

Threshold-style processing of Chinese characters for adult second-language learners

YING LIU

Liaoning Normal University, Liaoning, China
and University of Pittsburgh, Pittsburgh, Pennsylvania

MIN WANG

University of Maryland, College Park, Maryland

AND

CHARLES A. PERFETTI

University of Pittsburgh, Pittsburgh, Pennsylvania

To assess the learning of word form and meaning in an unfamiliar writing system, we carried out primed-naming experiments with learners of Chinese at the end of their first and second terms in a Chinese class at an American university. The subjects were required to name a target Chinese character after a prime character had been presented for 500 msec. There were three priming conditions defined by the relation between the prime and the target: orthographically similar, homophonic, and semantically related. At the end of the first term, there was a significant facilitation for naming speed in the orthographic condition, but not in the homophonic or semantic conditions. However, at the end of the second term, orthographical facilitation disappeared. Instead, naming speed was facilitated by semantically related primes. A threshold-style framework was proposed to illustrate the processing of Chinese orthography, phonology, and semantics by second-language learners.

Learning to read a second language in a new writing system presents complex challenges. An English speaker learning to read Chinese must acquire knowledge of the visual forms of characters, knowledge of the mappings of these forms to meaning and pronunciation, and knowledge of the language itself. Research with college learners suggests that acquisition of the first of these—learning the visual form of characters—is fairly rapid. Students can discriminate novel legal characters from illegal ones within at least the first 4 months of classroom learning (Wang, Perfetti, & Liu, 2003). Beyond this learning of form, it becomes a question of the acquisition of character representations that include orthographic, phonological, and semantic constituents that can be activated by the character form. This is the question we address here.

The question must be considered in the context of how Chinese differs from English and other alphabetic orthographies. Chinese is considered a logographic, or morpho-syllabic, system (see DeFrancis, 1989; Mattingly, 1992) in which the units of the orthography (characters) correspond to both syllables and morphemes. The typical Chinese character is a square-shaped symbol that, with some exceptions, represents one pronunciation and one morpheme. The characters are composed of *radicals*. Some radicals are characters by themselves, and some are not. Characters containing only one radical are called *simple*

characters, and those containing more than one radical are called *compound characters*, which can contain two to eight radicals.

An example of a simple character is 日, which is pronounced /ri/4 (PinYin system, the Chinese national standard alphabetic system that is used mainly to teach children the pronunciation of Chinese characters; the digit at the end represents one of the four tones in Chinese) and means *sun*. An example of a compound character is 青 (/qing/1, *green*), which is composed of one top and one bottom radical. When putting 日 to the left of 青, we have the character 晴, which is pronounced /qing/2 and means *sunshine*. In the examples above, the meanings of 晴 and 日 are highly related, and the pronunciations of 晴 and 青 have the same onset and rhyme. However, the relationships are not always consistent and can be misleading (Feldman & Siok, 1999a, 1999b).

The radicals can be further decomposed into strokes. There are five classes of strokes (一, |, J, 丿, and ㇇). A spatial combination of strokes in specified ways makes a radical, and a specific combination of radicals makes a character. There are two standard character printing form sets used in contemporary Chinese: simplified and traditional. In the two sets, some character forms are identical, but other forms in the traditional set have more radicals and strokes than do the same characters in the simplified

set. The simplified set was used in the present study. The number of strokes in all simplified Chinese characters ranges from 1 to 30.

Although the visual complexity of Chinese writing, as compared with linear alphabetic writing, is a striking feature, a deeper difference is that it allows form mappings that go directly from orthographic form to meaning. Although some evidence suggests that the reading of Chinese works by a direct route to meaning (Hoosain, 1991), research now clearly supports the involvement of phonology in reading for meaning (see the review by Tan & Perfetti, 1998). Phonological effects are not found as often in simple orthographic tasks that do not require reading for meaning (Shen & Forster, 1999; Zhou & Marslen-Wilson, 1996) as in semantic relation judgment and meaning categorization tasks (Chua, 1999; Perfetti & Zhang, 1995; Xu, Pollatsek, & Potter, 1999).

Although the access of phonology may be universal, the details of its use must depend on the writing system. A major difference between Chinese and English for orthographic processing has been reported by Perfetti and Tan (1998), who used a primed-naming paradigm with Chinese materials. They found a graphical facilitation effect at a 43-msec prime–target asynchrony (stimulus onset asynchrony [SOA]), but facilitation turned into inhibition when the SOA was longer (57 msec). More interesting, the onset of graphical inhibition was accompanied by a phonological facilitation, with a semantic facilitation at 85 msec. No semantic or phonological facilitation was observed at the short SOA (43 msec). In English, by contrast, orthographic and phonological priming develop together. Using a primed perceptual identification paradigm, Perfetti and Bell (1991) found that orthographic facilitation slightly preceded phonological facilitation at 35 msec and that the two increased together through 65 msec. Using a lexical decision task with French stimuli, Ferrand and Grainger (1994) also found orthographic priming to be slightly earlier than phonological priming (14 vs. 29 msec), and both effects kept rising until 67 msec, when orthographic priming reached its ceiling, relative to unrelated primes. This contrast between the coupling of orthography and phonology in English and their decoupling in Chinese suggests a difference in how phonology is activated in the two systems. In an alphabetic system, the word-level units do not wait for a complete specification of all letter units prior to activation of word-level phonology—hence, the *cascade* processing style (Berent & Perfetti, 1995; Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001; Seidenberg & McClelland, 1989). But in Chinese, the word-level phonology is not activated prior to a full orthographic specification of the character—hence, the *threshold* processing style (Perfetti, Liu, & Tan, 2002, 2005).

This threshold assumption of Chinese orthography is very important for the computational lexical constituency model (Perfetti et al., 2005). In this model, successful lexical access requires the activation of three constituents: orthography, phonology, and semantics. Each constituent has a corresponding level consisting of many processing units. Each orthographic unit represents one character in the mental lexicon. A phonological unit represents an

onset, a rhyme, or a tone. A semantic unit represents the meaning of one character. The process begins with the activation of one or multiple orthographic units (characters) that correspond to the input from radical-level units. Only when an orthographic unit reaches its activation threshold level does activation feed forward to phonological and semantic units.

Because the lexical constituency model concerns only skilled reading, it is silent on how the learning of the constituents proceeds. Certainly, a learner comes to acquire knowledge about the orthographic form of characters in connection with their meanings and pronunciations. However, the completeness of constituent learning and the accessibility of the constituents can vary over learning. It is possible that priming experiments can provide some window on the degree of orthographic, phonological, and semantic learning. In fact, a primed-naming study has been done with children (native speakers of Chinese) who were learning to read Chinese. Wu, Zhou, and Shu (1999) found differential effects: phonological effects in primed naming for third- and sixth-grade children, but semantic effects for sixth-grade children only.

With adults learning Chinese as a foreign language, there is little reason to expect this pattern of phonological facilitation before semantic facilitation that has been observed in studies of both adults and children who were native speakers. Native speakers learn the orthographic form of Chinese characters after they already know most of the phonological expressions and their meanings. Second-language learners do not have this language foundation but, rather, acquire the language along with the writing system. The unfamiliarity of Chinese characters and their visual-spatial differences from left-to-right linear alphabets may place a premium on learning orthographic form. Thus, the threshold assumption of Chinese character reading will apply to learners, as well as to skilled readers. The orthographic form of a character must be acquired as a functional unit, and connections to pronunciation and meaning will develop as the character form is acquired. It is likely also that the demands of translating meanings in second-language learning put a premium on the semantic connections of the character over the pronunciation connections. If these observations are correct, we would expect that adult learners, like Chinese skilled readers, will read learned characters in threshold style, showing an orthographic facilitation effect when the processing time for the prime is less than that needed to reach its threshold. We also would expect semantic primes, which preactivate characters with similar meanings, to be effective earlier in learning than are phonological primes.

To examine these issues, we turned to the primed-naming task, which has been successfully used to examine the character-processing speed of native Chinese speakers (Perfetti & Tan, 1998). By studying the priming effect as a function of prime type over two time points in learning, we could observe changes in the accessibility of orthographic, phonological, and semantic constituents of characters over learning. For example, the rapid acquisition of the orthographic forms of characters ob-

served in our previous study suggested an implicit perceptual learning that might pay off in the acquisition of specific characters (Wang et al., 2003). If so, we would expect that relatively early in learning, preceding a target character with an orthographically similar prime should produce facilitation. Because orthographic processing involves threshold-style processing, phonological and meaning processes will be delayed until the moment of orthographic recognition (at threshold). Thus, priming from characters related in meaning and pronunciation may have less effect early in learning, relative to orthographic primes. With additional learning, a brief exposure of a character can become sufficient for its recognition, since the character comes to reach its orthographic threshold more quickly. Thus, any advantage of a preceding orthographic prime is eliminated, due to the competition between the prime and the target. For native speakers, a long enough SOA and an orthographic prime will even produce inhibition, as was found by Perfetti and Tan (1998). Also, learning will lead to stronger connections from orthography to meaning and pronunciation, allowing both semantic and phonological primes to become facilitative.

In a consideration of the effects of prime types, the prime–target SOA is critical (Perfetti & Tan, 1998), and the very short SOAs that produce orthographic effects in skilled native Chinese readers will not necessarily produce interpretable priming effects in learners. In the present experiments, because the number of items in the curriculum was very limited, only one SOA could be used. In a naming experiment in which the same group of subjects had been used (Wang et al., 2003), the naming reaction times (RTs) ranged from 1,381 msec for high-frequency simple characters to 2,134 msec for low-frequency compound characters. Compared with the 700- to 800-msec naming RTs for native speakers (Hue, 1992; Perfetti & Zhang, 1991), the learners were nearly 1 sec slower. Because a longer SOA would allow more processing time for the prime and lead to possible facilitation effects, we used a 500-msec SOA, much longer than the four SOAs used by Perfetti and Tan (1998).

We will present the study below as two experiments, one at the end of the first term and one at the end of the second term. The Results section presents the results for both terms.

EXPERIMENT 1

Method

Subjects. Twenty-six undergraduate students enrolled in an elementary Chinese class at the University of Pittsburgh participated in the experiment at the end of their first term (12–15 weeks learning, 12 h/week). The subjects' ages ranged from 19 to 28 years. Twenty-three subjects had English as their native language, and the other 3 had alphabetic writing system languages (Spanish, Thai, and Vietnamese). None of the subjects had been formally exposed to any Chinese environment before taking the class. All the subjects had normal or corrected-to-normal vision. They were paid for their participation.

The subjects above also had participated in the lexical decision and naming experiments of Wang et al. (2003). These two experiments were carried out before the present experiment. However,

because all 261 items from the curriculum were used in the decision and naming experiments, any repetition effect should be equal across the items used in the present experiment.

Materials and Design. The Chinese-corpus-based character frequency (Li & Liu, 1988) would not apply to these subjects very well, although they tended to start from more commonly used characters, as Chinese native speakers do. Instead, we created a computerized curriculum file based on the text book (Barnes, unpublished manuscript) and tallied the number of appearances for each character with a computer program. In total, there were 261 characters in the first-term curriculum that occurred from 3 to 287 times. Any character that appeared fewer than 3 times did not enter further into material selection. The teaching method used at University of Pittsburgh discouraged any additional study of Chinese beyond the textbook, which enabled our curriculum-based frequency to serve as a corpus-based frequency and provide a very good estimate of the character familiarity level. Furthermore, the subjective familiarity assessed with the same group of subjects was highly correlated with the curriculum-based frequency (Wang et al., 2003). All the subjects were enrolled in reading and writing sections of the course, which required them to recognize and write individual characters. So the subjects were learning not only multicharacter words, but also the pronunciation and meaning of individual characters. Thus, curriculum-based frequency provided a good measure of character familiarity and was used to select materials.

There were three groups of Chinese prime–target character pairs: orthographically similar, homophonic, and semantically related. There were 40 pairs of characters in each group, with 20 related pairs and 20 unrelated control pairs. The related and control pairs in each group shared the same 20 target characters (see the Appendix for all the related primes, the control primes, and their common targets). In an orthographically similar pair, the prime either shared a radical with the target (such as 还 [hai/2, *still*] and 这 [zhe/4, *this*]; total, 15 pairs) or had similar strokes and structures (such as 人 [ren/2, *person*] and 八 [ba/1, *eight*]; total, 5 pairs), with no phonological or semantic relation. Homophone primes had the same onset and rhyme as the target but had no orthographical similarity or semantic relation (such as 工 [gong/1, *work*] and 公 [gong/1, *public*]). Due to the limited vocabulary of the subjects, two types of orthographical similarity (shared radical or similar stroke) were used in the orthographic condition, and the tones were not always the same (7 same and 13 different pairs) in the homophonic condition. Semantically related primes were either in the same category as the target (such as 兄 [xiong/1, *elder brother*] and 弟 [di/4, *younger brother*]; 10 pairs) or often appeared together as a compound word in the curriculum (such as 国 [guo/2, *country*] and 家 [jia/1, *home*]; 10 pairs). There was no orthographical or phonological relation in the semantic pairs.

Unrelated character primes were used as the control baseline for priming effects. Because of the limited vocabulary, it was not possible to use both unrelated character controls and noncharacter controls. However, Perfetti and Tan (1998) found no significant difference between these two control conditions. Restricted to a single control baseline, we chose the unrelated character, because it had all three lexical constituents and matched the character prime better. For example, if a neutral noncharacter control with no pronunciation and meaning had been used for the orthographic-priming condition, any inhibition effect from the prime could have been due either to the orthographic similarity or to the lack of competition at phonological and/or semantic level in the neutral condition. Even for a facilitation effect, it is still possible that the facilitation is a *general* effect that can be obtained from any character that has sound and meaning.

The unrelated control characters in each group were matched with the prime characters on average curriculum frequencies, radical numbers, and stroke numbers. Average frequencies in the three groups were as follows: orthographic group, related prime = 36, control prime = 36, and target = 21; homophonic group, related prime = 49, control prime = 50, and target = 23; semantic group, related prime = 40, control prime = 38, and target = 33. The vi-

sual similarity of orthographic pairs and the semantic relatedness of semantic pairs were evaluated by 10 native Chinese speakers on a 5-point scale. The mean standardized similarity for orthographically similar pairs was 2.94 ($z = .70$), and for control pairs, it was 1.17 ($z = -.70$). The mean standardized relatedness score for semantic related pairs was 4.16 ($z = .88$), and for control pairs, it was 1.26 ($z = -.89$).

Because the total number of characters needed in the study was 180 and there were only 261 characters in the stimulus pool of the first-term learners, it was unavoidable that some target characters were used repeatedly in the experiment, but the repetition was reduced as best as we could. The repetition was 11 characters at 2.1 times each for graphical targets, 13 at 2.2 times each for homophonic targets, and 11 at 2.4 times each for semantic targets.

Procedure. The subjects were required to sit 80 cm in front of a 17-in. computer monitor working at 60 Hz in a controlled illuminated room. The stimuli were shown in the center of the screen in black on a white background, and the size was 1.5 cm wide and 1.5 cm high. A trial began with a 500-msec fixation, followed by a 500-msec presentation of the first character (prime). The prime was replaced by the second character (target), with no interval between them; the latter stayed on the screen until the subject pronounced it aloud. Font size was the same for the prime and the target. However, because the written position and style of the strokes and radicals varied from character to character, there was very little physical overlapping between the prime and the target, even in the orthographically similar condition. The naming response was detected by a voice key, and the time from the onset of the target to the naming response was recorded. Accuracy of naming was recorded by a native Chinese speaker right away. Practice with 10 trials was given before the experiment. Because the targets were presented to each subject twice, once in the related condition and once in the control condition, the 120 trials in the experiment were pseudorandomized for each subject, to balance the order of stimuli. As a result, half of the targets were seen first in the priming condition, and the other half were seen first in the control condition. The experiment lasted about 10 min, with no break.

EXPERIMENT 2

We retested the students at the end of their second term (total of 27–30 weeks of learning Chinese, 12 h per week) with the same materials and procedure as those in Experiment 1. By this second test point, the average curriculum frequency of the characters used in the present study had increased from 36 to 68. The average frequencies of the related prime and the control prime were still matched after

this increase (orthographic group, related prime = 65, control prime = 73, and target = 42; homophonic group, related prime = 90, control prime = 89, and target = 43; semantic group, related prime = 75, control prime = 71, and target = 66).

Method

Subjects. Eighteen students from the second-term Chinese class at the University of Pittsburgh participated in the experiment. These students were a subset of the students in Experiment 1 who continued their learning of Chinese in the second term.

Materials, Design, and Procedure. The materials, design, and procedure were the same as those in Experiment 1, with the materials rerandomized.

RESULTS OF EXPERIMENTS 1 AND 2

Subject accuracy was calculated by dividing the number of correctly named trials by the total number of trials in each condition for each subject (excluding fewer than 5% for microphone failures). Item accuracy for each item was calculated by dividing the number of subjects who named this item correctly by the total number of subjects. Only the naming RTs for correct responses were included in the analysis. For each subject, RTs shorter than 300 msec or more than two standard deviations longer than the subject mean were omitted from further analysis (3.8% in the first term and 4.6% in the second term).

Mean RTs and accuracy across all the subjects are shown in Table 1, which contains the results for the 26 first-term subjects and for the 18 subjects who participated in both terms. The priming effects on RTs (Figure 1) were calculated by subtracting RTs in the related condition from those in the control condition, but the priming effects on accuracy (Figure 2) were calculated by subtracting accuracy in the control from accuracy in the related condition. So, a positive RT difference represents a shortened RT, and positive accuracy difference represents increased accuracy.

Overall, accuracy was much higher and naming times were much shorter in the second term (18 subjects) than in the first term (26 subjects). In the first term, the priming effects on accuracy were negligible for all three conditions (0.2%, 0.4%, and -1.0%). There was a large orthographic

Table 1
Mean Reaction Times (RTs, in Milliseconds) and Accuracy (as Percentages) in Experiments 1 and 2 (With Standard Errors)

	No. of Subjects	Condition	Prime Pairs				Control Pairs			
			RT		Accuracy		RT		Accuracy	
			<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Term 1	26	Graphic	1,241	75	87.2	2.7	1,302	70	87.0	2.9
		Homophonic	1,305	77	79.5	4.0	1,300	79	79.1	3.8
		Semantic	1,279	75	85.5	3.2	1,268	69	86.5	2.7
	18	Graphic	1,292	87	90.2	2.6	1,353	80	90.1	2.7
		Homophonic	1,358	96	84.2	4.2	1,337	92	82.2	4.5
		Semantic	1,323	88	88.4	3.6	1,307	76	88.7	2.9
Term 2	18	Graphic	1,137	65	93.6	1.7	1,130	51	94.3	2.0
		Homophonic	1,160	74	88.5	3.8	1,186	72	86.0	4.9
		Semantic	1,085	51	91.3	2.5	1,151	56	92.0	2.4

Note—The second set of results for Term 1 is for the 18 subjects who continued the study of Chinese into the second term.

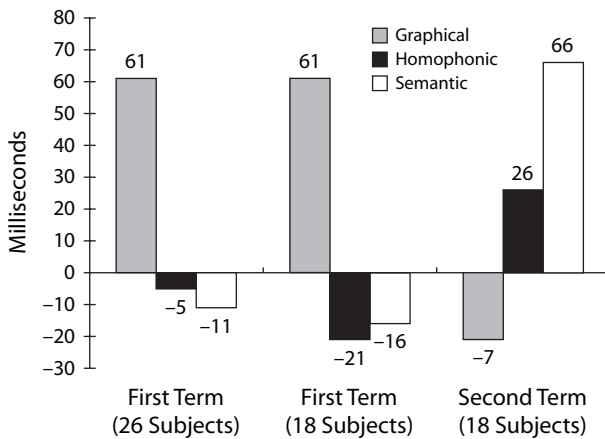


Figure 1. Priming effect on reaction times across terms.

facilitation (61 msec) for naming times, which represented faster naming when the prime and the target characters were orthographically similar. However, the homophonic and semantic effects on RTs were both negative and small (-5 and -11 msec).

In the second term, the priming effects on accuracy were also quite small (-0.7% and -0.7%), except in the homophonic condition (2.5%). For RTs, orthographic facilitation disappeared (-7 msec). Instead, there were homophonic (26 msec) and semantic (66 msec) facilitations.

To decide whether the observed effects were significant, subject and item ANOVAs were performed on the RTs and accuracy for both terms.

The 26 first-term subjects were analyzed with two within-subjects factors: condition (graphical, homophonic, or semantic) and priming (related or control). Greenhouse–Geisser adjustment for violation of variance equality was performed when needed. The ANOVA on accuracy showed a significant effect of condition [$F(2,50) = 7.605$, $MS_e = 0.012$, $p = .001$], but neither the priming effect [$F(1,25) = 0.033$, $MS_e = 0.002$, $p = .857$] nor the condition \times priming interaction [$F(2,50) = 0.639$, $MS_e = 0.001$, $p = .498$] was significant. The ANOVA on RTs

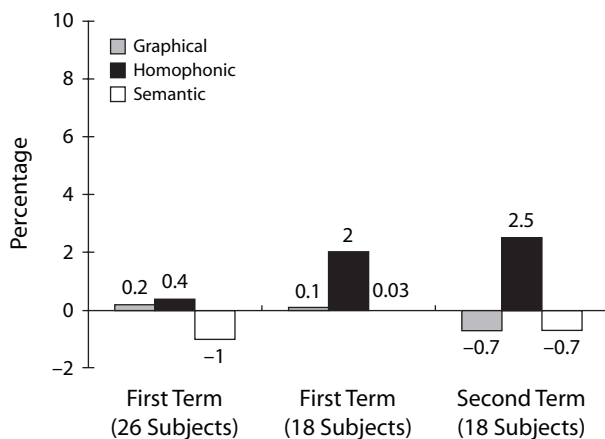


Figure 2. Priming effect on accuracy across terms.

showed a marginally significant condition \times priming interaction [$F(2,50) = 2.656$, $MS_e = 7,654.857$, $p = .080$].

Since we were interested mainly in the priming effects, three planned paired t tests were carried out between the three related conditions and their corresponding controls. It was found that only the 61-msec RT facilitation with orthographic priming was statistically reliable [$t(25) = -2.901$, $p = .008$]. An item analysis was consistent with the subject analysis, with a marginal graphical priming effect on RT [$t(19) = -1.897$, $p = .073$].

To look for changes across terms, we combined the 18 subjects who were in both the first and the second terms and did within-subjects ANOVAs with three factors: term (first or second), condition (graphical, homophonic, or semantic), and priming (related or control). Item ANOVAs were also carried out, using the data from the same 18 subjects, with two repeated factors (term and priming), because items were the same across terms. Since the three types of prime–target pairs used different target sets, condition was used as a between-items factor.

The subject ANOVA on accuracy showed a marginally significant effect of term [$F(1,17) = 3.377$, $MS_e = 0.07$, $p = .084$], a significant effect of condition [$F(2,34) = 7.129$, $MS_e = 0.02$, $p = .011$, $\epsilon = .612$], and a marginally significant condition \times priming interaction [$F(2,34) = 2.917$, $MS_e = 0.001$, $p = .095$, $\epsilon = .628$]. The item ANOVA on accuracy showed only a significant effect of term [$F(1,17) = 35.187$, $MS_e = 0.003$, $p = .000$].

Planned pairwise t tests were performed on accuracy between each related condition and its control condition. The only significant one was an item analysis of homophonic priming in the second term [$t(19) = 2.155$, $p = .044$].

The subject ANOVA on RTs showed a marginally significant effect of term [$F(1,17) = 4.253$, $MS_e = 442,302.570$, $p = .055$], a marginally significant effect of condition [$F(2,34) = 2.618$, $MS_e = 14,126.808$, $p = .088$], and a significant three-way term \times condition \times priming interaction [$F(2,34) = 3.914$, $MS_e = 7,149.924$, $p = .030$]. The item ANOVA on RTs showed a significant effect of term [$F(2,1) = 43.673$, $MS_e = 6,604.962$, $p = .000$] and a significant term \times priming \times type interaction [$F(57,2) = 3.209$, $MS_e = 9,879.772$, $p = .048$].

Planned paired t tests between each related condition and its control showed that in the first term, the 18 subjects who also participated in the second term had a pattern similar to that for all 26 subjects: a reliable 61-msec graphical facilitation [$t(17) = -2.146$, $p = .047$]. In the second term, the only significant difference was a 66-msec semantic facilitation effect [$t(17) = -2.880$, $p = .01$]. The 26-msec phonological facilitation did not reach a significant level [$t(17) = 0.701$, $p = .493$]. For item RTs, the result for the first term was consistent with that of the subject analysis. It showed a marginally significant orthographic priming effect on RT [$t(19) = -1.897$, $p = .073$]. However, no t test, not even that for semantic priming [$t(19) = -1.654$, $p = .115$], was significant for the second term.

A post hoc ANOVA was performed to compare the priming effects on RTs in the first term between the two

types of orthographically similar pairs: no shared radical versus shared radical. The mean priming effect for similar stroke and structure pairs that did not share radicals (five pairs; mean z score of relatedness, .78) was 155 msec. The mean priming effect for the shared radical orthographic pairs (15 items; mean z score of relatedness, .70) was 33 msec. There was a significant radical \times priming interaction [$F(1,25) = 5.113$, $MS_e = 18,870.74$, $p = .033$].

A second post hoc ANOVA on accuracy for the tone-same and tone-different homophonic pairs was performed in order to look for a potential effect of tone in the second term. Tone was used as a between-items factor, and priming was a within-items factor. The mean accuracy priming effect for the tone-same pairs (7 pairs) was 1.5%, and that for the tone-different pairs (13 pairs) was 3.1%. The interaction of tone and priming was not significant [$F(1,18) = 0.404$, $MS_e = 0.001$, $p = .533$].

DISCUSSION

The results showed a clear character-processing speed difference between the two consecutive terms. At the end of the first term, there was an orthographic facilitation of RT for naming when the prime and the target looked similar. However, at the end of the second term of learning, the orthographic facilitation disappeared. Instead, there was a facilitation of meaning for naming speed (shorter naming RT for the semantically related pairs) and a phonological facilitation for naming accuracy (higher accuracy for the homophonic pairs, but significant only in the item analysis). A similar graphical facilitation had been found in Chinese native speakers by Perfetti and Tan (1998), but only at the shortest SOA (43 msec) in their study. Shen and Forster (1999) found graphical facilitation at a 50-msec SOA. It is interesting to note that Chinese native speakers did not show graphical facilitation at longer SOAs (57, 86, and 115 msec in Perfetti & Tan, 1998) but actually showed inhibition effects at two of them (57 and 86 msec). According to the interactive constituency model (Perfetti et al., 2002, 2005), the speed with which the orthographic threshold is accessed determines for which SOA graphical facilitation can be observed. For our learners, a slower graphical analysis speed for Chinese characters allowed facilitative orthographic effects to emerge over a longer time window (500 msec).

Due to the limitation of materials and subjects, we did not manipulate SOA in the present study. The results showed that the fixed 500-msec SOA served as a short SOA for the first-term learners, because the graphical analysis of the prime character could not be completed within 500 msec, which facilitated the processing of the target character. Instead, in the second term, exactly the same SOA functioned as a relatively longer one, and the meaning of the prime could be accessed within that time. It is clear that there exists a critical time point that determines whether graphical preprocessing of the prime can facilitate target processing. Furthermore, this time point was different for the first-term learners, the second-term learners, and the native Chinese speakers. Even though we did not obtain any orthographic inhibition, the van-

ishing of the orthographic effect at the end of the second term suggested that the threshold-style processing applied also to those learners whose native writing systems were alphabetic. Accompanying the null effect of orthographic pairs were semantic and phonological effects. This overall pattern can be seen to reflect a threshold-sensitive processing by the learners. In the first term, 500 msec functioned as subthreshold, showing orthographic priming and only orthographic priming. By the second term, 500 msec functioned as above threshold, showing semantic and (less reliable) homophonic priming without orthographic priming. It is this pattern of disassociation of orthography from the effects of the other constituents that we take as the signature of a threshold-based lexical system.

The reasons that Chinese is a threshold system arise from the essential properties of the character system—specifically, the lack of a subsyllabic structure within the character. There is no part of the character that corresponds to the grapheme–phoneme mapping that is the essence of alphabetic writings. (It is a rather different matter that characters have constituent radicals that can function in reading. For further discussion, see Perfetti et al., 2005.) Thus, the orthographic character unit functions as an on–off gate to lexical processing, with no incremental buildup of grapheme–phoneme connections. Learners' developing lexical systems (as well as the developed lexical system of skilled readers) function in this threshold style because learners acquire characters as the unit of processing. They learn quickly that there is nothing in the character corresponding to the letter–sound connections they experience in the alphabetic systems of their native language. Rapid perceptual learning (Wang et al., 2003) gives them a sense of the visual form of characters, and practice with specific characters brings the characters, more or less one at a time, into a functional lexicon. The connections between each character and its pronunciation and meaning are learned with the character, but they may not reach the strength necessary to function automatically in character recognition.

Contrary to Chinese reading, it has been found that for second-language learning of alphabetic writing systems (French or Dutch), phonology has been as fast as orthography, which suggested a prelexical activation of phonological codes for the second-language words (Brysbart, Van Dyck, & Van de Poel, 1999; Van Wijndaele & Brysbart, 2002). So the separation of orthography and phonology by the orthographic threshold has been found only in nonalphabetic word recognition.

A Processing Framework for Chinese as Second-Language Learners

To illustrate these proposals about the course of learning, we will draw on a model of skilled Chinese reading developed by Perfetti et al. (2002, 2005), which simulates character naming, using the threshold assumption. An adaptation of this model for learners (the lexical constituency model) is similar in structure and processing assumptions to the version for native readers. Because we do not set parameters or report a simulation, we consider this to be

a framework for a learner model. The learner framework differs from the model for skilled reading in two major ways: a much smaller vocabulary and slightly stronger connection weights from orthography to meaning than from orthography to phonology. The second assumption reflects the slightly stronger semantic priming effect for the second term in our data; however, if learning was directed more at pronunciation than at meaning, the weight differences would be reversed. Unlike the critical threshold assumption, the connection weights can be considered to be empirically determined parameter settings.

In the absence of an implementation of a model, we will use the framework, as illustrated in Figure 3, as a way to picture a coarse-grained activation function for a learner. A visual input begins the process of character recognition, sending stroke and position information to the orthographic level. Each unit in the orthographic level corresponds to one learned character. Each input combination of strokes can send not only a full activation to the corresponding character unit in the orthographic level, but also partial activation to all the character units that are orthographically similar. After the threshold of any orthographic unit has been reached, it sends inhibition to all other orthographic units and activation to the phonological and semantic levels. An important feature of the orthographic level is that the number of units increases continuously with the number of characters learned.

The phonological level is a distributed representation level that contains 63 units according to the Chinese national standard PinYin system: 23 onsets (including a null onset), 34 vowels, and 5 tones (including a neutral tone). The combinations of the onsets, vowels, and tones above are sufficient to represent all the syllables of Mandarin Chinese. All the learners acquired this phonological knowledge during the first 2 weeks of their curriculum. As a result, all the phonological units were fully developed before the end of the first term. Because this level is a distributed representation, there are no within-level linkages, in contrast to the orthography and semantic levels.

Character meaning is a localized representation, each such representation corresponding to a unique meaning of a single character that has been learned (represented by English translation in the present model). The semantic level currently does not contain representations of sub-lexical semantic features. Semantic relations are reflected by between-character connections. There are bidirectional connections between the corresponding phonological and semantic units.

In a naming task, the visual input gradually activates a group of orthographic units, including the presented character and other orthographically similar characters. The orthographic unit of the presented character receives most of the activation, if the subject identifies the character correctly. As a result, that unit reaches the threshold before other orthographic units and starts to send inhibition to other units, which make the correct character unit a “winner.” Then the “winner” unit activates its corresponding phonological and semantic units. Because the semantic units have internal connections with semantically related

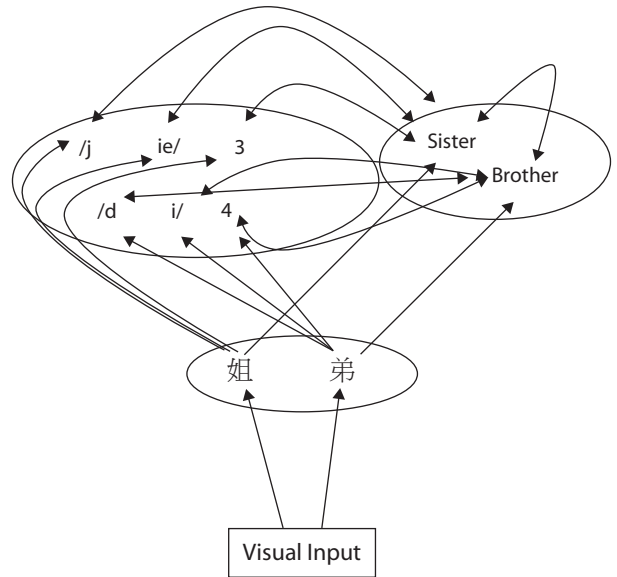


Figure 3. Framework for processing Chinese as a second language.

characters, a cohort of semantically related characters are activated and start to activate their phonology. Because the onset, rhyme, and tone units for the presented character receive most of the activation from the orthographic and semantic levels, they reach the threshold before the other phonological units do, and the subject does the naming correctly.

When the task is a primed-naming task, there is a short exposure of a prime character before presentation of the target character. The priming effect (both facilitation and inhibition) is a result of the preactivation of certain units in the model. For orthographically similar priming, the processing time needed for reaching the orthographic threshold is crucial for observing a priming effect. At the end of the first term, the orthographic threshold of the prime could not be reached within the 500-msec SOA. Because the target is orthographically similar to the prime, there is a preactivation of the orthographic unit for the target character before the target character is presented, which makes the activation speed for the target faster than that for the unrelated control. As a result, when processing speed for the prime is slow, the phonological units receive activation earlier and reach their threshold more quickly in the orthographically similar condition than in the control condition.

However, at the end of the second term, the processing speed at the orthographic level was much faster (130 msec faster than the RTs in the control conditions), and the orthographic unit of the prime could reach its threshold within 500 msec. At the moment the threshold is reached, it starts to send inhibition to all other units and depresses all the preactivations of orthographically similar units at the orthographic level. As a result, there is no facilitation effect on the target anymore. Furthermore, at a specific SOA, there can be inhibition effects, depending on how strong the competition is between the prime and the tar-

get orthographic units (Perfetti et al., 2002, 2005), even though it is not clear whether this competition is strong enough for inhibition to be observed in learners. It is possible that the strength of competition depends on the proficiency of reading.

After the orthographic threshold has been reached, the orthographic unit of the prime sends activation to its pronunciation and meaning units. In the condition in which the prime and the target had the same onset and rhyme (and sometimes, tone), a preactivation for the target pronunciation helped the target to be correctly named. A trend of shorter naming latencies (26 msec) was also observed. In the semantically related condition, the semantic unit of the prime partly activates the target unit; then the latter preactivates the exact onset, rhyme, and tone combination at the phonological level, which leads to a shorter naming latency.

Consistent with the fast learning of graphical structure (Wang et al., 2003), the learners had acquired graphical form information by the end of the first term, and their orthographic representations were sufficient to produce facilitation within a 500-msec SOA. Even though neither phonological nor semantic connections were functional within 500 msec during the first term, the learners did have access to pronunciations and meanings of the target character when allowed ample processing time, since these same subjects were able to perform the pronunciation spelling (PinYin) and translation tasks very well (Wang et al., 2003).

The present framework does not include a separate radical level, because radicals were not taught explicitly in the curriculum followed by our subjects. However, radicals not only play an important role in the reading of native Chinese speakers (Ding, Peng, & Taft, 2004; Taft & Zhu, 1997), but also help alphabetic users to learn Chinese (Taft & Chung, 1999; Wang, Liu, & Perfetti, 2004). The post hoc analysis showed that the five orthographically similar pairs without shared radicals had a stronger facilitation effect than did those with shared radicals. This result was opposite to that found with native speakers (Ding et al., 2004). It is possible that stroke and spatial analysis dominate the early processing of first-term learners. However, this result was based on five items only. There is no doubt that a learner will develop some knowledge of radicals at a certain learning stage, either explicitly or implicitly. But the length of learning needed is still an unresolved issue.

As is illustrated at the semantic level in Figure 3, it is quite feasible that English translations were used to mediate the accessing of the meaning of these Chinese characters, as with other second languages (Kroll, Michael, & Sankaranarayanan, 1998; Kroll & Sholl, 1992). It also has been found that Chinese speakers who were learning English showed significant priming from an English prime to a homophone of a semantically related Chinese character, which showed that Chinese translation mediated the access of meaning in the reading of English for English learners (Guo & Peng, 2003).

The present results do not provide clear evidence on whether phonology or semantics was processed faster for

these learners. (A post hoc *t* test between the priming effects of phonology and semantics in the second term was not significant; however, semantic priming effects were reliable for naming speed in the subject analysis, whereas phonological effects were reliable only for naming accuracy in the item analysis.) A reasonable assumption is that because the learners were exposed to connections from orthography to phonology and from orthography to meaning nearly simultaneously, they developed the two connections at a comparable rate. As their Chinese spoken proficiency reached a higher level, which relied on robust connections from phonology to meaning, we would expect more phonological involvement in meaning processes in word reading.

Our results were also partly confirmed by an event-related potential (ERP) study of learners from the same university population (Liu, Perfetti, & Wang, 2006). In a task of delayed naming, this study showed a larger N200 (orthographic) component for Chinese materials than for English at the occipital electrodes, but only in the first term of learning. By the second term, this orthographic effect was absent; instead, a larger N400 (interpreted as a semantic and phonological component) was observed.

A general picture has emerged on the basis of three studies of the same population of learners, using different methods: the present priming study, the ERP study (Liu et al., 2006), and an fMRI study of naming that showed similar activation patterns for Chinese native speakers and Chinese second-language learners (Nelson, Liu, Fiez, & Perfetti, 2003). Although learners cannot approximate the skill level of the native speaker in 1 year of learning, their character processing within this year begins to show the acquisition of the visual-graphic, semantic, and phonological components that function for the skilled native reader.

CONCLUSION

After one term of instruction (15 weeks of learning), adult learners of Chinese as a second language showed orthographic priming effects and only orthographic effects with 500 msec of prime exposure, suggesting that they had acquired functional orthographic representations of the characters. After another term, these orthographic effects gave their way to semantic and (weaker) phonological effects. This pattern, we suggest, reflects the threshold nature of the character lexicon, in which character recognition is achieved by an activation that reaches a character-specific threshold. First-term orthographic effects, on this account, reflected prethreshold activation of orthographic form; second-term semantic and phonological effects reflected lowering of the orthographic threshold, which allowed a postthreshold activation of semantics and phonology to occur within 500 msec. Like skilled readers, learners of Chinese must acquire character-specific representations that are accessed in threshold style. Learning—practice with specific characters—lowers the threshold of characters, allowing semantic and phonological connections to be accessed.

AUTHOR NOTE

This research was supported by National Science Foundation Grants BN0113243 and SBE-0354420 and National Natural Science Foundation of China Grant 30470567. We thank Dayle Barnes for allowing us to use his teaching materials in the present study. Correspondence concerning this article should be addressed to Y. Liu, Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Room 645, Pittsburgh, PA 15260 (e-mail: liuying@pitt.edu).

Note—This article was accepted by the previous editorial team, when Colin M. MacLeod was Editor.

REFERENCES

- BERENT, I., & PERFETTI, C. A. (1995). A rose is a REEZ: The two-cycles model of phonology assembly in reading English. *Psychological Review*, **102**, 146-184.
- BRYLSBAERT, M., VAN DYCK, G., & VAN DE POEL, M. (1999). Visual word recognition in bilinguals: Evidence from masked phonological priming. *Journal of Experimental Psychology: Human Perception & Performance*, **25**, 137-148.
- CHUA, F. K. (1999). Phonological recoding in Chinese logograph recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 876-891.
- COLTHEART, M., RASTLE, K., PERRY, C., LANGDON, R., & ZIEGLER, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**, 204-256.
- DEFRANCIS, J. (1989). *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii Press.
- DING, G., PENG, D., & TAFT, M. (2004). The nature of the mental representation of radicals in Chinese: A priming study. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **30**, 530-539.
- FELDMAN, L. B., & SIOK, W. W. T. (1999a). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory & Language*, **40**, 559-576.
- FELDMAN, L. B., & SIOK, W. W. T. (1999b). Semantic radicals in phonetic compounds: Implications for visual character recognition in Chinese. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 19-35). Mahwah, NJ: Erlbaum.
- FERRAND, L., & GRAINGER, J. (1994). Effects of orthography are independent of phonology in masked form priming. *Quarterly Journal of Experimental Psychology*, **47A**, 365-382.
- GUO, T., & PENG, D. (2003). The accessing mechanism of the less proficient Chinese-English bilinguals' conceptual representation. *Acta Psychologica Sinica*, **35**, 23-28.
- HOOSAIN, R. (1991). *Psycholinguistic implications for linguistic relativity: A case study of Chinese*. Hillsdale, NJ: Erlbaum.
- HUE, C. W. (1992). Recognition processes in character naming. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (Advances in Psychology, Vol. 90, pp. 93-107). Amsterdam: North-Holland.
- KROLL, J. F., MICHAEL, E., & SANKARANARAYANAN, A. (1998). A model of bilingual representation and its implications for second language acquisition. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 365-395). Mahwah, NJ: Erlbaum.
- KROLL, J. F., & SHOLL, A. (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (Advances in Psychology, Vol. 83, pp. 191-204). Amsterdam: North-Holland.
- LI, G., & LIU, R. (1988). *A dictionary of Chinese character information*. Beijing: Science Press.
- LIU, Y., PERFETTI, C. A., & WANG, M., (2006). Visual analysis and lexical access of Chinese characters by Chinese as second language readers. *Language & Linguistics*, **7**, 637-657.
- MATTINGLY, I. G. (1992). Linguistic awareness and orthographic form. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (Advances in Psychology, Vol. 94, pp. 11-26). Amsterdam: North-Holland.
- NELSON, J., LIU, Y., FIEZ, J. A., & PERFETTI, C. A. (2003, April). *Bilinguals' perception of words and word-like stimuli: A comparison of Chinese and English word form areas*. Poster presented at the 10th Annual Meeting of the Cognitive Neuroscience Society, New York.
- PERFETTI, C. A., & BELL, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and priming. *Journal of Memory & Language*, **30**, 473-485.
- PERFETTI, C. A., LIU, Y., & TAN, L. H. (2002). How the mind meets the brain in reading: A comparative writing systems approach. In H. S. R. Kao, C.-K. Leong, & D.-G. Gao (Eds.), *Cognitive neuroscience studies of the Chinese language* (pp. 36-60). Hong Kong: Hong Kong University Press.
- PERFETTI, C. A., LIU, Y., & TAN, L. H. (2005). The lexical constituency model: Some implications of research on Chinese for general theories of reading. *Psychological Review*, **112**, 43-59.
- PERFETTI, C. A., & TAN, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 101-118.
- PERFETTI, C. A., & ZHANG, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 633-643.
- PERFETTI, C. A., & ZHANG, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 24-33.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-568.
- SHEN, D., & FORSTER, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language & Cognitive Processes*, **14**, 429-459.
- TAFT, M., & CHUNG, K. (1999). Using radicals in teaching Chinese characters to second language learners. *Psychologia*, **42**, 243-251.
- TAFT, M., & ZHU, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **23**, 761-775.
- TAN, L.-H., & PERFETTI, C. A. (1998). Phonological codes as early sources of constraint in Chinese word identification: A review of current discoveries and theoretical accounts. *Reading & Writing*, **10**, 165-200.
- VAN WIJNENDAELE, I., & BRYLSBAERT, M. (2002). Visual word recognition in bilinguals: Phonological priming from the second to the first language. *Journal of Experimental Psychology: Human Perception & Performance*, **28**, 616-627.
- WANG, M., LIU, Y., & PERFETTI, C. A. (2004). The implicit and explicit learning of orthographic structure and function of a new writing system. *Scientific Studies of Reading*, **8**, 357-379.
- WANG, M., PERFETTI, C. A., & LIU, Y. (2003). Alphabetic readers quickly acquire orthographic structure in learning to read Chinese. *Scientific Studies of Reading*, **7**, 183-207.
- WU, N., ZHOU, X., & SHU, H. (1999). Sublexical processing in reading Chinese: A development study. *Language & Cognitive Processes*, **14**, 503-524.
- XU, Y., POLLATSEK, A., & POTTER, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 838-857.
- ZHOU, X., & MARSLER-WILSON, W. (1996). Direct visual access is the only way to access the Chinese mental lexicon. In G. Cottrell (Ed.), *Proceedings of the 18th Annual Conference of the Cognitive Science Society* (pp. 714-719). Mahwah, NJ: Erlbaum.

APPENDIX

Graphically Similar		Homophone		Semantically Related	
Prime (Control)	Target	Prime (Control)	Target	Prime (Control)	Target
子(左)	了	八(三)	吧	姐(马)	弟
还(好)	这	到(张)	道	教(级)	学
间(路)	问	德(韩)	得	说(他)	口
俄(师)	我	哥(京)	个	书(左)	写
小(下)	少	工(二)	公	父(分)	母
朋(法)	明	会(现)	回	飞(五)	走
吃(外)	吗	姐(朋)	结	明(韩)	亮
人(文)	八	里(学)	离	吃(外)	饿
姓(点)	她	两(您)	亮	好(们)	差
知(用)	和	六(五)	流	高(离)	大
大(几)	天	美(家)	没	北(只)	南
四(两)	回	年(姓)	念	老(那)	小
在(也)	车	是(你)	师	看(点)	见
上(方)	长	十(小)	时	年(太)	月
口(久)	只	有(他)	友	朋(俄)	友
亮(道)	京	在(国)	再	多(再)	少
忙(时)	快	坐(她)	昨	回(公)	去
走(因)	起	大(北)	打	忙(这)	累
学(海)	字	见(几)	间	你(是)	我
流(飞)	没	气(马)	起	家(也)	国

(Manuscript received May 18, 2005;
revision accepted for publication November 21, 2005.)