analysis of the errors also yielded no effect of SOA [$F(6,78) < 1$, $p = .80$]. However, the overall error rate did differ from zero [$F(1,13) = 5.51$, $p = .04$].

To summarize, a difference in effect was obtained between incongruent and congruent written color word distractors in spoken-color-word naming in the present experiment, whereas incongruent and congruent colors did not differentially affect spoken-word naming in Experiment 2. As in Experiment 2, only a few errors were made in the cross detection task, indicating that participants did look at the computer screen to see whether a cross or a written word was presented. Moreover, as in

Experiment 2, naming latencies increased with increasing distractor-second SOAs, suggesting that participants postponed onset of the naming response until they saw a written word or the Xs on the screen. However, in contrast to the results for colors and rectangles in Experiment 2, naming latencies were similar for incongruent written words and for the Xs in the present experiment. Finally, in both Experiments 2 and 3, naming of the spoken words took on average about 400–450 msec to initiate, whereas the duration of the auditory signal was about 700 msec. Thus, in both experiments, participants started naming the spoken words halfway through their perception, suggesting that the task of spoken-word naming was accomplished in similar ways in Experiments 2 and 3.

To conclude, written color word distractors yielded a Stroop effect in spoken-color-word naming in the present experiment, whereas colors did not affect spoken-word naming in Experiment 2. The Stroop effect of written words in the present experiment suggests that spoken-word naming is not immune to interference from visual stimuli. The pattern of the complete absence of a Stroop effect of colors in spoken-word naming (Experiment 2), the presence of a Stroop effect of written words in spoken-word naming (Experiment 3), and the presence of a Stroop effect of spoken words in color naming (Experiment 1) supports the word production architecture account of the color-word Stroop asymmetry.

EXPERIMENT 4

Written distractor words yield a Stroop effect in written-word naming (W. R. Glaser & Glaser, 1989). Experiment 3 showed that written distractor words also yield a Stroop effect in spoken-word naming. The aim of Experi-

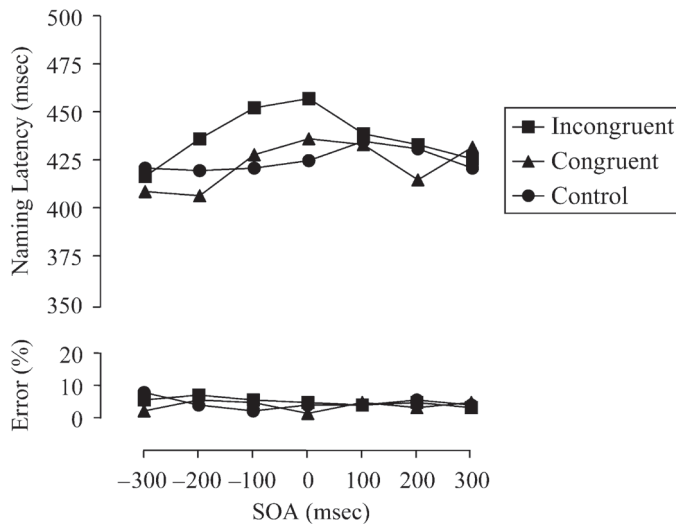


Figure 4. Mean latencies and error percentages for written-color-word naming while hearing spoken color word distractors in the incongruent, congruent, and control conditions in Experiment 4. The within-participants 95% confidence interval for the latencies is ± 17.6 msec.

ment 4 was to test for the presence of a Stroop effect of spoken words on the naming of written words.

Method

Participants. The experiment was conducted with 14 paid participants from the pool of the Max Planck Institute. All participants were young adult native speakers of Dutch. None of the participants took part in Experiments 1–3.

Materials, Design, Procedure, Apparatus, and Analyses. These were all similar to their counterparts in Experiment 1. The stimuli consisted of the spoken Dutch color words “rood,” “groen,” and “blauw” used in Experiment 1 and their written counterparts. The written words were presented in 36-point lowercase Arial font, as in Experiment 3. The control condition consisted of a written word without a spoken distractor, and in some trials a spoken word was paired with a cross. The participants were instructed to name the written words as quickly as possible while trying to make no mistakes and to ignore the spoken color words. When the cross appeared, they simply had to press the response button instead.

Results and Discussion

Figure 4 gives the mean written-word-naming latencies and the mean error percentages for each distractor condition by SOA. Table 1 gives the means of the latencies and the error percentages for the manual responses to the cross.

The statistical analysis of the written-word-naming latencies yielded a main effect of distractor [$F(2,26) = 3.65$, $MS_e = 1,472$, $p < .04$] but not of SOA [$F(6,78) = 1.19$, $MS_e = 2,557$, $p = .32$]. There was no interaction of distractor and SOA [$F(12,156) = 1.20$, $MS_e = 1,130$, $p = .28$]. Pairwise comparisons between distractor conditions showed that neither the incongruent condition [$F(1,13) = 2.96$, $MS_e = 2,208$, $p = .109$] nor the congruent condition [$F(1,13) < 1$, $MS_e = 1,401$, $p = .68$] differed from the control condition. However, the incongruent condi-

tion was slower than the congruent condition [$F(1,13) = 11.60$, $MS_e = 806$, $p < .005$]. The statistical analysis of reading errors yielded no main effect of distractor [$F(2,26) < 1$, $p = .53$] or of SOA [$F(6,78) < 1$, $p = .91$], and no interaction of distractor and SOA [$F(12,156) < 1$, $p = .53$].

The statistical analysis of the manual-response latencies yielded no effect of SOA [$F(6,78) < 1$, $MS_e = 1,373$, $p = .53$], and neither did the statistical analysis of the errors [$F(6,78) < 1$, $p = .75$]. The overall error rate did not differ from zero [$F(1,13) = 2.15$, $p = .17$].

Stroop effects were obtained for written distractor words in spoken-word naming (Experiment 3) and for spoken distractor words in written-word naming (Experiment 4). However, unlike in the results for Experiment 3, the written-word-naming latencies did not increase with increasing distractor-second SOAs (in line with the results of Experiment 1). Whereas participants had to postpone the naming of the spoken word until they saw a color or an empty rectangle on the screen in Experiment 3, this was not required in the present experiment, and this difference explains the different effects of SOA in the two experiments.

GENERAL DISCUSSION

Four experiments examined crossmodal versions of the Stroop task (1) to look for Stroop asymmetries in color naming, spoken-word naming, and reading and to evaluate the time course of these asymmetries and (2) to compare these findings to current models of the Stroop effect. In Experiment 1, participants named color patches while ignoring spoken color words presented at various SOAs. Either the onset of the spoken word was 300, 200, or

100 msec earlier than that of the color patch; the onsets of the spoken word and the color coincided; or the onset of the spoken word was 100, 200, or 300 msec later than that of the color patch. In Experiment 2, participants named spoken words while ignoring color patches presented at the same SOAs as in Experiment 1, with a secondary visual detection task to assure that they looked at the colors in both tasks. In Experiment 3, participants named spoken words while ignoring written color words, and finally, in Experiment 4, participants read aloud written while ignoring spoken color words.

In Experiment 1, differential effects of incongruent and congruent spoken-word distractors were obtained on color naming. The largest effect of spoken words was around $SOA = 0$ msec, replicating the classic findings that M. O. Glaser and W. R. Glaser (1982; W. R. Glaser & Glaser, 1989) obtained with written-word distractors. In contrast with Shimada's (1990) results, the longest naming latencies were also obtained around $SOA = 0$ msec, thus replicating the findings of Elliott et al. (1998) and M. O. Glaser and W. R. Glaser (1982; W. R. Glaser & Glaser, 1989). In Experiment 2, no differential effects of incongruent and congruent color distractors on spoken-word naming were obtained at any SOA. Only a few errors were made in the secondary cross detection task, indicating that participants looked at the computer screen to see whether a cross or a color was presented. Moreover, production latencies in Experiment 2 increased with increasing distractor-second SOAs, suggesting that participants postponed naming a spoken word until they saw a color or an empty rectangle on the screen. These results indicate that participants saw the distractor colors, so whereas spoken color words yielded a Stroop effect on color naming in Experiment 1, there was no such effect from colors on spoken-word naming in Experiment 2, demonstrating a visual-auditory color-word Stroop asymmetry. A comparison of the latencies of color naming (Experiment 1) and spoken-word naming (Experiment 2) in the control conditions revealed that these latencies were very similar, and this was confirmed in an independent examination of the latencies. Thus, the visual-auditory color-word Stroop asymmetry was obtained with similar latencies for color naming and word naming. Taken together, the results of Experiments 1 and 2 support the architectural rather than the discriminability or the pathway strength account of the asymmetry.

The present results do not preclude that differences in dimensional discriminability may contribute to the size of Stroop effects or cause Stroop asymmetries in the absence of architectural differences. For example, differences in font size (Melara & Algom, 2003) and spatial differences between a word and a color patch (Gatti & Egeth, 1978) influence the size of Stroop effects; architectural and discriminability factors can actually complement one another. However, the present findings suggest that when the color and word dimensions are matched in discriminability, a color-word Stroop asymmetry is still obtained, suggesting that an architectural difference rather than a

difference in discriminability causes the basic color-word Stroop asymmetry.

In Experiment 3, a difference in effect was obtained between incongruent and congruent written color words on spoken-word naming, whereas incongruent and congruent colors did not differentially affect spoken-word naming in Experiment 2. The Stroop effect of written words in Experiment 3 suggests that spoken-word naming is not immune to interference from visual stimuli. In Experiment 4, a difference in effect was also obtained between incongruent and congruent spoken color words on written-word naming. The absence of an effect of colors (Experiment 2) and the presence of an effect of written words (Experiment 3) on spoken-word naming (and vice versa, as in Experiment 4), and the presence of an effect of spoken words on color naming (Experiment 1), all support the architectural account of the color-word Stroop asymmetry. According to this account, both written words and spoken words have priority access to pronunciation before meaning, whereas the reverse holds for colors.

The present results are also relevant for evaluating the classic "translation" account of the color-word Stroop asymmetry advanced by Virzi and Egeth (1985). The translation account holds that colors and words are mapped from one internal code to another by a translation device. According to this account, interference is obtained when the relevant stimulus type (e.g., a color stimulus) does not match the response type (e.g., a word response)—that is, when the task requires a translation between types. This could explain why written words interfere with color naming (which requires a color-to-word translation) and why color patches do not interfere with word reading (which does not require a translation). Spoken-word naming, however, would seem to be a non-translation task par excellence, because it requires mapping speech onto speech. Some theories even go so far as to assume that the very same representations used for producing speech are also involved in recognizing speech (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). However, although spoken-word naming is unaffected by color patches, it is affected by written words. This suggests that the tight link between speech perception and production can be broken into by written words. Thus, even though spoken-word naming involves no translation according to the translation account, it is still affected by distractors that require a translation between input and output modalities—namely, written words.

To conclude, performing the visual-auditory color-word Stroop task yielded a color-word asymmetry in effects between the tasks of color and word naming similar to the classic visual color-word Stroop situation. Naming colors was differentially affected by incongruent and congruent spoken color words, but naming spoken color words was not differentially affected by incongruent and congruent color patches at any SOA. The visual-auditory color-word Stroop asymmetry was obtained with equivalent latencies for color and spoken-word naming. Naming spoken color words was differentially affected by incongruent and

congruent written color words, showing that spoken-word naming is not immune to interference from visual stimuli, and spoken distractor words yielded a Stroop effect in written-word naming. These results support the architectural account of the color–word Stroop asymmetry.

REFERENCES

- 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.
- CATTELL, J. M. (1886a). The time it takes to see and name objects. *Mind*, **11**, 63-65.
- CATTELL, J. M. (1886b). The time taken up by cerebral operations. *Mind*, **11**, 524-538.
- 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.
- COHEN, J. D., & HUSTON, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 453-476). Cambridge, MA: MIT Press, Bradford Books.
- COWAN, N. (1989a). The reality of cross-modal Stroop effects. *Perception & Psychophysics*, **45**, 87-88.
- COWAN, N. (1989b). A reply to Miles, Madden, and Jones (1989): Mistakes and other flaws in the challenge to the cross-modal Stroop effect. *Perception & Psychophysics*, **45**, 82-84.
- COWAN, N., & BARRON, A. (1987). Cross-modal, auditory–visual Stroop interference and possible implications for speech memory. *Perception & Psychophysics*, **41**, 393-401.
- DAMIAN, M. K., & MARTIN, R. C. (1999). Semantic and phonological codes interact in single word production. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 345-361.
- ELLIOTT, E. M., & COWAN, N. (2001). Habituation to auditory distractors in a cross-modal, color–word interference task. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **27**, 654-667.
- ELLIOTT, E. M., COWAN, N., & VALLE-INCLAN, F. (1998). The nature of cross-modal color–word interference effects. *Perception & Psychophysics*, **60**, 761-767.
- GATTI, S. V., & EGETH, H. E. (1978). Failure of spatial selectivity in vision. *Bulletin of the Psychonomic Society*, **11**, 181-184.
- GLASER, M. O., & GLASER, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 875-894.
- GLASER, W. R., & DÜNGELHOFF, F.-J. (1984). The time course of picture–word interference. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 640-654.
- GLASER, W. R., & GLASER, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General*, **118**, 13-42.
- HANAUER, J. B., & BROOKS, P. J. (2003). Developmental change in the cross-modal Stroop effect. *Perception & Psychophysics*, **65**, 359-366.
- LEVELT, W. J. M., ROELOFS, A., & MEYER, A. S. (1999). A theory of lexical access in speech production. *Behavioral & Brain Sciences*, **22**, 1-38.
- LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P., & STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychological Review*, **74**, 431-461.
- LONG, G. M., & LYMAN, B. J. (1987). Foveal and parafoveal processing of asynchronous Stroop stimuli. *British Journal of Psychology*, **78**, 151-162.
- MACLEOD, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.
- MACLEOD, C. M. (1992). The Stroop task: The “gold standard” of attentional measures. *Journal of Experimental Psychology: General*, **121**, 12-14.
- MACLEOD, C. M., & MACDONALD, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, **4**, 383-391.
- MELARA, R. D., & ALGOM, D. (2003). Driven by information: A tectonic theory of Stroop effects. *Psychological Review*, **110**, 422-471.
- MEYER, A. S., & SCHRIEFERS, H. (1991). Phonological facilitation in picture–word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 1146-1160.
- MEYER, A. S., & VAN DER MEULEN, F. F. (2000). Phonological priming effects on speech onset latencies and viewing times in object naming. *Psychonomic Bulletin & Review*, **7**, 314-319.
- MILES, C., & JONES, D. M. (1989). The fallacy of the cross-modal Stroop effect: A rejoinder to Cowan (1989). *Perception & Psychophysics*, **45**, 85-86.
- MILES, C., MADDEN, C., & JONES, D. M. (1989). Cross-modal, auditory–visual Stroop interference: A reply to Cowan and Barron (1987). *Perception & Psychophysics*, **45**, 77-81.
- POSNER, M. I., & BOIES, S. J. (1971). Components of attention. *Psychological Review*, **78**, 391-408.
- ROELOFS, A. (1992). A spreading-activation theory of lemma retrieval in speaking. *Cognition*, **42**, 107-142.
- ROELOFS, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, **64**, 249-284.
- ROELOFS, A. (2003). Goal-referenced selection of verbal action: Modeling attentional control in the Stroop task. *Psychological Review*, **110**, 88-125.
- ROELOFS, A., & HAGOORT, P. (2002). Control of language use: Cognitive modeling of the hemodynamics of Stroop task performance. *Cognitive Brain Research*, **15**, 85-97.
- SCHRIEFERS, H., MEYER, A. S., & LEVELT, W. J. M. (1990). Exploring the time course of lexical access in language production: Picture–word interference studies. *Journal of Memory & Language*, **29**, 86-102.
- SHIMADA, H. (1990). Effect of auditory presentation of words on color naming: The intermodal Stroop effect. *Perceptual & Motor Skills*, **70**, 1155-1161.
- SMITH, M. C., & MAGEE, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, **109**, 373-392.
- STROOP, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.
- STUART, D. M., & CARRASCO, M. (1993). Semantic component of a cross-modal Stroop-like task. *American Journal of Psychology*, **106**, 383-405.
- VRIZI, R. A., & EGETH, H. E. (1985). Toward a translational model of Stroop interference. *Memory & Cognition*, **13**, 304-319.

(Manuscript received November 4, 2003;
revision accepted for publication November 23, 2004.)