Attention and Facilitation: Converging Information Versus Inadvertent Reading in Stroop Task Performance

Ardi Roelofs
Radboud University Nijmegen

Disagreement exists about whether color–word Stroop facilitation is caused by converging information (e.g., Cohen et al., 1990; Roelofs, 2003) or inadvertent reading (MacLeod & MacDonald, 2000). Four experiments tested between these hypotheses by examining Stroop effects on response time (RT) both within and between languages. Words cannot be read aloud to produce facilitation between languages. Dutch–English bilingual participants named color patches while trying to ignore Dutch or English color words presented at a wide range of preexposure and postexposure stimulus onset asynchronies. The color patches were named in Dutch (Experiments 1 and 2) or English (Experiments 3 and 4). In all experiments, Stroop facilitation and interference effects were obtained in mean RTs with similar time courses within and between languages. Facilitation was generally present throughout the entire RT distributions. These results suggest that Stroop interference and facilitation have a common locus within and between languages, supporting the converging information hypothesis of Stroop facilitation.

Keywords: attention, facilitation, naming, reading, Stroop

Attentional control includes the ability to formulate goals and plans of action and to follow these while resisting distraction. This ability is critical to normal human functioning, and it is central to intelligence (e.g., Kane & Engle, 2002). A widely employed task in studying attentional control is the color–word Stroop task (Stroop, 1935), sometimes called the "gold standard" of attentional measures (MacLeod, 1992). This task involves naming the ink color of incongruent or congruent color words (e.g., the word RED or GREEN printed in green; say "green") or a nonverbal control stimulus (e.g., a series of *X*s in green). Alternatively, participants name color patches with superimposed words rather than their ink colors (e.g., M. O. Glaser & Glaser, 1982). Response time (RT) is typically longer in the incongruent than the control condition, descriptively called *interference*, and often shorter in the congruent than the control condition, descriptively called *facilitation* (see MacLeod, 1991, for a review). Whereas researchers generally agree that Stroop interference reflects the processing effect of diverging color and word information, they have found no agreement on the source of Stroop facilitation.

According to the converging information hypothesis (e.g., Cohen, Dunbar, & McClelland, 1990; Melara & Algom, 2003; Roelofs, 2003), converging information on the word and color dimensions combines to produce facilitation on congruent trials,

just as diverging information causes interference on incongruent trials. Under this hypothesis, interference and facilitation arise from the same underlying mechanism. In contrast, according to the inadvertent reading hypothesis (Kane & Engle, 2003; MacLeod & MacDonald, 2000), the congruent condition does not measure true facilitation, but instead is contaminated by inadvertent word reading. Kane and Engle (2002, 2003) suggested that because of occasional lapses of attention, the goal of color naming is lost on some of the trials. If the goal is not recovered in time and the word is inadvertently read, this leads to a naming error on incongruent trials but a correct response on congruent trials. In saying "green" to the word GREEN in green ink, it is unclear whether participants name the color or erroneously read the word. Consequently, accidental reading errors may be included in the calculation of the color naming latency in the congruent condition. This yields apparent facilitation because reading is some 100–200 ms faster than color naming (e.g., MacLeod, 1991). If the inadvertent reading hypothesis is correct, then theories that predict true facilitation accompanying interference (e.g., Cohen et al., 1990; Melara & Algom, 2003; Roelofs, 2003) are challenged.

If reading errors are the source of Stroop facilitation, the effect should disappear when the errors can be prevented or removed. Reading errors can be filtered in a between-language version of the Stroop task (MacLeod & MacDonald, 2000). For example, in saying "green" to the French word VERT in green ink, the participants must have named the color. If Stroop facilitation stems from converging information, then facilitation should occur whenever word and color information converge. Thus, bilingual participants should show facilitation within and between languages. In contrast, if Stroop facilitation stems from unobservable reading errors, then facilitation should disappear when such errors can be prevented or removed. Bilingual individuals should show facilitation within a language but not between languages.

In an English–French bilingual Stroop study described by MacLeod and MacDonald (2000),[1] facilitation was obtained from congruent trials within language (e.g., GREEN printed in green; say "green") and interference from congruent trials between languages (e.g., VERT printed in green; say "green"). Between-language interference on congruent trials had also earlier been obtained by Dalrymple-Alford (1968), who asked Arabic–English bilingual speakers to name the presentation color of English color words or series of *X*s using their first language, Arabic. The colors were red (*ahmar* in Arabic), green (*qahdar*), blue (*azrak*), yellow (*asfar*), and black (*aswad*). More recently, Abunuwara (1992) asked trilingual participants to name Stroop color words in their first language (Arabic) and their second and third languages (Hebrew and English). The colors were red, green, blue, and yellow, and the distractor words were the corresponding names in Arabic, Hebrew (*ah-dom*, *yah-rok*, *kah-khol*, and *tsah-hov*, respectively), and English. A geometrical shape consisting of two segmented parallel lines served as control item. Relative to control, the congruent words yielded, on average, a 45-ms facilitation effect within language but a 58-ms interference effect between languages. The facilitation and interference effects may arise if participants occasionally read the word aloud in the within-language condition (which speeds up responding) but covertly repair such reading errors in the between-language condition (which delays responding). To summarize, within-language facilitation and between-language interference are obtained in the congruent condition, which supports the inadvertent reading hypothesis of Stroop facilitation.

The evidence for inadvertent reading is not conclusive, however. The phonological forms of the color names of English, French, Arabic, and Hebrew are very different. Thus, the color words provide converging conceptual information but diverging phonological information in the between-language congruent condition. For example, the French word VERT denotes the same color as the English word GREEN, but the phonological properties of the words are very different. In contrast, in the congruent condition in the standard within-language version of the Stroop task, the word and color (e.g., GREEN in green ink) are identical not only conceptually but also phonologically. If facilitation from converging conceptual information is offset by an opposing effect of diverging phonological information, this may yield between-language interference in the congruent condition. This explanation of between-language interference in terms of divergent phonological information is in agreement with the converging information hypothesis of standard Stroop facilitation.

Figure 1 illustrates how facilitation and interference effects may occur as a result of converging and diverging information in performing the Stroop task. The figure shows the processing stages assumed by the WEAVER++ model, which is a computational model of spoken word planning and its attentional control (Levelt, Roelofs, & Meyer, 1999; Roelofs, 1992, 1997, 2003). The model distinguishes between conceptual identification, lemma retrieval, and word–form encoding, with the encoding of forms further divided into morphological, phonological, and phonetic encoding. Information is retrieved from a lexical network by spreading activation. During conceptual identification in color naming, a concept is selected for the perceived color. In lemma retrieval, a selected concept is used to retrieve a lemma from memory, which is a representation of the syntactic properties of a word (not
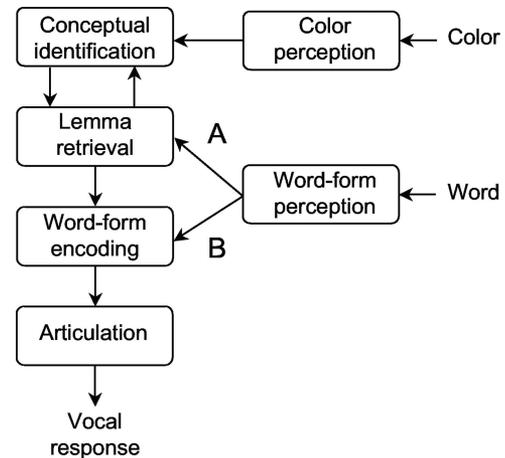

Figure 1. The WEAVER++ model applied to Stroop task performance. Perceived words may affect conceptual identification and lemma retrieval (route A) and word–form encoding (route B) in parallel.

relevant for the Stroop task, but crucial for the normal use of the word in sentences). For example, the lemma of the word *green* says that it can be used as an adjective. In word–form encoding, the lemma is used to retrieve the morphophonological properties of the word from memory in order to construct an appropriate articulatory program. For example, for *green* the morpheme <green> and the speech segments /g/, /r/, /i/, and /n/ are retrieved and a motor program is generated. Finally, articulation processes execute the motor program, which yields overt speech. The model assumes that a perceived word activates corresponding lemmas and concepts (route A in Figure 1) as well as word forms (route B) in parallel. The literature on bilingual word recognition suggests that word forms are activated across languages to the extent that their orthography overlaps (e.g., Dijkstra & Van Heuven, 2002, for a review). For example, the English word GREEN partly activates the form of the Dutch word *groen* (converging information) but not of the French word *vert* (diverging information).

The model has successfully been applied to Stroop task performance (Roelofs, 2003). In performance on incongruent trials, such as naming the presentation color of the English word GREEN or French word VERT in red ink (say "red"), the information activated by the word GREEN or VERT via routes A and B does not correspond to the target word *red*, which yields interference. In performance on within-language congruent trials, however, such as naming the presentation color of the English word GREEN in green ink (say "green"), the information activated by the word GREEN via routes A and B corresponds to the target word *green*. Because planning the word receives converging activation at both conceptual and word–form encoding levels, the naming of the color is facilitated compared with a control condition. Computer simulations reported in Roelofs (2003) revealed that WEAVER++ not only successfully accounts for within- and between-language Stroop interference and within-language facilitation but also for several other classic findings,

---

[1] MacLeod and MacDonald (2000) describe the results but not the methods and analyses in detail.

mostly taken from the review by MacLeod (1991). With only three free parameters, the model accounted for 96% of the variance of 16 classic studies (250 data points). The model can also explain why for some language combinations (e.g., English, French, Arabic, and Hebrew), Stroop interference is obtained between languages on congruent trials. For example, in naming the presentation color of the French word VERT in green ink using English (say "green"), the word VERT activates overlapping conceptual information via route A, but it activates diverging word–form (i.e., morphological, phonological, and phonetic) information via route B. If facilitation from converging conceptual information (route A) is opposed by an interfering effect of diverging phonological information (route B), this may yield between-language interference in the congruent condition.

This interpretation of between-language interference receives some support from a Spanish–Catalan bilingual Stroop study by Costa, Albareda, and Santesteban (2008). They observed between-language facilitation from congruent words compared with color-unrelated words. Spanish and Catalan color words provide converging conceptual information (through route A in Figure 1) but diverging phonological information (through route B) in the between-language congruent condition. However, different from a series of Xs (Dalrymple-Alford, 1968) or geometrical shapes (Abunuwara, 1992), unrelated words also provide diverging phonological information in the control condition. Comparing the between-language effect of congruent color words and unrelated control words factors out the diverging phonological information and reveals the effect of converging conceptual information in the congruent condition. This may explain why Costa et al. (2008) obtained between-language facilitation from congruent words relative to control.

Although the results of Costa et al. (2008) are suggestive, converging evidence is needed. Such evidence may be obtained by using languages with color words that have highly similar phonological forms, such as Dutch and English. The Dutch words for the colors red, green, and blue are *rood*, *groen*, and *blauw*. The pronunciations of the Dutch and English words are similar but they differ, so inadvertent reading responses can be detected in the between-language condition. In phonetic terms, the pronunciations are /red/, /grin/, and /blu:/ for the English color words and /rot/, /grun/, and /blau/ for the Dutch words (e.g., Booij, 1995). Because the amount of diverging phonological information is small, between-language Stroop facilitation should be obtained even compared with nonverbal control stimuli, which are often used in standard within-language Stroop experiments. For example, previous research used swastikas (Stroop, 1935), a series of *V*s (Dyer, 1971), and a series of *I*s (M. O. Glaser & Glaser, 1982). This prediction concerning facilitation derived from the converging information hypothesis was tested in four experiments reported in the present article. Facilitation in the Stroop task is often observed only with preexposure of the word (e.g., M. O. Glaser & Glaser, 1982). Thus, the between-language facilitation predicted by the converging information hypothesis may occur at preexposure stimulus onset asynchronies (SOAs) even if facilitation is not present at zero SOA. In order not to miss between-language facilitation if it exists, effects were assessed using a wide range of SOAs. In addition, assessing within- and between-language Stroop effects at several preexposure and postexposure SOAs allows for a more extensive comparison of within- and between-language effects

than assessing performance at zero SOA only, as was done in previous studies (e.g., Abunuwara, 1992; Dalrymple-Alford, 1968; MacLeod & MacDonald, 2000). In contrast to the converging information hypothesis, the inadvertent reading hypothesis predicts within-language facilitation and between-language interference (if reading errors are covertly repaired) in the congruent condition relative to control.

In Experiments 1 and 2, Dutch–English bilingual individuals named color patches in their first language Dutch while trying to ignore Dutch and English color words. Although the Dutch participants were fluent in their second language, English, they were linguistically unbalanced in that English was spoken less fluently than Dutch. It is possible that inadvertent reading is more likely to occur in second-language color naming with first-language word distractors than in first-language color naming with second-language distractors, because of the difference in relative strength of the languages. If so, between-language interference may be obtained in second-language color naming even if it is not obtained in first-language color naming. This possibility was examined in Experiments 3 and 4, where the participants named the stimuli in their second language, English, while trying to ignore Dutch and English color words. According to the converging information hypothesis, between-language facilitation should be obtained regardless of the naming language, Dutch (Experiments 1 and 2) or English (Experiments 3 and 4).

The Dutch and English distractor words were blocked by language (Experiment 1) or languages were randomly mixed (Experiments 2–4). The blocking of languages in Experiment 1 resembles the situation in standard Stroop experiments, where the language of the written color words also does not vary from trial to trial. Still, blocking trials by distractor languages may evoke special processing strategies. If the distractor language does not correspond to the target language in a block of trials, the distractor words may perhaps be blocked out more strongly on the basis of language than if languages match, reducing between-language distractor effects. The trials of the different distractor languages were randomly intermixed in Experiments 2–4 to prevent such processing strategies.

In addition to analyzing mean RTs, the entire RT distributions of responding on congruent and control trials were examined in the present experiments (cf. Luce, 1986). The inadvertent reading hypothesis needs to assume that oral reading responses occur on only a few trials, because if inadvertent reading occurs on all congruent trials, the magnitude of facilitation should be much bigger than is usually observed. Whereas oral reading is typically some 100–200 ms faster than color naming, the facilitation in Stroop experiments is typically some 25 ms (e.g., M. O. Glaser & Glaser, 1982; W. R. Glaser & Glaser, 1989). Under the converging information hypothesis, this facilitation effect reflects a constant speed-up of some 25 ms on most of the congruent trials compared with control trials. Consequently, the distribution curve for RTs on congruent trials should be just that for control trials, except that the former is shifted toward shorter RTs by some 25 ms (cf. Roelofs, 2008). In contrast, according to the inadvertent reading hypothesis, the facilitation stems from a 100- to 200-ms speed-up that occurs only on a small proportion of the congruent trials. Consequently, the effect on the mean RT might be the same 25 ms, but the effect on the RT distributions should be different. The distribution curves of RTs in the congruent and control conditions should then differ

among the short RTs but converge at the longer RTs. These predictions regarding RT distributions will be evaluated by examining the entire RT distributions in the congruent and control conditions.

Experiment 1

In the first experiment, participants named color patches in Dutch while trying to ignore Dutch or English color words. Trials were blocked by distractor language. The converging information hypothesis predicts facilitation both within and between languages, whereas the inadvertent reading hypothesis predicts facilitation within language and between-language interference if reading errors are covertly repaired. Moreover, given that the magnitude of facilitation tends to be largest when words are presented in time before the color patches (e.g., M. O. Glaser & Glaser, 1982), the facilitation is expected at preexposure SOAs.

Method

Participants. The experiment was carried out with 28 participants (paid €5), who were young adult students at Radboud University Nijmegen. The participants in this experiment and the other ones were native speakers of Dutch, who spoke English rather fluently. They had ages ranging from 18 to 30 years with a mean of 22 years. A self-rating questionnaire was used to obtain proficiency scores for the participants. The scores were on a 5-point scale, in which 1 represents equal proficiency in English and Dutch and 5 represents much lower proficiency in English than Dutch. Participants rated their proficiency for English compared with Dutch as 2.8, on average. Scores for English use were also measured at a 5-point scale, where 1 represents less than 1 hr per week and 5 represents more than 10 hr per week. Participants used English for 1.6 hr per week, on average. Age of English onset refers to the age at which participants started learning English, which was approximately at the age of 11, on average.

Materials and design. The stimuli consisted of red, green, and blue color patches and the corresponding Dutch color words ROOD, GROEN, and BLAUW and the English color words RED, GREEN, and BLUE. The colors were presented as colored rectangles of 1.5 cm high and 4.5 cm wide. The written words were presented in 36-point lowercase Arial font. A row of five *X*s served as control stimulus.

There were four independent variables. The first variable was *Stroop condition*. Within each language, there were three congruent pairings (Dutch ROOD–red, GROEN–green, BLAUW–blue, and English RED–red, GREEN–green, BLUE–blue), three incongruent pairings (Dutch ROOD–blue, GROEN–red, BLAUW–green, and English RED–blue, GREEN–red, BLUE–green), and three control stimuli (the color patches combined with the row of *X*s). The second variable was *SOA* with seven levels: –300, –200, –100, 0, 100, 200, and 300 ms. A minus sign (e.g., –300 ms) denotes preexposure of the distractor word. The order of SOA blocks was counterbalanced across participants using a Latin square. The third variable concerned the relation between the target and distractor languages (within-language vs. between-language), referred to as *status*. The color word distractors were either Dutch or English words. The fourth variable was *language order*. Half the participants first received the Dutch distractor

words and then the English ones, and vice versa for the other half of the participants. Each of the nine stimuli occurred four times in random order within an SOA block.

Procedure and apparatus. The participants were tested individually. They were seated in front of a CRT monitor (NEC Multisync) and a Sennheiser microphone connected to an electronic voice key. The distance between participant and screen was approximately 50 cm. The participants were asked to name the color patches as quickly as possible while trying to make no mistakes. A trial started with the presentation of the color–word stimulus with the appropriate SOA. The stimuli remained visible for 1.5 s after color onset. Following stimulus presentation, the screen was blank for 1 s, after which the next trial began. An IBM-compatible computer controlled the data collection.

Analysis. After each trial, the experimenter coded the response for errors: wrong response word, wrong pronunciation of the word, disfluency, triggering of the voice key by a nonspeech sound, and failure to respond within 1.5 s after color presentation. Incorrect responses were excluded from the statistical analyses of the naming latencies. The RTs and errors were submitted to analyses of variance with the crossed variables Stroop condition, SOA, status, and language order. Stroop condition, SOA, and status were tested within participants, and language order was tested between participants. Paired *t* tests assessed Stroop facilitation (i.e., congruent vs. control) and Stroop interference (i.e., incongruent vs. control) for each SOA, using a Bonferroni corrected α of .007.

To obtain the RT distribution curves in the experiment, I divided the rank-ordered RTs for each participant into 20% quantiles and computed mean RTs for each quantile, separately for the congruent and control Stroop conditions. By averaging these quantile means across participants, Vincentized cumulative distribution curves are obtained (Ratcliff, 1979). Vincentizing the latency data across individual participants provides a way of averaging data while preserving the shapes of the individual distributions.

Results and Discussion

Figure 2 gives the mean Dutch naming latencies for Stroop condition, SOA, and status. Responding was slowest in the incongruent condition and fastest in the congruent condition. Facilitation (i.e., faster responses in the congruent than control condition) was obtained both within and between languages. The time course of the Stroop effects did not depend much on whether the distractor language was Dutch (within-language) or English (between-languages). Figure 2 also gives the error rates, which did not differ much among conditions.

The error rates were below 2%. Statistical analysis of the errors yielded no significant effects (all ps > .20). Statistical analysis of the naming latencies yielded effects of Stroop condition, $F(2, 52) = 104.34$, $p < .001$; SOA, $F(6, 156) = 7.70$, $p < .001$; and status, $F(1, 26) = 8.66$, $p < .007$. Means for the Dutch (within-language) and English (between-language) distractors were 561 and 581 ms, respectively. There was an interaction of Stroop condition and SOA, $F(12, 312) = 5.56$, $p < .001$, which was independent of status, $F(12, 312) < 1$, $p = .59$. Responding was overall faster in the congruent than in the control condition, $F(1, 27) = 20.30$, $p < .001$. The facilitation did not vary with SOA, $F(6, 162) = 1.53$, $p = .17$, and this was independent of status, $F(6,$
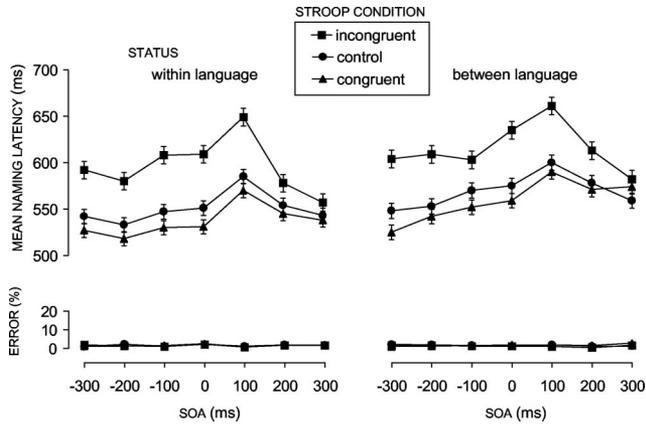
Figure 2. Mean Dutch naming latencies and error percentages per stimulus onset asynchrony (SOA) and Stroop condition for the Dutch distractor words (within language) and the English distractor words (between language) in Experiment 1. The error bars indicate one standard error.

$162) < 1$, $p = .87$. Planned comparisons revealed that the facilitation reached significance at the SOAs of −300 ms, −200 ms, and 0 ms (ps $< .001$), but not at the other SOAs. Responding was overall slower in the incongruent than in the control condition, $F(1, 27) = 86.25$, $p < .001$. The interference varied with SOA, $F(6, 162) = 5.04$, $p < .001$, independent of status, $F(6, 162) = 1.20$, $p = .31$. Planned comparisons revealed that the interference was significant at all SOAs (ps $< .001$), except at SOA = 300 ms. To summarize, Stroop interference and facilitation effects were obtained both within and between languages. The magnitude of the effects and their time course did not differ between distractor languages (i.e., the within- and between-language effects were the same).

Kane and Engle (2002, 2003) suggested that inadvertent reading occurs because of lapses of attention on some of the trials. If the word is inadvertently read before attention is regained, this leads to a naming error on incongruent trials but to a correct and fast response on congruent trials. In this view, the mean RT of the congruent condition is made up by a few fast reading responses and many normal color naming responses, whereas the mean RT of the control condition is determined by normal color naming responses only. Thus, facilitation should be confined to the leading edge of the RT distribution. To determine whether this was the case, the latency distributions in the congruent and control conditions were examined. Figure 3 gives the distributional plots for the congruent and control trials for the preexposure SOAs. In the mean RTs, facilitation reached significance at the preexposure SOAs of −300 ms and −200 ms. The figure shows that the facilitation effect existed throughout almost the entire latency range both within and between distractor languages. Thus, the latency distributions suggest that facilitation does not result from the occasional odd trial and that facilitation is not present for the relatively fast responses only. To conclude, the consistent facilitation throughout the RT distributions presents a challenge to the hypothesis that Stroop facilitation arises from inadvertent reading.

The participants receiving the Dutch distractor words first had longer color naming latencies than those receiving the English distractor words first, 599 and 543 ms, respectively, $F(1, 26) = 5.25$, $p < .03$. Moreover, the effect of Stroop condition differed depending on the language order, $F(2, 52) = 7.44$, $p < .001$, but this order effect was not modulated by SOA, $F(12, 312) < 1$, $p = .88$, or status, $F(2, 52) < 1$, $p = .76$. The magnitude of the Stroop interference was larger when participants received the Dutch distractors before the English ones (i.e., 58 ms) than vice versa (i.e., 33 ms), $F(1, 26) = 8.16$, $p < .008$, independent of status, $F(1, 26) < 1$, $p = .74$. The magnitude of Stroop facilitation did not vary
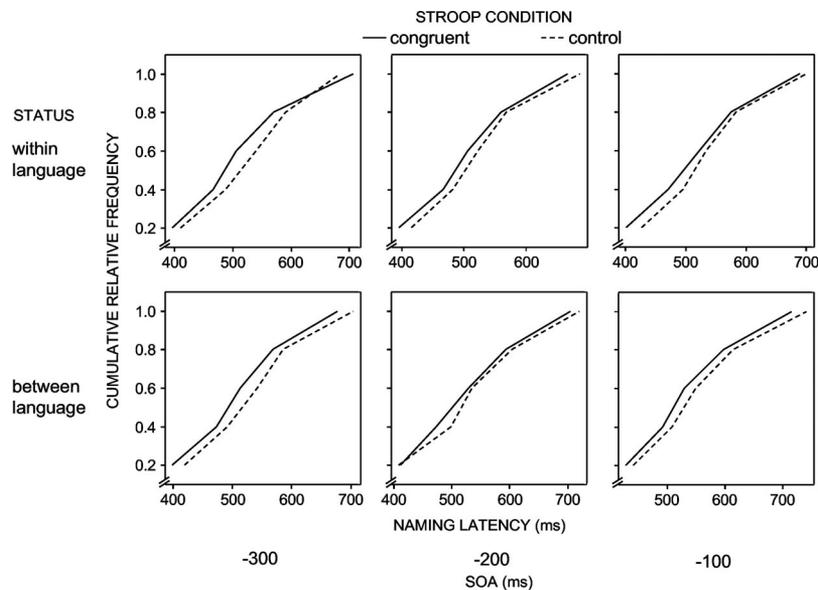

Figure 3. Vincentized cumulative distribution curves for the response latencies in the congruent and control conditions within and between languages in Experiment 1. The distributions are shown for the preexposure stimulus onset asynchronies (SOAs) of −300, −200, and −100 ms.

with language order, $F(1, 26) = 1.08$, $p = .31$, independent of status, $F(1, 26) < 1$, $p = .67$. To summarize, the magnitude of interference but not facilitation depended on the order of receiving the distractor languages. This effect of language order was independent of status. Perhaps participants adjusted their response criterion depending on language order. Elsewhere, I demonstrated using WEAVER++ how differences in response criterion may affect Stroop interference without affecting facilitation (Roelofs, 2003), as observed in the present experiment.

To conclude, Stroop facilitation and interference effects were obtained not only within language, as expected, but also between languages. The facilitation was present throughout the entire RT distribution, both within language and between languages. Because reading cross-language distractor words rather than naming color patches would not produce a facilitation effect but an error in the between-language condition, the appearance of a Stroop facilitation effect in both the within-language and between-language conditions provides evidence for the converging information hypothesis of Stroop facilitation and suggests that Stroop interference and facilitation effects have a common locus.

Experiment 2

In Experiment 1, trials were blocked by distractor language, which may have induced strategic adjustments, as suggested by the interaction between Stroop condition and language order. The distractor languages were randomly intermixed in the second experiment to avoid such strategies. Again, participants named color patches in Dutch while trying to ignore Dutch or English color words. The converging information hypothesis predicts facilitation both within and between languages, whereas the inadvertent reading hypothesis predicts within-language facilitation and between-language interference. Facilitation is again expected at preexposure SOAs.

Method

The method was the same as in Experiment 1, except that the distractor language (i.e., within-language vs. between-language status) now randomly varied from trial to trial. The naming latencies and errors were submitted to analyses of variance with the crossed within-participant variables Stroop condition, SOA, and status. The experiment was carried out with 21 new paid participants from the same population as in the previous experiment.

Results and Discussion

Figure 4 gives the mean Dutch naming latencies. Responding was slowest on incongruent trials and fastest on congruent trials. Facilitation was obtained at preexposure SOAs, both within and between languages. The time course of the Stroop effects did not depend much on whether the distractor language was Dutch or English. Error rates were higher for the incongruent trials than the other ones.

The statistical analysis of the errors yielded effects of Stroop condition, $F(2, 40) = 18.17$, $p < .001$; SOA, $F(6, 120) = 2.27$, $p < .04$; and status, $F(1, 20) = 7.81$, $p < .011$. Stroop condition and SOA interacted, $F(12, 240) = 2.23$, $p < .011$, which was independent of status, $F(12, 240) = 1.38$, $p = .18$. Most errors
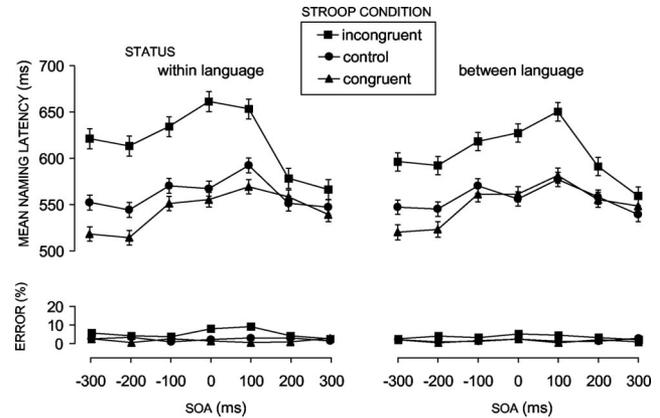

Figure 4. Mean Dutch naming latencies and error percentages per stimulus onset asynchrony (SOA) and Stroop condition for the Dutch distractor words (within language) and the English distractor words (between language) in Experiment 2. The error bars indicate one standard error.

were made in the slowest condition (i.e., on incongruent trials), suggesting that there was no speed–accuracy trade-off in the data.

The statistical analysis of the latencies yielded effects of Stroop condition, $F(2, 40) = 73.21$, $p < .001$, and SOA, $F(6, 120) = 6.15$, $p < .001$, and a marginal effect of status, $F(1, 20) = 4.00$, $p < .06$. There was an interaction of Stroop condition and SOA, $F(12, 240) = 7.90$, $p < .001$, which was independent of status, $F(12, 240) = 1.05$, $p = .41$. Responding was overall faster on congruent than control trials, $F(1, 20) = 12.00$, $p < .002$. The facilitation varied with SOA, $F(6, 120) = 3.91$, $p < .001$, independent of status, $F(6, 120) < 1$, $p = .44$. Planned comparisons revealed that the facilitation was significant at the SOAs of –300 ms and –200 ms ($ps < .001$), but not at the other SOAs. Responding was overall slower on incongruent than control trials, $F(1, 20) = 69.94$, $p < .001$. The interference varied with SOA, $F(6, 120) = 7.35$, $p < .001$, independent of status, $F(6, 120) < 1$, $p = .61$. Planned comparisons revealed that the interference was significant at all SOAs ($ps < .001$), except SOA = 300 ms. To summarize, Stroop interference and facilitation occurred both within and between languages.

Figure 5 gives the distributional plots for the congruent and control trials for the preexposure SOAs, as in Experiment 1. In the mean RTs, facilitation was obtained at the SOAs of –300 ms and –200 ms. The figure shows that the facilitation effect existed throughout almost the entire RT range both within and between distractor languages. Thus, the RT distributions indicate that facilitation does not result from the occasional odd trial and that it is not present for the relatively fast responses only. Thus, the RT distributions also do not support the hypothesis that Stroop facilitation arises from inadvertent reading.

To conclude, Stroop facilitation and interference effects were obtained not only within language, but also between languages, as in Experiment 1. The pattern of results obtained in Experiment 1 was replicated with language manipulated randomly across trials in order to prevent participants from developing response strategies. Facilitation was present throughout the entire RT distribution, both within language and between languages. Because reading distractor words would not yield facilitation but errors in the
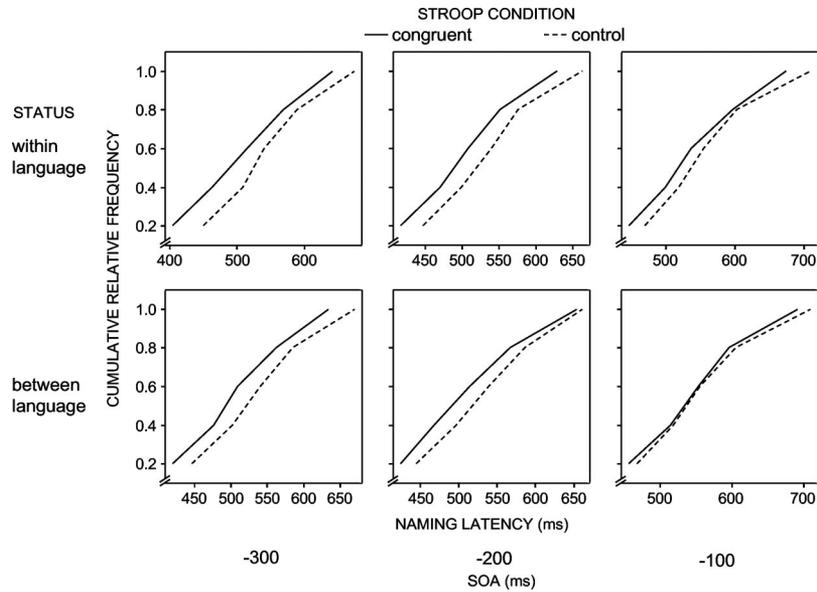
Figure 5. Vincentized cumulative distribution curves for the response latencies in the congruent and control conditions within and between languages in Experiment 2. The distributions are shown for the preexposure stimulus onset asynchronies (SOAs) of –300, –200, and –100 ms.

between-language condition, the facilitation in both the within-language and between-language conditions supports the converging information hypothesis of Stroop facilitation.

Experiment 3

Because of greater expertise in first- than second-language use, it is possible that inadvertent reading is more likely to occur in second-language than in first-language color naming and that the inadvertent reading hypothesis holds for second-language naming. If so, between-language interference on congruent trials should be obtained in second-language naming. This is because reading errors need to be repaired covertly on between-language congruent trials, which delays responding. In contrast, according to the converging information hypothesis, it should not matter whether colors are named in the first or second language for facilitation to occur. The third experiment tested between the converging information and the inadvertent reading hypotheses in second-language color naming. Dutch–English bilingual participants named color patches in English while trying to ignore Dutch or English color words. As in Experiment 2, the distractor language varied randomly from trial to trial.

Method

The method was the same as in Experiment 2, except that the color patches were now named in English. The experiment was carried out with 21 new paid participants from the same population as in the previous experiments.

Results and Discussion

Figure 6 gives the mean English naming latencies. Responding was slowest on incongruent trials and fastest on congruent trials.

Facilitation was obtained both within and between languages. Facilitation was especially prominent at preexposure SOAs. The time course of the Stroop effects did not depend much on whether the distractor language was Dutch or English. Error rates were highest for the incongruent trials.

The statistical analysis of the errors yielded an effect of Stroop condition, $F(2, 40) = 9.11$, $p < .001$; a marginal effect of SOA, $F(6, 120) = 1.86$, $p = .09$; and no effect of status, $F(1, 20) < 1$, $p > .79$. There was an interaction of Stroop condition and SOA, $F(12, 240) = 2.15$, $p < .02$, which varied with status, $F(12, 240) = 2.57$, $p = .003$. Most errors were made in the slowest (i.e.,
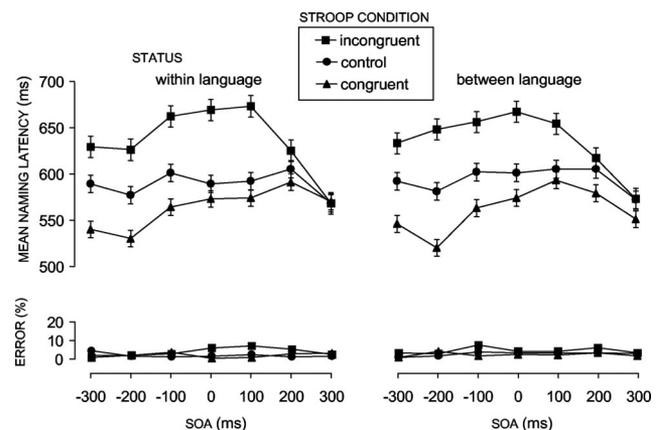

Figure 6. Mean English naming latencies and error percentages per stimulus onset asynchrony (SOA) and Stroop condition for the English distractor words (within language) and the Dutch distractor words (between language) in Experiment 3. The error bars indicate one standard error.

incongruent) condition, which suggests that there was no speed–accuracy trade-off in the data.

The statistical analysis of the latencies yielded effects of Stroop condition, $F(2, 40) = 104.94$, $p < .001$, and SOA, $F(6, 120) = 6.42$, $p < .001$, but not of status, $F(1, 20) < 1$, $p > .54$. There was an interaction of Stroop condition and SOA, $F(12, 240) = 9.73$, $p < .001$, which was independent of status, $F(12, 240) = 1.38$, $p = .18$. Responding was overall faster on congruent than control trials, $F(1, 20) = 97.82$, $p < .001$. The facilitation varied with SOA, $F(6, 120) = 5.36$, $p < .001$, independent of status, $F(6, 120) < 1$, $p = .55$. Planned comparisons revealed that the facilitation was significant at all SOAs ($ps < .001$), except SOA = 300 ms. Responding was overall slower on incongruent than control trials, $F(1, 20) = 77.54$, $p < .001$. The interference varied with SOA, $F(6, 120) = 8.60$, $p < .001$, independent of status, $F(6, 120) = 1.07$, $p = .38$. Planned comparisons revealed that the interference was significant at all SOAs ($ps < .001$), except at the SOAs of 200 ms and 300 ms. To summarize, Stroop interference and facilitation occurred both within and between languages.

Figure 7 gives the distributional plots for the congruent and control trials for the preexposure SOAs, as in the previous experiments. In the mean RTs, facilitation was obtained at all these SOAs. The figure shows that the facilitation effect existed throughout the entire RT range both within and between distractor languages. Thus, the RT distributions indicate that facilitation does not result from the occasional odd trial and that it is not only present in the leading edge of the distribution. To conclude, the RT distributions also do not support the hypothesis that Stroop facilitation arises from inadvertent reading.

To assess whether the magnitude of Stroop effects differed between and within languages depending on the naming language (Dutch or English), the latencies of Experiments 2 and 3 were submitted to a combined analysis of variance with naming language as between-participant variable. The statistical analysis of the latencies yielded no effect of naming language, $F(1, 40) < 1$, $p > .39$, nor were there interactions with status, $F(1, 40) = 1.01$, $p > .32$; SOA, $F(6, 240) < 1$, $p > .68$; and Stroop condition, $F(2, 80) = 2.89$, $p = .06$. There were also no higher order interactions. This suggests that the between- and within-language effects and their time course do not depend on the naming language, Dutch or English.

To conclude, Stroop facilitation and interference effects were obtained not only within language but also between languages, as in the previous experiments. The effects were obtained even though the color patches were named in the second language, English. Facilitation was present throughout the entire RT distribution, both within language and between languages. The facilitation in both the within-language and between-language conditions supports the converging information hypothesis of Stroop facilitation.

Experiment 4

With three color patches, three color words, and a series of *X*s, there are six possible color–word combinations in the incongruent condition but only three combinations in the congruent condition and three in the control condition. In order to have an equal number of stimuli in each of the Stroop conditions in Experiments 1–3, incongruent trials were constructed by repeatedly pairing one color word with one color patch (e.g., RED–blue, GREEN–red, BLUE–green). Thus, each word was combined with two color patches, one incongruent and one congruent patch. It is possible that this may have contributed to a specific strategy on the part of the participants. A participant saw each pairing 28 times, and this may have been frequent enough to realize that when the word RED appeared (for example), the color patch had to be red or blue.
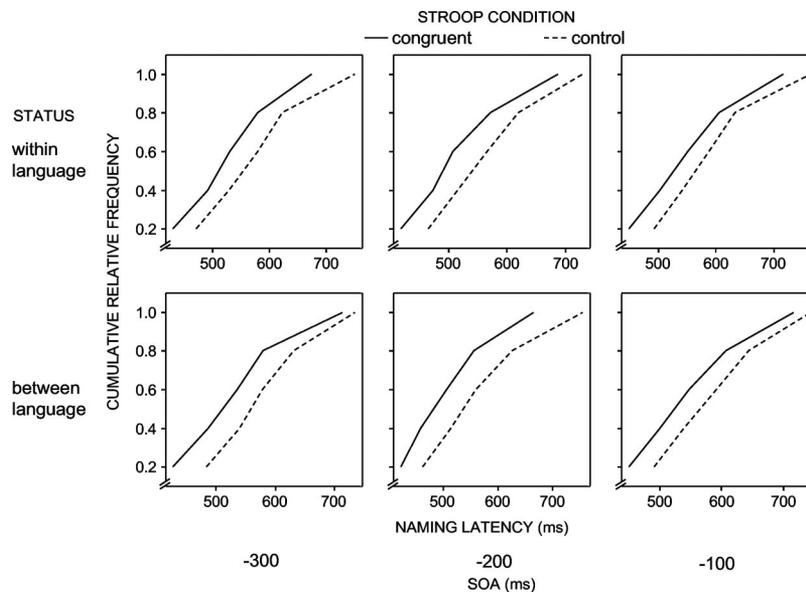

Figure 7. Vincentized cumulative distribution curves for the response latencies in the congruent and control conditions within and between languages in Experiment 3. The distributions are shown for the preexposure stimulus onset asynchronies (SOAs) of –300, –200, and –100 ms.

When a word appeared, participants had only two possible responses, whereas when a control stimulus appeared, all three colors were possible. Consequently, facilitation may reflect a slow control condition rather than a speed-up on congruent trials because of converging information.

The aim of the fourth experiment was to examine this possibility by completely crossing the items. Thus, each color word was combined with all three colors, making up two incongruent stimuli and one congruent stimulus. Similarly, the control series of Xs was combined with all three colors. Consequently, both when a word and a series of Xs appeared, participants had three possible responses. If facilitation is still obtained, it must reflect a speed-up on congruent trials because of convergent information rather than a slow control condition. Because items were completely crossed, there was a greater number of incongruent than congruent trials. Stroop effects tend to become smaller with an increasing number of incongruent trials (e.g., Kane & Engle, 2003). Thus, if between-language facilitation persists in the present experiment (with more incongruent than congruent trials), this would strongly support the converging information hypothesis.

Method

The method was the same as in Experiments 2 and 3, except that the color words and color patches were now completely crossed. The participants named color patches in English while trying to ignore Dutch or English distractor words. The latencies and errors were submitted to analyses of variance with the crossed within-participant variables Stroop condition, SOA, and status. The experiment was carried out with 14 new paid participants from the same population as in the previous experiments.

Results and Discussion

Figure 8 gives the mean English naming latencies. Responding was slowest on incongruent trials and fastest on congruent trials. Facilitation was obtained both within and between languages,
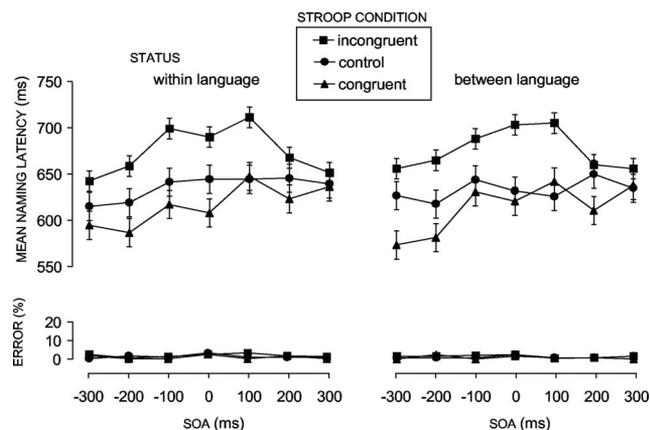

Figure 8. Mean English naming latencies and error percentages per stimulus onset asynchrony (SOA) and Stroop condition for the English distractor words (within language) and the Dutch distractor words (between language) in Experiment 4. The error bars indicate one standard error.

especially at preexposure SOAs. The time course of the Stroop effects did not depend much on whether the distractor language was Dutch or English. Error rates were highest for the incongruent trials.

The statistical analysis of the errors yielded an effect of Stroop condition, $F(2, 26) = 3.73$, $p < .04$; a marginal effect of SOA, $F(6, 78) = 1.93$, $p = .09$; and no effect of status, $F(1, 13) < 1$, $p > .76$. There was no interaction of Stroop condition and SOA, $F(12, 156) < 1$, $p > .81$, independent of status, $F(12, 156) = 1.17$, $p = .31$. Most errors were made in the slowest (i.e., incongruent) condition, which suggests that there was no speed–accuracy trade-off in the data.

The statistical analysis of the latencies yielded effects of Stroop condition, $F(2, 26) = 70.86$, $p < .001$; a marginal effect of SOA, $F(6, 78) = 2.01$, $p = .08$; and no effect of status, $F(1, 13) < 1$, $p > .94$. There was an interaction of Stroop condition and SOA, $F(12, 156) = 3.77$, $p < .001$, which was independent of status, $F(12, 156) < 1$, $p = .87$. Responding was overall faster on congruent than control trials, $F(1, 13) = 23.53$, $p < .001$. The facilitation varied with SOA, $F(6, 78) = 3.41$, $p < .005$, independent of status, $F(6, 78) < 1$, $p = .60$. Planned comparisons revealed that the facilitation was significant at the preexposure SOAs of –300 ms and –200 ms, and somewhat surprisingly, at SOA = 200 ms (ps < .001). Responding was overall slower on incongruent than control trials, $F(1, 13) = 45.28$, $p < .001$. The interference varied with SOA, $F(6, 78) = 3.69$, $p < .003$, independent of status, $F(6, 78) < 1$, $p = .89$. Planned comparisons revealed that the interference was significant at the SOAs ranging from –200 ms preexposure to 100 ms postexposure (ps < .001). To summarize, Stroop interference and facilitation occurred both within and between languages.

Figure 9 gives the distributional plots for the congruent and control trials for the preexposure SOAs, as in the previous experiments. In the mean RTs, facilitation was obtained at the preexposure SOAs of –300 ms and –200 ms. The figure shows that the facilitation effect existed throughout almost the entire RT range both within and between distractor languages. Thus, the RT distributions indicate that facilitation does not result from the occasional odd trial and that facilitation is not present for the relatively fast responses only. To conclude, the RT distributions also do not support the hypothesis that Stroop facilitation arises from inadvertent reading.

In summary, between-language facilitation was obtained even though the items were completely crossed (making the color unpredictable from the word) and there was a larger number of incongruent than congruent trials. This finding strongly supports the convergent information hypothesis.

General Discussion

Whereas researchers generally agree that Stroop interference reflects the processing effect of diverging color and word information, they have found no agreement on the source of Stroop facilitation. According to the converging information hypothesis (e.g., Cohen et al., 1990; Melara & Algom, 2003; Roelofs, 2003), facilitation arises from converging information on the word and color dimensions on congruent trials in the Stroop task. In contrast, according to the inadvertent reading hypothesis (Kane & Engle, 2003; MacLeod & MacDonald, 2000), facilitation in the congruent condition occurs when participants inadvertently read the color
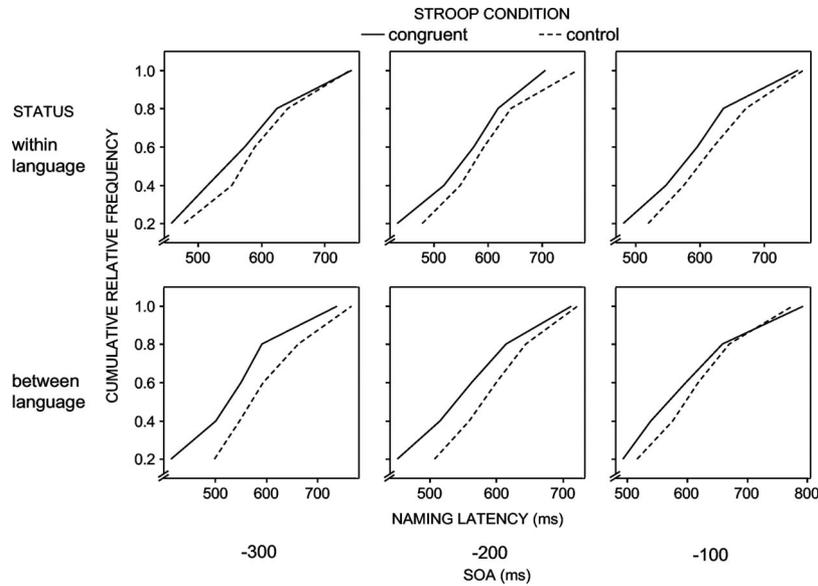
Figure 9.   Vincentized cumulative distribution curves for the response latencies in the congruent and control conditions within and between languages in Experiment 4. The distributions are shown for the preexposure stimulus onset asynchronies (SOAs) of –300, –200, and –100 ms.

word instead of performing the color naming task. These accidental reading errors are impossible to detect in a within-language Stroop task and may be included in the calculation of the color naming latency. If reading errors are the cause of Stroop facilitation, the effect should disappear when the errors can be prevented or removed. Reading errors can be filtered in a between-language version of the Stroop task. Because reading cross-language distractor words rather than naming color patches would not produce a facilitation effect in the between-language condition, the appearance of a Stroop facilitation effect in both the within-language and between-language conditions would provide evidence for the converging information hypothesis of Stroop facilitation and would suggest that Stroop interference and facilitation effects have a common locus.

Previous studies reported between-language interference in the congruent Stroop condition (Abunuwara, 1992; Dalrymple-Alford, 1968; MacLeod & MacDonald, 2000), supporting the inadvertent reading hypothesis. However, the phonological forms of the color words differed between languages in these studies. The diverging phonological information may have yielded the interference in the congruent condition. In line with this interpretation, between-language facilitation has been observed for congruent words compared with color-unrelated words rather than nonverbal items as control (Costa et al., 2008). The diverging-phonology account predicts between-language Stroop facilitation rather than interference for languages whose color words have highly similar phonological forms, such as Dutch and English, even when a series of *X*s is used as control. This prediction was tested in four experiments.

In the experiments, color patches were named in Dutch (Experiments 1 and 2) or English (Experiment 3 and 4). The distractor words were blocked by language (Experiment 1) or languages were randomly mixed (Experiments 2–4). Moreover, given that earlier studies suggest that facilitation effects typically occur with

preexposure of the word, the Stroop effects were assessed at a wide range of preexposure and postexposure SOAs. This also allowed for a more thorough comparison of within- and between-language Stroop effects than in earlier studies, which used zero SOA only.

In all four experiments, Stroop facilitation and interference effects were obtained both within and between languages. Facilitation was especially prominent at preexposure SOAs. The time course of the effects was similar within and between languages. Facilitation was generally present throughout the RT distributions. These results suggest that Stroop interference and facilitation have a common locus within and between languages, supporting the converging information hypothesis and challenging the inadvertent reading hypothesis of Stroop facilitation.

Facilitation was present at preexposure SOAs in all four experiments, as is typically observed in the Stroop task (e.g., M. O. Glaser & Glaser, 1982). If only zero SOA had been tested in the present experiments, facilitation would have been missed in Experiments 2 and 4, incorrectly suggesting that cross-language facilitation is not a robust phenomenon. However, facilitation was consistently obtained at preexposure SOAs in all four experiments. The dependence of facilitation on SOA observed in the present experiments stresses the importance of testing for effects using a range of SOAs rather than zero SOA only. Moreover, assessing within- and between-language Stroop effects at several preexposure and postexposure SOAs allows for a more extensive comparison of within- and between-language effects than assessing performance at zero SOA only. The present results suggest that the time course of Stroop interference and facilitation effects is very similar within and between languages, suggesting a common locus of the effects.

Experiments 1 and 2 yielded between-language facilitation in first-language color naming. It is possible that inadvertent reading occurs in second-language color naming even if it does not occur

in first-language color naming because of the greater expertise in first- than second-language use. If so, between-language interference on congruent trials should be obtained in second-language naming. In contrast, according to the converging information hypothesis, it should not matter whether colors are named in the first or second language for facilitation to occur. In Experiments 3 and 4, Dutch–English bilingual participants named color patches in their second language, English, while trying to ignore Dutch or English color words. Again, between-language facilitation was obtained. Moreover, a combined analysis of the data of Experiments 2 (first-language naming) and Experiment 3 (second-language naming) revealed that there were no differences in Stroop effects and their time course depending on the naming language, first or second. These results suggest that converging information is the source of Stroop facilitation regardless of the naming language.

The converging information hypothesis was further supported by analyses of the RT distributions. Under the converging information hypothesis, facilitation reflects a constant speed-up of processing on most of the congruent trials compared with control trials (cf. Roelofs, 2008). Consequently, the distribution curve for RTs on congruent trials should be just that for control trials, except that the former is shifted toward shorter RTs by a constant. In contrast, according to the inadvertent reading hypothesis, the facilitation stems from a large speed-up that occurs only on a small proportion of the congruent trials. Consequently, the effect on the mean RT might be the same, but the effect on the RT distributions should be different. The distribution curves of RTs in the congruent and control conditions should then differ among the short RTs but converge at the longer RTs. The analyses of the RT distributions showed that the facilitation effect was present throughout the distributions, supporting the converging information hypothesis.

The results regards facilitation of the present experiments agree with the findings on the effect of phonologically related noncolor words in Stroop color naming reported by Roelofs (2003). Dutch participants were asked to name color patches while trying to ignore distractor words that were related in form to color words. In the incongruent condition, the distractor word was phonologically related to the name of another color. For example, the word GROEP [group] was presented in red ink. The Dutch word GROEP is phonologically related to the color term *groen* [green]. In the congruent condition, the distractor word was phonologically related to the name of the color patch. For example, the word GROEP [group] was presented in green ink. Relative to neutral noncolor words, the incongruent distractor yielded interference and the congruent distractors yielded facilitation. In saying "*groen*" to the word GROEP in green ink, it is clear that the participants named the color, because reading the word aloud would have yielded a wrong response. Thus, facilitation must have arisen from converging phonological information. If phonologically related distractor words yield facilitation, it is plausible to assume that color words themselves also yield facilitation in the congruent condition in a standard Stroop experiment. The present Experiments 1–4 provide evidence that color words indeed yield such facilitation.

Stroop interference is often smaller between than within languages (e.g., Francis, 1999). This was not observed in the present experiments. The Stroop literature suggests that the relative magnitude of between- and within-language Stroop interference de-

pends on the phonological similarity of the color words in the two languages (see MacLeod, 1991, for a review). Color words of Dutch and English are phonologically very similar, which explains why Stroop interference did not really differ between the within-language and between-language conditions. Similarly, the relative magnitude of between- and within-language Stroop facilitation should depend on the phonological similarity of the color words. This explains why Stroop facilitation was obtained between Dutch and English in the present experiments, but not in previous experiments using French, English, Hebrew, and Arabic (Abunuwara, 1992; Dalrymple-Alford, 1968; MacLeod & MacDonald, 2000).

To conclude, Stroop facilitation and interference effects were obtained both within and between languages. Facilitation was generally present throughout the entire RT distributions. These results suggest that Stroop interference and facilitation have a common locus within and between languages. The findings support the converging information hypothesis and challenge the inadvertent reading hypothesis of Stroop facilitation.

References

Abunuwara, E. (1992). The structure of the trilingual lexicon. *European Journal of Cognitive Psychology, 4,* 311–322.

Booij, G. E. (1995). *The phonology of Dutch.* Oxford, England: Clarendon Press.

Cohen, J., Dunbar, K., & McClelland, J. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97,* 332–361.

Costa, A., Albareda, B., & Santesteban, M. (2008). Assessing the presence of lexical competition across languages: Evidence from the Stroop task. *Bilingualism: Language and Cognition, 11,* 121–131.

Dalrymple-Alford, E. C. (1968). Interlingual interference in a color-naming task. *Psychonomic Science, 10,* 215–216.

Dijkstra, A. F. J., & Van Heuven, W. J. B. (2002). The architecture of the bilingual word recognition system: From identification to decision. *Bilingualism: Language and Cognition, 5,* 175–197.

Dyer, F. N. (1971). The duration of word meaning responses: Stroop interference for different preexposures of the word. *Psychonomic Science, 25,* 229–231.

Francis, W. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin, 125,* 193–222.

Glaser, M. O., & Glaser, W. R. (1982). Time course analysis of the Stroop phenomenon. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 875–894.

Glaser, W. R., & Glaser, M. O. (1989). Context effects in Stroop-like word and picture processing. *Journal of Experimental Psychology: General, 118,* 13–42.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9,* 637–671.

Kane, M. J., & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General, 132,* 47–70.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1–38.

Luce, R. D. (1986). *Response times: Their role in inferring elementary mental organization.* New York, NY: Oxford University Press.

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 atten-

tional 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.

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin, 86,* 446–461.

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. (2008). Dynamics of the attentional control of word retrieval: Analyses of response time distributions. *Journal of Experimental Psychology: General, 137,* 303–323.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643–662.