INTRA- VERSUS INTER-LANGUAGE STROOP EFFECTS IN TWO TYPES OF WRITING SYSTEMS*

Sheng-Ping Fang,+ Ovid J. L. Tzeng,++ and Liz Alva+++ 

Abstract. The relation between word processing strategy and the orthographic structure of a written language was explored in the present study. Three experiments were conducted using Chinese-English, Spanish-English, and Japanese-English bilinguals, respectively. Each subject was asked to perform a modified Stroop color-naming task where the stimulus and the response language were either the same or different. The magnitude of Stroop effect was greater in the intra-language condition than in the inter-language condition. When the magnitude of reduction of Stroop interference from the intra- to the inter-language condition was compared across all bilingual groups, an inverse relationship was found between the magnitude of reduction and the degree of similarity between the orthographic structures of the two written languages. It is concluded that reading logographic and phonologic symbols entails different processing mechanisms and that controversial issues in bilingual processing cannot be resolved without taking into account the effect of orthographic variations on the information processing system.

The invention of written symbols to represent spoken language is undoubtedly one of the most important achievements in the history of mankind. The written symbol has enabled us to overcome the limitations of space and time imposed by oral communication and has allowed us to extend our thoughts across centuries as well as continents.

There have been many different types of writing systems invented to represent various types of spoken languages. The designing principles for writing systems can be divided into two different categories. The first type of orthography evolved from the earlier semasiography, which expresses a

---

*This paper is to appear in Memory & Cognition, in press.
+University of California, Riverside.
++Also University of California, Riverside.

Acknowledgment. We thank Jessie S. H. Deou and Susan Stone for their great help in conducting the third experiment, and Katherine Malley for her helpful comments on earlier drafts. The final version of this paper has also benefited from comments made by Irvine Biederman and the reviewers. The present study was supported in part by an Inter-Campus Research Fund from the Academic Senate of the University of California, Riverside to the second author, and in part by a trainingship from the National Institute of Mental Health to the third author (supervised by Dr. Robert Singer).


213
general idea in picture drawings rather than a sequence of words in a sentence, to logographs with each symbol expressing a single particular morpheme. The concept underlying the development of this type of orthography is to map the written symbols directly onto words, from which meaning is generated. The second type of orthography evolved from the rebus (a representation of a word or phrase by pictures that suggest how a word is pronounced in the spoken language, e.g., 🌟 for idea) to the syllabary and then to the alphabet. The concept behind it is sound writing. That is, the relation of sign to meaning is meant to be mediated through the sound system of the spoken language. This difference in how lexical units may be recovered from written symbols raises an important and interesting question: Do our visual information processing strategies differ when the information is presented in different formats? In recent years, this question has become of major concern among many cognitive psychologists (Biederman & Tsao, 1979; Gleitman & Rozin, 1977; Park & Arbuckle, 1977; Tzeng, Hung, & Garro, 1978).

That reading different writing systems may entail different information processing strategies is supported by some recent clinical and experimental observations. Sasanuma (1974) reported that the ability of Japanese aphasic patients to use logographic (kanji) and phonologic (kana) scripts can be selectively impaired. Parallel to this finding, in visual hemi-field experiments in which stimuli are presented to the right or left visual field briefly via a tachistoscope, a right visual field (i.e., left hemisphere) advantage is usually found for the recognition of phonologically based symbols such as English words or Japanese kana scripts, while a left visual field advantage is found for the recognition of single Chinese characters (Tzeng, Hung, Cotton, & Wang, 1979). Furthermore, in a cross-language study that investigated the effects of language (Chinese vs. English) and mode of stimulus presentation (visual vs. auditory), Turnage & McGinnies (1973) found that visual input facilitated the learning for Chinese subjects whereas auditory input produced superior recall performance for American subjects. All these results seem to point out that readers of different scripts may have developed different processing strategies in order to achieve efficient reading. It is of utmost importance for cognitive psychologists to find out at which level of information processing these differences due to orthographic variations occur.

A recent study of Biederman and Tsao (1979) shed light on the issue of the orthographic variations by using a Stroop (1935) interference paradigm. It is an established fact that in the Stroop color-word test, it requires more time to name a series of color patches when the patches are themselves incongruent color names (e.g., GREEN in red ink) than when the patches are simple colored rectangles. Biederman and Tsao (1979) found a greater interference effect for Chinese subjects in a Chinese version Stroop color-naming task than for American subjects in an English version. They attributed this difference to the possibility that there may be fundamental differences in the perceptual demands of reading Chinese and English. Since the perception of color and the direct accessing of meaning from a pattern's configuration are functions that have been assigned to the right hemisphere, it was suggested that during the Stroop test these two functions might be competing for the same perceptual capacity of the right hemisphere. This competition could have been avoided in the English Stroop test because reading English and naming color are executed by different hemispheric mechanisms. Biederman and Tsao further speculated that there may be some fundamental differences in the
obligatory processing of Chinese and English prints. They suggested that a reader of alphabetic writing cannot refrain from applying an abstract rule system to the word whereas a reader of Chinese may not be able to refrain from configurational processing of the logograph.

The conceptualization that reading different types of scripts automatically activates different types of perceptual constraints is an intriguing one. It leads to a unique prediction concerning the bilingual processing in a modified Stroop task. Suppose a Spanish–English bilingual subject is asked to name colors once in each of the two languages for color stimuli that are either Spanish color words, English color words, or control patches. Based on previous empirical findings (Dyer, 1971; Preston & Lambert, 1969), we can predict that color naming speed will be relatively slower when the naming language and the language of the color words are the same than when they are different. In other words, we can predict that the Stroop interference effect will be reduced in the inter-language condition as compared with that in the intra-language condition. But since both Spanish and English are alphabetic scripts that tend to activate similar obligatory processing strategies, the magnitude of reduction in the Stroop interference would not be much. Now suppose we ask a group of Chinese–English bilingual subjects to perform the inter- and intra-language Stroop tasks in which the interfering and the naming languages are either Chinese or English. It is again reasonable to predict that the inter-language condition will produce less Stroop interference than the intra-language condition. However, the most important question is whether the magnitude of reduction (from the intra- to the inter-language condition) will be greater, equivalent, or less for the Chinese–English bilinguals, as compared to that for the Spanish–English bilinguals. According to Biederman and Tsao's (1979) conjecture that reading alphabetic and logographic scripts entails different perceptual demands, one would predict that the magnitude of reduction (i.e., from the intra- to the inter-language condition) should be greater for the Chinese–English bilinguals than for the Spanish–English bilinguals. This expectation results from the assumption that while English and Spanish scripts activate similar obligatory processing strategies and thus are competing for the same perceptual demands, the Chinese and English scripts activate different obligatory processing strategies that do not interfere with each other. Experiments 1 and 2 were conducted to test this unique prediction generated from the considerations of orthographic variations and their relations to human information processing. Experiment 3 was conducted to further test this prediction while holding the phonological factor constant by using Japanese–English bilingual subjects.

**METHOD**

**Experiment 1**

**Subjects.** Thirty Chinese–English (C–E) bilinguals with normal color vision served as subjects. All were students at the University of California. Twenty of them were recruited from the Riverside campus and the remaining ten were from the Berkeley campus. All subjects had learned Chinese as their first language. All of them passed TOEFL (Test of English as a Foreign Language) before they were admitted into the University of California. Based upon their naming latencies of English and Chinese color terms (printed in black ink), all of them should be classified as Chinese dominant.
Materials. Three stimulus boards were prepared: one control board, one color-word board in English, and one color-word board in Chinese. Each board measured 40.6 x 50.8 cm².

The control board was constructed with six rows of ten 3 x 3 cm² patches, the colors of which were either red, blue, green, or brown. The patches were spaced 2 cm apart within each row and the rows were spaced 3 cm apart. Among the 60 patches, each of the four colors appeared 15 times in a random arrangement except that no color ever appeared twice in succession.

On the English board, the color arrangement was identical to that on the control board while each patch was replaced with an English word indicating an incongruent color name. Due to the physical nature of English words, each color word was 1.5 cm tall and up to 3 cm wide, centered in the place where the patch would have been. Words and colors used on this board were red, blue, green, and brown (Note: they are all monosyllabic words). Each word and color appeared 15 times randomly and no word or color appeared twice in succession.

The Chinese board resembled the English version in all aspects except that each English word was transformed into its corresponding Chinese character and measured 3 x 3 cm². The characters used on the Chinese board were красно (red), сине (blue), зеленно (green), and коричнево (brown), representing red, blue, green, and brown, respectively. The Chinese characters are monosyllabic in nature.

Design and Procedure. Each subject was given six tasks: (1) color naming of patches in English, (2) color naming of patches in Chinese, (3) color naming of English color-words in English, (4) color naming of English color-words in Chinese, (5) color naming of Chinese color-words in English, (6) color naming of Chinese color-words in Chinese. The order of administration was random.

Before the experiment started, the subject sat in front of a table while the stimulus board was placed on it, covered with a heavy blank paper sheet. The experimenter first explained the task and procedure to the subject. The subject was asked to perform each task as accurately and as quickly as possible, and to correct mistakes wherever possible. The subject was also asked not to point at the items while naming their colors. It was especially emphasized not to read the words but to name the colors of them instead. The subject was then asked to respond to two practice items, one Chinese character 猩红 (representing yellow) in pink ink and another character 紫 (representing purple) in yellow ink. After proper responses were made, the experiment started. Each time a stimulus board was to be displayed, the subject was informed of the type of task to be performed. The stimulus board was covered again as soon as the task was completed. Color naming times for entire boards were recorded with a stopwatch to the nearest tenth of a second. Time between tasks was minimal, representing only the delay required to record data and obtain the new stimulus board.
Experiment 2

Subjects. Thirty Spanish-English (S-E) bilinguals with normal color vision served as subjects. All had learned Spanish as their first language with half of them Spanish dominant and the other half English dominant by their own estimates. However, based upon their naming latencies of English and Spanish color words (printed in black), all of them should be classified as Spanish dominant.

Materials. Three stimulus boards were used in Experiment 2, namely, one control board, one English color-word board, and one Spanish color-word board. Both the control board and the English board were identical to those used in Experiment 1. The Spanish board resembled its English counterpart in all aspects except that each English word was transformed into its Spanish equivalent. The Spanish equivalents were rojo, azul, verde, and cafe.

Design and Procedure. Each subject was given six tasks: (1) color naming of squares in English, (2) color naming of squares in Spanish, (3) color naming of English color-words in English, (4) color naming of English color-words in Spanish, (5) color naming of Spanish color-words in English, (6) color naming of Spanish color-words in Spanish. The order of administration was random. The instruction and procedure were the same as those in Experiment 1. Color naming times for entire boards were recorded with a stopwatch to the nearest tenth of a second.

RESULTS AND DISCUSSION

For each subject, the color naming time for the entire board was transformed into the naming time for a single item in milliseconds. This transformation procedure was applied to each of the six tasks and then the mean color-naming time for each of the six tasks was calculated based upon these transformed scores across the whole group. The data of the C-E bilinguals are presented in Table 1 (Experiment 1) and the data of the S-E bilinguals are presented in Table 2 (Experiment 2). Note that scores in parentheses represent the magnitude of the Stroop interference effect.

At first glance, the data presented in Table 1 seem to suggest that English color words produce greater Stroop interference (492 msec) than Chinese color characters (402 msec), a result at odds with that obtained by Biederman and Tsao (1979). However, careful reflection reveals that this comparison between our data and those of Biederman and Tsao may not be a valid one. In the present experiment, English is the second language for our subjects whereas in Biederman and Tsao's experiment, English is the native language for their American subjects. Thus, the data, as shown in Table 1, should not be taken as an instance of failure to replicate Biederman and Tsao. In fact, our concern here is not to compare the degrees of interference between the Chinese Stroop task and the English Stroop task. Rather, the concern is with whether or not English and Spanish words (being both alphabetic scripts) would activate the same processing mechanism such that switching languages in a bilingual Stroop task should not reduce the amount of interference as much as in the case of switching between English and Chinese (a logographic script).
Table 1

Mean Color Naming Times (msec per item) for C-E Bilinguals on the Stroop Tasks (N = 30).

<table>
<thead>
<tr>
<th></th>
<th>English color-word</th>
<th>Chinese color-word</th>
<th>Control square</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Response</td>
<td>1431 (605)</td>
<td>1128 (302)</td>
<td>826</td>
<td>(454)</td>
</tr>
<tr>
<td>Chinese Response</td>
<td>1098 (378)</td>
<td>1221 (501)</td>
<td>728</td>
<td>(440)</td>
</tr>
<tr>
<td>Mean</td>
<td>(492)</td>
<td>(402)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses indicate the amount of interference (color-word minus control square).

Table 2

Mean Color Naming Times (msec per item) for S-E Bilinguals on the Stroop Tasks (N = 30).

<table>
<thead>
<tr>
<th></th>
<th>English color-word</th>
<th>Spanish color-word</th>
<th>Control square</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Response</td>
<td>1169 (495)</td>
<td>1017 (343)</td>
<td>674</td>
<td>(419)</td>
</tr>
<tr>
<td>Spanish Response</td>
<td>1166 (446)</td>
<td>1110 (398)</td>
<td>720</td>
<td>(418)</td>
</tr>
<tr>
<td>Mean</td>
<td>(470)</td>
<td>(366)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses indicate the amount of interference (color-word minus control square).
But before we examine the data pertinent to the above concern, let us clarify one particular point about the rationale behind the methodology. It can be argued that in no situation do subjects visually process words in the two languages simultaneously and that we may have a confusion between input (reading) and output (naming) mechanisms. Consequently, one may ask on what basis we can expect reading and naming to engage in one similar set of mechanisms. This question can be answered quite easily on empirical grounds. First, an automatic speech recoding of visually presented words is an established fact and it occurs in processing words written in alphabetic as well as non-alphabetic (such as Chinese, Japanese, etc.) scripts (Erickson, Mattingly, & Turvey, 1977; Tzeng, Hung, & Wang, 1977). Second, an automatic graphemic recoding of auditorily presented words has recently been established in a series of experiments by Seidenberg and Tanenhaus (1979) and by Nolan, Tanenhaus, and Seidenberg (1981). More importantly and interestingly, further studies on the graphemic recoding phenomenon by Tanenhaus, Flanigan, and Seidenberg (in press) demonstrated that such an automatic graphemic-recoding was responsible for slowing down color-naming responses in a Stroop-like task. Similar findings were also reported by Conrad (1978). Therefore, our assumption that the orthographic factor is involved in a color-naming task is completely justified.

Let us now examine the data presented in Tables 1 and 2 with respect to predictions made earlier in this paper. First of all, the Stroop interference effect was indeed reduced in the inter-language condition as compared with that in the intra-language condition. There was a 213 msec per item reduction for the C-E bilinguals and a 48 msec per item reduction for the S-E bilinguals. And indeed, the magnitude of reduction appeared greater for the former than for the latter.

A one-tailed planned comparison between inter- and intra-language Stroop effects was made for both bilingual groups. The magnitude of shift-language reduction was significant for the C-E subjects but not for the S-E subjects, t (29) = 6.08, p < .0001 and t (29) = 1.48, p < .10, respectively. Thus, the main prediction was confirmed. That is, the reduction scores of the two groups did differ significantly, and the magnitude of reduction was greater for the C-E bilinguals than for the S-E bilinguals.

For each bilingual group, a repeated-measures analysis of variance was also performed with the stimulus language as one factor and the response language as the second factor. For the C-E subjects, the main effect for the stimulus language was significant, F (1,29) = 6.35, MSE = 38225, p < .05, whereas the main effect for the response language was not, F (1,29) < 1. Also significant was the interaction between the two factors, F (1,29) = 36.94, MSE = 36697, p < .001. Further analysis of simple effects showed that there was significantly less interference whenever response and stimulus languages were different compared to the cases when they were the same. For the S-E subjects, the only significant effect found was the main effect of the stimulus language, F (1,29) = 13.52, MSE = 24031, p < .001, with English color-words resulting in greater interference than Spanish color-words in both response conditions.

For both S-E and C-E subjects, the stimulus language had much stronger control over the degree of interference effect as compared to the response
language. Both groups exhibited a significant main effect of the stimulus languages while, in both groups, response languages accounted for essentially zero percent of the total variance. These results suggest that the bilingual Stroop effect is more likely to be at the perceptual level than at the response level. The emphasis on the stimulus factor is in line with Biederman and Tsao's conjecture that the orthographic structure in the written language may play an important role in determining the magnitude of the Stroop effect. They also localize such an orthographic effect at the perceptual stage. They reason that different orthographic structures may impose different task demands such that different perceptual mechanisms are activated to meet these demands. This conceptualization also helps to explain the results of the two bilingual groups. Since both English and Spanish are alphabetic scripts, the perceptual mechanisms activated to process them are similar. Consequently, switching languages would not reduce the Stroop effect. On the other hand, Chinese logographs and English letters are two different scripts, and switching language means turning off one perceptual mechanism and turning on another one such that little interference would occur.

Based upon the above observations, we may induce a more generalized statement about the effect of the orthographic structure on the bilingual Stroop interference. That is, for any group of bilingual subjects, the magnitude of reduction from the intra- to the inter-language Stroop interference effect is a linearly decreasing function of the degree of similarity between the orthographic structures of the two languages. The validity of such an assertion can be tested by examining the patterns of the bilingual Stroop effects in the existing literature. To do this, we recalculated from the results of the present experiment and two other different bilingual experiments the magnitude of reduction of the Stroop interference from the intra- to the inter-language condition (Dyer, 1971, Experiment II, session 1; Preston & Lambert, 1969). All together, there were five types of bilingual subjects, namely, Chinese-English, Hungarian-English, Spanish-English, German-English, and French-English bilinguals. Wherever more than one experiment was run with respect to a certain type of bilingual, data were combined for that bilingual condition. We ranked these reduction scores according to their magnitude and obtained the following results (Table 3): Chinese-English bilinguals revealed a reduction of 213 msec; Hungarian-English, 112 msec; Spanish-English, 68 msec; German-English, 36 msec; and French-English, 33 msec per item. The ordering of the last three categories is particularly revealing. Why should switching between Spanish and English produce a greater reduction of interference than that between French and English or between German and English? It is certainly not intuitively obvious why Spanish and English are more orthographically dissimilar than French and English (or German and English). However, if we examine the spellings of color terms across these languages, then the deviation of Spanish becomes immediately clear. For example, red, blue, green, and brown are translated and spelled as rot, blau, grün, and braun in German; as rouge, bleu, vert, and brun in French; but as rojo, azul, verde, and café, respectively, in Spanish. Clearly, with respect to the color terms used in all these studies, Spanish color terms are orthographically more dissimilar to English color terms than both French and German. Correspondingly, we also observed a greater reduction of Stroop interference. This pattern confirms our expectation that the magnitude of reduction is a negative function of the degree of similarity between the orthographic structures of the two written languages. In other
words, the greater the orthographic similarity between the two languages, the stronger the competition for the same processing mechanisms and thus the smaller the reduction of Stroop interference from the intra- to the inter-language condition.

<table>
<thead>
<tr>
<th>Language Pair</th>
<th>Reduction (msec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese-English</td>
<td>213</td>
</tr>
<tr>
<td>Kanji-English</td>
<td>121</td>
</tr>
<tr>
<td>Hungarian-English</td>
<td>112</td>
</tr>
<tr>
<td>Hirakana-English</td>
<td>108</td>
</tr>
<tr>
<td>Spanish-English</td>
<td>68</td>
</tr>
<tr>
<td>German-English</td>
<td>36</td>
</tr>
<tr>
<td>French-English</td>
<td>33</td>
</tr>
</tbody>
</table>

Data from Experiment 3.

However, since orthographic similarity is highly correlated with phonological similarity, an alternative explanation is to attribute the effect of switching language to the phonological factor instead of the orthographic factor. Even though these two explanations are not necessarily mutually exclusive, it is important to determine which factor (orthographic or phonological) contributes more to the reduction of the Stroop interference. Experiment 3 was conducted to weigh the importance of the orthographic factor while holding the phonological factor constant.

**EXPERIMENT 3**

To answer the question whether the orthographic difference alone can account for the lexical processing and consequently the differential shift-language effects observed in the last two experiments, Japanese-English bilingual subjects were tested in Experiment 3.

Japanese is unique in the sense that three different types of scripts are concurrently used to represent the spoken language. Among the three types of scripts, Chinese logographs, referred to as kanji, are generally used to write the content words. The other two kinds of scripts, which are referred to as
hirakana and katakana and are syllabic in nature, are used for writing grammatical particles and foreign words, respectively. Though these three types of scripts differ in their writing styles, the words written with any one of the scripts are read in exactly the same pronunciation. This unique aspect of Japanese writing enables us to vary the orthographic structures while holding the phonological factor constant.

In this experiment, color-words were written in either kanji, hirakana, or English. With respect to the script/speech relationship embedded in the orthographic structure of the writing system, the hirakana script as a sound-writing system bears closer relation to the English script than the kanji logograph does. Following the arguments advanced by Biederman and Tsao (1979), it is reasonable to assume that the hirakana and English scripts are more likely to share a common processing mechanism than the kanji and English scripts. Accordingly, if the orthographic factor alone can effectively account for the differential reduction scores observed in Experiments 1 and 2, then the magnitude of reduction (from the intra- to the inter-language condition) should be significantly greater for the kanji-English condition than for the hirakana-English condition. On the other hand, if the phonological factor plays a more important role, then little difference in the magnitude of reduction should be observed between the kanji-English and the hirakana-English condition. Of course, there is always the possibility that both factors may play determinant roles in the bilingual Stroop effect.

What about the direct comparison between the pure cases (i.e., no language switching) of kanji and hirakana conditions? Biederman and Tsao (1979) demonstrated that more Stroop-type interference occurred in logographic than in alphabetic scripts. However, their demonstration has been criticized on the grounds of a possible confounding by two very different subject populations (Tzeng et al., 1978). In the present experiment, with kanji and hirakana scripts as the experimental materials, we were able to draw subjects from the same population and assign them randomly to two different conditions. Any demonstrated effect of orthography on the magnitude of the Stroop interference, therefore, should not be attributed to the subject factor.

Method

Subjects. Fifty Japanese-English bilingual students with normal color vision served as subjects. They were all natives of Japan and had at least six years of formal training in English as a second language. Most of them were enrolled in the ESL (English as a Second Language) Extension program and had been in the U.S. for less than one year. Thirty-eight subjects were tested at the University of California, Riverside campus and the remaining twelve were tested at the University of California, Berkeley campus. Subjects at both campuses were randomly divided into two groups. Group 1 was exposed to color-words in kanji and English while Group 2 was exposed to color-words in hirakana and English.

Materials. Four stimulus boards were prepared: one control board, one color-word board in English, one color-word board in hirakana, and one color-word board in kanji. For the consistency of grammatical form in Japanese, the four colors and color-names used in this experiment were red, blue, green, and purple. Both the control board and the English board resembled those used in
Experiments 1 and 2 except that the color and the word brown were replaced with purple in all cases. The hirakana board resembled the English version in all aspects except that each English word was transformed into hirakana. The hirakana equivalents were あか (AKA), あつ (AUO), みどり (MIDORI), and みずき (MURASAKI), representing red, blue, green, and purple. Their kanji counterparts were 3x3 cm² large and were the characters つ (red), す (blue), み (green), and く (purple). The control board, the English board, and the kanji version composed the stimuli for Group 1. The control board, the English board, and the kana version composed the stimuli for Group 2.

Design and Procedure. Subjects were randomly divided into two groups. All subjects were asked to perform the following four tasks: (1) color naming of squares in English, (2) color naming of squares in Japanese, (3) color naming of English color-words in English, and (4) color naming of English color-words in Japanese. Two additional tasks were assigned to Group 1 subjects: (5) color naming of kanji in English, and (6) color naming of kanji in Japanese. Similarly, subjects in Group 2 were asked to perform two additional tasks: (5) color naming of hirakana in English, and (6) color naming of hirakana in Japanese. The order of administration was random within each group and yoked between groups. The instruction and procedures were the same as those in Experiments 1 and 2. Color naming times for entire boards were recorded with a stopwatch to the nearest tenth of a second.

Results and Discussion

Color naming times for the entire card board were again transformed into reaction times of naming a single item in milliseconds. Table 4 shows the mean reaction times required for performing the six tasks. The scores of the Stroop effect shown in parentheses were analyzed separately for Group 1 and Group 2.

The scores of Stroop interference obtained from Group 1 were subjected to a repeated two-way ANOVA that examined the effect of the stimulus language and that of the response language. Statistical analysis revealed that the main effect of the stimulus language is significant, $F(1,24) = 8.11, MSe = 20083, p < .01$, whereas the main effect of the response language was not, $F(1,24) = 3.03, MSe = 32514$. There was also a significant interaction effect between the stimulus and response languages, $F(1,24) = 13.67, MSe = 27016, p < .005$. Further analysis suggested that the interaction resulted mainly from kanji scripts being exceptionally interfering when subjects are naming in Japanese.

A similar ANOVA was carried out on data of Group 2 subjects. The statistical analyses revealed neither an effect of the stimulus language nor an effect of the response language, $F(1,24) = 3.11, MSe = 16795$, and $F(1,24) = 2.00, MSe = 44964$, respectively. However, there was a significant interaction between these two factors, $F(1,24) = 9.50, MSe = 30645, p < .01$. Post-hoc analysis of simple effects showed that when subjects were naming in English, English scripts interfered more than hirakana, $F(1,48) = 4.98, MSe = 9930, p < .05$, and when subjects were naming in Japanese, hirakana interfered more than English, $F(1,48) = 30.04, MSe = 9930, p < .005$. In the presence of hirakana, naming colors in Japanese was more difficult than in English, $F(1,48) = 9.49, MSe = 37804, p < .005$, while naming colors in one language was not more difficult than in the other when English words were presented, $F(1,48) < 1$. 223
Table 4

Mean Color Naming Times (msec per item) for Japanese-English Bilinguals on the Stroop Tasks

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N = 25)</td>
<td>(N = 25)</td>
</tr>
<tr>
<td></td>
<td>Eng. color-word</td>
<td>Eng. color-word</td>
</tr>
<tr>
<td></td>
<td>Kanji color-word</td>
<td>Kana color-word</td>
</tr>
<tr>
<td></td>
<td>Control square</td>
<td>Control square</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>English</td>
<td>994</td>
<td>990</td>
</tr>
<tr>
<td></td>
<td>954</td>
<td>928</td>
</tr>
<tr>
<td></td>
<td>704</td>
<td>721</td>
</tr>
<tr>
<td>Response</td>
<td>(290)</td>
<td>(269)</td>
</tr>
<tr>
<td></td>
<td>(250)</td>
<td>(207)</td>
</tr>
<tr>
<td></td>
<td>(270)</td>
<td>(238)</td>
</tr>
<tr>
<td>Japanese</td>
<td>913</td>
<td>910</td>
</tr>
<tr>
<td></td>
<td>1115</td>
<td>1064</td>
</tr>
<tr>
<td></td>
<td>681</td>
<td>689</td>
</tr>
<tr>
<td>Response</td>
<td>(232)</td>
<td>(221)</td>
</tr>
<tr>
<td></td>
<td>(434)</td>
<td>(375)</td>
</tr>
<tr>
<td></td>
<td>(333)</td>
<td>(298)</td>
</tr>
<tr>
<td>Mean</td>
<td>(261)</td>
<td>(245)</td>
</tr>
<tr>
<td></td>
<td>(342)</td>
<td>(291)</td>
</tr>
</tbody>
</table>

Note. Numbers in parentheses indicate the amount of interference (color-words minus control square).
Of particular concern is whether differences in the orthographic structure play a decisive role in the magnitude of Stroop interference in a mixed-language condition. A one-tailed planned comparison between the intra- and the inter-language condition was made for each of these two groups. The magnitude of shift-language reduction was highly significant for both groups. There was a 121 msec per item reduction for Group 1 (kanji), \( t (24) = 3.68, p < .005 \), and a 108 msec per item reduction for Group 2 (hirakana), \( t (24) = 3.08, p < .005 \). However, the reduction scores of the two groups did not differ significantly, even though the direction of the difference was consistent with our expectation, \( t (48) = .28, \text{ns} \). Apparently, the phonological factors contribute more to the reduction of Stroop interference in the mixed-language condition than the orthographic factor does.

Another comparison was made between the two conditions where both stimulus and naming languages were Japanese. Shimamura and Hunt (Note 1) conducted a Stroop experiment with color-words written either in kana or in kanji (a within-subject factor). They found a higher Stroop effect for kanji than for kana script with Japanese subjects. In the present experiment, color naming in Japanese did appear more difficult for the kanji version than for the kana version (434 vs. 375). Again, the difference is in the right direction. However, the difference was not statistically significant, \( t (48) = .23, \text{ns} \).

According to the above results, it does not seem that a strong explanation based upon variations in orthography has gained support in Experiment 3. Yet, the orthographic factor cannot be totally dismissed without some cautious comments. In all comparisons made between kanji and hirakana processing, the direction of differences exhibited an expected pattern but the differences failed to reach a statistically significant level. However, we have noted that similar studies carried out in other laboratories (Shimamura & Hunt, Note 1; Biederman, personal communication) with a more powerful design (within-subject instead of between-subject) and with other dependent measures (e.g., error rates) did report significant differences. Therefore, we think the orthographic factor does play a role, but may not be as important as the phonological factor, in the bilingual Stroop experiment.

A criticism has always been raised against the comparison of kanji and kana symbols in the color naming task. For fluent readers of Japanese, the color terms they read in everyday life are usually expressed in kanji script and rarely in kana. Hence, the greater interference observed for the kanji script may be attributable to this familiarity factor. To counter such an argument, Shimamura and Hunt (Note 1) and Biederman (personal communication) presented further evidence showing that in a simple word naming experiment (naming words printed in black), color terms written in kana were actually named much faster than color terms written in kanji. Similar findings were reported by Feldman and Turvey (1980). So, although color terms are more frequently written in the kanji form and although kanji are more compact graphic representations of words in general, naming time was consistently less for the kana. Thus, familiarity seems not to be a major factor in this case.
GENERAL DISCUSSION

In recent years, reading research has become a significant interdisciplinary endeavor with contributions from such diverse fields as anthropology, artificial intelligence, cognitive psychology, educational psychology, linguistics, and neuropsychology. The present study tackles the issue of word processing from a cross-language perspective. Since the way a spoken language is represented graphemically varies from language to language, it is essential to find out whether such orthographic variations impose different processing requirements on readers of different written scripts. Two questions are of particular concern in the present study. First, would different processing mechanisms be activated in reading the logographic and the alphabetic scripts? Second, does the particular pair of languages that a bilingual individual knows have a specific effect on the degree of language overlap? For instance, should Chinese-English bilinguals be considered as qualitatively different from Spanish-English bilinguals with respect to their lexical representations?

The first question can be answered more or less in an affirmative manner. Indeed, the idea that reading logographic and phonologic symbols entails different cognitive strategies and processing mechanisms has been supported by studies concerning aphasia (Sasanuma, 1974), visual lateralization effects (Tzeng et al., 1979), quantity-comparison tasks (Besner & Coltheart, 1979), and serial recall (Turnage & McGinnies, 1973). Biederman and Tsao have suggested that there may be fundamental differences in the obligatory processing of alphabetic and logographic print. A reader of alphabet writing cannot refrain from applying an abstract rule system to the word, whereas a reader of Chinese cannot refrain from configurational processing of the logograph.

Answers to the second question are less unequivocal. On the one hand, we see that a rough estimate of the magnitude of reduction in the Stroop effect in mixed-language conditions (as compared to pure-language conditions) from among seven different types of bilingual subjects exhibits an orderly relationship between the orthographic structure and the amount of reduction. On the other hand, experiments with the two types of Japanese scripts only provide minimal support for the predictions generated from the consideration of orthography. Nevertheless, we also noted that data from other similar studies did provide much stronger support. Thus, we may conclude that the orthographic structure does play an important role, independent of the phonological factor, in the lexical formation of a bilingual subject.

The implication of such orthographic and phonological effects for research in bilingual processing is clear. We simply cannot, or should not, lump data of different types of bilingual subjects together and attempt to come up with a general statement about the processing mechanism. It has been the common practice of investigators of bilingualism to talk about L1 (first language) and L2 (second language) without paying much attention to the degree of orthographic and phonological similarities between the two languages. No wonder there is so much inconsistency from one bilingual study to another. For example, there is currently a controversy as to the pattern of the hemispheric dominance in L1 and L2 of a bilingual subject. It is conceivable that a Spanish-English bilingual should show a very different cerebral lateralization pattern from that of a Chinese-English bilingual (Tzeng et al., 1979). Thus, without taking into account the influence of the orthographic
structure, many controversial issues in bilingual processing are difficult to resolve.

The relation between language and thought has been a topic of intensive investigation for hundreds of years. Delineation of script/speech relationships and discovery of how the orthographic variations affect our information processing system will no doubt open up a new possibility for specifying the nature of symbol/thought interactions.

REFERENCE NOTE


REFERENCES


Feldman, L. B., & Turvey, M. T. Words written in Kana are named faster than the same words written in Kanji. Language and Speech, 1980, 23, 141-147.


**FOOTNOTE**

Biederman also suggested that we examine the error rates across kanji and kana conditions. We did keep the records of errors in each condition. Because of the tremendous amount of individual differences and the uncertainty of the nature of these errors, we did not analyze them systematically. However, the overall pattern is consistent with the argument that the kanji Stroop task is much more difficult than the kana Stroop task. The mean errors committed in the kanji and kana conditions are 5.42 and 2.75, respectively.