Journal of Experimental Psychology:
Learning, Memory, and Cognition
1999, Vol. 25, No. 4, 858-875

Whole Word, Frequency-General Phonology in Semantic Processing of Chinese Characters

Sulan Zhang, Charles A. Perfetti and Hui Yang
University of Pittsburgh

Two experiments were conducted to examine the locus of the phonological and semantic interference effects in Chinese reported by C. A. Perfetti and S. Zhang (1995a). Both experiments found that phonological effects in semantic judgments were general across word frequency and independent of component phonology, thus localizing phonological interference at the whole-character level. However, effects of component phonology were obtained for low-frequency characters when pronunciation judgments were made. The results are consistent with the assumption that phonology is activated at 2 levels in Chinese reading, 1 at the whole character level and 1 at the phonetic component level. A strong semantic interference was also found in both experiments, implying that both phonology and semantics are automatically activated. The constituency model provides a framework to explain the pervasive word-level phonology seen across writing systems.

Evidence for phonology in Chinese word reading has accumulated in recent years across a variety of word reading tasks (Chua, 1995; Lam, Perfetti, & Bell, 1991; Leck, Weekes, & Chen, 1995; Perfetti & Tan, 1998; Perfetti & Zhang, 1991, 1995a; Tan, Hoosain, & Peng, 1995; Tan, Hoosain, & Siok, 1996; Tan & Perfetti, 1997; Zhang, Feng, & He, 1994). Of particular theoretical interest has been the demonstration of phonological involvement in semantic tasks for which it is unneeded and, in fact, counterproductive to the goals of the task. Perfetti and Zhang (1995a) provided such a demonstration in a task that required participants to make judgments of meaning similarity. In this task, participants were presented with two single-character words, one after another, and required to judge whether they had similar meanings (synonyms). On key foil trials, the two characters were not semantically related but were segmental homophones—that is, characters with the same sequence of phonemes without regard for tone. Compared with unrelated foil trials, these homophone foils produced large interference effects, observable both in decision times and error rates. These phonological interference effects were observed in two experiments over a range of stimulus onset asynchronies (SOAs) between 90 and 310 ms. In a parallel task, participants made judgments about pronunciation sameness. Critical trials were foils in which the two characters had no phonological similarity but were semantically similar (synonyms). Interference effects—semantic interferences in this case—were found in this task as well. These semantic interference effects, however, were slower to emerge, at 140 ms, rather than at 90 ms SOA. Perfetti and Zhang (1995a) cautioned against making conclusions about the relative time course of semantic and phonological information by comparisons across these tasks; however, they concluded that the data were clear in demonstrating that phonological information was rapidly and automatically activated by the presentation of single characters.

Unanswered in these experiments is a question of some theoretical interest. The phonological effects observed by Perfetti and Zhang (1995a) were based on whole-character pronunciation. That is, the interference from phonology was defined to refer to the pronunciation of each character, not the pronunciation of its components. However, most single characters in modern Chinese usage are compound characters, containing two or more distinct components. A recent dictionary study found that 85% of the 9,641 characters in the Modern Chinese Dictionary (1992) are phonetic compounds, composed of two functionally independent components (Perfetti & Tan, 1999). One component (often called the "significate" or the semantic radical) may signal the meaning category of the whole compound, while a second component (the "phonetic") may signal something about the pronunciation of the compound. Significates are typically, but not always, on the left half of the compound; phonetics, typically but not always, are on the right. In most cases, the phonetic is also a character that can stand alone and thus has its own pronunciation and meanings. Recent research suggests that both a component's function—its status as a significate or a phonetic—and its position may play roles in lexical decision tasks, although the exact nature of these roles is far from clear (Feldman & Siok, 1997; Taft & Zhu, 1997).

Although most characters in the Perfetti and Zhang

Sulan Zhang, Charles A. Perfetti, and Hui Yang, Department of Psychology and Learning Research Development Center, University of Pittsburgh. Sulan Zhang is now at Monterey Technologies. Hui Yang is now at Beijing Normal University.

The research reported in this article was supported in part by National Science Foundation Grants SBR-9223125 and SBR-9616519 to Charles A. Perfetti. Experiment 1 was part of Sulan Zhang's doctoral dissertation. The authors wish to acknowledge the help of Liu Ying in the final preparation of the manuscript.

Correspondence concerning this article should be addressed to Charles A. Perfetti, 644 LRDC, 3939 O'Hara Street, Pittsburgh, Pennsylvania 15260. Electronic mail may be sent to perfetti@vms.cis.pitt.edu.

(1995a) experiments were compounds of this type, the relationship between the pronunciation of the phonetic component and the compound was not systematically varied. Thus, their conclusion generalizes over the most common type of single character word, but it is silent on the possible role of phonetic components in producing phonological interference. In fact, effects of phonetic components have been observed in naming experiments, at least for low-frequency characters (Fang, Horng & Tzeng, 1986; Hue, 1992; Seidenberg, 1985; Yang & Peng, 1997). And a study by Flores D'Arcais, Saito, and Kawakami (1995) found that when the components of compound characters were presented asynchronously, naming was facilitated only when the phonetic (and not the significate) was presented first, suggesting that the name of the phonetic, which was the same as the name of the character, became available before the name of the whole character. It is not at all clear, however, whether such phonetic component effects should occur when the reader's task is to evaluate meaning. The discovery of such effects in a meaning task would certainly implicate a very general "sublexical" phonological process in reading.

In the two experiments reported below, we address these questions of component phonology and character frequency. We ask whether interference effects can be localized at the component or whole-character level and whether they are restricted to characters of low-frequency or are general across frequency.

Experiment 1

A major goal in Experiment 1 was to examine two important questions about the role of phonology in semantic decisions. First, is the main source of the phonological interference effect at the level of whole-character phonology or at the level of phonetic components of characters? Second, is phonological interference (at either the whole-character or the component level) general over word frequency or is it restricted to low-frequency characters? An important secondary goal is to examine whether component phonology, whatever its role in semantic processing, plays a role when readers must perform an implicit phonological task, judgments of pronunciation.

The key manipulation in both tasks involves the characteristics of the foil trials. The logic follows Perfetti and Zhang (1995a): In meaning judgment, critical foil trials consist of pairs of characters that have no meaning similarity, but are homophones. However, within this set of critical homophone trials there are two kinds of homophones: Some homophone characters are valid phonograms, compound characters that contain a component (typically one that can stand alone as a character) having the same pronunciation as the character as a whole. Following Perfetti, Zhang, and Berent (1992), we refer to this class of characters as valid phonetic compounds. Other homophones are characters without a phonetic component that shares their pronunciation. These characters can have one of several configurations: They can be phonograms, containing a phonetic component (an invalid phonetic) that happens not to have the same pronunciation as the whole character in this particular case; they can be semantic compounds, containing two components, neither of which is classified as a phonetic nor provides any cue to pronunciation; or they can be simple characters, having no components. Thus, this set of characters without valid phonetic components comprises a varied class of characters whose character-level pronunciations must be retrieved on the basis of the whole character. We refer to these collectively as characters with uncued phonology.

Table 1 illustrates the structure of characters with valid phonetic components. The leftmost character is a whole character consisting of two components, one of which is a valid phonetic. To the right of the whole character, the phonetic is repeated in parentheses for clarity. The top rows of the table show two homophones with valid phonetics. Thus, the character 俘 and the character 腹 are both pronounced "fu"; the first character is translatable as *to capture* and the second, as *belly*. In a meaning judgment, the correct response would be "no" (whereas in a homophone judgment, the correct response would be "yes"). But both characters have a component (the phonetic) on the right side that, as a free standing character, is also pronounced "fu." Note that this component is not the same (visually) in the two cases. Thus, the two characters have two different phonetics, both pronounced "fu" (disregarding tone). The middle row of Table 1 shows the control foils, with the core character "fu" (*to capture*) now paired with a control character pronounced "ma." The bottom row shows the core character paired with its synonym cohort, "qin." Finally, note that the phonetic components in Table 1 are on the right side of the compound character. This left–right arrangement

Table 1
Experiment 1. Examples of Materials With Valid Components

Character type	Whole character (component)	Pronunciation: whole character (component)	Translation: whole character	Correct synonym judgment	Correct homophone judgment
Homophone	俘 (孚)	/fu/ (/fu/)	to capture		
	腹 (复)	/fu/ (/fu/)	belly	no	yes
Control	俘 (孚)	/fu/ (/fu/)	to capture		
	玛 (马)	/ma/ (/ma/)	agate	no	no
Synonym	俘 (孚)	/fu/ (/fu/)	to capture		
	擒 (禽)	/qin/(/qin/)	to capture	yes	no

Table 2
Experiment 1. Examples of Materials of Uncued Phonology

Character type	Whole character (component)	Pronunciation: whole character (component)	Translation: whole character	Correct synonym judgment	Correct homophone judgment
Homophone	(谷)	/yu/ (/gu/)	bath		
	郁 (有)	/yu/ (/you/)	gloomy	no	yes
Control	浴 (谷)	/yu/ (/gu/)	bath		
	杜 (土)	/du/ (/tu/)	to prevent	no	no
Synonym	浴 (谷)	/yu/ (/gu/)	bath		
	澡 (澡)	/zao/	bath	yes	no
Homophone	累	/lei/	tired		
	泪	/lei/	tears	no	yes
Control	累	/lei/	tired		
	朝	/chao/	forward	no	no
Synonym	累	/lei/	tired		
	乏	/fa/	tired	yes	no

with the phonetic on the right is the most frequent configuration of components, with a minority of compounds having either vertical configurations or horizontal configurations with the phonetic on the left.

Table 2 follows the same form, but here the characters are of uncued phonology type. The top half of the table shows a pair of homophones with invalid phonetics, components whose pronunciations diverge from that of the whole character. Thus 浴 (/yu/, *bath*) and 郁 (/yu/, *gloomy*) are homophones. But the first contains a phonetic component that by itself is pronounced "gu," and the second contains a phonetic component that by itself is pronounced "you." In the lower half of Table 2 is an example of two different types of uncued phonology, semantic compounds and simple characters. For example, the core character 累 (/lei/, *tired*), its homophone 泪 (/lei/, *tears*) and its control character 朝 (/chao/, *forward*) are all semantic compound characters; the synonym character 乏 (/fa/, *tired*) is a simple character. In each case, the phonology is uncued by a component and is available only from the whole character.

Thus, if phonological processes are carried by components, the characters with valid components should behave differently from characters with uncued phonology. In particular, if phonological interference is due to component phonology, then interference should be greater for homophones containing valid phonetic components than for homophones with uncued phonology. If the phonological interference effect is due to whole-character phonology, the presence of a valid phonetic should not matter.

Consider now the task in which participants judge pronunciation rather than meaning sameness. This task has been used to expose meaning interference; however, for the present issue of component phonology, it takes on an additional role: If the pronunciation of a phonetic component is available and useful during these judgments of pronunciation, then characters with valid components should be evaluated more quickly than characters with uncued phonology. Such a result would be interesting in comparison to naming task results, because it would confirm a component effect when phonology is only implicitly activated rather than required for articulation.

Finally, we examine the frequency of the characters in these tasks. There are two questions to ask: Is phonology activated at the character-level for high- as well as low-frequency characters? Is the involvement of component phonology, if any, observable only in lower-frequency characters? A general answer to both questions might be that only low-frequency characters allow sufficient activation of phonology at either the whole-character level or the component level. This prediction is based on the assumption that low-frequency characters should take longer to identify, thus increasing the probability of both component phonology and whole-character phonology being activated prior to a meaning decision. This prediction would appear to be in accord with results of both English and Chinese naming tasks that show regularity effects only for low-frequency words. However, it is important to note that phonetic validity does not map onto either regularity or consistency in the sense that they are applied to alphabetic writing systems.[1]

By contrast, a theory that claims that the identification of any character arouses its phonological word form (e.g., Perfetti et al., 1992) would predict that character-level phonology would be activated for all words, regardless of frequency. Whether component phonology effects are frequency-general is a less direct question for this theory, which assumes that phonology can be activated at many levels but always at the word level. In fact, the statistical consistency of component phonology (i.e., the extent to

---

[1] In English word naming, effects of spelling-pronunciation consistency typically have been found only for low-frequency words, as summarized by the well-known Frequency × Regularity interaction (e.g., Seidenberg, Waters, Barnes, & Tanenhaus, 1984). However, consistency effects can be observed for high-frequency as well as low-frequency words, especially when neighborhood characteristics of words are taken into account (Jared, 1997). In the Chinese case, however, phonetic validity does not map onto consistency. Validity is defined by whether the compound character in question has the same pronunciation as its phonetic does when the phonetic stands alone. Valid phonetics are not consistent when other characters containing the phonetic are pronounced differently.

which all characters that contain the component have the same pronunciation) may control the extent to which a component phonology is functional for a particular character in which it occurs (Tan & Perfetti, 1999).

Because we used sequential presentation of the characters in both meaning judgment and pronunciation judgments, it is useful to note that interference effects in these tasks do not depend on sequential presentation. If the effect were found only in sequential presentation, one might argue that the effect arises from sequential memory demands. Zhang (1996), however, found interference effects with simultaneous (0 SOA) presentation that were nearly identical (47 ms) to the effects reported in Perfetti and Zhang (1995a). This result, although not eliminating a role for memory in the judgment task, does establish that the interference effect is robust across display conditions and is not fundamentally dependent on a requirement to remember the first character. This fact allows the use of sequential displays, which bring more control of the viewing situation.

Method

Materials and Design

The materials were 192 characters, 48 core characters, each paired with three cohorts: a synonym, a homophone, and an unrelated control. The pairing of each core with three cohorts resulted in 144 experimental trials. A core character and its three cohorts were matched in character frequency and the presence or absence of cued phonology. Frequency and component validity were defined between sets of core characters and their cohorts. Half of the core characters and their cohorts were high frequency, with an average of 120 per million; the other half were low frequency, with an average of 7.2 per million, according to the Modern Chinese Frequency Dictionary (1986). Within each frequency category, half of the core characters and their cohorts had valid phonetic components, and half had uncued phonology. Also within each frequency category, the frequency of the characters with valid phonetic components was closely matched with the frequency of characters with uncued phonology. Among the 48 homophone pairs, all were homophonic at the phonemic (segmental) level, and 30 pairs were also identical in tones. Finally, among the set of phonetic compounds, 70% were left–right compounds with the phonetic on the right in all cases. In summary, there were 12 core characters (plus their cohorts) to represent each combination of Frequency (high/low) × Component Phonology (valid phonetic/uncued phonology; see Appendix A).

Each participant received all 144 trials (48 core characters × 3 cohorts) in continuous blocks of 48 trials that presented each core character once. Thus, within a block of 48, each core character was paired with just one cohort. A different random ordering of core characters was generated for each participant.

Synonym ratings. Prior to conducting the main experiment, ratings of synonym pairs were obtained from an independent sample of 15 Mandarin speakers. The raters responded to 120 pairs, half of which were pairs that the experimenters regarded as synonyms, and from which we expected to use 48 for the main experiment. The remaining 60 pairs represented a range of meaning similarity, from quite unrelated through partial similarity. This range was used so that participants would be encouraged to use the entire 5-point rating scale. From these ratings, we chose the 48 (of 60 candidates) pairs with the highest mean similarity rating, which was 4.16 (SD = 0.49). (For comparison, the unrelated pairs had a mean of 1.11 [SD = 0.14], and the partial meaning pairs had a mean similarity rating of 2.91 [SD = 0.81].) These 48 character pairs defined the core character and its synonym cohort for the main experiment.

Participants

Seventy-eight Mandarin speakers attending the University of Pittsburgh or Carnegie Mellon University participated for payment. Half of these (39) participated in meaning judgments, and half participated in pronunciation (homophone) judgments.

Procedure

The experimental session was located in a windowless laboratory housing a computer controlled tachistoscopic shutter. The experimental session began with 18 practice trials followed by 144 uninterrupted experimental trials. Each experimental trial consisted of a core character exposed for 180 ms with a 10-ms interval before the appearance of a cohort, which remained displayed until the participant responded or until 3 s elapsed. This 190-ms SOA was chosen to be well within the observed range of both phonological (90–310 ms) and semantic (140–310) interference effects. In meaning judgment, participants were instructed to respond "yes" when the two characters had similar meanings and "no" otherwise. In homophone judgment, participants responded "yes" when the two characters had similar pronunciations (were homophones except for tone), "no" otherwise.

Results

Meaning Judgment

The main question was under what conditions phonological interference would occur on "no" trials. As Table 3 shows for decision times, phonological interference emerged strongly and generally. Participants required more time to reach a correct "no" decision on homophone trials (M = 954 ms) than on control trials (M = 902 ms) across all conditions. This 52-ms effect was reliable by participants, $F_1(1, 38) = 115.73, p < .001, MSE = 1{,}794$, and by items, $F_2(1, 11) = 22.88, p < .001, MSE = 4{,}310$.

Component phonology did not produce a reliable main effect, $F_1(1, 38) = 1.83, p = .18, MSE = 3{,}609$; $F_2(1, 11) = 1.21, p > .05, MSE = 2{,}884$. More important, the phonological interference effect did not interact with component phonology; this effect was 52 ms for both characters with valid phonetics and for characters with uncued phonology.

There was a 32-ms frequency effect, $F_1 = (1, 38) = 28.67, p < .001, MSE = 2{,}739$, and $F_2(1, 11) = 4.56, p = .056, MSE = 6{,}043$. However, specific interest is in the generality of the phonological interference effect across frequency. The interaction of Frequency × Component Phonology was significant by participants, $F_1(1, 38) = 7.17, p < .05, MSE = 2{,}723$, and only marginally so by items, $F_2(1, 11) = 3.11, p = .11, MSE = 2{,}114$. The direction of this effect suggests that phonological interference was somewhat larger in the judgment of low-frequency characters (67 ms) than in the judgment of high-frequency characters (36 ms); however this difference was reliable only by participants, $t_1(266) = 2.84, p < .05$; $t_2(77) = 1.46, p =$

Table 3
Experiment 1. Mean Correct Decision Times and Error Rates for "No" Decisions in Synonym Judgment Task

Trial type	Decision time (ms)	PI effect (ms)	Error rate (%)	PI effect (%)
High frequency				
Component-valid		23		0.86
Homophone foils	937		1.07	
Control foils	914		0.21	
Uncued phonology		49		0.85
Homophone foils	924		1.28	
Control foils	875		0.43	
M	912	36	0.75	0.86
Low frequency				
Component-valid		80		4.92
Homophone foils	980		5.13	
Control foils	900		0.21	
Uncued phonology		55		3.42
Homophone foils	976		3.63	
Control foils	921		0.21	
M	944	67.5	2.30	4.17
Overall *M*		52		2.52

Note. PI (phonological interference) effect = (homophone foils − control foils).

.15. A Dunn's multiple comparison showed a general phonological effect for both frequency categories; for high-frequency characters, $t_1(266) = 4.54, p < .01$; $t_2(77) = 2.96, p < .01$; for low-frequency characters, $t_1(266) = 8.54, p < .001$; $t_2(77) = 3.55, p < .02$. Thus, the phonological interference effect was quite general over variations in both component phonology and frequency.

Error rates were low but showed the expected complementary pattern: higher for homophone trials (2.78%) than control trials (0.27%), $F_1(1, 38) = 36.88, p < .001, MSE = 0.19$; $F_2(1, 11) = 5.07, p < .05, MSE = 0.01$. This phonological interference effect was not affected by component phonology (2.89% for valid phonetics, 2.13% for uncued phonology). Error rates were higher for low-frequency (2.30%) than high-frequency characters (0.75%), although this difference was reliable only by participants, $F_1(1, 38) = 13.34, p < .001, MSE = 0.192$; $F_2 < 1$. As with decision times, the interaction of frequency with the phonological interference effect was reliable by participants, $F_1(1, 38) = 16.04, p < .01$, but not items, $F_2 < 1$. The trend of this interaction is that phonological interference effect was larger for low-frequency (4.17%) than for high-frequency (0.86%) characters.

"Yes" trials are of secondary interest in this task. For completeness, the decision times and error rates are shown in Table 4 along with the "yes" trials for the homophone judgment task. Decisions were faster for high-frequency trials than for low-frequency trials, and low-frequency characters with valid components produced faster decisions than low-frequency characters with uncued phonology. However, this pattern was reversed for high-frequency characters. It is likely that this reversal was due to the fact that high-frequency characters with uncued phonology included a large number of simple characters (9 of 48) that could be identified more rapidly.

Homophone Judgments

Semantic interference, the difference between synonym foil trials and control foil trials, was observed in the decision times, as shown in Table 5. This 38-ms effect was reliable by participants, $F_1(1, 38) = 57.35, p < .001, MSE = 1{,}952$, and by items, $F_2(1, 11) = 23.15, p < .001, MSE = 1{,}773$. There was also a 55-ms frequency effect, $F_1(1, 38) = 56.12, p < .001, MSE = 4{,}146$; $F_2(1, 11) = 14.52, p < .01, MSE = 4{,}315$. The frequency and semantic interference effects interacted reliably by participants only, $F_1(1, 38) = 5.43, p < .05, MSE = 1{,}702$; $F_2(1, 11) < 1$. The trend in this interaction is that semantic interference was larger for low-frequency (49 ms) than for high-frequency (27 ms) items. There was no effect of phonetic validity, F_1 and $F_2 < 1$.

Error rates were very low but converge with the conclusions based on decision times. Error rates were higher to synonym foils than control foils by an average of 3%, $F_1(1,$

Table 4
Experiment 1. Mean Correct Decision Times (in Milliseconds) and Error Rates (%) for "Yes" Decisions

Trial type	High frequency	Low frequency	*M*
Synonym judgments			
Component-valid	966 (5.98%)	968 (6.20%)	967 (6.09%)
Uncued phonology	915 (2.35%)	998 (5.56%)	956.5 (3.96%)
M	940.5 (4.17%)	983 (5.88%)	961.75 (5.03%)
Homophone judgments			
Component-valid	908 (3.85%)	930 (1.92%)	919 (2.89%)
Uncued phonology	913 (2.35%)	954 (5.34%)	933.5 (3.85%)
M	910.5 (3.1%)	942 (3.63%)	926.25 (3.37%)

Table 5

Experiment 1. Mean Correct Decision Times and Error Rates for "No" Decisions in Homophone Judgment Task

Trial type	Decision time (ms)	SI effect (ms)	Error rate (%)	SI effect (%)
		High frequency		
Component-valid		16		1.92
Synonym foils	892		2.35	
Control foils	876		0.43	
Uncued phonology		38		1.71
Synonym foils	906		1.92	
Control foils	868		0.21	
M	885.5	27	1.23	1.82
		Low frequency		
Component-valid		46		3.84
Synonym foils	964		4.91	
Control foils	918		1.07	
Uncued phonology		52		5.13
Synonym foils	966		5.56	
Control foils	914		0.43	
M	940.5	49	2.99	4.49
Overall *M*		38		3.16

Note. SI (semantic interference) effect = (synonym foils − control foils).

38) = 35.43, $p < .001$, $MSE = 0.315$ and $F_2(1, 11) = 11.88$, $p < .01$, $MSE = 0.0013$. And they were higher to low-frequency than high-frequency characters, although reliably so only by participants, $F_1(1, 38) = 11.91$, $p < .01$, $MSE = 0.293$; $F_2(1, 11) = 2.62$, $p = .13$, $MSE = 0.002$. The interaction of frequency with the synonym interference effect was also reliable by participants only, $F_1(1, 38) = 9.37$, $p < .01$, $MSE = 0.214$; $F_2(1, 11) = 2.04$, $p = 18$, $MSE = 0.002$. The trend of this interaction is that synonym interference was higher for low-frequency characters (error rate = 4.49%) than for high-frequency characters (1.82%). Again, there was no effect of component phonology, F_1 and $F_2 < 1$.

"Yes" trials. "Yes" trials are of specific interest in the homophone judgment task, because a valid phonetic might be expected to facilitate evaluation of the character name. In fact, as shown in Table 4, decisions were faster for characters with valid phonetic components than for characters with uncued phonology, an effect reliable by participants only, $F_1(1, 38) = 4.94$, $p = .03$, $MSE = 2{,}790$; $F_2(1, 11) < 1$. However, this effect was restricted to low-frequency characters, as suggested by the interaction of frequency and component phonology, also significant by participants only, $F_1(1, 38) = 8.45$, $p < .01$, $MSE = 2{,}016$, but not items, $F_2(1, 11) = 1.35$, $p > .10$, $MSE = 4{,}461$. Error rates showed no reliable main effects of frequency and phonetic validity; although low-frequency characters with invalid phonetics produced higher error rates than other combinations of frequency and validity, this interaction was significant by participants only, $F_1(1, 38) = 11.88$, $p < .01$, $F_2(1, 11) < 1$.[2]

Discussion

Experiment 1 establishes several interesting facts about the phonological interference effect. First, it confirms, with a different set of materials, both the phonological and the semantic interference effects reported by Perfetti and Zhang (1995a). The effect sizes are comparable, although slightly larger here than in the original studies: The phonological interference effect was 52 ms in Experiment 1 and 46 ms in Perfetti and Zhang (1995a) for 190-ms SOA. The semantic interference effect was 38 ms in Experiment 1, compared with 26 ms (at 190 ms) in Perfetti and Zhang (1995a). Also replicating the original studies is the finding that "yes" responses were faster in homophone judgments (926 ms)

---

[2] In considering the conclusions from these results, a few questions concerning the materials are relevant. For example, although most pairs of homophones with valid phonetics had different phonetic components, one homophone pair shared a phonetic, creating a potential confound between partial visual similarity and complete phonetic identity. Other potential questions are noted in Appendix A. To assess whether the conclusions would survive a more conservative inclusion of materials, we carried out new items analyses that included only core and cohort pairs that were free of any question. This resulted in the elimination of two core characters and their cohorts in each condition. The results of this new analysis confirmed the reported results. For example, the analysis of decision times for semantic judgments confirmed a general phonological interference effect, $F_2(1, 9) = 51.48$, $p < .001$, $MSE = 2{,}017$. Component phonology, as in the original analysis, produced no main effect, $F_2 < 1$, nor an interaction with phonological interference. Frequency did become less reliable by items, $F_2(1, 9) = 3.54$, $p = .09$, $MSE = 25{,}205$, and its interaction with the phonological interference was reduced and significant by participants only, $F_2 < 1$. The interference effect was 59.5 ms for low-frequency characters and 44 ms for high-frequency characters. The pattern of error rate effects was unaffected by the reanalysis. For homophone judgments, the pattern of significant results was unchanged for both "no" and "yes" decision times and for error rates.

than in meaning judgments (962 ms). Although the reliability of this between-task difference was not specifically tested here, this difference is predictable on the assumption that meaning evaluation and comparison is more complex than pronunciation retrieval and comparison.

What Experiment 1 specifically adds to the picture is the conclusion that phonological interference effects in the meaning judgment task arise at the level of the whole-character name, rather than strictly at the component level. The phonological interference effect occurred both for characters with valid phonetic components and for characters lacking valid component phonology. Taking the error data and the decision times together, the results indicate that whole-character phonology is sufficient to produce interference, whether that phonology is cued by a phonetic component or not. Notice that finding little or no evidence for a component effect does not mean that there is no component phonology involved in character identification. We return to this point in the General Discussion and Conclusion.

The results also confirm a general frequency effect in both pronunciation and meaning judgments. More interesting is that the phonological interference effect occurred for high-frequency as well as low-frequency characters. This suggests that, in a meaning decision task, the no-semantics-without-phonology effect is not a matter of slow retrieval of low-frequency items. It is also consistent with the conclusion that the interference effect is carried at the whole-character level, where there is less reason to expect that word frequency should interact with phonology. Nevertheless, there were indications in the data that the phonological effects were larger for low-frequency items than for high-frequency items, although in both decision times and error rates these differences were reliable only by participants. However, this interaction—allowing its reliability for the moment—has a different interpretation than is usually given to Frequency × Regularity or Frequency × Consistency interactions in word reading. Because the 36-ms effect for high-frequency words was reliable, any frequency interaction is a matter of relative magnitude, not one of all-or-none occurrence of a phonological effect.

Finally, the discovery of a component phonology effect in homophone "yes" judgments, at least for low-frequency characters, suggests that component phonology assists the access to whole-character pronunciations. When the reader's task requires attention to the phonological forms associated with the character, a convergence between the pronunciation of the component phonetic and the character as a whole may facilitate performance. This result converges with results from naming experiments showing effects of phonetic components on the naming of low-frequency phonograms.

Although Experiment 1 demonstrates that uncued (whole character) phonology is sufficient for phonological interference, it does not provide a clear contrast between valid and invalid phonetic compounds. The characters that did not contain valid phonetic components, the uncued phonology characters, included several varieties—compounds (phonograms) with invalid phonetics, semantic compounds, and simple characters. For example, among the 24 pairs of homophones with uncued phonology characters, 10 pairs were phonogram–phonogram, 2 pairs were semantic compound–semantic compound, 2 pairs were single character–single character, and 10 pairs were mixes of the three types. Although each of these varieties satisfies the uncued phonology requirement, one might wonder whether their variability poses problems of interpretation. Perhaps the presence of simple characters and semantic compounds in the materials somehow affected the readers' tendency to use information within a compound. However, the implication of this possibility is that valid phonetics could be strategically used in the task, since the variability in the uncued set meant that there were more valid phonetics than invalid phonetics. The results, of course, were otherwise. Nevertheless, it is clear that a simple contrast that involves only phonograms—characters with either valid or invalid phonetics—would be informative and would reduce item variability. Experiment 2 provides this contrast, and by the use of a new set of materials, further tests the generality of the result favoring whole-character phonology in producing the phonological interference effect.

Experiment 2

Experiment 2 is a replication of Experiment 1 with a new set of materials that contained only phonograms. Some phonograms had the same pronunciations as their phonetics, whereas other phonograms had pronunciations different from their phonetics. We refer to the first as *valid phonograms,* and the second as *invalid phonograms.* The structure of the valid phonograms is the same as in Experiment 1. The structure of the invalid phonograms of Experiment 2 is shown in Table 6. As in Experiment 1, the tasks were homophone judgment and synonym judgment. Unlike Experi-

Table 6
Experiment 2. Examples of Invalid Phonograms

Character type	Whole character (component)	Pronunciation: whole character (component)	Translation: whole character	Correct synonym judgment	Correct homophone judgment
Homophone	遗 (贵)	/yi/ (/gui/)	to lose		
	移 (多)	/yi/ (/duo/)	to move	no	yes
Control	遗 (贵)	/yi/ (/gui/)	to lose		
	创 (仓)	/chuang/ (/cang/)	to create	no	no
Synonym	遗 (贵)	/yi/ (/gui/)	to lose		
	掉 (卓)	/diao/ (/zhuo/)	to lose	yes	no

ment 1, each core character was viewed only one time by each participant.

Method

Materials and Design

The materials were 192 phonograms, 48 core characters each with three cohorts, paired to form 144 experimental trials. The 48 core characters included only 5 of those used in Experiment 2; only 3 of the 192 core–cohort pairings were the same as those in Experiment 1. A core character and its three cohorts were matched in character frequency, phonetic frequency, phonetic consistency and phonetic validity. As in Experiment 1, frequency and phonetic validity were defined over sets of core characters and their cohorts. Half of the core characters and their cohorts were high-frequency, average 266.9 per million, and half were low-frequency, average 10.8 per million, according to the *Modern Chinese Frequency Dictionary* (1986). Within each frequency category, half of the core characters and their cohorts had valid phonetics, and half had invalid phonetics. As in Experiment 1 and in the writing system generally, most of the characters that were phonograms (77%) were of the left–right type, with the phonetic on the right in nearly all cases (93%). Also within each frequency category, the frequency of the characters with valid phonetics was closely matched with the frequency of characters with invalid phonetics. Among the 48 pairs of core characters and their phonogram cohorts, all pairs were phonemic (segment level) homophones, and 23 pairs also had identical tones. In summary, there were 12 core characters (plus their cohorts) to represent each combination of frequency and phonetic validity, for a total of 144 (12 × 3 × 2 × 2) unique character pairings (See Appendix B).

A counterbalanced design ensured that each participant viewed the core character only once. Each participant thus had 48 experimental trials, 16 "yes" and 32 "no." In addition, there were 16 filler pairs in each condition, 16 homophone pairs for the homophone judgments and 16 synonym pairs for the synonym judgment trials, equating the number of "yes" and "no" trials.

Additional controls involved the frequency and the consistency of the phonetics. The frequency of the phonetics was matched across conditions (see Appendix B), with an average of 486 per million. Also matched across conditions was phonetic consistency, defined as the ratio of the number of characters containing a given phonetic that have the same whole-character pronunciation to the total number of all characters that contain the phonetic. For example, for the character 剧 (/ju/, *intense*), all the characters with the phonetic 居 (/ju/, *live in*) are pronounced as /ju/, so the consistency value is 1. For the character 依 (/yi/, *to rely on*), the phonograms with the phonetic 衣 (/yi/, *clothes*) vary: 依 (/yi/, *to rely on*), 裔 (/yi/, *descendants*) and 袭 (/xi/, *to raid*), so the consistency value of 依 (/yi/, *to rely on*) is ⅔ (0.67). Consistency value was calculated according to *ChangYong XingShengZi* (The Common Chinese Phonetic Compound Characters, 1982).

Synonym ratings. To obtain 48 pairs of synonyms, ratings of 108 word pairs were obtained from an independent sample of 11 Mandarin speakers prior to the experiment. Half of the pairs were judged to be synonyms by the experimenters, and the remaining 54 pairs represented a range of meaning similarity, from unrelated through partial similarity. The raters were asked to indicate the degree of meaning similarity of each pair of words on a scale of 1 to 5, in which 1 = *identical meaning* and 5 = *unrelated,* the opposite anchoring of Experiment 1. In order to report scalar values that are comparable across the experiments, the similarity scores were transformed to allow 1 to define unrelated pairs and 5 to define meaning identity. This resulted in an average meaning similarity for the 48 synonym pairs of 2.84 ($SD = 0.69$). (For comparison, the unrelated pairs had a mean of 0.44, $SD = 0.23$, and the partial meaning pairs had a mean similarity rating of 1.13, $SD = 0.89$.) These 48 pairs defined the core characters and their synonym cohorts for the experiment.

Subjects

Sixty native Mandarin speakers from the same population as Experiment 1 participated with payment. None had participated in Experiment 1. Thirty participants made synonym judgments, and 30 different participants made homophone judgments.

Procedure

The procedure was similar to that of Experiment 1, except that words were presented on a computer screen rather than in a tachistoscopic display and participants made a manual response rather than a spoken one. A core character was exposed for 182 ms (compared with 180 ms in Experiment 1), based on the 14-ms refresh cycle of the computer screen. An experimental trial consisted of a core character exposed for 182 ms with a 10-ms interstimulus interval prior to the appearance of a cohort, which remained displayed until the participant responded. Instructions for synonym judgment and homophone judgment were the same as in Experiment 1. Participants pressed a "yes" button when the two characters were synonyms (in meaning judgments) or had the same pronunciations (in phonological judgments), and pressed a "no" button otherwise. As in Experiment 1, tone differences were to be ignored. The experimental session began with 20 practice trials.

Results

Meaning Judgment

The important result in meaning judgment was that phonological interference emerged strongly and generally once again. Tables 7 and 8 show the decision times for correct "no" answers and Figure 1 shows the net phonological interference effect. Participants required more time to reject homophones than controls did, across all conditions. This 41.5-ms effect was reliable by participants, $F_1(1, 29) = 5.54$, $p < .05$, $MSE = 18{,}498$, and by items, $F_2(1, 22) = 5.03$, $p < .05$, $MSE = 21{,}114$.

Phonetic validity was not reliable as a main effect, $F_1(1, 29) = 1.10$, $p = .30$, $MSE = 27{,}375$; $F_2 < 1$, nor, most important, did phonetic validity interact with foil type, F_1 and F_2 both <1. Thus, the phonological interference effect was not greater for valid than for invalid phonograms. As shown in Figure 1, the magnitude of the phonological interference effect (homophone minus controls) varied only modestly (from 32 to 60 ms) around the overall mean effect of 41.5 ms, and this nonsignificant variability actually favored invalid phonetics (50 ms) over valid phonetics (33 ms).

The frequency effect was nearly the same (31 ms) as in Experiment 1 (32 ms), but it was not reliable, $F_1(1, 29) = 2.51$, $p = .12$, $MSE = 23{,}849$; $F_2(1, 22) = 2.40$, $p = .14$, $MSE = 10{,}911$. Of main interest is the interaction of Frequency × Foil Type, which was not reliable either by participants or by items, F_1 and $F_2 < 1$. The phonological

Table 7
Experiment 2. Mean Correct Decision Times and Error Rates for "No" Decisions in Synonym Judgment Task

Trial type	Decision time (ms)	PI effect (ms)	Error rate (%)	PI effect (%)
		High frequency		
Valid		32		5.83
Homophone foils	1045		8.33	
Control foils	1013		2.50	
Invalid		60		6.67
Homophone foils	1051		8.33	
Control foils	991		1.67	
M	1025	46	5.21	6.25
		Low frequency		
Valid		34		6.67
Homophone foils	1047		15.00	
Control foils	1013		8.33	
Invalid		40		9.17
Homophone foils	1103		13.33	
Control foils	1063		4.16	
M	1056.55	37	10.21	7.92
Overall *M*		41.5		7.08

Note. PI (phonological interference) effect = (homophone foils − control foils).

interference effect was general across both high-frequency characters (46 ms) and low-frequency characters (37 ms).

Error rates, although somewhat higher than in Experiment 1, produced a similar picture: more errors on homophone pairs (11.25%) than controls (4.17%), $F_1(1, 29) = 13.38$, $p < .01$, $MSE = 0.34$; $F_2(1, 22) = 6.40$, $p < .05$, $MSE = 1.67$. This phonological interference effect was not influenced by whether the phonetic was valid (6.3%) or invalid (7.9%) or by the frequency of the characters (7.92% for low- and 6.40% for high-frequency characters), all Fs < 1. Validity was not significant as a main effect, $F_1(1, 29) = 3.01$, $p = .09$, $MSE = 0.17$; $F_2 < 1$, nor did it interact with frequency, F_1 and $F_2 < 1$. Unlike the decision time results, the main effect of frequency was reliable, $F_1(1, 29) = 10.56$, $p < .01$, $MSE = 0.29$; $F_2(1, 22) = 7.98$, $p < .01$, $MSE = 0.88$.

Although "no" trials are of primary interest, the decision times and error rates of correct "yes" decisions are shown in Table 8, along with the results for homophone judgments. Decision times were faster for high-frequency than for low-frequency items, $F_1(1, 29) = 5.99$, $p < .05$, $MSE = 29{,}391$; $F_2(1, 11) = 8.47$, $p < .05$, $MSE = 8{,}710$. No other effects were reliable (all Fs less than 1). In error rates, there was no main effect of either frequency or validity, all Fs < 1. The interaction of frequency and validity was not reliable, $F_1(1, 29) = 3.46$, $p = .073$, $MSE = 0.62$; $F_2(1, 11) = 1.94$, $p = .19$, $MSE = 3.88$, despite a tendency for low-frequency items only to show a validity effect (5%).

Homophone Judgment

As shown in Table 9, correct "no" decision times for deciding that two characters did not have the same pronunciation were longer for synonym pairs than for control pairs. This 33.5-ms effect was reliable by participants, $F_1(1, 29) = 5.16$, $p < .05$, $MSE = 13{,}024$, and by items, $F_2(1, 22) = 5.90$, $p < .05$, $MSE = 11{,}088$.

Table 8
Experiment 2. Mean Correct Decision Times (in Milliseconds) and Error Rates (%) for "Yes" Decisions

Trial type	High frequency	Low frequency	*M*
	Synonym judgment		
Phonetic valid	952 (10.0%)	1014 (7.5%)	982.5 (8.75%)
Phonetic invalid	941 (7.5%)	1031 (12.5%)	985.5 (10.0%)
M	946.5 (8.75%)	1022.5 (10.0%)	984 (9.38%)
	Homophone judgment		
Phonetic valid	843 (12.5%)	865 (17.5%)	854 (10.0%)
Phonetic invalid	824 (14.25%)	978 (32.5%)	901 (23.4%)
M	833.5 (13.4%)	921.5 (25.0%)	877.5 (16.7%)

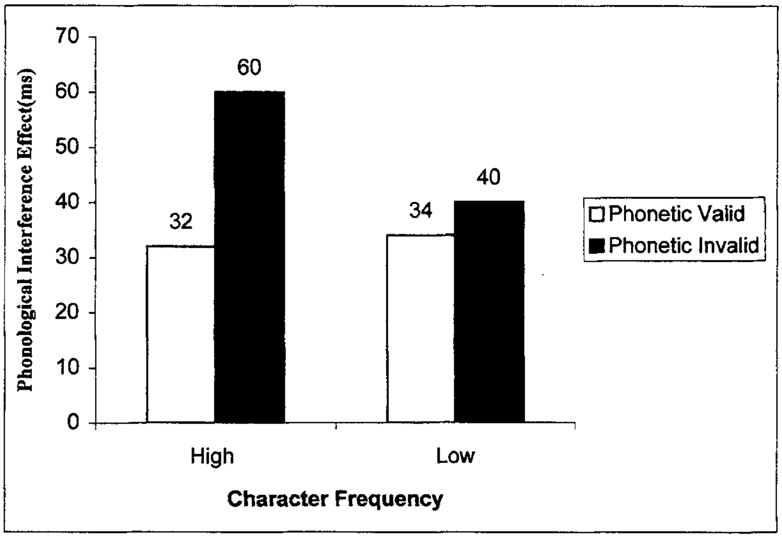

Figure 1. The phonological interference effect in decision times (homophone minus control) for synonym judgment in experiment 2.

"No" decision times were longer for invalid phonograms (mean = 961.5 ms) than for valid phonograms (mean = 930.5 ms). This 31-ms effect was marginally reliable by participants only, $F_1(1, 29) = 3.88$, $p = .059$, $MSE = 14{,}628$; $F_2(1, 22) = 1.10$, $p = .30$, $MSE = 9{,}560$. Frequency was reliable both by participants and by items, $F_1(1, 29) = 31.25$, $p < .001$, $MSE = 17{,}993$; $F_2(1, 22) = 11.38$, $p < .01$, $MSE = 18{,}065$. No interactions were reliable for Phonetic Validity × Frequency, $F_1(1, 29) = 1.41$, $p > .20$; $F_2(1, 22) = 1.72$, $p > .21$. For Phonetic Validity × Foil Type, Foil Type × Frequency, and Phonetic Validity × Foil Type × Frequency, all F_1s and F_2s < 1. Thus, the tendency for faster decisions on trials with valid phonetics was general across control and synonym foils and general across frequency.

Error rates showed the same synonym interference trend as decision times, with higher error rates on synonym trials (12.45%) than on control trials (6.85%), $F_1(1, 29) = 9.18$, $p < .01$, $MSE = 0.331$; $F_2(1, 22) = 6.47$, $p < .05$, $MSE = 1.007$. Frequency was marginally reliable by participants, $F_1(1, 29) = 3.92$, $p = .057$, $MSE = 0.308$, $F_2(1, 22) = 2.25$, $p = .15$, $MSE = 1.67$. There was no main effect of phonetic validity, F_1 and $F_2 < 1$.

Table 9
Experiment 2. Mean Correct Decision Times and Error Rates for "No" Decision in Homophone Judgment Task

Trial type	Decision time (ms)	SI effect (ms)	Error rate (%)	SI effect (%)
		High frequency		
Valid		33		5.83
Synonym foils	908		10.83	
Control foils	875		5.0	
Invalid		27		7.35
Synonym foils	917		11.5	
Control foils	890		4.15	
M	897.5	30	7.87	6.59
		Low frequency		
Valid		37		4.15
Synonym foils	988		14.15	
Control foils	951		10	
Invalid		37		5.08
Synonym foils	1038		13.33	
Control foils	1001		8.25	
M	994.5	37	11.43	4.62
Overall *M*		33.5		5.6

Note. SI (semantic interference) effect = (synonym foils − control foils).

"Yes" trials. Responses on "yes" trials are of interest in homophone judgments, because such trials give direct opportunity to observe the use of the phonetic in carrying out the task. As shown in Table 8, valid phonetics produced faster decision times than did invalid phonetics, but only for low-frequency characters. There was an interaction of Frequency × Phonetic Validity, reliable by participants only, $F_1(1, 29) = 8.82, p < .01, MSE = 14{,}820$; $F_2(1, 11) = 3.54, p = .087, MSE = 13{,}830$. For high-frequency characters there was a nonsignificant slowing of decision times for characters with valid phonetics, F_1 and $F_2 < 1$. However, for low-frequency characters, phonetic validity speeded decision times (865 ms vs. 978 ms), $F_1(1, 29) = 9.76, p < .01, MSE = 19{,}586$; $F_2(1, 11) = 5.63, p < .05, MSE = 19{,}992$. Error rates tell a slightly different story: For both high and low frequency, valid phonetics produced lower error rates than did invalid phonetics, a difference reliable by participants, $F_1(1, 29) = 7.91, p < .01, MSE = 0.83$, and by items, $F_2(1, 11) = 8.82, p < .01, MSE = 1.47$. Also, error rates were lower for high-frequency than for low-frequency trials, $F_1(1, 29) = 5.04, p < .05, MSE = 0.66$; $F_2(1, 11) = 8.09, p < .05, MSE = 2.47$. Although the interaction between frequency and phonetic validity was marginally reliable by participants only, $F_1(1, 29) = 3.56, p = .069, MSE = 0.599$; $F_2(1, 11) = 3.26, p = .09, MSE = 3.38$; the larger validity effect for low-frequency characters (15%) compared with high-frequency characters (1.75%) suggests that validity matters more for low-frequency characters. A post-hoc analysis showed a reliable validity effect for low-frequency characters, $F_1(1, 29) = 9.02, p < .01, MSE = 0.89$; $F_2(1, 11) = 6.29, p < .05, MSE = 3.82$, but not high-frequency characters, $F_1(1, 29) = 1.13, p > .29, MSE = 0.531$; $F_2 < 1$.

Finally, because validity effects seemed to be stronger for low-frequency than for high-frequency characters over a number of conditions and measures, we summarize these effects in Table 10. Although the statistical reliability of these effects are variable and confined to by-participants analyses, the generality of the pattern is suggestive. As is apparent in Table 10, validity effects were absent in both decision times and error rates for high-frequency characters. For low-frequency characters, validity effects were observed in all conditions either in decision times, error rates, or both.

Discussion

Experiment 2 provides further evidence for highly general phonological interference in a meaning judgment task. This size of the phonological interference effect (41.5 ms, SOA = 192 ms) is comparable to that observed in Experiment 1 (52 ms, SOA = 190 ms) with different materials. Furthermore, as in Experiment 1, the phonological interference effect was not dependent on component phonetic validity or on character frequency. This result was obtained with careful control over two important component factors, the frequency and consistency of the phonetic.

Although there was no evidence that the phonetic is responsible for the phonological interference effect, equally interesting is evidence that the presence of a valid phonetic facilitated judgment of pronunciation, at least for low-frequency characters. Replicating the pattern of Experiment 1, readers were faster and more accurate for low-frequency characters when the pronunciation of the character was the same as its phonetic component. Furthermore, although not statistically reliable, this tendency for the phonetic to facilitate performance for low-frequency characters was present even in semantic judgments.

Thus, we have two different glimpses of the relationship between phonology and frequency in this experiment. Phonological interference was observed equally for high- and low-frequency characters. However, there was a tendency for phonetic validity effects to be larger for low-frequency characters. These two results imply that phonology arises at two different levels in character identification. The phonological interference effect is primarily a product of word-level phonology, not component-level phonology. Both high- and low-frequency words have their word-level phonology activated, and this produces the interference effect. However, component-level phonology is also activated, and at least in low-frequency words, this activation plays a role in character identification, affecting both speed and accuracy of judging meaning and pronunciation.

Although the results from Experiment 2 replicated the critical results of Experiment 1 regarding general phonological interference across variation in phonetic validity and frequency, there is a visible difference in the results: The overall decision times were slower in Experiment 2, and the

Table 10
Experiment 2. Phonetic Validity Effects (Invalid Phonograms − Valid Phonograms)

Task and decision type	Decision time (ms)		Error rate (%)	
	High frequency	Low frequency	High frequency	Low frequency
	Synonym judgments			
Yes decisions	−11	17	−2.5	5*
No decisions	−8	53*	−0.42	−2.92
	Homophone judgments			
Yes decisions	−19	113***	1.75	15**
No decisions	12	50**	−0.09	−1.29

*$.05 < p < .10$. **$p < .05$. ***$p < .01$.

error rates were higher. It is not possible to localize the source of the overall speed and accuracy differences, although three classes of possibilities are salient. One is that different materials were used in the two experiments. In Experiment 2, the careful selection of character compounds constrained by phonetic validity, phonetic frequency, and phonetic consistency, as well as whole-character frequency might have led to materials that imposed more difficult processing demands on the reader. One specific possibility is that the exclusive use of phonograms, with many invalid, created difficulty. Another possibility for the meaning judgment task is that the synonym pairs were less similar in Experiment 2 than in Experiment 1. This could have led to higher error rates and slower decision times in the semantic judgment task. However, this would not explain the slower decision times and higher rates that were observed in the homophone judgment task. An alternative possibility is the laboratory display conditions. There are a number of differences between the tachistoscopic lab of Experiment 1 and the computer display lab of Experiment 2, including timing precision, luminance contrast, visual angle, and font characteristics. Such differences can produce variation in the functional visual display that affects performance. Finally, Experiment 1 required an oral "yes" or "no" response, whereas Experiment 2 required a manual response. Such differences in materials and laboratories could have affected response measures, but the pattern of critical replication of effects stands in contrast to such variability.

General Discussion and Conclusion

These experiments add to the evidence that reading for meaning in Chinese involves phonology; they also illuminate the interpretation of the meaning judgment paradigm that has produced this evidence. The most important results concern the generality of the phonological interference effect: Phonological interference is highly general over character frequency and over component phonology. Also important is the generality of the semantic interference effect and its apparent independence from component phonology. We discuss the implications of the generality-over-frequency result, then the component phonology result, and finally the semantic interference effect.

Phonology and Word Frequency

The evidence for phonology in high-frequency words partially contrasts to results typically found in naming for Chinese (Fang, Horng, & Tzeng, 1986; Hue, 1992; Seidenberg, 1985) and English (Paap, Chen, & Noel, 1987; Seidenberg, Waters, Barnes, & Tanenhaus, 1984). On the other hand, it aligns with evidence in English from identification paradigms using backward masking (Berent & Perfetti, 1995; Perfetti & Bell, 1991) and priming (Perfetti & Bell, 1991), which have found phonological effects across both high- and low-frequency words. Moreover, even in English naming, it is now clear that even high-frequency words are affected by phonological consistency of spellings (Jared, 1997). In Chinese naming, similar consistency effects can be obtained for high-frequency words (Tan & Perfetti, 1999). Thus, we conclude that phonological involvement is quite general across word frequency in a number of word processing tasks.

This conclusion on the generality of phonology, however, may appear to be at odds with the semantic categorization data in English. Following van Orden's (1987) demonstration of homophone interference in semantic categorization of high- as well as low-frequency foils, Jared and Seidenberg (1991) found homophone interference for low-frequency words (foils) only when the semantic category was "specific," (e.g., car part—*break*) and not when it was "general" (e.g., living thing—*break*). They interpreted this difference as owing to a priming effect that occurs with specific categories, which can prime a high-frequency exemplar (e.g., car part—*brake*) that can interfere with the "no" decision required by the homophone. Thus, on the account of Jared and Seidenberg (1991), phonological activation is genuine and prior to meaning activation for low-frequency words but not for high-frequency words. In semantic categorization experiments with Japanese Kanji (Chinese characters), Wydell, Patterson, and Humphreys (1993) found homophone interference for high-frequency characters as well as low. However, because they used specific categories, they pointed out that their result for high-frequency words could be explained by the priming mechanism postulated by Jared and Seidenberg: Priming by a category name can eliminate frequency effects that would otherwise control phonology. In the meaning judgment task used in the present experiments, this kind of priming cannot be a factor. The reader is not given a category name that can prime a homophone exemplar. Rather the reader is given first one character and then a second that is a homophone. The only kind of priming that could occur here is phonological—from a character to one of its many (not just one, as in English) homophones. There is no value to this priming, however, in a semantic task. Its occurrence would simply be an example of the more general proposition that the pronunciation of characters is activated automatically. Of course, one may argue that what is important in the meaning judgment task is the requirement to compare meanings, and that this requirement encourages phonological coding of the character. This is plausible. But it implies a fine distinction: meaning can be *accessed* without phonology, but it cannot be *used* without phonology.

Aside from the interpretation of the present phonological interference results, we think there is more to learn about the role of phonology in the meaning activation of high-frequency words. Van Orden and Goldinger (1994) suggested that for high-frequency words, "visual-phonologic resonance" can appear so quickly as to escape detection. Methods that depend on 500 and more milliseconds of response time might well fail to expose phonological effects and lead to unwarranted null-result conclusions. Consistent with this idea is a study by Xu (1997) using a response signal linked to the onset of the category exemplar to control the timing of the participant's semantic category decision. He found that high-frequency words (foils) actually produced more interference at short SOAs (200 ms) than did low-frequency words; at 400 ms SOA, low-frequency words showed the same magnitude of interference as high-frequency SOAs.

These questions about phonological mediation and high-frequency words, however, are beyond what the present results can address. We do not conclude that our phonological interference result for high-frequency words must be explained by a mediation hypothesis, in which the meanings of the characters are accessed on the basis of their pronunciations. This is because, more generally, we do not interpret even the results for low-frequency words to require a phonological mediation explanation. What the results require is that phonology is activated automatically. This conclusion, in turn, raises the question of why this happens in Chinese.

Whole Characters and Phonetic Components

The evidence of the present experiments is that the phonological interference effect requires a whole-character explanation: Phonetics that gave valid information about the pronunciation of the character produced no more interference than characters without such information. However, the evidence also showed that the component phonology may be used when phonology is required by the task, as it was in homophone judgments. Even in semantic judgments, there was a tendency (Experiment 2) for valid phonetics to lead to generally faster semantic decisions on "no" trials for low-frequency characters. Thus, we would not necessarily conclude that phonological components play no role in the access or use of meaning information.

We can summarize our conclusion on component phonology this way: Component phonology is available to processing as part of the character representation system. Reading a character involves the activation of its component phonology as well as its character-level phonology. When the component phonology is consistent with the phonology of the character as a whole, the convergent activation speeds identification of the character and access to its meaning. However, the duration of the component activation is short, and character-level phonology and semantics come to dominate the identification process. Thus, the effect of the component activation is more observable in facilitative situations, such as naming and pronunciation judgments, where the convergent phonology assists in the task. In the meaning judgment task, the interfering phonology arises at the character level, as the meaning and the phonological form associated with the character are automatically retrieved. Any prior activation of component phonology has become irrelevant.

A full account of component phonology in character reading must consider the role of frequency. Whether the component activation is relevant only for low-frequency characters, which take longer to achieve word-level access, cannot be settled in these experiments. There is certainly a suggestion across various conditions and measures of Experiment 2 that validity effects—the difference between phonograms with valid components and phonograms with invalid components—were more visible in low-frequency than in high-frequency characters (See Table 10). It may be tempting to conclude that a "race" between component processes and character-level processes occurs, and the component processes win the race only for slow-process (low-frequency) characters. Of course, this sort of explanation is familiar from the Frequency × Regularity interaction in naming. However, we believe there is little reason to accept this explanation until more work is done to sort out important character and component effects, including, for example, the frequency of the component relative to that of the character. Although the standard assumption is that it is the frequency of the whole (word or character) that modulates the effect of component phonology, the relative frequency of the component, which is likely to be higher than a low-frequency character but lower than a high-frequency character, could be a critical factor. The fact that component validity effects for high-frequency words have been observed in recent naming experiments (Tan & Perfetti, 1999) also require a cautious conclusion concerning the assumption that character frequency limits the role of a character component. In sorting out the effects of components and whole characters, the relative frequency of the phonetic, its validity, its consistency, and possibly its combinability—the number of different compounds that contain a given phonetic (Feldman & Siok, 1997; Taft & Zhu, 1997)—may need to be examined.

Semantic Interference

It is also important that these experiments give further evidence for semantic interference when readers make judgments about pronunciation. Both Experiments 1 and 2 showed reliable semantic interference effects, replicating Perfetti and Zhang (1995a). Such a pattern—interference of semantics by phonology and interference of phonology by semantics—requires an interactive account of lexical processing that links meaning and form at the character level. It is not that reading goes from orthography-to-phonology-to meaning (although it may often do that). It is that the reading of an orthographic form activates both phonological forms and semantics and neither can be easily suppressed.

Why is There Phonology in Reading Chinese for Meaning?

Beyond the interpretation of these interference effects is the deeper question of why phonology is involved at the word level in reading for meaning. The Chinese writing system is best characterized not as logographic, but as morpho-syllabic or semantic-phonological, reflecting the multiple layers represented by the characters and their components. In its phonetics, the system indeed has graphic-phonological elements. The argument that all writing systems represent phonology to some extent (De Francis, 1989; Mattingly, 1992) seems correct and such a universal has implications for reading and writing work across language systems. However, at the level of meaning analysis—the whole character level—the relevant constituents are the meaning and the pronunciation of the character, rather than its components.

It must be emphasized that even with its component phonology, the Chinese writing system has an in-principle potential for ignoring phonology by directly encoding character meaning. Against this potential, the evidence that readers do not ignore phonology is especially interesting. We must look beyond computation of sublexical elements in understanding this phenomenon: When readers, in alphabetic or nonalphabetic systems read words, they retrieve their phonological forms, not just their meanings. This phonological process may be universal, automatic, and multilayered (involving any level of graphic-phonological mappings allowed by the system), according to the universal phonological principle (Perfetti et al., 1992).

The central role of phonology can be understood by considering word reading in terms of the constituency model (Perfetti & Tan, 1998, 1999; Perfetti & Zhang, 1995b). Word identification is the momentary convergence of the three constituents—orthographic form, phonological form, and some nonformal (meaning) semantic substance, which are the definitive constituents of word representations. This understanding of word identification does not leave phonology to some optional "code" (that may or may not mediate meaning retrieval or lexical access). Instead, it places phonology as a necessary constituent of the identified object—the word such that it has orthography *X*, phonology *Y*, and meaning range *Z*. Because the mapping of orthography to phonology, at the word level, is more deterministic (more nearly one-to-one) than the mapping from orthography to meaning, there is a simple basis for expecting phonological involvement in reading across writing systems. Given a single printed word form, it usually has a single corresponding phonological word form, allowing a strong learned connection to develop between the two forms.

All fully represented alphabetic systems appear to behave in this largely deterministic way.[3] Given a full spelling, a pronunciation is largely determined, occasional exceptions noted (e.g., *read* and *wind* in English). For computational models with both feedforward and feedback activation among orthographic, phonological, and semantic representations, orthographic–phonological connections must be stronger than the orthographic–meaning connections. This in fact is the idea underlying the recurrent network resonance models of Van Orden and Goldinger (1994).

Chinese, at the whole character level, has a high degree of one-to-many mappings from phonology to orthography; there are many homophones. However, it is *relatively* more deterministic in the reverse direction, the mapping of orthography to phonology. For about 90% of characters in modern use, there is only one pronunciation, according to Zhou (1979). Thus, there appears to be an important shared feature between the Chinese writing system and alphabetic systems that allows both to have spoken word forms more-or-less reliably connect with printed word forms.

This conceptual framing of word identification may help in understanding the pervasive role of phonology in word identification. And it applies as well to Chinese as to English or Italian. Quite aside from the usual question of "prelexical" versus "postlexical" phonology, the evidence is increasingly supportive of the hypothesis that encounters with printed words lead to activation of word-level phonology in any writing system.

---

[3] In its typical unvoweled form, Hebrew does not behave this way and is more deterministic from phonology to orthography than the other way around.

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.

ChangYong XingShengZi [The common Chinese phonetic compound characters]. (1982). SiChuan, China: SiChuan People's Press.

Chua, F. K. (1995, December). *The role of phonological information in logographic processing.* Paper presented at the International Conference on the Cognitive Processing of Chinese and Other Asian Languages, Hong Kong.

DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii.

Fang, S. P., Horng, R. Y., & Tzeng, O. J. L. (1986). Consistency effect in the Chinese character and pseudo-character naming tasks. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology, and the Chinese language* (pp. 11–21). Hong Kong: Center of Asian Studies.

Feldman, L., & Siok, W. T. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 776–781.

Flores D'Arcais, G. B., Saito, H., & Kawakami, M. (1995). Phonological and semantic activation in reading Kanji characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 34–42.

Hue, C. W. (1992). Recognition processes in character naming. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93–107). Amsterdam: North-Holland.

Jared, D. (1997). Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory and Language, 36,* 505–529.

Jared, D., & Seidenberg, M. S. (1991). Does word identification proceed from spelling to sound to meaning? *Journal of Experimental Psychology: General, 120,* 358–394.

Lam, A., Perfetti, C. A., & Bell, L. (1991). Automatic phonetic transfer in bidialectal reading. *Applied Psycholinguistics, 12,* 299–311.

Leck, K. J., Weekes, B. S., & Chen, M. J. (1995). Visual and phonological pathways to the lexicon. *Memory and Cognition, 23,* 468–476.

Mattingly, I. G. (1992). Linguistic awareness and orthographic form. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 11–26). Amsterdam: North-Holland.

Paap, K. R., Chen, L., & Noel, R. W. (1987). *Word recognition: Is the sky falling on top-down processing?* Paper presented at the annual meeting of the Psychonomic Society, Seattle, Washington.

Perfetti, C. A., & Bell, L. (1991). Phonemic activation during the first 40 ms of word identification: Evidence from backward masking and masked priming. *Journal of Memory and Language, 30,* 473–485.

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, and Cognition, 24*(1), 101–118.

Perfetti, C. A., & Tan, L. H. (1999). The constituency model of Chinese word identification. In A. Inhoff, J. Wang, & H. C. Chen (Eds.), *Reading Chinese scripts: A cognitive analysis.* Mahweh, NJ: Erlbaum.

Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 633–643.

Perfetti, C. A., & Zhang, S. (1995a). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 24–33.

Perfetti, C. A., & Zhang, S. (1995b). The universal word identification reflex. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 33, pp. 159–189). San Diego: Academic Press.

Perfetti, C. A., Zhang, S., & Berent, I. (1992). Reading in English and Chinese: Evidence for a "universal" phonological principle. In R. Frost & L. Katz (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 227–248). Amsterdam: North-Holland.

Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition, 19,* 1–30.

Seidenberg, M. S., Waters, G. S., Barnes, M. A., & Tanenhaus, M. K. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior, 23,* 383–404.

Taft, M., & Zhu, X. (1997). Sub-morphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 761–775.

Tan, L. H., Hoosain, R., & Peng, D.-L. (1995). Role of early presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 43–54.

Tan, L. H., Hoosain, R., & Siok, W. W. T. (1996). The activation of phonological codes before access to character meaning in written Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 865–883.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37,* 41–57.

Tan, L. H., & Perfetti, C. A. (1999). Phonetic component validity and consistency both affect Chinese character naming. Manuscript in preparation.

van Orden, G. C. (1987). A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition, 15,* 181–198.

van Orden, G. C., & Goldinger, S. D. (1994). The interdependence of form and function in cognitive systems explains perception of printed words. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 1269–1291.

Wydell, T. N., Patterson, K. E., & Humphreys, G. W. (1993). Phonologically mediated access to meaning for Kanji: Is a *rows* still a *rose* in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1082–1093.

XianDai HanYu CiDian [Modern Chinese dictionary] (2nd ed.). (1992). Beijing, China: Shangwu Press.

XianDai HanYu PinLu CiDian [Modern Chinese frequency dictionary]. (1986). Beijing, China: Beijing Language Institute Press.

Xu, B. (1997). *How much is phonological information worth: Evidence of functional dependency on phonology in the perception of word meaning.* Unpublished doctoral dissertation. University of Pittsburgh.

Yang, H., & Peng, D. L. (1997). The learning and naming of Chinese characters of elementary school children. In H. C. Chen. (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 323–346). Hong Kong: The Chinese University Press.

Zhang, S. (1996). *The universal characteristics of phonological processes in writing system comparisons.* Unpublished doctoral dissertation. University of Pittsburgh.

Zhang, W. T., Feng, L., & He, H. D. (1994). The activation of phonological and semantic information in Chinese character recognition. In H.-W. Chang, J.-T. Huang, C.-W. Hue, & O. J. L. Tzeng (Eds.) *Language processing in Chinese* (pp. 185–198). Amsterdam: Elsevier.

Zhou, Y. G. (1979). Characters of multiple pronunciation in modern Chinese. *Zhongguo Yuwen, 153,* 401–405.

Appendix follows on next page

Appendix A

Materials for Experiment 1

No.	Core	Homophone	Synonym	Control
		High-frequency phonetic-valid characters		
1	嫁	架	婚	固
2	神	审	佛	格
3	抬	态	担	歌
4	睁	征	瞪	供
5	浮[a]	肤	漂	纺
6	盒	荷	箱	慌
7	镇	侦	村	唱
8	惊	镜	吓[b]	顶
9	织	枝	缝	楼
10	挺	停	伸	编
11	植	职	种	攻
12	置	智	放	渡
		High-frequency uncued phonology characters		
1	污	屋	脏	测
2	累	泪	乏	朝
3	洞	冬	孔	付
4	侧	策	哥	冲
5	兄	凶	旁	初
6	足	族	脚	称
7	讲	奖	谈	抽
8	母	亩	娘	层
9	喊	汗	嚷	江
10	吊	掉	挂	富
11	凉	粮[c]	冷	查
12	挥	恢[c]	摇	独
		Low-frequency phonetic-valid characters		
1	茎	颈	梗	汪
2	悸	嫉	骇	菇
3	愚	榆	笨	俱
4	仆[d]	蒲	佣	玛
5	绞	蕉	拧	泳
6	涝	姥	淹	瞄
7	篓	搂	箩	叼
8	饲	嘶	哺	枫
9	馍	蘑	馒	玲
10	犁	漓	耪	捻
11	俘[e]	袱	擒	棋
12	煨	蔚	焖	逻
		Low-frequency uncued phonology characters		
1	盈	莹	赚	雁
2	晾	粱	晒	仁
3	宰	崽	屠	舵
4	裕	郁	澡	杜
5	罐	冠	坛	驯
6	淌	唐	漾	炊
7	浆	疆	汁	拐
8	羞	嗅	臊	苏
9	卸	泻[f]	拆	虹
10	祭	忌[f]	悼	杏
11	雇	箍	聘	肾
12	斩	盏	铡	慧[f]

Note. [a]Phonetic of character 浮 is low in frequency; [b]Character 吓 has two possible pronunciations, which are /xia/ and /he/; [c]Characters 粮 and 恢 have same pronunciations with their phonetics; [d]Character 仆 only shares the vowel with its phonetic 卜; [e]Phonetic of 俘 is low in frequency; [f]Characters 泻, 忌, and 慧 share some phonological information with their phonetics.

Appendix B

Materials for Experiment 2

No.	Core	Homophone	Synonym	Control
High-frequency phonetic-valid phonograms				
1	剧	距	烈	描
2	议	艺	论	胞
3	管	惯	筒	程
4	悟	误	懂	洋
5	记	极	忆	神
6	愿	圆	望	格
7	惊	镜	慌	粒
8	浮	肤	漂	刚
9	模	沫	仪	抗
10	依	议	附	杆
11	置	指	装	浓
12	缝	疯	补	证
Character frequency (*SD*)	270.71 (213)	253.04 (229)	256.91 (273)	258.01 (190)
Phonetic frequency (*SD*)	513.81 (755)	476.80 (718)	508.29 (976)	509.39 (534)
Consistency value (*SD*)	0.77 (0.20)	0.78 (0.25)	0.82 (0.27)	0.70 (0.28)
High-frequency phonetic-invalid phonograms				
1	视	室	读	板
2	遗	移	掉	创
3	积	技	聚	抽
4	播	波	撤	判
5	济	基	救	影
6	始	适	初	梯
7	输	殊	送	刻
8	途	图	路	错
9	破	迫	损	测
10	接	洁	连	矿
11	持	吃	握	沙
12	瞎	峡	盲	效
Character frequency (*SD*)	276.24 (215)	276.24 (264)	275.69 (331)	267.96 (147)
Phonetic frequency (*SD*)	428.18 (730)	539.23 (871)	479.56 (382)	469.61 (423)
Consistency value (*SD*)	0.38 (0.30)	0.30 (0.16)	0.43 (0.29)	0.38 (0.22)
Low-frequency phonetic-valid phonograms				
1	悸	忌	骇	喧
2	禄	麓	俸	沐
3	馍	蘑	饼	痒
4	讥	汲	讽	址
5	涧	溅	溪	薪
6	媚	镁	妩	溢
7	狡	礁	诈	莲
8	苞	褓	蕾	矜
9	拭	狮	拂	熄
10	俐	沥	伶	狰
11	厢	饷	坊	茵
12	愚	屿	笨	殉
Character frequency (*SD*)	10.05 (4.78)	11.10 (5.97)	11.10 (7.88)	10.66 (7.55)
Phonetic frequency (*SD*)	521.55 (512)	485.64 (448)	543.65 (689)	492.82 (392)
Consistency value (*SD*)	0.84 (0.19)	0.84 (0.21)	0.78 (0.28)	0.82 (0.28)

(*Appendix continues*)

Appendix B (*continued*)

Materials for Experiment 2

No.	Core	Homophone	Synonym	Control
Low-frequency phonetic-invalid phonograms				
1	睹	杜	瞥	烘
2	懈	泄	怠	洽
3	浊	灼	垢	挛
4	魂	馄	魄	谦
5	烹	抨	煮	靴
6	恕	墅	饶	倩
7	懦	诺	慑	押
8	橙	瞠	桔	勘
9	摞	裸	砌	砰
10	稠	酬	糯	倔
11	歼	煎	剿	揩
12	愧	渍	惭	抠
Character frequency (*SD*)	10.44 (8.83)	10.99 (8.83)	11.55 (9.02)	10.61 (7.73)
Phonetic frequency (*SD*)	448.62 (420)	489.50 (600)	411.05 (377)	464.64 (461)
Consistency value (*SD*)	0.42 (0.29)	0.48 (0.26)	0.38 (0.26)	0.40 (0.27)

Received June 25, 1997
Revision received January 13, 1999
Accepted January 13, 1999 ■