The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t716100704

Increased Stroop interference with better second-language reading skill

Wouter Braet[a]; Nele Noppe[b]; Johan Wagemans[a]; Hans Op de Beeck[ac]
[a] Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium [b] Department of Japanese Studies, University of Leuven, Leuven, Belgium [c] Laboratory of Biological Psychology, University of Leuven, Leuven, Belgium

First published on: 04 October 2010

To cite this Article Braet, Wouter , Noppe, Nele , Wagemans, Johan and Op de Beeck, Hans(2011) 'Increased Stroop interference with better second-language reading skill', The Quarterly Journal of Experimental Psychology, 64: 3, 596 — 607, First published on: 04 October 2010 (iFirst)

To link to this Article: DOI: 10.1080/17470218.2010.513735

URL: http://dx.doi.org/10.1080/17470218.2010.513735

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2011, 64 (3), 596–607

Psychology Press
Taylor & Francis Group

Increased Stroop interference with better second-language reading skill

Wouter Braet
Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

Nele Noppe
Department of Japanese Studies, University of Leuven, Leuven, Belgium

Johan Wagemans
Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium

Hans Op de Beeck
Laboratory of Experimental Psychology, University of Leuven, Leuven, Belgium;
Laboratory of Biological Psychology, University of Leuven, Leuven, Belgium

Skilled readers demonstrate remarkable efficiency in processing written words, unlike beginning readers for whom reading occurs more serially and places higher demands on visual attention. In the present study, we used the Stroop paradigm to investigate the relationship between reading skill and automaticity, in individuals learning a second language with a different orthographic system. Prior studies using this paradigm have presented a mixed picture, finding a positive, a negative, or no relationship between the size of Stroop interference and reading skills. Our results show that Stroop interference in the second language was positively related to reading skill (when controlled for interference in the first language). Furthermore, interference was positively related to objective but not subjective indices of the amount of exposure to the second language. We suggest that the lack of consistency in the results of earlier studies may be due, at least in part, to these studies looking at Stroop interference in isolation, rather than comparing interference between languages.

Keywords: Visual word recognition; Stroop effect; Automaticity; Second-language acquisition.

The human visual system, in skilled readers, demonstrates remarkable efficiency when reading written words. For example, word length effects are minimal for high-frequency words of up to six letters (Ferrand & New, 2003), and both orthographic priming and semantic priming have been observed with presentation times of only tens of milliseconds (Evett & Humphreys, 1981;

Correspondence should be addressed to Wouter Braet, Laboratory of Experimental Psychology, University of Leuven, Tiensestraat 102, B 3000 Leuven, Belgium. E-mail: wouter.braet@gmail.com

The authors wish to extend their thanks to the following persons: Frank Amand for technical support, Stephanie Poot for administrative support, and Naoko Sakurai and Stefanie Nijs for their invaluable help with participant recruitment. W.B. is funded by the Research Foundation–Flanders (FWO-Vlaanderen); N.N. is funded by the Faculty of Arts (University of Leuven); J.W. is funded by METH/08/02 and IMPH/06/GHW; H.O.D.B. is funded by CREA/07/004 and the Human Frontier Science Program Organization.

http://www.psypress.com/qjep
DOI:10.1080/17470218.2010.513735

H. W. Lee, Rayner, & Pollatsek, 1999). Reading is assumed to rely largely on automatic processes (e.g., Brown, Gore, & Carr, 2002a; Rawson & Middleton, 2009), which under normal conditions place little or no demands on visual attention (L. Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008; though also see Risko, Stolz, & Besner, 2005). It has been proposed that this efficiency is largely due to experience with written words, through which neural representations are acquired for increasingly complex word features (Dehaene, Cohen, Sigman, & Vinckier, 2005; Vinckier et al., 2007) or even for individual words (Glezer, Jiang, & Riesenhuber, 2009).

Furthermore, dual-task studies that show gradual increases in automaticity in reading throughout the lifespan (e.g., Lien et al., 2006), as well as a relationship between the degree of automaticity and reading proficiency in young adults (Ruthruff, Allen, Lien, & Grabbe, 2008), provide corroborative evidence for the role of experience in skilled reading.

While these studies suggest a positive relationship between reading skills and automaticity, a number of studies using the Stroop paradigm (Stroop, 1935) appear to suggest the opposite pattern. In this paradigm, participants see colour words printed in a different colour as the semantic meaning of the word (e.g., "red" printed in blue) and are asked to report on the physical colour of the words (while ignoring their meaning). Performance on these stimuli is then compared to a baseline of colour naming where there is no mismatch between the physical colour and semantic meaning of the stimuli. Stroop interference—that is, the difference in response times between these two types of stimuli—is typically thought to reflect the degree of automatic processing in reading (Brown, Joneleit, Robinson, & Brown, 2002b; Logan, 1997; also see MacLeod & Dunbar, 1988). The most straightforward prediction would, therefore, be that there should be a positive relationship between reading skills and Stroop interference, with greater interference for skilled readers. Several lines of evidence suggest otherwise, however. First, comparisons between different age-groups suggest a ∼-shaped function[1] between age and Stroop interference: Stroop interference initially increases as children learn to read (Peru, Faccioli, & Tassinari, 2006; Schadler & Thissen, 1981; Schiller, 1966). This is followed by a gradual decline of Stroop interference into adulthood, up to the age of 60, after which interference increases again (Comalli, Wapner, & Werner, 1962; Roelofs & Hagoort, 2002). Secondly, comparisons between groups with reading disorders and controls have consistently shown less Stroop interference in the control group, despite their greater reading skills (Everatt, Warner, Miles, & Thomson, 1997; Faccioli, Peru, Rubini, & Tassinari, 2008; Hicks & Jackson, 2005; Kapoula et al., 2010; Protopapas, Archonti, & Skaloumbakas, 2007). Thirdly, negative relationships between Stroop interference and several indices of reading proficiency have also been reported in the general population, in the absence of reading disorders (Protopapas et al., 2007). Finally, Protopapas and colleagues confirmed their finding of a negative relationship between reading skills and Stroop interference using simulations based on two computational models designed for the Stroop task (J. D. Cohen, Dunbar, & McClelland, 1990; J. D. Cohen, Usher, & McClelland, 1998; Roelofs, 2003).

These findings are based on comparisons between groups or individuals, making it difficult to distinguish between the effects of reading proficiency and those of other individual differences. However, several studies using related paradigms (such as naming pictures that have congruent or incongruent words superimposed on them, or naming colours of noncolour words) have

[1] This may reflect a ∩-shaped relationship between automaticity (and/or executive functions) and interference (given that reading skills tend to decline in old age). Such a relationship can be explained based on models of the dynamics of temporal overlap between two processes (colour naming and word reading; e.g., Zmigrod & Hommel, 2010). Interference is assumed to be maximal when both processes occur in close temporal proximity (i.e., at intermediate proficiency) and less at either high or low proficiency levels, as word processing occurs either too early or too late to interfere with colour naming.

demonstrated that low-frequency words induce more interference than do high-frequency words (the same pattern emerges for nonwords compared to words), further suggesting a negative relationship between interference and the level of reading automaticity (Burt, 2002; Dhooge & Hartsuiker, 2010; Miozzo & Caramazza, 2003; Monsell, Taylor, & Murphy, 2001).

In sum, all these results demonstrate a negative relationship between reading automaticity and Stroop interference in a variety of different contexts, contrary to what would be predicted from training studies that found increasing interference when individuals were trained to associate shapes with colours (MacLeod & Dunbar, 1988; also see Cattell, 1886; Stroop, 1935).

A negative relationship between reading skills and Stroop interference can be explained by making two assumptions. First, reading automaticity should quickly become sufficient to ensure obligatory processing of the word form, even for novice readers. Otherwise, Stroop interference would primarily reflect the proportion of words that are read, probably resulting in a positive correlation with reading proficiency (this mechanism may explain the positive relationship that is typically observed in young children learning to read). Secondly, colour naming should not be able to proceed before processing of the word form has been completed (as the response to the word form has to be suppressed at the output stage, prior to selection of the response to the colour—this process is believed to reflect bottlenecks in response selection; see Protopapas et al., 2007), or both processes should use a shared pool of attentional resources (which are prioritized for word reading, compared to colour naming; see, e.g., Burt, 2002). Under these assumptions, proficient readers will finish processing the word form more quickly than poor readers, leading to less interference in the Stroop task.

When considered in the context of individuals who are learning to read, a negative relationship between reading skill and Stroop interference suggests that there is maximal interference early in the learning process, after which interference gradually decreases. Furthermore, individuals learning a second language (L2) should show greater interference in L2 than in L1 (their first language), once reading is sufficiently obligatory (and this difference between L1 and L2 interference should decrease with increasing L2 proficiency).

However, two studies comparing Japanese–English bilinguals observed the opposite pattern (also see Mohamed Zied et al., 2004, for similar findings with French/Arabic bilinguals): Stroop interference was greater when the words were presented in participants' native language for both English (Sumiya & Healy, 2008) and Japanese speakers (Sumiya & Healy, 2004), though there was no reliable relationship between the size of the interference effect and self-report measures of reading skill.[2] Other studies found no differences in Stroop interference between the dominant and nondominant language (e.g., T. M. Lee & Chan, 2000).

In the present paper, we reexamine the relationship between reading skills and Stroop interference (as an index of automatic processing of the word form), in the context of language learning, by comparing Stroop interference in Dutch (L1) and Japanese (L2) for native-Dutch individuals learning Japanese. We specifically focused on the relationship between training-induced reading skill in L2 and the difference in Stroop interference between L1 (where automaticity can be assumed to be near ceiling) and L2. This comparison allows us to control for individual differences that may affect performance on the Stroop task in addition to automaticity-of-reading, such as response strategies, speed-of-processing, response inhibition (which is conceptually similar to the

[2] Both papers also report within-language interference scores, though no analyses on these scores are reported. These results suggest equivalent Stroop interference for responding to L1 stimuli in L1, compared to responding to L2 stimuli in L2, in Sumiya and Healy (2004); conversely, Stroop interference seems larger for L1 than for L2 in Sumiya and Healy (2008). It is interesting to note that the authors further report higher L2 proficiency in their 2004 study, whereas no difference is apparent between L1 and L2 interference scores.

notion of cognitive control, see, e.g., Verguts & Notebaert, 2009), or phonological and naming skills. For example, response inhibition/cognitive control can vary between individuals (if the stimulus material is kept constant) or between tasks/based on stimulus properties (e.g., dependent on the proportion of congruent trials in a block of the Stroop task, see Tzelgov, Henik, & Berger, 1992). In the context of second-language acquisition, there is greater cognitive control for L1 than for L2, as individuals are better able to adapt their level of Stroop interference, based on the proportion of congruent trials, in their first language (Tzelgov, Henik, & Leiser, 1990). As we do not manipulate the proportion of congruent trials in our design, we are unable to directly assess the relative contribution of automaticity and control, or how both may evolve in relationship to L2-skill levels. Nevertheless, the observation that Stroop interference is typically greater in L1 than in L2 suggests that the effects of language proficiency on cognitive control are secondary to its effects on automaticity (see Tzelgov & Kadosh, 2009; Tzelgov et al., 1990). The relative contribution of individual differences in response inhibition/cognitive control on the Stroop effect (compared to differences in reading automaticity) is probably significant, as evident from the large number of clinical studies that use the Stroop task primarily to assess response inhibition. Thus, a *relative* interference score is less likely to reflect differences in response inhibition/cognitive control, as both individual (each participant was tested in both languages) and task (the proportion of congruent trials was the same in both blocks) parameters are kept constant, but is instead assumed to indicate the relative difference in automaticity between the two languages.

Different predictions emerge based on whether Stroop interference is driven primarily by the level of reading automaticity (MacLeod, 1991; MacLeod & Dunbar, 1988; Stroop, 1935), or by competition in attentional resources between colour naming and reading (Burt, 2002; Protopapas et al., 2007). Under the first account, we predict that, when individual differences are controlled for, Stroop interference will show a positive, rather than a negative, relationship with reading skill and/or experience. In addition to this, we expect interference effects to be larger in participants' native language. Alternatively, if Stroop interference is driven primarily by both processes (colour naming and word reading) relying on shared resources, the opposite pattern should occur: Interference should be greater in L2 than in L1, and there should be a negative relationship between (relative) Stroop interference and (second-language) reading ability.

Method

Participants

Twenty-nine participants (of whom 10 were female) in the age range 18–38 years (mean = 22.6, $SD = 4.4$) were involved in this study and were paid for their participation. At the time of the study all but 2 were enrolled as full-time undergraduate or postgraduate students in Japanese Studies (the remaining 2 had previously studied Japanese, though neither was currently enrolled in the Japanese Studies programme; 1 speaks Japanese at home). All participants were native speakers of Dutch, had normal or corrected-to-normal vision, and reported no reading disorders. This study was approved by the local ethics committee.

The curriculum of the Japanese Studies programme includes intensive language study, with an estimated 338 contact hours per academic year in the first year and 286 contact hours/year in the second, evenly divided between writing, reading, and conversation lessons. The number of contact hours drops in the later two years of the programme (200/year in the third and 150/year in the fourth year), at which point students are expected to conduct their language study in a much more independent fashion. Many students spend approximately a year in Japan between the third and fourth years. As the study was conducted near the end of the academic year, participants would have had a minimum of approximately 300 contact hours (for first-year students) at the time of testing.

Procedure

Participants performed a modified version of the Stroop task (Stroop, 1935). Colour words were presented on a computer monitor in either the corresponding (congruent trials) or a different colour (incongruent trials), and participants were asked to report the colour in which the words were shown, irrespective of their meaning. Participants responded using a button-box (this response was chosen over a verbal response, in order to avoid confounds[3] from the response language, i.e., differences related to naming between L1 and L2; manual responses typically lead to an underestimation of the Stroop effect, when compared to naming; see MacLeod, 1991; Repovš, 2004). The mapping of colours to the four buttons was kept the same for all participants, who received a short (40 trials) training block in which on every trial, a circle was displayed in one of the four colours, and participants were asked to press the corresponding button as fast and accurately as possible.

Participants performed two blocks of the Stroop task, one in Dutch and one in Japanese, the order of which was counterbalanced. Each block consisted of 240 trials, of which half[4] were congruent (e.g., "green" printed in green), and half were incongruent (e.g., "green" printed in red). The following colour words were used: groen, geel, rood, bruin (Dutch block), or 緑, 黄, 赤, 茶 (Japanese block; green, yellow, red, and brown). Participants were asked to report whether they were familiar with all four kanji, and data from 2 participants who were not familiar with all four colour words were excluded from the analyses; the remaining 27 knew all kanji that were used in this study). The timeline of a single trial is shown in Figure 1.

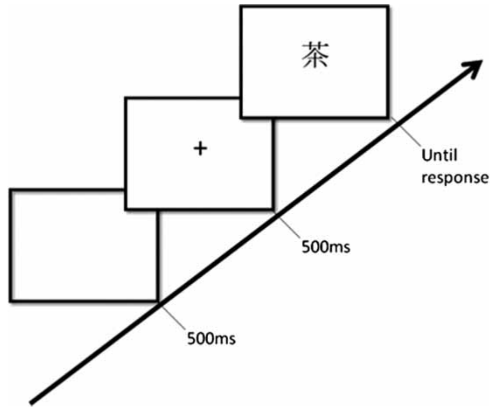

Figure 1. *Timeline of a single trial: A blank screen was displayed for 500 ms, followed by a fixation cross presented for 500 ms. The stimulus word was subsequently presented until the participant made a response. Stimuli were presented against a grey background.*

The choice of presenting Japanese words in kanji, rather than in either of the hiragana or katakana syllabaries, was motivated primarily by functional imaging studies that show different processing of these scripts: Kanji reading (like word reading in Dutch) involves direct lexical access and orthographic retrieval through the ventral route, while kana reading is thought to rely on the more indirect dorsal route (Thuy et al., 2004). Furthermore, while the colour words can be written using either kanji and kana, the most common form (and therefore likely the most visually familiar) is in kanji.

In addition to the Stroop task, participants were asked to complete a brief questionnaire, as well as a 30-item kanji test. In the questionnaire, participants were asked to provide information about the number of years they had spent studying Japanese, the length of time they had resided in

[3] For example, studies using verbal responses suggest a complex relationship between within- and between-language interference, which depends on the orthographic similarity between the words in the two languages (see e.g., Fang, Tzeng, & Alva, 1981; though also see Smith & Kirsner, 1982). While it is unclear to what extent this effect still exists when using manual responses, it is unlikely that in our study there is any "carry-over" interference from one language to the other, as there is no visual similarity between the same word in Japanese (in kanji) and in Dutch.

[4] The proportion of congruent/incongruent trials has been related to the size of the Stroop effect in numerous studies (with greater interference as the proportion of congruent trials increases). A recent study (Schmidt & Besner, 2008) suggests that this relationship emerges due to response contingencies and not because the proportion of congruent trials modulates attention to the written word form (i.e., the extent to which reading is obligatory). An equal proportion of congruent and incongruent trials is also likely to lead to more reliable estimates of reaction times/accuracy for the congruent condition.

Japan, and whether they had ever taken part in the Japanese Language Proficiency Test (JLPT) or the Kanji Aptitude Test (standardized tests to measure Japanese proficiency that are administered worldwide; they require an extensive amount of extracurricular kanji study, meaning that a participant might score significantly higher on the kanji test in this study by virtue of having studied for one of the two standardized tests). They were also asked to estimate the average amount of time they spent reading, writing, or listening to Japanese per week inside and outside of lessons[5] (also see Appendix).

Kanji for the test were divided into four levels of difficulty to correspond roughly with the level of ability expected from students in each respective year of study at the Japanese Studies programme at the University of Leuven. Items for the test were chosen by comparing kanji appearing in established learning materials for the four separate years of the programme with kanji appearing in each of the four levels of the JLPT, a standardized test administered by the Japanese Ministry of Education that is used to measure the Japanese ability of non-native speakers. Only kanji appearing both in the learning materials and in the corresponding JLPT level were selected for use in the test. Participants were asked to choose the correct Dutch translation of each kanji from four possible answers. Each participant was presented with 30 kanji chosen from the four levels of difficulty.

Results

Mean[6] reaction times (RTs; of correct trials only) and accuracy scores were analysed separately, as repeated measures analyses of variance (ANOVAs), with language (Dutch or Japanese) and congruency (congruent or incongruent) as factors.

Participants (see Figure 2) were slower to respond on incongruent trials than on congruent trials—main effect for congruency: $F(1, 26) = 52.01$, $p < .001$—and the size of this effect was smaller in Japanese than in Dutch: interaction between language and congruency, $F(1, 26) = 10.43$, $p = .003$. The main effect of language was not significant ($F < 1$, *ns*). In terms of accuracy, participants made more errors on incongruent trials, $F(1, 26) = 10.28$, $p = .004$, but neither the main effect of language nor the interaction between language and congruency was significant (both $F \leq 1.27$, *ns*). The remaining analyses focus on RTs.

There was also a significant correlation between the size of the Stroop effect (the difference between congruent and incongruent trials) in Dutch and Japanese, $r(27) = .54$, $p = .004$, which suggests significant individual differences that affect Stroop interference in the native language as well as in a second language acquired in adulthood. Thus, we normalized individual scores by dividing the effect size in Japanese by the effect size in Dutch: $\Delta\text{Stroop} = \text{Stroop(L2)}/\text{Stroop(L1)}$. A value of 1 indicates equivalent interference in the two languages, while values below 1 indicate greater interference in Dutch than in Japanese. It should be noted that difference scores (including Stroop interference) are considered less reliable than RT scores (Cronbach & Furby, 1970; Strauss, Allen, Jorgensen, & Cramer, 2005) and that this problem can potentially be exacerbated in our ΔStroop measure. We therefore computed reliability measures for both interference scores in Dutch and Japanese, as well as for ΔStroop, by calculating split-half correlations, corrected by Spearman–Brown's formula. Reliability for ΔStroop was indeed lower than that for both interference scores (.59, compared to .71 for both Dutch and Japanese), but sufficiently high to observe significant correlations between this measure and measures of

[5] One participant did not complete either the questionnaire or the kanji test and was excluded from analyses that used these variables. Participants who were not currently enrolled in the Japanese Studies programme were excluded from analyses that used the variable "year of study".

[6] The same pattern of results emerges when using median RTs instead of means.

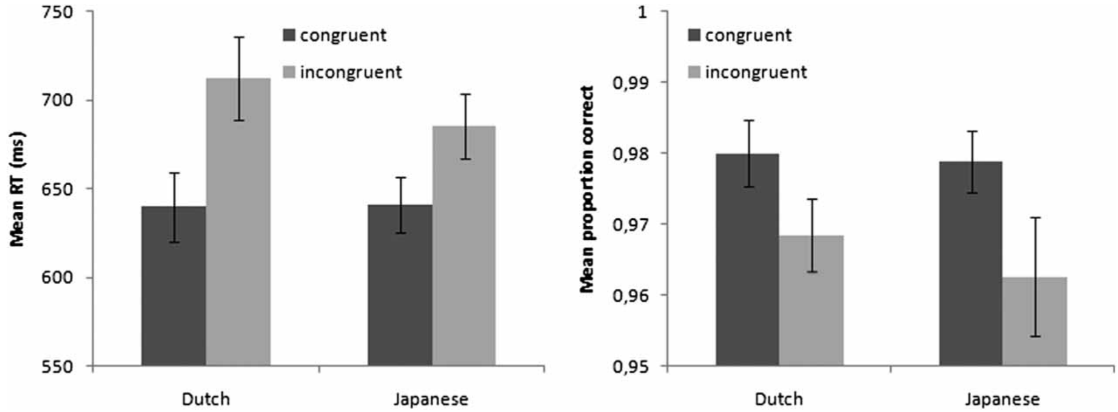

Figure 2. *Reaction times (RTs; left) and accuracy scores (right) for congruent and incongruent trials in Dutch (L1) and Japanese (L2). Error bars represent 1 standard error.*

language proficiency (given the number of participants, correlations equal to or larger than .38 are significant). We then investigated the relationship between ΔStroop and those measures that relate to L2 expertise.

We observed significant positive correlations between ΔStroop and participants' scores on the kanji test, $r(24) = .48$, $p = .012$. This confirms the prediction that the degree of Stroop interference in L2, normalized by the Stroop interference in L1, is directly related to the proficiency in L2 (see Figure 3), but is inconsistent with the notion that colour naming can only proceed after processing of the word form has been completed (as processing of the word form should finish earlier with greater reading skills, leading to less

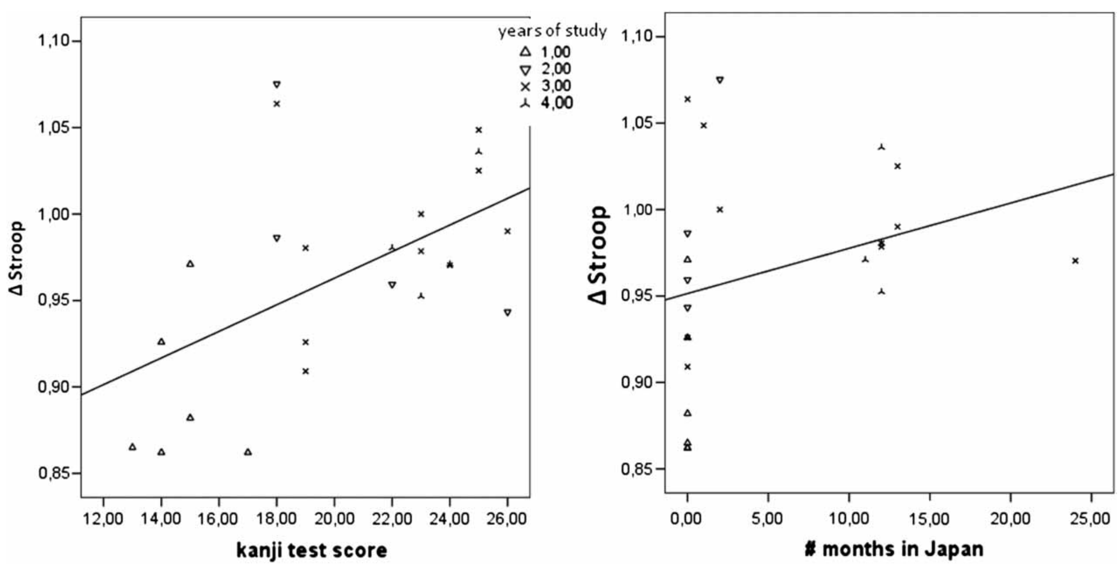

Figure 3. *ΔStroop scores show a positive relationship to test scores on the kanji test (left) and to the number of months participants had lived in Japan (right).*

interference). The normalization is important, as argued the introduction: Without normalization, neither the size of the Stroop effect in Japanese nor that in Dutch correlated with language proficiency, $r(24) = -.07$, $p = .73$; and $r(24) = -.23$, $p = .26$, respectively. In fact, without normalization there was a nonsignificant trend for Stroop interference in both L1 and L2 to correlate negatively with proficiency, similar to what was observed by Protopapas and colleagues (2007). Furthermore, if we divide participants into two groups based on study year (Years 1 and 2, $N = 11$, vs. Years 3 and 4, $N = 15$), ΔStroop only differed significantly from 1 for the first group, $t(10) = 3.65$, $p = .004$, who showed greater interference in Dutch. The second group demonstrated equivalent interference in L2 and L1, $t(14) = 1.11$, $p = .29$.

The scores on the kanji test are only one of many ways to assess second-language proficiency. We expect that this language proficiency is related to the amount of training with the second language. Indeed, the scores on the kanji test were correlated with the number of years the participants had been studying Japanese (Spearman's rho = .69, $p < .001$) and their length of stay in Japan (Spearman's rho = .54, $p = .005$). Furthermore, ΔStroop was correlated with the number of years they had been studying Japanese (Spearman's rho = .42, $p = .032$) and their length of stay in Japan (Spearman's rho = .56, $p = .003$). These correlations emphasize that the kanji test-scores are very much driven by the amount of training and exposure that participants have had with/to the second language. More subjective variables that relate to the intensity of language training did not correlate with ΔStroop, such as participants' estimates of their "average time/week spent on activities in Japanese" either for study, $r(24) = .03$, $p = .88$, or leisure, $r(24) = .02$, $p = .93$.

Discussion

In this study, we investigated the relationship between second-language skill-level and relative Stroop interference, in individuals who were learning Japanese. Participants showed interference on incongruent trials (in terms of slower response times, and more errors, than on congruent trials) in both Dutch and Japanese. This suggests that word recognition was sufficiently automatic (as well as obligatory) in both languages to impair performance on the colour naming task. Interference (in terms of response times, but not error rates) was also greater in Dutch (the participants' native language) than in Japanese, which is consistent with the notion that Stroop interference is positively related to practice (MacLeod, 1991; MacLeod & Dunbar, 1988), but which has not been found in all relevant studies.

Most importantly, we observed significant relations between the relative size of Stroop interference in Japanese and several indices of Japanese skill-level/experience (kanji test-score, length of stay in Japan, and years-of-study, though not the average weekly hours practice), showing greater interference with greater skill. This finding was opposite to the relationship that was found by Protopapas and colleagues (2007), but consistent with the notion that the relationship between automaticity and skill levels in reading is monotonically increasing. The results were also in line with prior studies that observed smaller Stroop interference in a second language (Mohamed Zied et al., 2004; Sumiya & Healy, 2004, 2008; but also see T. M. Lee & Chan, 2000), though these studies found no relationship with (self-reported) skill levels.

This pattern of results is difficult to reconcile with the notion that both word reading and colour naming rely on a limited and shared pool of resources. Skilled readers require less attentional resources (and/or time) to process the word form than do individuals with lower reading scores (e.g., Ruthruff et al., 2008). Thus, more resources should be available for colour naming and to inhibit conflicting information from the word form, leading to a negative relationship between reading skill and Stroop interference. Instead, we observed a positive relationship, which suggests that both dimensions are processed relatively independent of one another and that Stroop interference probably emerges at the stage of response selection (also see, e.g., Atkinson, Drysdale, & Fulham, 2003).

Alternatively, a positive relationship between (second-language) reading skills and (relative) Stroop interference might indicate learning-induced changes in the obligatory nature of word recognition. As reading skills improve, there could be an increase in the proportion of trials on which the word form is processed early enough to affect colour naming, and hence greater interference. The data further suggest that reading automaticity and/or its obligatory character develop relatively fast, as interference scores between L1 and L2 only differed significantly for individuals in the first two years of the Japanese study programme, after which automaticity could no longer be distinguished from that in participants' first language. This supports the notion that visual representations to novel words are formed quickly. While this has previously been demonstrated with training studies with new words (Share, 1999) or pseudowords (Clay, Bowers, Davis, & Hanley, 2005) in participants' L1, the current study extends these findings to L2 learning in a different orthographic system.

The present results do not allow us to distinguish between the effects of reading automaticity and cognitive control (the ability to inhibit the inappropriate response on incongruent trials), as we did not manipulate the proportion of congruent trials. This is unfortunate, as both have opposing effects on Stroop interference scores, and both probably differ between L1 and L2 (though cognitive control was controlled for at the level of individual and item-based variation). However, as additional cognitive control emerges as a consequence of increased automaticity, and its effects are probably smaller when comparing Stroop interference between two languages (Tzelgov & Kadosh, 2009), our results are likely to predominantly reflect the effects of differences in automaticity between the two languages.

Conclusions

To summarize, the present study investigated the relationship between reading automaticity and second-language skill/experience, in individuals learning Japanese. The results show that automaticity is positively related to second-language reading skill, when controlled for automaticity in the first language. We suggest that, when using the Stroop paradigm to assess the relative automaticity of two processes (such as reading in the native and in a second language), such a normalization is important, to avoid undue influence of individual differences in, for example, response inhibition/cognitive control. To our knowledge, this is the first study that shows that Stroop interference and reading skill are positively related, by controlling for individual differences by comparing L2 Stroop effects with L1 Stroop effects. Additionally, the present study indicates that reading automaticity (in terms of relative interference on the Stroop task) in the second language quickly becomes comparable to that of the first language, at least for high-frequency words such as those used in this study. If automatic reading depends on (neural) representations that correspond to visual word forms (Dehaene et al., 2005; Vinckier et al., 2007), this suggests that such representations may be formed relatively fast, even in adults learning a language with a different orthography from that of their native language.

Original manuscript received 17 March 2010
Accepted revision received 28 July 2010
First published online 4 October 2010

REFERENCES

Atkinson, C. M., Drysdale, K. A., & Fulham, W. R. (2003). Event-related potentials to Stroop and reverse Stroop stimuli. *International Journal of Psychophysiology*, *47*, 1–21.

Brown, T. L., Gore, C. L., & Carr, T. H. (2002a). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic"? *Journal of Experimental Psychology: General*, *131*, 220–240.

Brown, T. L., Joneleit, K., Robinson, C. S., & Brown, C. R. (2002b). Automaticity in reading and the

Stroop task: Testing the limits of involuntary word processing. *American Journal of Psychology*, *115,* 515–543.

Burt, J. S. (2002). Why do non-color words interfere with color naming? *Journal of Experimental Psychology: Human Perception and Performance*, *28,* 1019–1038.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *11,* 63–65.

Clay, F., Bowers, J. S., Davis, C. J., & Hanley, D. A. (2005). Teaching adults new words: The role of practice and consolidation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *33,* 970–976.

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., Usher, M., & McClelland, J. L. (1998). A PDP approach to set size effects within the Stroop task: Reply to Kanne, Balota, Spieler, and Faust (1998). *Psychological Review, 105,* 188–194.

Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: Contribution of the dorsal and ventral visual pathways. *NeuroImage*, *40,* 353–366.

Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood and aging. *The Journal of Genetic Psychology*, *100,* 47–53.

Cronbach, L. J., & Furby, L. (1970). How should we measure "change" – or should we? *Psychological Bulletin, 74*, 68–80.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: A proposal. *Trends in Cognitive Sciences*, *9,* 335–341.

Dhooge, E., & Hartsuiker, R. J. (2010). The distractor frequency effect in picture-word interference: Evidence for response exclusion. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *36,* 878–891.

Everatt, J., Warner, J., Miles, T. R., & Thomson, M. E. (1997). The incidence of Stroop interference in dyslexia. *Dyslexia*, *3,* 222–228.

Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, *33,* 225–250.

Faccioli, C., Peru, A., Rubini, E., & Tassinari, G. (2008). Poor readers but compelled to read: Stroop effects in developmental dyslexia. *Child Neuropsychology*, *14,* 277–283.

Fang, S. P., Tzeng, O. J. L., & Alva, L. (1981). Intralanguage vs. interlanguage Stroop effects in two types of writing systems. *Memory & Cognition*, *9,* 609–617.

Ferrand, L., & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Acta Psychologica*, *113,* 167–183.

Glezer, L. S., Jiang, X., & Riesenhuber, M. (2009). Evidence for highly selective neuronal tuning to whole words in the "visual word form area". *Neuron*, *62,* 199–204.

Hicks, C., & Jackson, P. (2005). A study of the relationship between the Stroop effect and reading age in dyslexic subjects. *Journal of Research in Reading*, *4,* 29–33.

Kapoula, Z., Lê, T. T., Bonnet, A., Bourtoire, P., Demule, E., Fauvel, C., et al. (2010). Poor Stroop performances in 15-year-old dyslexic teenagers. *Experimental Brain Research*, *203,* 419–425.

Lee, H. W., Rayner, K., & Pollatsek, A. (1999). The time course of phonological, semantic, and orthographic coding in reading: Evidence from the fast-priming technique. *Psychonomic Bulletin & Review*, *6,* 624–634.

Lee, T. M., & Chan, C. C. (2000). Stroop interference in Chinese and English. *Journal of Clinical and Experimental Neuropsychology*, *22,* 465–471.

Lien, M. C., Allen, P. A., Ruthruff, E., Grabbe, J., McCann, R. S., & Remington, R. W. (2006). Visual word recognition without central attention: Evidence for greater automaticity with advancing age. *Psychology and Aging*, *21,* 431–447.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly*, *13,* 123–146.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109,* 163–203.

MacLeod, C. M., & Dunbar, K. (1988). Training and Stroop-like interference: Evidence for a continuum of automaticity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *14,* 126–135.

Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General*, *132,* 228–252.

Mohamed Zied, K., Phillipe, A., Pinon, K., Havet-Thomassin, V., Aubin, G., Roy, A., et al. (2004). Bilingualism and adult differences in inhibitory mechanisms: Evidence from a bilingual Stroop task. *Brain & Cognition*, *54,* 254–256.

Monsell, S., Taylor, T. J., & Murphy, K. (2001). Naming the color of a word: Is it responses or task sets that compete? *Memory & Cognition*, *29,* 137–151.

Peru, A., Faccioli, C., & Tassinari, G. (2006). Stroop effects from 3 to 10 years: The critical role of reading acquisition. *Archives Italiennes de Biologie*, *144,* 45–62.

Protopapas, A., Archonti, A., & Skaloumbakas, C. (2007). Reading ability is negatively related to Stroop interference. *Cognitive Psychology*, *54,* 251–282.

Rawson, K. A., & Middleton, E. L. (2009). Memory-based processing as a mechanism of automaticity in text comprehension. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35,* 353–370.

Repovš, G. (2004). The mode of response and the Stroop effect: A reaction time analysis. *Horizons of Psychology*, *13,* 105–114.

Risko, E. F., Stolz, J. A., & Besner, D. (2005). Basic processes in reading: Is visual word recognition obligatory? *Psychonomic Bulletin & Review*, *12,* 119–124.

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.

Ruthruff, E., Allen, P. A., Lien, M. C., & Grabbe, J. (2008). Visual word recognition without central attention: Evidence for greater automaticity with greater reading ability. *Psychonomic Bulletin and Review*, *15,* 337–343.

Schadler, M., & Thissen, D. (1981). The development of automatic word recognition and reading skill. *Memory & Cognition*, *9,* 132–141.

Schiller, P. H. (1966). Developmental study of color-word interference. *Journal of Experimental Psychology*, *72,* 105–108.

Schmidt, J. R., & Besner, D. (2008). The Stroop effect: Why proportion congruent has nothing to do with congruency and everything to do with contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34,* 514–523.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72,* 95–129.

Smith, M. C., & Kirsner, K. (1982). Language and orthography as irrelevant features in colour-word and picture-word Stroop interference. *Quarterly Journal of Experimental Psychology Section A*, *34,* 153–170.

Strauss, G. P., Allen, D. N., Jorgensen, M. L., & Cramer, S. L. (2005). Test–retest reliability of standard and emotional Stroop tasks: An investigation of color–word and picture–word versions. *Assessment*, *12,* 330–337.

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

Sumiya, H., & Healy, A. F. (2004). Phonology in the bilingual Stroop effect. *Memory & Cognition*, *32,* 752–758.

Sumiya, H., & Healy, A. F. (2008). The Stroop effect in English–Japanese bilinguals: The effect of phonological similarity. *Experimental Psychology*, *55,* 93–101.

Thuy, D. H., Matsuo, K., Nakamura, K., Toma, K., Oga, T., Nakai, T., et al. (2004). Implicit and explicit processing of kanji and kana words and non-words studied with fMRI. *NeuroImage*, *23,* 878–889.

Tzelgov, J., Henik, A., & Berger, J. (1992). Controlling Stroop effects by manipulating expectations for color words. *Memory & Cognition*, *20,* 727–735.

Tzelgov, J., Henik, A., & Leiser, D. (1990). Controlling Stroop interference: Evidence from a bilingual task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16,* 760–771.

Tzelgov, J., & Kadosh, R. C. (2009). From automaticity to control in bilinguals. *Trends in Cognitive Sciences*, *13,* 455.

Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Science*, *13,* 252–257.

Vinckier, F., Dehaene, S., Jobert, A., Dubus, J. P., Sigman, M., & Cohen, L. (2007). Hierarchical coding of letter strings in the ventral stream: Dissecting the inner organization of the visual word-form system. *Neuron*, *55,* 143–156.

Zmigrod, S., & Hommel, B. (2010). Temporal dynamics of unimodal and multimodal feature binding. *Attention, Perception & Psychophysics*, *72,* 142–152.

APPENDIX

Participant characteristics

Table A1. *Results of questionnaire*

Kanji-test score	*Use of Japanese (hours per week)*		*Study year*	*Months in Japan*
	Leisure	*Work*		
14	24	5	1	0
15	9	4	1	0
15	20	10	1	0
17	20	2	1	0
14	27	0	1	0
13	20	5	1	0
26	8	40	2	0
18	5	6	2	0
22	18	10	2	0
18	30	1	2	2
23	8	2	3	12
25	12	7	3	13
24	10	2	3	24
19	10	2	3	0
19	4	5	3	0
26	15	10	3	13
18	40	10	3	0
19	10	1	3	12
25	9	10	3	1
23	25	5	3	2
22	12	1	4	12
23	3	3	4	12
24	4	1	4	11
25	6	8	4	12
20	13	6		1
12	0	10		6