The feedback consistency effect in Chinese character recognition: Evidence from a psycholinguistic norm

Chia-ying Lee[1,2,3], Chun-hsien Hsu[1*], Ya-ning Chang[1], Wei-fan Chen[1], and Pei-chun Chao[2]

[1]Academia Sinica
[2]National Yang-Ming University
[3]National Central University

* Corresponding author: Dr. Chun-hsien Hsu
Postal address: Institute of Linguistics, Academia Sinica, No. 128, Section 2, Academia Road, 115 Taipei, Taiwan
Phone: +886-2-2652-5000 # 6132
Fax: +886-2-2785-6622
E-mail: kevinhsu@gate.sinica.edu.tw

Abstract

Researchers have clearly shown the importance of phonology in literacy acquisition and in visual word recognition. For example, the spelling-to-sound consistency effect has been observed in visual word recognition tasks, in which the naming responses are faster and more accurate for words with the same letters that also have the same pronunciation (e.g., -ean is always pronounced /in/ as in *lean*, *dean*, *bean*, etc.). In addition, some studies have reported a much less intuitive feedback consistency effect when a rime can be spelled in different ways (e.g., /ip/ in *heap* and *deep*) in lexical decision tasks. Such findings suggest that with activation flowing back and forth between orthographic and phonological units during word processing, any inconsistency in the mappings between orthography and phonology should weaken the stability of the feedback loop and, thus, should delay recognition. However, several studies have failed to show reliable feedback consistency in printed word recognition. One possible reason is that the feedback consistency is naturally confounded with many other variables, such as orthographic neighborhood or bigram frequency, as these variables are difficult to tease apart. Furthermore, there are challenges in designing factorial experiments that perfectly balance lexical stimuli on all factors besides feedback consistency. This study aims to examine the feedback consistency effect in reading Chinese characters by using a normative data of 3423 Chinese phonograms. We collected the lexical decision time from 180 college students. A linear mixed model (LMM) analysis was used to examine the feedback consistency effect by taking into account additional properties that may be confounded with feedback consistency, including character frequency, number of strokes, phonetic combinability, semantic combinability, sematic ambiguity, phonetic consistency, noun-to-verb ratios, and morphological boundedness. Some typical effects were observed, such as the more frequent and familiar a character is, the faster one can decide it is a real character. More importantly, the LMM analysis revealed a significant feedback consistency effect while controlling for the other factors, which indicated that the pronunciation of phonograms might accommodate the organization of Chinese orthographic representation. Our study disentangled the feedback consistency from the many other factors and supports the view that phonological activation would reverberate to orthographic representation in visual word recognition.

Key words: feedback consistency, lexical decision, interactive activation models, linear mixed model

1. Introduction

Understanding the nature of the interactions among orthography, phonology, and semantics has been a central issue in developing models of visual word recognition. Speech has primacy over written language both in the history of mankind and in the life of an individual (Liberman 1992). Acquiring spoken language primarily involves mastering the linkage between phonology and semantics. To become literate, one must also develop efficient phonological recoding strategies for mapping visual symbols onto speech sounds to access meaning. Therefore, phonology plays a critical role in literacy acquisition. Evidence from alphabetic languages has shown that heterographic homophones (i.e. orthographically and semantically distinct words that share the same pronunciation, e.g., MAID-MADE) require more time and cognitive effort to process (Ferrand & Grainger 2003; Pexman et al. 2001) and are more prone to semantic confusion (Van Orden 1987) than non-homophonic words. This disadvantage is taken to reflect competition driven by mandatory phonological processing during visual word recognition. Studies also found a feedforward consistency effect (i.e., whether orthography is pronounced consistently) in reading aloud. That is, naming responses are faster and more accurate for words that have orthographic and phonological consistency. For example, since -EAN in the final position of a word in English is always pronounced /in/ (e.g. lean, dean, bean, etc.), processing such words would have advantages in naming tasks (Jared 1997, 2002; Taraban & McClelland 1987). These findings support the automatic activation of phonological codes in visual word recognition.

While there is a general agreement that visual word recognition is influenced by feedforward consistency, the bi-directional interactive activation model (BIAM) (Grainger & Ferrand 1994; Grainger & Ziegler 2008) propose that the *feedback* consistency, which refers to whether phonology is spelled consistently, may also affect visual word recognition. The BIAM assumes that visual word recognition relies on bidirectional co-activation of orthographic, phonological and semantic units (Grainger & Ziegler 2008). The cross-code consistency represents the level of compatibility or coherence of all co-activated representations across the orthographic, phonological, and semantic codes. The mapping consistency between orthography (O) and phonology (P) can be measured bi-directionally, namely O-to-P consistency and P-to-O consistency. Any inconsistency in the mappings between orthography and phonology would weaken the stability of the feedback loop. Thus, in addition to an O-to-P (feedforward) consistency effect on visual word recognition, the BIAM also predicts a P-to-O (feedback) consistency effect. However, it remains controversial on whether the feedback consistency effect can be reliably found in visual word recognition. The present study attempts to address this question by collecting and analyzing lexical decision times for 3423 Chinese phonograms. The unique contribution of feedback consistency in Chinese character recognition will be examined by using a linear mixed model analysis (Baayen et al. 2008) that takes into account potential confounding factors including character frequency, familiarity, number of strokes, phonetic combinability, semantic combinability, sematic ambiguity, phonetic consistency, noun-to-verb ratios, and morphological boundedness.

In English, the feedback consistency is defined as whether words contain rimes that could be spelled in multiple ways (i.e., /ʌk/ is consistent in that it is always spelled -UCK, while /ip/ is inconsistent since it could be spelled either -EEP or –EAP) (Stone et al. 1997). Stone et al. found responses in lexical decision to feedback inconsistent words were much longer than those to consistent words. Moreover, more errors were made in recognizing feedback inconsistent words than feedback consistent words (9.8% versus 3.3%). Ziegler et al. (1997) replicated the feedback consistency effect in French, a language with a high degree of feedback inconsistency. These findings suggest a role for feedback from phonology to orthography and thus posit a major theoretical challenge to more traditional bottom-up theories of word recognition (Norris et al. 2000).

However, there have been some challenges to whether the feedback consistency effect can be reliably found across languages and in different tasks. For example, in French, Peereman and Bonin (1998) found the feedback consistency effect in writing but not in lexical decisions. While they were able to replicate the feedback consistency effect using stimuli of Ziegler et al. (1997), the feedback consistency effect disappeared when subjective familiarity was partialled out (Peereman & Bonin 1998). After orthogonally manipulating the feedforwad and feedback consistency at both rime and onset levels, Ziegler et al. (2008) failed to find the feedback consistency effect in visual lexical decision task. In English, however, the feedback consistency effect seems to be more robust. For example, Lacruz and Folk (2004) reported stable feedforward and feedback consistency effects in both lexical decision and naming tasks when controlling for word frequency, subjective familiarity, and a number of other variables, such as word length, orthographic neighborhood, bigram frequency, and summed frequency of friends. By doing so, only eight items were available for each cell at the cost of matching variables.

A possible reason for these discrepancies is that, most studies only look at one or two linguistic properties, such as feedback consistency or feedforward consistency, in a factorial design, in order to determine the effect of a particular variable on response latencies. However, most psycholinguistic variables are continuous in nature, and many co-vary in complicated ways. By dichotomizing the continuous variables into categories (e.g., high versus low consistency) and trying to match other variables (e.g., word frequency), factorial designs might result in a substantial loss of statistical power and require selection of atypical stimuli. A potential solution is to investigate the existence of feedforward and feedback consistency effects in a large-scale database. For example, Balota and colleagues (2004) collected speeded naming and lexical decision responses for 2438 single-syllable English words and used hierarchical regression to examine the unique contribution of feedforward and feedback consistency at both rime and onset levels. Their study demonstrated reliable feedback consistency effects at the rime level on naming and lexical decision performance, especially for slow participants, after controlling for word frequency, familiarity, along with other related variables (e.g., feedforward consistency, neighborhood size and word length, etc.). It appears that feedback consistency effects in visual word recognition are less robust in languages, English being one exception. One possible explanation for this exception is that the rime plays a special role not only in phonology but also in orthography in English (Goswami et al. 2005; Ziegler et al. 2008). In light of these controversial findings, the role of feedback consistency in visual word recognition deserves further study as this theoretically intriguing effect has major implications about bidirectional coupling between orthography and phonology. Specifically, although phonology would be co-activated whenever a visual word is processed, it remains to be decided whether there would be reverberation or resonance (feedback loop) influences from phonology to orthography.

—insert Figure 1 here—

Chinese, a non-alphabetic system, may offer an interesting insight into the issue. Originally, Chinese characters were mainly designed to resemble objects that they represented either literally (pictographs) or metaphorically (ideographs). However, not all concepts or meanings are concrete enough to be conveyed by their pictorial resemblance to a physical object. By the Shang dynasty (1600–1046 BC), many phonograms had been created to solve this problem by compounding a semantic radical (i.e., 木, mu4, wood) and a phonetic radical (i.e., 風, fon1, wind) to form a phonogram (DeFrancis 1989). In modern Chinese, over 80% of Chinese characters are phonograms. Very often (for around 90% of compound characters), the semantic radical appears on the left and the phonetic radical on the right (Lo et al. 2007). The semantic radical may give a hint for character's meaning and the phonetic radical may provide information for character's pronunciation.

Given the majority of Chinese characters are phonograms, studies have attempted to measure the bidirectional mapping consistency between Chinese orthography and phonology with the representation of statistical relationships between orthographic forms and their pronunciations. For example, the O-to-P consistencies for Chinese characters refer to the reliability of a phonetic radical (e.g., and in Figure 1) provide a phonological clue defined by phonetic consistency, or whether the pronunciation of a character agrees with those of its orthographic neighbors containing th䍃same 㐬onetic radical. A series of behavioral studies have demonstrated that Chinese readers capture the mapping consistency between character and sound (Fang et al. 1986; Lee et al. 2005; Tzeng et al. 1995) and the neural correlates responsible for the Chinese O-to-P transformation are very similar to what has been suggested for reading alphabetic writing systems (Lee et al. 2010; Lee et al. 2004). Meanwhile, ERP evidence suggests that the consistency effect in reading Chinese impacts early sublexical phonological computation and later lexical semantic competition (Lee et al. 2007; Lee et al. 2006). Most importantly, Hsu et al., (2009) examined if the consistency effect would be modulated by the orthographic neighborhood size of phonetic radical, which is defined as the number of phonograms sharing the same phonetic radical and termed as phonetic combinability (Feldman and Siok, 1997). They found that the consistency effect was mainly found in phonograms with large phonetic combinability (e.g., "碗" /wan3/, its phonetic radical 宛 is shared by 6 phonograms, such as "腕", "婉", "琬", "惋", "豌", and "蜿"), but not in those with small phonetic combinability (e.g., "綴" /zhui4/, sharing its phonetic radical with only three phonograms, including "掇" /duo2/, "輟" /chuo4/, and "啜" /chuo4/). It seems that characters with larger phonetic combinability are likely to exert more pressure toward awareness of phonological validity represented by phonetic radicals than do those with smaller phonetic combinability. This suggests an interplay between phonetic combinability and the mapping from orthography to phonology in the different stages of lexical processing (Hsu et al. 2009).

Compared to the extensive evidence of the consistency in O-to-P mapping in reading Chinese, very few studies have addressed the influence of P-to-O mapping consistency effect on Chinese visual word recognition (Chen et al. 2009). Unlike English in which words are composed of letters that represent phonemes, Chinese characters map onto single-syllable morphemes (Myers 2010). Chinese syllables have a relatively simple form, with the majority have a consonant-vowel structure, and only two consonants, the nasal consonants /n/ and /ŋ/, can follow the vowel in Mandarin (Hua & Dodd 2000). As a consequence, based on Academia Sinica Balanced Corpus of Modern Chinese (the Sinica Corpus, hereafter) (Chen et al. 1996), there are only about 1100 syllables, including tones, distributed over about 5,000 characters. Given the pervasive homophony in Chinese, a greater impact from orthography may be expected during Chinese spoken word recognition. There are two ways to measure the mapping consistency from phonology to orthography in Chinese. The first is homophone density, which is defined as the number of characters sharing exactly the same pronunciation (including tone). The second is orthographic consistency (P-O feedback consistency), which is defined as whether a set of homophones can be divided into subgroups based on their sublexical orthographic units (phonetic radicals). For example, in Figure 1, homophone density of the syllable /ma3/ is 5, given it can be corresponded to five characters, "碼", "瑪", "馬", "螞", and "嗎". Moreover, /ma3/ is a orthographically consistent syllable because these five homophones sharing the same phonetic radical 馬. On the other hand, the syllable /pu4/ can be mapped onto 5 homophones (“暴”, “曝”, “瀑”, “鋪” and “舖”), yet these homophones can be subdivided into two groups: one sharing the phonetic radical “暴” and the other one sharing the phonetic radical “甫”. Thus, /pu4/ is an orthographically inconsistent syllable. Studies have demonstrated that homophone density plays a role on visual word cognition (Chen et al. 2009; Liu et al. 2007; Tan et al. 2001), although whether the effect is facilitative or inhibitory remains debatable. For example, Chen et al (2009) demonstrated a facilitative homophone density effect on a Chinese character recognition task, suggesting that a single representation for homophones at the phonological word form level while the null effect of homophone density was reported by Liu et al. (2007). In addition, whether the orthographic consistency, which is defined at the sublexical level, would affect visual word recognition has not been systematically studied yet.

This study aims to examine the role of P-O feedback consistency in Chinese visual word recognition. Research on visual word recognition has traditionally involved factorial designs by manipulating a target variable. Words are selected and divided into groups, which differ in specific target variables (e.g., word frequency, spelling-to-sound regularity, and neighborhood density). The effects of these variables on the speed and accuracy of naming or making lexical decisions are measured. However, both feedback and feedforward consistency may naturally confound with many other variables, such as frequency and orthographic neighborhood. It is difficult to categorize each item in an experiment on the multiple dimensions necessary for factorial crossing of relevant variables. Therefore, this study aims to collect lexical decision times for a large-scale database of 3423 Chinese phonograms and to perform linear mixed model (LMM) analyses (Baayen et al. 2008) on the reaction time. Traditionally, psycholinguists and cognitive psychologists have used analysis of variance (ANOVA) or hierarchical regression techniques to estimate the effects of treatments, and have computed F-statistics by measuring the variation around the means from participants or items separately. Compared to traditional ANOVA or hierarchical regression analysis, LMM analysis has several advantages. First, multiple random variables can be included simultaneously. Second, parameter estimates are fitted by a restricted maximum likelihood criterion, which is not biased to the means of participants or items. Therefore, LMM analysis can handle unbalanced data sets and missing observations. Third, F-statistics assume that variance of treatments is homogenous (the sphericity assumption), an assumption that is sometimes violated in response latencies. Bagiella et al. (2000) have demonstrated that LMM analysis can provide more conservative results than F-statistics when the data is non-spherical. Finally, as Baayen, Davidson, and Bates (2008) have argued, LMM analysis makes fewer incorrect rejections of a true null hypothesis (i.e., the Type I error) and is less biased by number of observations than F-statistics. To look for the feedback consistency effect in Chinese character recognition, in addition to P-to-O mapping consistency factors, a set of potential confounding factors were selected on the basis of previous studies (Baayen et al. 2008; Liu, Shu, and Li, 2007) and were incorporated as fixed factors in LMM analysis. The factors included seven lexical variables (i.e., character frequency, number of strokes, familiarity, phonetic combinability, semantic combinability, homophone density, and feedforward consistencies) and three morphological variables (i.e., noun-to-verb ratios, morphological boundedness and semantic ambiguity). The details of these variables will be further elaborated in the Predictors section.

2. Method

2.1 Materials

The Sinica Corpus (Chen et al. 1996) is based on more than five million words (approximately 10 million characters) culled from written materials, including textbooks, newspapers, works of literature, popular works of fiction and nonfiction, and transcripts. There are 5640 characters in the Sinica Corpus, including 3967 phonograms. Of the phonograms, 544 are homographs. In the lexical decision study, homographs were excluded. In total, 3423 phonograms were used in the lexical decision study. The phonograms and predictors can be accessed via an online inquiry system (http://ball.ling.sinica.edu.tw/namingdatabase/index.html) (Chang et al. in press).

2.2 Predictors

Ten predictors were included in the study. The correlation coefficients between predictors can be seen in Table 1. The definitions of each predictor are as follows.

(1) *Log frequency*. The log-transformed frequency of each character was calculated based on character frequency taken from the Sinica Corpus. The corpus consists of more than five million words (approximately 10 million characters) culled from textbooks, newspapers, works of literature, popular works of fiction and nonfiction, and transcripts. The character frequency is defined as the number of times that a character appears per million.

(2) *Feedforward consistency*. For character recognition, feedforward consistency here refers to whether the pronunciation of a character agrees with those of its orthographic neighbors, which contain the same phonetic radical, regardless of tonal differences (Lee et al., 2004, 2007, 2010). In the literature, there are two ways to measure Chinese phonetic consistency. The first consistency index (CI) is to measure the proportion of a specific pronunciation within a set of orthographic neighbors that share the same phonetic radical. The second CI considers character frequency, also known as the frequency-weighted consistency index. It is defined as the ratio of the summed frequencies of characters sharing a phonetic radical that have the same pronunciation to the summed frequencies of characters sharing that phonetic. Studies of English and Chinese have demonstrated the importance of the frequency-weighted consistency index in naming responses (Jared et al., 1990; Lee et al., 2005). Thus, this study used the frequency-weighted consistency measure for analysis.

(3) *Feedback consistency*. This index is defined as the ratio of the summed frequency of homophones with the same phonetic radical to the summed frequency of homophones.

(4) *Homophone density*. Homophone density refers to the number of characters sharing the same syllable structure and tone. In the Sinica Corpus, about 1100 syllables are represented, 80% of which were shared by more than one character. The score of homophone density varied from one to 69.

(5) *Familiarity*. Values of familiarity were taken from Chang et al., (submitted), which were based on college students' subjective ratings of how frequently they encounter a character. In their study, participants were asked to rate each character on a 7-point scale from "never encountered" to "encountered several times a day" (i.e., 1: never; 2: once a year; 3: once a month; 4: once a week; 5: once every two days; 6: once a day; 7: several times a day).

(6) *Number of strokes*. The number of strokes indicates the level of visual complexity of the character. In the present study, the number of strokes ranged from four (e.g., 仍 /reng1/) to 30 (e.g., 鸝 /li2/).

(7) *Radical combinability (phonetic and semantic).* In studies of English, the orthographic neighborhood size of a word is defined as the number of words that can be created by changing a single letter of the word while maintaining the original positions of the letters (Coltheart et al. 1977). For example, the word *cheat* has four neighbors (i.e. cheap, chest, cleat, and wheat). In Chinese, an analogy of orthographic neighborhood size is radical combinability defined by Feldman and Siok (1997). Radical combinability measures the

number of characters that share the same radical. It can be further divided into phonetic combinability and semantic combinability, which are defined as the number of phonograms that share the same phonetic and semantic radical, respectively. Parallel to the effect of orthographic neighborhood size in English (Andrews 1989, 1992, 1997), several studies have demonstrated facilitative combinability effects for both semantic and phonetic radicals—characters with large combinability revealed a faster response latency in the character decision task than characters with small combinability (Chen & Weekes 2004; Feldman & Siok 1997; Hsiao et al. 2006; Wu et al. 2012). The metrics of both phonetic combinability and semantic combinability were based on the norm reported by Hsu et al. (2009). The scales of phonetic combinability and semantic combinability ranged from one to 20 and from one to 226, respectively.

(8) *Residual semantic ambiguity*. A character is defined as semantically ambiguous if it is associated with multiple meanings. The scores used here were based on subjective ratings taken from those used in Huang et al. (2011) and Hsu et al. (2011). In their studies, each character had a mean rating score ranged from one to five to indicate the degree of semantic ambiguity (i.e. one meaning to highly diverse meanings). Specifically, the rating 1 represents that the meaning of the character is explicit and only one meaning is available (e.g. 糖 /tang2/, sweet) or that the character is a constituent of monomorphic disyllabic words (e.g. 垃 /le4/ is always used with 圾 /se4/ as 垃圾 which means garbage). The rating 2 indicates that there are two or three different meanings available for the character (e.g. 程 /cheng2/, rules or a journey). The rating 3 indicates the character have three or four different meanings (e.g. 沙 /sha1/, sand, granules, and hoarse). The rating 4 represents that the meaning of the character is fairly diverse, and there are four to five different meanings available for the character (e.g. 風 /feng1/, wind, style, landscape, and disease). The rating 5 represents that the meaning of the character is highly diverse and can have more than 5 different meanings (e.g. 花 /hua1/, flower, spend, colorful, blurred, pattern, trick, peanut and etc.). Because scores of semantic ambiguity were highly correlated with log frequency (r = .71) and familiarity (r = .68), we regressed character frequency and familiarity out of semantic ambiguity scores to reduce collinearity. The residualized score was called residual semantic ambiguity and could be interpreted in the same way as semantic ambiguity. The correlations between log-transformed frequency, familiarity, and residual semantic ambiguity are shown in Table 1.

(9) *Morphological boundedness*. In this study, the concept of morphological boundedness indicated whether the character could be able to stand alone as a word based on Sinica Corpus. For instance, neither “蚯” nor “蚓” is a morpheme when used in the typical sense of “蚯蚓”, “an earthworm”. Scores of morphological boundedness were categorized into two numerical values: 0 indicates the character could not be a monomorphemic word, and 1 indicates the character is an attested monomorphemic word. It is worthy noting that in the present study, the morphological boundedness score does not directly reflect the productivity of morphemes, which indicates the ability to combine with many different morphemes.

(10) *Noun-to-verb ratio*. Although the procedure of lexical decision task does not provide contextual information, behavioral studies have demonstrated that noun-to-verb ratios are relevant for the processing of morphologically simple words (Baayen & Moscoso del Prado Martín 2005; Feldman & Basnight-Brown 2005; Tabak et al. 2005). This score was taken from Hsu et al. (2011). Based on the token frequency of words that consist of target characters, noun-to-verb ratio is calculated by estimating the ratio of the sum of frequencies of nominal words to the sum of frequencies of verbal words. A large noun-to-verb ratio indicates that a character is frequently used in nominal words, whereas a small noun-to-verb ratio indicates that a character is frequently used in verbs.

—insert Table 1 here—

2.3 Procedure

One hundred and eighty college students (18 to 30 years old) volunteered to participate as paid participants in the lexical decision experiment. All participants are Chinese native speakers (111 females and 69 males) and were college students recruited from Chung Yuan Christian University. The experiment was done in a behavioral testing room. In the lexical decision task, 3423 characters were randomly separated into six subsets. Five subsets had 570 phonograms, and the remaining one had 573 phonograms. Each participant only received one subset of phonograms. Therefore, every six participants would complete a corpus. To balance "yes" and "no" responses during the lexical decision, radicals of real phonograms were used to create 573 "no" items by combining two radicals that do not co-occur in Chinese (pseudo-characters) or by combining two radicals and changing one stroke (non-characters). Each participant would either received 1146 trials or 1140 trials. In the experiment, trials were randomized and separated into ten blocks. Participants would take a break after finishing each block, and they took about 40 minutes to finish the experiment.

2.4 LMM analysis

In the lexical decision task, the average error rate was 16.6% (SD = 3.8%). Before performing LMM analyses, incorrect responses were excluded. For correct responses, trials with extreme responses (longer than 1500 ms or shorter than 300 ms) were also excluded. The average lexical decision latency was 546 ms (SD = 57) for real characters. The response latencies were log transformed for subsequent LMM analyses following Kuperman et al. (2009). The same dependent variable was analyzed in two sessions with different purposes. In the first session, three predictors of interest, including log frequency, feedback consistency and homophone density, and the two-way interactions involving log frequency were considered as fixed effects. A likelihood ratio test was used to evaluate the linear mixed models. By comparing models, these analyses should indicate whether feedback consistency, homophone density and their interactions with log frequency account for the changes in response latencies. In the second session, we utilized all predictors to evaluate whether the results in the first session would remain significant.

The LMM analysis was done by using the lmer program of the lme4 package (Bates & Sarkar 2007). To evaluate the significance of each fixed effect, we employed the pvals.fnc program of the languageR package to perform a Markov Chain Monte Carlo simulation, using 1000 simulations to obtain the p values. These packages are supplied in the R system for statistical computing (Ver. 2.9.1; R Development Core Team, 2009).

2.5 Data analysis

2.5.1 Designs of LMM models with log frequency, feedback consistency and homophone density

A Control Model was first estimated by including log frequency as a fixed effect and participants and items as crossed random effects, following previous studies that showed frequency notably predicts response latencies of lexical decision (Allen et al. 2005; Balota et al. 2004) and that estimating random intercepts for participants and items is necessary for analyzing data of linguistic experiments (Baayen et al. 2008). From the Control Model, Model 1, 2, and 3 were fashioned by incorporating additional predictors and two-way interactions as fixed effect. Model 1 included homophone density and the interaction between homophone density and log frequency. Model 2 included feedback consistency and the interaction between feedback consistency and log frequency. Model 3 included homophone density, feedback consistency, the interaction between homophone density and log frequency, and the interaction between feedback consistency and log frequency. Each of these models was compared to the Control Model to estimate the likelihood ratio and the χ^2. The significance of each predictor was defined by reaching a significant criterion at the .05 level. The *p* values were obtained through Markov Chain Monte Carlo (MCMC) sampling supported by the pvals.fnc program of the language R package (Baayen, Davidson, & Bates, 2008). While any two-way interaction was not significant, the model was refitted without the interaction.

2.5.2 Design of the LMM model with all predictors

To further evaluate effects of feedback consistency and homophone density, all lexical predictors were included as fixed effects. Based on the suggestion of Baayen et al. (2006), a backward stepwise procedure was adopted here to identify significant effects of polynomial terms and two-way interactions between predictors of interest, and the trial numbers were also included as predictors. Previous studies have suggested that frequency appears to modulate effects of orthographic neighborhood size, consistency, word length, etc (Andrews 1989, 1992; Balota et al. 2004; Lee et al. 2005). Therefore, only two-way interactions including log frequency were tested in Model 4. Once any interaction effect was not significant, the non-significant interaction term was excluded and the LMM model was refitted. This is due to that non-significant interaction terms might decrease the influence of predictors.

3. Results
3.1 LMM models with log frequency, feedback consistency and homophone density

—insert Table 2 here—

Table 2 shows that, in comparison to the Control Model, the three different sets of predictors involved in Models 1, 2 and 3 significantly increased the likelihood ratios of LMM models. The interaction of homophone density and log frequency was excluded from Model 1 and 3 because this effect was not significant. In Model 1, both log frequency and homophone density were significant predictors (*p*s $< .001$). Model 2 focused on feedback consistency effects; both log frequency and feedback consistency were significant predictors (*p*s $< .001$). Additionally, the interaction between log frequency and consistency was also significant ($p < .05$). When the predictors in Model 1 and Model 2 were combined in Model 3, all predictors remained significant (*p*s $< .001$). Results across models indicated a robust effect of log frequency on lexical decision latencies. That is, high frequency characters revealed faster responses than low frequency characters did. In addition, Model 3 indicated significant effects of homophone density, feedback consistency and the interaction between consistency and log frequency on latencies of lexical decision. Moreover, Model 3 had higher likelihood ratio than Model 1 ($\chi^2 = 21.28$, $p < .001$) and Model 2 ($\chi^2 = 18.09$, $p < .001$). The beta values of Model 3 indicated that characters with many homophones elicited longer response latencies than those with few homophones. High consistency characters yielded shorter response latencies than low consistency characters. Regarding the interaction predictor, only low frequency characters showed the facilitative effect of feedback consistency. The details of this interaction will be addressed in the next section.

3.2 LMM models with all predictors

—insert Table 3 here—

In Model 4, the LMM analysis started by estimating all predictors, their polynomial terms, and two-way interactions. Then, non-significant polynomial terms and interactions were excluded and the model was refitted. Table 3 shows the results of Model 4 with all predictors and significant effects of four sets of two-way interactions and three polynomial terms. Two predictors, feedforward consistency and homophone density, did not show significant effects (ps >.1). The likelihood ratio test indicated a significant improvement after including the polynomial terms and the two-way interactions (χ^2 =104.04, $p < .001$). Like the results in Model 3, Model 4 also showed facilitative effects of log frequency and feedback consistency. As expected, familiar characters revealed faster responses than unfamiliar characters did. Both semantic and phonetic radical combinability also revealed facilitative effects. That is, characters with large radical combinability were associated with decreased in response latencies. As for the number of strokes, the results indicated that complex characters elicited slower responses than simple characters did. Predictors of semantic and morphological properties also showed facilitative effects on response latencies. Semantically ambiguous characters (i.e. those with multiple meanings) elicited faster responses than unambiguous characters. The effect of morphological boundedness indicated that responses to monomorphemic characters were faster than those could not stand as a word by themselves. Finally, the noun-to-verb ratio also showed a significant effect. Nominal characters had faster responses than verbal characters.

—insert Figure 2 here—

Model 4 showed three significant interactions: (1) feedback consistency by log frequency, (2) phonetic combinability by log frequency, (3) residual semantic ambiguity by log frequency. Figure 2a shows feedback consistency effects in high, median, and low levels of log frequency. In low frequency characters, high feedback consistency characters showed faster response than low consistency characters did. In the highest level of log frequency, interestingly, an opposite pattern of feedback consistency was found. Finally, both phonetic combinability and semantic ambiguity effects interacted with log frequency. Figures 2b and 2c show facilitative effects of combinability and semantic ambiguity in every level of log frequency.

4. Discussion

This study aimed to explore whether phonological activation would reverberate to orthographic representation in Chinese character recognition. To overcome the limitation of factorial designs (Balota et al. 2004), lexical decision times of 3424 Chinese phonograms were collected and LMM analyses were performed with various lexical and sublexical variables of Chinese characters. First of all, the present results replicate several robust phenomena in the studies of visual word recognition. For example, Model 4 showed that both familiarity and character frequency yield facilitative effects on lexical decision latencies. Similar findings have been demonstrated in normative lexical decision and naming data for English single-syllable words (Balota et al., 2004) and in the naming data for 2423 Chinese characters (Liu et al., 2007). For the characteristics that are specific for Chinese characters, our data revealed that reading characters with more strokes elicits longer response latencies than reading those with fewer strokes and such an effect for number of strokes is consistent with the results in Liu et al. (2007) and other behavioral studies (Leong et al. 1987). In addition, both phonetic combinability and semantic combinability showed facilitative effects on lexical decision latencies. In particular, the LMM analysis revealed a significant interaction between log frequency and phonetic combinability, as in Figure 2(b). Although the facilitative phonetic combinability effects could be obtained in all frequency ranges, but such an effect tends to be stronger in reading low frequency characters. These results are in general consistent with the findings in Feldman and Siok (1997, 1999) and the orthographic neighborhood size effects found in alphabetic writing systems (Andrews, 1989; 1992). In summary, these findings indicated that using LMM analyses with normative data could corroborate the results observed by using factorial designs.

The most important contribution of this study is to separate two types of orthographic consistency effects, homophone density and feedback consistency, on Chinese visual word recognition. The result of Model 3 showed an inhibitory effect of homophone density. However, when other lexical variables and two-way interactions were considered in Model 4; the homophone density effect was no longer significant. The inhibitory homophone density effect evident in Model 3 appears to be inconsistent with previous studies that have either reported a facilitative effect of homophone density (Chen et al. 2009; Ziegler et al. 2000) or a null effect of homophone density (Liu et al. 2007) on Chinese visual word recognition. However, the present finding is congruent with studies of English and French, which have consistently shown prolonged response latencies in reading words with higher homophone density (Pexman et al. 2001; Rubenstein et al. 1971). The inhibitory homophone density effect may be due to the competition from the homophone mates that were activated from the shared pronunciation, and this supports that visual word recognition would also receive feedback activation from phonology.

Unlike the homophone density that was only survive in Model 3, both Model 3 and Model 4 showed the facilitatory feedback consistency effect. Most importantly, the feedback consistency effect persisted when other lexical variables were considered in Model 4; however it is not the case for homophone density. According to the bi-directional interactive activation model (Grainger & Ferrand 1994; Grainger & Ziegler 2008), visual word recognition is influenced not only by the spelling-to-sound consistency, but also by the sound-to-spelling consistency. It is possible that reading a Chinese character may be facilitated by feedback connections from phonology to orthography if a character shares the similar orthographic unit, i.e., phonetic radical, with its homophone mates; otherwise the inhibitory effect may be expected. This might explain the contradictory findings of homophone density effects in previous studies on Chinese character recognition (Chen et al. 2009; Liu et al. 2007), since neither study considered feedback consistency.

There have been controversial findings about whether the feedback consistency can be reliably found in visual lexical decision and one of the major debates is whether the word frequency and subjective familiarity are well matched. Although Stone et al., (1997) reported a facilitative feedback consistency effect, word frequency ratings used in their study were taken from the Brown Corpus (Kučera & Francis 1967), which has a relatively limited size and is based on about one million words. However, when Peereaman et al., (1998) used a much larger CELEX corpus (14 million words) to calculate word frequencies of Stone et al.'s stimuli and found their consistent words were significantly more frequent than were their inconsistent words. A question raised by Peereaman et al. (1998) is whether the feedback consistency effect resists to a better control for word frequency. To resolve this, Balota et al., (2004) demonstrated facilitative feedback consistency effects in lexical decision and naming data from a large-scale database, while controlling for word frequency, subjective familiarity, and a number of other variables. With a similar research approach, our LMM analyses demonstrated a facilitative feedback consistency effect and its interaction with log frequency in the normative data of lexical decision times. A high feedback consistency character, by definition, has a large set of homophones and most of these homophones share the same phonetic radical. Our data suggest that high feedback consistency characters tend to yield a faster lexical decision time than low feedback consistency characters do. However, such an effect was only reliable for low frequency characters. The facilitative feedback consistency effect in reading low frequency characters is congruent with the feedback consistency effect in English (Balota et al., 2004; Lacruz and Folk, 2004), supporting the reverberation (a feedback loop) from phonology to orthography. To be more specific, reading a character would activate its phonology and which in turn sends feedbacks to orthography and activate a set of homophones. The target character would receive better facilitation if these homophones share similar orthographic features, such as phonetic radical.

As can be seen in Figure 2(a), the interaction between feedback consistency and log frequency might indicate two types of relationships between phonological neighbors for high versus low frequency characters. The facilitative feedback consistency effect in reading low frequency characters might reflect co-activation among homophones, which in turn sends feedback to the orthographic level and helps access low frequency characters. In contrast, in reading high frequency characters, co-activation among homophones might somehow cause interference during lexical retrieval. Since the feedback consistency was estimated by weighting the frequency of homophones, the high frequency and high feedback consistency characters might trigger representations of many high frequency homophones. Therefore, the inhibitory effect of feedback consistency in reading high frequency characters might reflect competition among homophones.

The present results also provide supportive evidence for semantic involvement in lexical decisions. Although the task did not require one to consider the meaning of a character explicitly, lexical decision latencies were affected by semantic ambiguity, noun-to-verb ratio, and morphological boundedness. The facilitative effect of semantic ambiguity is consistent with findings in Rodd et al. (2002) and Baayen et al. (2006), which shows that the rich semantic representations associated with words facilitate their recognition. Additionally, LMM analyses also showed an interaction between semantic ambiguity and log frequency. This interaction indicates that the facilitative semantic ambiguity effect is strong in reading low frequency characters, yet this effect diminishes in reading high frequency characters. . The facilitative effect of noun-to-verb ratio agreed with the conventional finding that reading nominal words is easier than reading verbal words (Baayen et al. 2006). Finally, the morphological boundedness (whether the character could be able to stand alone as a word) also affects the lexical decision time and suggests that reading characters involves the processing of their morphological properties.

5. Conclusion

In conclusion, the present study provides an LMM analysis with lexical decision latencies collected from a relative large sample of participants and items. The main finding demonstrates a facilitative effect of feedback consistency in lexical decision latencies, supporting the view that phonological activation would reverberate to orthographic representation in visual word recognition.

Acknowledgement

This research was supported by grants under the National Science Council NSC96-2628-H-001-058-MY3 and NSC102-2420-H-001-006-MY2, and Academia Sinica AS-102-

TP-C06. We are grateful to Wen-Hsuan Chan, Shu-Min Huang, and Angela Ku-Yuan Tzeng for their helps in conducting the lexical decision experiment.

Reference

Allen, Philip A., Albert F. Smith, Mei-Ching Lien, Jeremy Grabbe & Martin D. Murphy. 2005. Evidence for an Activation Locus of the Word-Frequency Effect in Lexical Decision. *Journal of Experimental Psychology: Human Perception and Performance* 31.4:713-721.

Andrews, Sally. 1989. Frequency and neighborhood effects on lexical access: activation or search? *Journal of Experimental Psychology. Learning, Memory, and Cognition* 15.5:802-841.

Andrews, Sally. 1992. Frequency and neighborhood effects on lexical access: Lexical similarity or orthographic redundancy? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18.2:234-254.

Andrews, Sally. 1997. The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review* 4.4:439-461.

Baayen, R. Harald, Douglas J. Davidson & Douglas M. Bates. 2008. Mixed-Effects Modeling with Crossed Random Effects for Subjects and Items. *Journal of Memory and Language* 59.4:390-412.

Baayen, R. Harald, Laurie B. Feldman & Robert Schreuder. 2006. Morphological Influences on the Recognition of Monosyllabic Monomorphemic Words. *Journal of Memory and Language* 55.2:290-313.

Baayen, R. Harald & Fermín Moscoso del Prado Martín. 2005. Semantic density and past-tense formation in three Germanic languages." (2005): 666-698. *Language* 666-698.

Bagiella, Emilia, Richard P. Sloan & Daniel F. Heitjan. 2000. Mixed-effects models in psychophysiology. *Psychophysiology* 37.01:13-20.

Balota, David A., Michael J. Cortese, Susan D. Sergent-Marshall, Daniel H. Spieler & Melvin J. Yap. 2004. Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General* 133.2:283-316.

Bates, Douglas & Deepayan Sarkar. 2007. lme4: Linear mixed-effects models using S4 classes.

Chang, Ya-ning, Chun-hsien Hsu, Jie-li Tsai, Chien-liang Chen & Chia-ying Lee. in press. The Psycholinguistic Database for Chinese Character Naming. *Behavior research methods*. doi: 10.3758/s13428-014-0559-7

Chen, Hsin-chin, Jyotsna Vaid & Jei-tun Wu. 2009. Homophone density and phonological frequency in Chinese word recognition. *Language and Cognitive Processes* 24.7-8:967-982.

Chen, Keh-jiann, Chu-ren Huang, Li-ping Chang & Hui-li Hsu. 1996. Sinica Corpus: Design Methodology for Balanced Corpora. *Proceeding of the 11th Pacific Asia Conference on Language, Information and Computation*, 167-176. Seoul: Kyung Hee University.

Chen, May-jane & Brendan S. Weekes. 2004. Effects of semantic radicals on Chinese character categorization and character decision. *Chinese Journal of Psychology* 46.2&3:181-196.

Coltheart, Max, Eileen Davelaar, Jon Torfi Jonasson & Derek Besner. 1977. Access to the internal lexicon. *Attention and Performance VI*, ed. by Stanislav Dornič, 535-555. Hillsdale, N.J: Lawrence Erlbaum.

DeFrancis, John. 1989. *Visible speech: The diverse oneness of writing systems.* Honolulu: University of Hawaii Press.
Fang, Sheng-ping, Ruey-yun Horng & Ovid J. L. Tzeng. 1986. Consistency effects in the Chinese character and pseudo-character naming tasks. *Linguistics, psychology, and the Chinese language*, 11–21.
Feldman, Laurie B. & Dana M. Basnight-Brown. 2005. Semantic influences on morphological processing: As revealed by the forward masked priming task. Paper presented at 4th International Workshop on Morphology, Cambridge, England.
Feldman, Laurie B. & Witina W.T. Siok. 1997. The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 23.3:776-781.
Feldman, Laurie B. & Witina W.T. Siok. 1999. Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language* 40.4:559-576.
Ferrand, Ludovic & Jonathan Grainger. 2003. Homophone interference effects in visual word recognition. *Quarterly Journal of Experimental Psychology* 56A:403-419.
Goswami, Usha, Johannes C. Ziegler & Ulla Richardson. 2005. The effects of spelling consistency on phonological awareness: a comparison of English and German. *Journal of Experimental Child Psychology* 92.4:345-365.
Grainger, Jonathan & Ludovic Ferrand. 1994. Phonology and orthography in visual word recognition: Effects of masked homophone primes. *Journal of Memory and Language* 33:218-233.
Grainger, Jonathan & Johannes C. Ziegler. 2008. Cross-code consistency in a functional architecture for word recognition. *Single-word reading: Behavioral and biological perspectives.*, ed. by Elena L. Grigorenko & Adam J. Naples, 129-157. Mahwah, NJ US: Lawrence Erlbaum Associates Publishers.
Hsiao, Janet Hui-wen, Richard Shillcock & Michal Lavidor. 2006. A TMS examination of semantic radical combinability effects in Chinese character recognition. *Brain Research* 1078.1:159-167.
Hsu, Chun-hsien, Chia-ying Lee & Alec Marantz. 2011. Effects of visual complexity and sublexical information in the occipitotemporal cortex in the reading of Chinese phonograms: a single-trial analysis with MEG. *Brain and Language* 117.1:1.
Hsu, Chun-hsien, Jie-li Tsai, Chia-ying Lee & Ovid J.-L. Tzeng. 2009. Orthographic combinability and phonological consistency effects in reading Chinese phonograms: an event-related potential study. *Brain and Language* 108.1:56-66.
Hua, Zhu & Barbara Dodd. 2000. Development and change in the phonology of Putonghua-speaking children with speech difficulties. *Clinical Linguistics & Phonetics* 14.5:351-368.
Huang, Hsu-wen, Chia-ying Lee, Jie-li Tsai & Ovid J.-L. Tzeng. 2011. The sublexical semantic ambiguity effects for reading Chinese disyllabic compounds. *Brain and Language* 17.2:77-87.
Jared, Debra. 1997. Spelling-sound consistency affects the naming of high-frequency words. *Journal of Memory & Language* 36:505-529.
Jared, Debra. 2002. Spelling-sound consistency and regularity effects in word naming. *Journal of Memory and Language* 46:723-750.
Kučera, Henry & W. Nelson Francis 1967. *Computational analysis of present-day American English.* Providence, RI: Brown University Press.
Kuperman, Victor, Robert Schreuder, Raymond Bertram & R. Harald Baayen. 2009.

Reading Polymorphemic Dutch Compounds: Toward a Multiple Route Model of Lexical Processing. *Journal of Experimental Psychology: Human Perception and Performance* 35.3:876-895.

Lacruz, Isabel & Jocelyn R. Folk. 2004. Feedforward and feedback consistency effects for high-and low-frequency words in lexical decision and naming. *The Quarterly Journal of Experimental Psychology Section A* 57.7:1261-1284.

Lee, Chia-ying, Hsu-wen Huang, Wen-jui Kuo, Jie-li Tsai & Ovid J.-L. Tzeng. 2010. Cognitive and neural basis of the consistency and lexicality effects in reading Chinese. *Journal of Neurolinguistics* 23.1:10-27.

Lee, Chia-ying, Jie-li Tsai, Hsu-wen Huang, Chun-hsien Hsu, Ovid J.-L. Tzeng & Daisy L. Hung. 2007. Temporal dynamics of the consistency effect in reading Chinese: an event-related potentials study. *NeuroReport* 18.2:147-151.

Lee, Chia-ying, Jie-li Tsai, Hsu-wen Huang, Ovid J.-L. Tzeng & Daisy L. Hung. 2006. The temporal signatures of semantic and phonological activations for Chinese sublexical processing: An event-related potential study. *Brain Research* 1121.1:150-159.

Lee, Chia-ying, Jie-li Tsai, Wen-jui Kuo, Daisy L. Hung, Ovid J.-L. Tzeng, Tzu-chen Yeh, Low-tone Ho & Jen-chuen Hsieh. 2004. Neuronal correlates of consistency and frequency effects on Chinese character naming: an event-related fMRI study. *NeuroImage* 23.4:1235-1245.

Lee, Chia-ying, Jie-li Tsai, Erica Chung-i Su, Ovid J.-L. Tzeng & Daisy L. Hung. 2005. Consistency, Regularity, and Frequency Effects in Naming Chinese Characters. *Language and Linguistics* 6.1:75-107.

Leong, Che Kan, Pui-Wan Cheng & Robert Mulcahy. 1987. Automatic processing of morphemic orthography by mature readers. *Language and Speech* 30.2:181-196.

Liberman, Alvin M. 1992. The Relation of Speech to Reading and Writing. *Advances in Psychology*, ed. by Frost Ram & Katz Leonard, 167-178: North-Holland.

Liu, Youyi, Hua Shu & Ping Li. 2007. Word naming and psycholinguistic norms: Chinese. *Behavior research methods* 39.2:192-198.

Lo, Ming, Chih-Wei Hue & Fang-Zhi Tsai. 2007. Chinese Readers' Knowledge of How Chinese Orthography Represents Phonology. *Chinese Journal of Psychology* 49.4:315-334.

Myers, James. 2010. Chinese as a natural experiment. *The Mental Lexicon* 5.3:421-435.

Norris, Dennis, James M. McQueen & Anne Cutler. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23.3:299-325.

Peereman, Ronald & Patrick Bonin. 1998. Is perception a two-way street? The case of feedback consistency in visual word recognition. *Journal of Memory and Language* 39.2:151-174.

Pexman, Penny M., Stephen J. Lupker & Debra Jared. 2001. Homophone effects in lexical decision. *Journal of Experimental Psychology: Learning Memory and Cognition* 27.1:139-156.

Rodd, Jennifer, Gareth Gaskell & William Marslen-Wilson. 2002. Making sense of semantic ambiguity: Semantic competition in lexical access. *Journal of Memory & Language* 46.2:245-266.

Rubenstein, Herbert, Spafford S. Lewis & Mollie A. Rubenstein. 1971. Evidence for phonemic recoding in visual word recognition. *Journal of Verbal Learning and Verbal Behavior* 10.6:645-657.

Stone, Gregory O., Mickie Vanhoy & Guy C. Van Orden. 1997. Perception is a two-street:

feedforward and feedback phonology in visual word recognition. *Journal of Memory & Language* 36.3:337-359.
Tabak, Wieke, Robert Schreuder & R. Harald Baayen. 2005. Lexical statistics and lexical processing: semantic density, information complexity, sex, and irregularity in Dutch. *Linguistic evidence—Empirical, theoretical, and computational perspectives*, ed. by Stephan Kepser & Marga Reis, 529-555. Berlin: Mouton de Gruyter.
Tan, Li-hai, Ho-ling Liu, Charles A. Perfetti, John A. Spinks, Peter T. Fox & Jia-hong Gao. 2001. The neural system underlying Chinese logograph reading. *NeuroImage* 13.5:836-846.
Taraban, Roman & James L. McClelland. 1987. Conspiracy effects in word pronunciation. *Journal of Memory & Language* 26.6:608-631.
Tzeng, Ovid J. L., Zhong Hui Lin, Daisy L. Hung & Wei Ling Lee. 1995. Learning to be a conspirator: A tale of becoming a good Chinese reader. *Speech and reading: A comparative approach.*, ed. by Beatrice De Gelder & Jose Morais, 227-246. Hove, England UK: Erlbaum (Uk) Taylor & Francis, Publishers.
Van Orden, Guy C. 1987. A ROWS is a ROSE: Spelling, sound, and reading. *Memory and Cognition* 15:181-198.
Wu, Yan, Deyuan Mo, Yiu-kei Tsang & Hsuan-chih Chen. 2012. ERPs reveal sub-lexical processing in Chinese character recognition. *Neuroscience letters* 514.2:164-168.
Ziegler, Johannes C., Marie Montant & Arthur M. Jacobs. 1997. The feedback consistency effect in lexical decision and naming. *Journal of Memory and Language* 37:533-554.
Ziegler, Johannes C., Ana Petrova & Ludovic Ferrand. 2008. Feedback consistency effects in visual and auditory word recognition: where do we stand after more than a decade? *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34.3:643-661.
Ziegler, Johannes C., Li-hai Tan, Conrad Perry & Marie Montant. 2000. Phonology Matters: The Phonological Frequency Effect in Written Chinese. *Psycholoical Science* 11.3:234-238.

Legend

Table 1. Mean and S.D. of predictors, and the correlation coefficients between predictors.

Table 2. Row 2 to 5 indicates beta values of predictors estimated in each model. The bottom row shows the χ^2 by comparing performance of Control Model with that of Model 1–3.

Table 3. Summary of Fixed Effects of Model 4 for lexical decision latencies

Figure 1. Illustrations of two types of mapping consistency between orthography and phonology in Chinese**.**

Figure 2. The lines plot the effect of (a) feedback consistency, (b) phonetic combinability, and (c) semantic ambiguity for three levels of log frequency. *P* values indicate the significance of effects of each variable at each frequency level. Dotted lines: high frequency level (3rd quartile of log frequency). Dashed lines: mean frequency level (mean of log frequency). Solid lines: low frequency level (first quartile of log frequency).

音形一致性對中文字辨識之影響

李佳穎[1,2,3]、徐峻賢[1]、張亞寧[1]，陳薇帆[1]、趙培君[2]
[1]中央研究院
[2]國立陽明大學
[3]國立中央大學

過去研究已指出語音表徵對閱讀習得與文字辨識的重要性。比如，形音一致性效果顯示，讀音一致的字母串（比如 －ean 在 lean, dean 與 bean 的讀音皆一樣）引發較快且正確的唸名反應。另外有研究指出，拼字形態較多樣的韻母（比如包含韻母 /ip/ 的詞可拼為 heap 或 deep）所構成的詞在詞彙辨識作業中需要較長的辨識時間，此效果稱為音形一致性。這些現象顯示文字訊息激發相關的語音、字形單位，不一致的對應降低交互激發之穩定性，並影響詞彙辨識的效率。然而後續研究未重複驗證音形一致性效果，其可能原因是音形一致性與其他變項共變，比如頻率、字形鄰項個數等等。詞彙變項之間的共變特性導致因素設計所需要之實驗材料受到限制，即研究者無法僅操弄音形一致性但控制所有的詞彙特性。本研究以大量詞彙庫與行為實驗來檢驗中文詞彙辨識中的音形一致性效應。180 位大學生參與詞彙辨識作業，實驗材料包含 3423 個形聲字。反應時間以線性混合模型進行分析，除了音形一致性，另納入其他詞彙特徵為共變項，包含主/客觀頻率、筆畫數，部件結合度、語意歧義性、形音一致性…等。分析結果呈現典型的詞彙辨識效果，比如，高頻、以及高熟悉度的字引發較快的辨識時間。此外，線性混合模型分析也顯示，估計了各種詞彙變項之影響的同時，音形一致性俱有顯著影響力，越一致的字引發越快的辨識速度。此現象顯示語音表徵會影響字形表徵的聯結方式，而且在文字辨識歷程中，兩種表徵會交互激發。

關鍵字：音形一致性、詞彙判斷作業、交互激發模型、線性混合模型

Table 1. Mean and S.D. of predictors, and the correlation coefficients between predictors.

	Mean (S.D.)	FREG	FB	HD	FF	FAM	PCOM	SCOM	NS	SA	MB
log frequency (FREQ)	2.11 (1)										
feedback consistency (FB)	.41 (.39)	.17***									
homophone density (HD)	1.93 (9.19)	–.03	–.39***								
feedforward consistency (FF)	.54 (.32)	–.01	.17***	.1***							
Familiarity (FAM)	4.18 (1.04)	.82***	.15***	–.08***	–.02						
phonetic combinability (PCOM)	7.24 (3.98)	.05**	.11***	.003	–.38***	.04*					
semantic combinability (SCOM)	85.08 (66.62)	.02	.04*	–.05**	.14***	.03	–.09***				
number of stroke (NS)	13.11 (4.06)	–.13***	–.04*	.04*	.11***	–.17***	–.15***	–.18***			
semantic ambiguity (SA)	–.04 (1.03)	.06***	.02	–.02	–.05**	.07***	.01	.03	–.04*		
morphological boundedness (MB)	.66 (.47)	.5***	.08***	–.09***	–.06***	.55***	.02	.004	–.04*	.09***	
noun-to-verb ratio	.36 (1.05)	.09***	.02	.06***	.04*	.002	.04*	–.08***	.01	–.15***	–.01

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 2. Row 2 to 5 indicates beta values of predictors estimated in each model. The bottom row shows the χ^2 by comparing performance of Control Model with that of Model 1–3.

Predictors	Control Model	Model 1	Model 2	Model 3
log frequency	–.068***	–.068***	–.075***	–.075***
homophone density		.0005***		.0006***
feedback consistency			–.032***	–.028***
consistency by log frequency			.0135***	.014***
χ^2 comparing with Control Model		16.099***	19.306***	37.397***

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 3. Summary of Fixed Effects of Model 4 for lexical decision latencies

Model 4	beta	std. error	t value	pMCMC
(Intercept)	6.711	.018	367.1	< .001
trial numbers	.000	.000	-6.1	< .001
log frequency	–.028	.005	-5.4	< .001
log frequency^2	.002	.001	2.2	.024
feedback consistency	–.012	.006	-2	.034
familiarity	–.114	.008	-13.5	< .001
familiarity^2	.006	.001	5.6	< .001
phonetic combinability	–.002	.001	-4.2	< .001
semantic combinability	.000	.000	-3	.002
semantic combinability^2	.000	.000	2.8	.008
number of strokes	.001	.000	3.8	< .001
residual semantic ambiguity	–.012	.003	-4	< .001
morphological boundedness	–.007	.002	-2.9	.002
noun-to-verb ratios	–.007	.001	-7.4	< .001
homophone density	.000	.000	1.7	.078
feedforward consistency	–.003	.003	–.8	.396
feedback consistency by log frequency	.006	.003	2.4	.02
phonetic combinability by log frequency	.001	.000	2.5	.006
residual semantic ambiguity by log frequency	.002	.001	2.5	.004

Note. MCMC = Monte Carlo Markov chain; ^2: polynomial terms

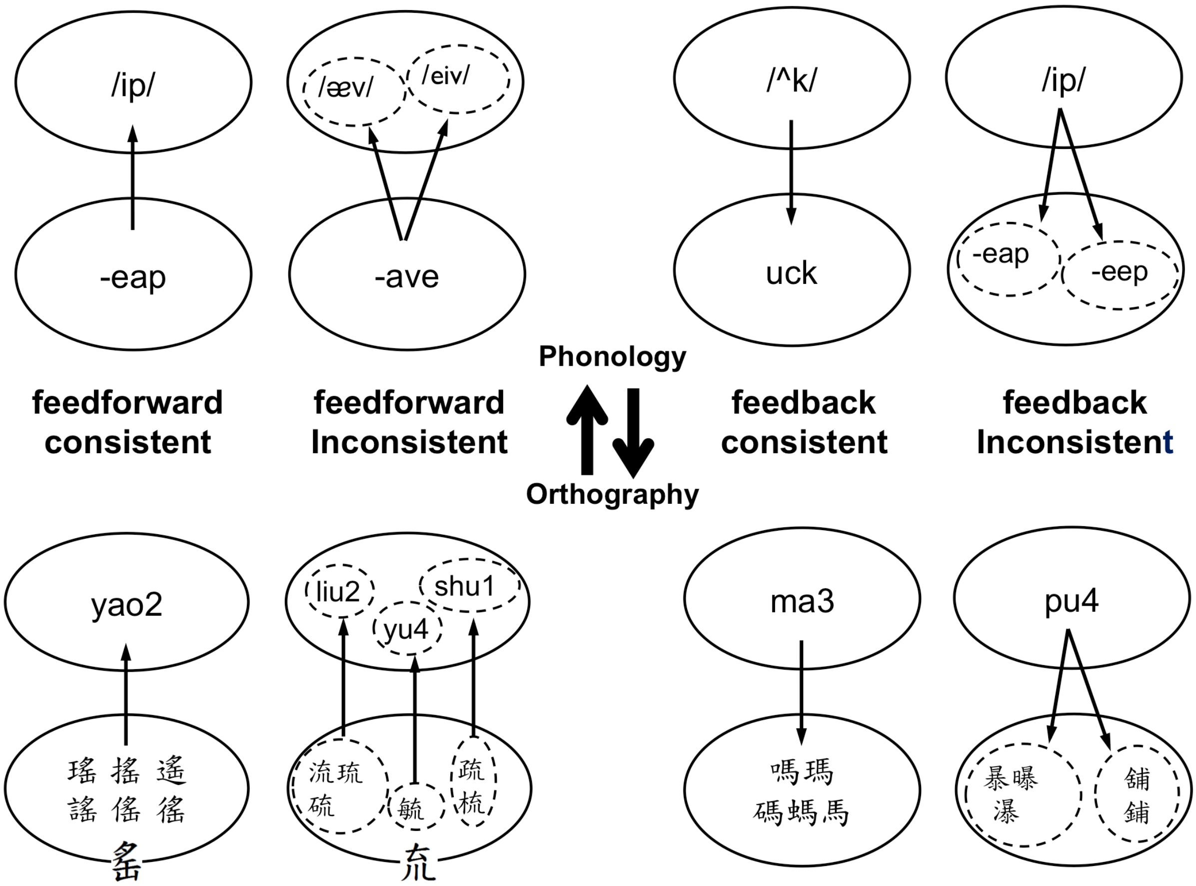

Figure 1. Illustrations of two types of mapping consistency between orthography and phonology in Chinese.

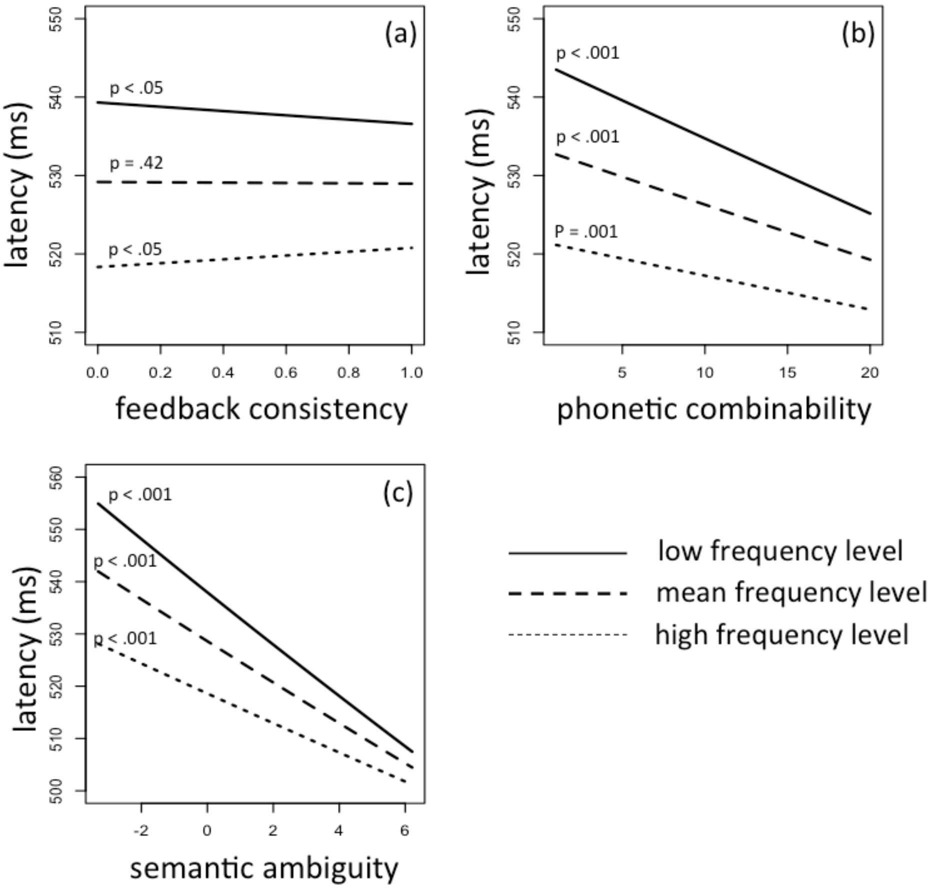

Figure 2. The lines plot the effect of (a) feedback consistency, (b) phonetic combinability, and (c) semantic ambiguity for three levels of log frequency. *P* values indicate the significance of effects of each variable at each frequency level. Dotted lines: high frequency level (3rd quartile of log frequency). Dashed lines: mean frequency level (mean of log frequency). Solid lines: low frequency level (first quartile of log frequency).