THE RECOGNITION OF SPOKEN MONO-MORPHEMIC COMPOUNDS IN CHINESE[*]

Yu-da Lai[a] and James Myers[b]
Providence University[a]
National Chung Cheng University[b]

ABSTRACT

This paper explores the auditory lexical access of mono-morphemic compounds in Chinese as a way of understanding the role of orthography in the recognition of spoken words. In traditional Chinese linguistics, a compound is a word written with two or more characters whether or not they are morphemic. A mono-morphemic compound may either be a binding word, written with characters that only appear in this one word, or a non-binding word, written with characters that are chosen for their pronunciation but that also appear in other words. Our goal was to determine if this purely orthographic difference affects auditory lexical access by conducting a series of four experiments with materials matched by whole-word frequency, syllable frequency, cross-syllable predictability, cohort size, and acoustic duration, but differing in binding. An auditory lexical decision task (LDT) found an orthographic effect: binding words were recognized more quickly than non-binding words. However, this effect disappeared in an auditory repetition and in a visual LDT with the same materials, implying that the orthographic effect during auditory lexical access was localized to the decision component and involved the influence of cross-character predictability without the activation of orthographic representations. This claim was further confirmed

by overall faster recognition of spoken binding words in a cross-modal LDT with different types of visual interference. The theoretical and practical consequences of these findings are discussed.

Key words: spoken word recognition, mono‐morphemic compound, orthography

1. INTRODUCTION

The mental lexicon has long been of central interest in psycholinguistic research, but, as pointed out by Henderson (1985), most of the focus has been on visual word recognition. This is particularly true for Chinese, where aside from a few studies on spoken word recognition, lexical research has consisted of word reading tasks, or even the reading of isolated characters (Myers 2006). The reasons for this are primarily practical, of course, but in the case of Chinese the preference given to reading in psycholinguistic studies also mirrors the primacy given to orthography in the "folk linguistics" of naïve native speakers. The purpose of this paper is to examine the possibility that the native speakers are right in some sense, and that Chinese orthography does indeed exert an influence on the processing of spoken words. This possibility is explored by taking advantage of the existence of mono-morphemic compounds, oxymoronic entities arising from the peculiarities of Chinese orthography.

1.1 Orthography and Spoken Word Recognition

There is evidence from a variety of languages and experimental paradigms for the activation of orthographic representations during the auditory word recognition of alphabetic languages: spelling and orthographic consistency influence auditory rhyme judgments (Seidenberg and Tanenhaus 1979; Zecker, Tanenhaus, Alderman, and Siqueland 1986; McPherson, Ackerman, and Dykman 1997), auditory lexical decisions (Ziegler and Ferrand 1998; Miller and Swick 2003; Ventura, Morais, Pattamadilok, and Kolinsky 2004; Pattamadilok, Morais, Ventura and Kolinsky 2007), semantic categorization (Pattamadilok, Perre, Dufau, and Ziegler 2009), the effects of Stroop

interference in auditory-to-visual priming (Tanenhaus, Flanigan, and Seidenberg 1980), syllable monitoring (Taft and Hambly 1985), auditory-to-auditory priming (Jakimik, Cole, and Rudnicky 1985; Hallé, Chéreau, and Segui 2000), and phoneme detection (Frauenfelder, Segui, and Dijkstra 1990; Dijkstra, Roelofs, and Fieuws 1995). For example, in a task that required participants to decide whether two words rhymed or not, Seidenberg and Tanenhaus (1979) found that participants made faster responses to words that had the same spelling for the rhyme, e.g., *toast* vs. *roast*, than to those with a different spelling, e.g., *toast* vs. *ghost*. In another study using a lexical decision task (LDT), Ziegler and Ferrand (1998) found that the consistency in orthography, i.e., for words ending with a consistent rhyme, rather than those ending with an inconsistent rhyme, resulted in faster decision times. More recently, accumulating data from brain imaging studies (e.g., Booth, Burman, Meyer, Gitelman, Parrish, and Mesulam 2004; Orfanidou, Marslen-Wilson, and Davis 2006) has shown evidence consistent with the above observations of orthographic activation during auditory word recognition. For example, Orfanidou et al. (2006) found that when participants were involved in an auditory lexical decision process, the activated area of the left posterior fusiform gyrus was close to the so-called visual word form area (McCandliss, Cohen, and Dehaene 2003).

The currently existing items of evidence collectively support the idea of an orthographic influence on spoken word recognition, but researchers have not reached a consensus on the timing of its emergence, i.e., before or after lexical access. For example, in a phoneme monitoring experiment in Dutch, Dijkstra et al. (1995) found that detection times for the phoneme /k/ were slower for spoken words where this sound is spelled with the less commonly used grapheme "c" than for words with the more commonly used "k." They concluded that this effect was mediated by lexical access, since the size of the effect depended on the location of the phoneme in the word: a greater effect was found when the target phonemes were located after the uniqueness point[1] (henceforth UP). A similar interference effect from orthography was also reported in

---

[1] Radeau, Mousty, and Bertelson (1989) defined the UP as the point where listeners had enough information to downsize the cohort candidates to one, i.e., a point of possible full identification.

another study on French words (Hallé et al. 2000) but interpreted using a pre-lexical account. Other researchers localize the source of the orthographic effect in the phonological lexicon itself. For example, Taft and Hambly (1985) interpret their finding that the Basic Orthographic Syllabic Structure (BOSS) influenced auditory syllable monitoring as showing that phonological forms in the lexicon have themselves been modified through the learning of English orthographic conventions. The debate among the previous behavioral studies remains unresolved because it was not possible to use decision times and accuracy measures, which reflect the ultimate decisions of word recognition, to tease apart the lexical and post-lexical interpretations (Pattamadilok et al. 2009: 22). It is only recently that the locus of the orthographic consistency effect been localized during the time-course of spoken word recognition in studies with ERPs (Perre and Ziegler 2007; Pattamadilok et al. 2009). For example, in a semantic task which required participants to decide whether a given spoken French word belonged to a certain semantic category, Pattamadilok et al. (2009) manipulated the orthographical consistency of either the first or the second syllable as well as the word frequency, e.g., high frequency word *bouteille* vs. low frequency word *astuce* where the second but not the first syllable has more than one possible spelling. Their ERPs data showed that the orthographic effect was time-locked to the point where the inconsistency occurred, and that the onset of the effect took place even before the onset of the word frequency effect. Pattamadilok et al. thus concluded that orthographic information was activated fairly early.

The question of orthographic influences on spoken word processing does not seem to have been studied in a non-alphabetic language, but the nature of Chinese orthography is such that one might expect such influence to be particularly strong. The basic unit is the character, and it virtually always corresponds to one syllable. Syntactically free words, however, generally have more than one character; Zhou and Marslen-Wilson (1995: 547) estimate that the proportion of disyllabic compounds in the Academia Sinica Balanced Corpus of Modern Chinese (Chen, Huang, Chang, and Hsu 1996) is about 74% by type and 34% by token, and that the proportion of bound characters (i.e., characters that never appear in isolation) is around 36%. Nevertheless, a character is usually

44

regarded by native speakers as a word-level linguistic unit (Chao 1968) and virtually all dictionaries are character-based. This notion of character as word arises partly from the fact that spaces are not used to separate words in written text, but another important factor is that characters are generally morphemic, and do seem to play an active role in word reading (e.g., Hoosain 1992; Zhang and Peng 1992; Liu and Peng 1997; Zhou, Marslen-Wilson, Taft, and Shu 1999; see reviews in Taft, Liu, and Zhu 1999, and Myers 2006). Some have argued that even spoken word recognition is essentially morpheme-based (e.g., Zhou and Marslen-Wilson 1994, 1995), but this seems unlikely: the high proportion of homophonous morphemes (there are only about 1,200 distinct syllables, including tones, distributed over more than 6,000 characters) makes a morpheme-based algorithm for spoken word access eminently impractical (Packard 1999; Myers and Gong 2002).

Thus, finding that characters act as units in spoken word recognition, even if they are not morphemic, would represent strong evidence of orthographic influence in auditory lexical access. Although it is relatively uncommon, characters are not always morphemic in Chinese, but instead may represent just part of a multi-character morpheme. The clearest type of multi-character morpheme is the *binding word*, where the component characters only (or virtually only) appear in this one word; the term binding (adopted from Taft and Zhu 1995) is thus more restrictive than bound, which includes characters that may appear in many words. Binding words are morphemes that have existed in Chinese for such a long time that their unitary status has been codified by the creation of new characters that share the same semantic radical. A typical example of a binding word is *zhang1lang2* 蟑螂 (cockroach); both characters have the "crawling creature" radical at the left (the digits used in the Pinyin transcriptions represent the four Mandarin tones). Another type of apparently mono-morphemic words consists of what we call *opaque non-binding words*. These include more recent borrowings, written with characters that also appear as morphemes in other words or in isolation, but which are chosen for their phonetic similarity to the appropriate portions of the borrowed word. A typical example is *han4bao3* 漢堡 (hamburger), borrowed from English; by themselves, the character *han4* means "Chinese" and the character *bao3* means "fort."

Like binding words, opaque non-binding words are also assumed to be mono-morphemic (DeFrancis 1984; Starosta, Kuiper, Ng, and Wu 1998; Zhou et al. 1999).

However, because of the dominant notion among native speakers that characters are inherently meaningful, both types of mono-morphemic words are still called "compounds" (*fu4he2ci2* 複合詞) and even the meaningless components of binding words are given separate entries in dictionaries (defined as if they meant the same as the whole word). The basic difference between the two types of words is orthographic: due to the fact that binding words are composed of unique characters sharing a radical, they give the appearance of an orthographic whole, whereas opaque non-binding words do not. If one knows the written forms, a non-binding word may seem less semantically opaque than a binding word, but this impression is driven solely by the orthography. By contrast, there seems to be nothing in the phonology that may distinguish binding and opaque non-binding words, not even distributional patterns and sound-meaning correspondences of the sort that allow pre-literate children to identify the components of transparent compounds. The next section discusses how the orthographic wholeness through bindingness is quantified and its role in Chinese word recognition.

1.2 Cross-character Predictability and Bindingness in Chinese Compounds

The degree to which a compound is an orthographic whole can be expressed through the notion of *cross-character predictability*: binding words have a high cross-character predictability (in fact the maximum possible in the Chinese lexicon), while mono-morphemic non-binding words have a much lower cross-character predictability. In this study we quantify predictability using the measure of *mutual information* (MI) as used in corpus linguistics (Church and Hanks 1990). MI compares the probability of observing components x and y together (the joint probability) with the probability of observing x and y independently (chance), thus describing how common a collocation is between components when the lexical frequency of each constituent has been factored out. If there is a genuine collocational relation between

components x and y, the joint probability $P(x, y)$ will be larger than chance $P(x)P(y)$, thus causing larger MI values (MI $(x, y) > 0$). More precisely, MI is formulated as MI $(x, y) = \log_2 [P(x, y)/(P(x)P(y))]$. Note that MI (predictability) is positively correlated with word frequency but negatively correlated with character frequency. Thus if the word frequencies of binding and non-binding words are matched, the character frequencies must necessarily be different (lower in binding words, higher in non-binding words). The relevance of this will be seen below.

Given the orthographically holistic nature of binding words in comparison to non-binding words, it should not be surprising that readers treat characters in binding words differently from those in non-binding words. In one experiment, Taft and Zhu (1995) asked participants to pronounce binding characters that occur either only in the first position or only in the second position of binding words. The results showed that second-position characters took longer to name than first-position characters, as if the phonological form for the entire word had to be retrieved before speech could begin. However, in another experiment they performed with non-binding characters that always occur in a fixed position (i.e., always in first or always in second place), no such positional effect was found, as if the pronunciation for each character could be accessed independently. In a related finding, Taft (2003) showed readers isolated characters (*zi4*) and asked them to judge if the status of the given characters were words (*ci2*); participants in this task found it harder to reject bound but non-binding characters than binding characters, even though neither are truly words. Thus the mere fact that characters can appear in more than one context gives them a status closer to free words.

Opaque non-binding words also show some processing similarities with transparent compounds. Zhou et al. (1999) found that lexical decisions for compound targets sharing the first character with a masked compound prime were faster regardless of whether the target was a genuine morphemic compound (Experiment 1) or an opaque non-binding word like *sha1fa1* (Experiment 4). Moreover, while semantically transparent compounds show consistent character frequency effects in LDTs (e.g., Liang 1992; Taft, Huang, and Zhu 1994; Lee 1995; Peng, Liu, and Wang 1999), semantically opaque compounds may show no

effect of character frequency (Liang 1992; Taft et al. 1994; Lee 1995) or even a reversed frequency effect, with higher-frequency characters leading to slower response times for words matched in frequency (Peng et al. 1999). As proposed by Peng et al. (1999), this pattern may arise because the characters are automatically activated during lexical decisions. If the characters happen to be related in meaning to the whole word, they will facilitate word judgments, but, if they are unrelated (i.e., compete with word meaning), they may make word judgments more difficult. This view would seem to predict that compounds matched for both character and word frequency would be easier to access if semantically transparent, but this is not consistently found: in some studies, semantically transparent compounds are responded to more slowly (and less accurately) than semantically idiosyncratic or opaque compounds (Liang 1992; Lee 1995; Su 1998), while others find the reverse pattern (Tsai 1994; Lü 1996) or no difference (Chen 1993). A possible explanation for the inconsistency is confusion on the part of the participants over what level is being judged, characters or compounds; the activated component characters of the transparent compounds may thus become competitors of a sort.

While not denying the role of semantics, cross-character predictability may be an additional factor leading to variable negative frequency effects for opaque compounds and slower response times for transparent compounds, assuming that predictability and opacity are to some extent confounded. Cross-character predictability is expected to aid word access: recognition of one character will readily facilitate recognition of the other in high-predictability compounds. Since character frequency is inversely related to cross-character predictability, opaque compounds with higher-frequency characters will have lower predictability, and without the additional assistance from knowledge of the semantic relationships between the characters and the whole word, this may lead to negative character frequency effects. Meanwhile, in the case of word frequency matched across opaque and transparent compounds, the presumed higher predictability of the characters in opaque compounds may make them easier to access overall.

Can cross-character predictability affect spoken word processing as well? Myers and Gong (2002) claimed that it can, but their experiments

examining character frequency effects on auditory lexical access involved transparent compounds instead of mono‑morphemic compounds, so the effects were not necessarily orthographic. In this paper we thus describe experiments comparing the auditory access of binding words and opaque non-binding words. If orthography does have an effect, we would expect that binding words, being orthographic wholes, will be easier to recognize than non-binding words. Tsai (1998), in an unpublished master's thesis, claims to have found just such a result in an auditory LDT: there was a faster response to binding words, matched in whole-word frequency, than to opaque non‑binding words. However, this result, while suggestive, cannot be taken as conclusive, since key phonological variables (including syllable frequency and syllable MI) were not taken into consideration. Moreover, no attempt was made to determine the source of the apparent orthographic effect. Did the effect reflect modification of phonological representations in the lexicon due to experience with Chinese orthography, or was it a post-lexical effect, due to processes taking place after word access had been completed? Or even alternatively, could it be the case that the so-called orthographic effect during auditory lexical access was due to the activation of the semantics of the compounds? That is, the semantics of the characters in non‑binding words compete with the semantics of the whole word while those in binding words do not. The three experiments described below use materials strictly controlled for phonological properties, and vary the nature of the task to determine the source of orthographic influence.

2. EXPERIMENT 1: AUDITORY LEXICAL DECISION

In order to ensure that any differences in response times in our auditory LDT were due to orthography alone, we controlled all variables, other than the mono‑morphemic compound type, that might be expected to influence response times. In addition to word frequency, these variables were syllable frequency (found by Zhou and Marslen-Wilson 1994, to influence lexical decision times to spoken transparent compounds), first-syllable cohort size, cross-syllable predictability, and

phonetic duration. In addition, to reduce any possible influence of variation in the UP (which tends to be already quite late in Chinese disyllabic words anyway due to the small number of syllable types), the non‑word foils were designed to be identical to real words except for the tone on the second syllable, thus forcing participants to listen until the end of the word before making a decision.

2.1 Method

2.1.1 Participants

Twenty students, ten males and ten females, at National Chung Cheng University in southern Taiwan, were paid for their participation in this experiment.

2.1.2 Design and materials

The task was an auditory LDT. Twenty-one opaque non‑binding words were chosen from the *Guoyu ribao wailaiyu cidian* published by the Mandarin Daily News Association (1981) and twenty-one binding words were selected both from Fu's (1986) *Lianmian zidian* and from Tsai (1998). Both sets were matched by whole-word frequency, syllable frequency in both first and second position, cross-syllable predictability, first-syllable cohort size, and acoustic duration, but differed by cross-character predictability (higher in binding words) and character frequency (higher in non-binding words).

Word frequency was controlled (binding: token frequency $Mdn = 2858.77$, log frequency $M = 11.92$ ($SD = 1.53$); non-binding: token frequency $Mdn = 5431.66$, log frequency $M = 12.56$ ($SD = 1.27$); $t(40) = 1.48$, ns), frequency counts of which came from the *Zhongwen shumianyu pinlü cidian* (Chinese Knowledge Information Processing Group 1994). Syllable frequency was calculated by counting all of the occurrences of a given syllable in the *Mandarin Chinese Character Frequency List Based on National Phonetic Alphabets* (Chinese Knowledge Information Processing Group 1995), frequency counts of which came from a large corpus of written Mandarin (the Academia

Sinica Balanced Corpus). Since semi-homophones, i.e., segmentally identical homophones that have different tones, are found to prime Chinese lexical processing especially via auditory presentation (Zhou 2000) and for monosyllabic words (Yip 2001), the syllable frequency for the present research was calculated by adding up the occurrences of homophones of the given syllable regardless of its tone types. There were no significant difference in the frequencies of the first syllables across the two conditions (binding: Mdn = 8981, log frequency M = 12.91 (SD = 1.68); non-binding: Mdn = 8596, log frequency M = 12.72 (SD = 1.84); $t(40)$ = 0.36, ns). To control the cohort sizes of the two conditions, the numbers of words sharing the same first syllable with our experimental items in the *Zui xin Lin Yutang han-ying cidian* (Li and Lin 1987) were calculated. The cohort sizes for the two conditions were not significantly different from each other in an unpaired t-test for equal variances (binding: M = 48, SD = 27.85; non-binding: M = 34, SD = 23.39; $t(40)$ = 1.74, ns). Cross-syllable predictability was quantified using syllable MI, computed using whole-word frequencies (there are virtually no disyllabic homophones in Chinese) and syllable frequencies; the syllable MI was matched across the two conditions (binding: M = 10.48, SD = 3.64; non-binding: M = 9.91, SD = 2.82; $t(40)$ = 0.56, ns). The key independent variable that remained was cross-character predictability; the mean character MI for the binding words was significantly higher than that for the opaque non-binding words (binding: M = 23.87, SD = 2.13; non-binding: M = 14.03, SD = 4.22; $t(30)$ = 9.52, p < .01). However, with these strict controls in place and the limited choice of the stimulus items available, even though a great majority of the materials belonged to the "concrete" category, the numbers of the sub-semantic categories in the two sets were not perfectly counter-balanced. That is, more binding words were of the "animal category". To avoid any potential confound, the target items were interweaved among a total of 42 non-word foils and presented in a way that no more than four consecutively presented targets were pairs of real words and their corresponding non-word foils.

For each of the 42 real-word targets, a non-word foil was created by keeping the first syllable and changing the tone of the second syllable (following Tsai 1998; Myers and Gong 2002). For example, for the non-

binding word *han4bao3* 漢堡 (hamburger), which has tone 4 (a falling tone) on the first syllable and tone 3 (a contour tone) on the second one, was paired with the non-word *han4bao4* with tone 4 (a falling tone) on the second syllable. The tone of the second syllable was not confined to any particular type because, on the one hand, these foils could be written out with commonly-seen characters (e.g., 漢抱, a meaningless non-lexical combination of "Chinese" and "embrace") for the purpose of Experiment 3 (visual LDT). On the other, given the design with various tone types in the second syllable, participants had to pay further attention until the end of all of the auditory stimuli before they could make a decision. Due to rampant homophony in Chinese and the negative syllable frequency effect in the first position of compounds (Zhou and Marslen-Wilson 1994), the benefit of this particular design might be that participants could quickly learn that even the non‑word foils were very close to real words and could thus develop a strategy to reduce the interference from real-word neighbours or competition between similar-sounding words during auditory lexical access of real words (Myers and Gong 2002). In addition, one could reasonably expect the potential difference in accuracy of (non)word recognition in tasks of different modalities, as will be seen later.

All the experimental items, i.e., the whole words and non-word foils, were produced and recorded by a female native speaker of Mandarin and were digitized at a sampling rate of 22.05 kHz for auditory presentation. The acoustic durations of the auditory stimuli were measured and edited using the Kay Elemetrics Computerized Speech Lab; the lengths of auditory items across the two conditions were not significantly different (binding: $M = 873$ ms, $SD = 70.26$; non-binding: $M = 841$ ms, $SD = 71.28$; $t(40) = 1.47$, *ns*). The list of all our experimental items with their statistical information is given in Appendix I.

2.1.3 Procedure

Participants were tested individually before PC compatible computers in a sound-attenuated room. To reduce possible strategic semantic priming to minimum, the stimuli were presented over headphones in pseudo-random order (different for each participant),

52

where no more than four consecutively presented targets were pairs of real words and their corresponding non-word foils. In each trial, participants were first shown the English word "Attention" in the center of the screen, followed 1000 ms later by the auditory stimulus. Participants had to give as rapid and accurate a response as possible as to whether the stimulus was a real word, by pressing either the *zhen1ci2* 真詞 (real word) button on the right side or *fei1ci2* 非詞 (non-word) button on the left side of the response box. The response time limit was set to 3000 ms and the response time (RT) was measured from the onset of the stimuli to the onset of the participants' responses. The whole of the experimental procedure was co-ordinated using the E-prime software (Schneider, Eschman, and Zucolotto 2002). There were 14 practice trials before the main experiment began. The whole experiment took approximately 10 minutes per participant.

2.2 Results and Discussion

The data from one of the participants were discarded because of an error rate of 16% for real-word items. Errors included both incorrect categorization of words and non-words and RTs above the time limit of 3000 ms. About 6.5% of all responses were excluded using this criterion. The mean by-participant RTs and error rates, computed for the data of the remaining participants, are reported in Table 1 (also see Appendix II for the mean RTs and error rates for each item).

Table 1. Experiment 1: Mean lexical decision times (ms) and error rates (%)

	Mean RT (*SD*)	Percentage error	Example
Binding	1473 (207.28)	3.5%	蟑螂 "cockroach"
Opaque non-binding	1540 (216.98)	3%	漢堡 "hamburger"

As Table 1 shows, RTs for binding words were about 67 ms faster than RTs for opaque non-binding words. This difference was significant

both by participant ($t_1(18) = 8.56$, $p < .05$) and by item ($t_2(40) = 2.82$, $p < .05$). The error rates showed no significant difference by participant or by item ($t_1(18) = 0.46$, *ns*; $t_2(40) = 0.34$, *ns*).

In an attempt to factor out potential nuisance variables like the difference in the overall response speeds, we divided the participants by their mean RTs to all the items, including foils, categorizing them as slow or fast responders relative to the median of the mean RTs. We found that the type of mono-morphemic compound still had significant effects, both by participant and by item, for both the slower group (binding: $M = 1605$ ms, $SD = 205.95$; non-binding: $M = 1679$ ms, $SD = 248.11$; $t_1(8) = 4.71$, $p < .05$, $t_2(40) = 2.33$, $p < .05$) and the faster group (binding: $M = 1345$ ms, $SD = 30.8$; non-binding: $M = 1415$ ms, $SD = 50.19$; $t_1(9) = 6.41$, $p < .05$, $t_2(40) = 3.00$, $p < .05$).

As robust as the pattern is, the effect reported above, one may argue, was not necessarily due to information derived from orthography, since it could instead be due to variation in the UP of the auditory stimuli or to the semantic conflict. For the first possibility, since the non-word foils were designed to differ from the target words only in the tone of the second syllable, the measurement of the RT from the onset of the whole word may not be able to reflect the real process of spoken word recognition but just the differences in the locations of the UP of the binding and non-binding words. Regarding the latter possibility, listeners may imagine that the components of opaque non-binding words are actually related to other words, e.g., *han4bao3* 漢堡 (hamburger) as somehow being related to *han4* 漢 (Chinese) and *bao3* 堡 (castle), so that the "hypothesized" semantic relatedness derived from the individual characters may conflict the whole word meaning, hence resulting in slower RTs relative to binding words.

To test whether the observed pattern was due to the above possibilities, we conducted repeated-measures regression analyses on the results of Experiment 1 using the simplest form of the procedure recommended in Lorch and Myers (1990). Since we assume that the UP, acoustic durations, and/or the component semantics may help predict RTs, a measurement of the length of the UP in each auditory stimulus and a semantic transparency test on the same materials, mixed with fillers, were conducted. For the measurement of UP length, we assumed

that the onset of the pitch of the second syllable was the UP since this is the point at which the tone information of the second syllable becomes available; in addition, the non‑word foils were designed to differ from the targets only in the tone of the second syllable. Therefore, we measured the UP length (in milliseconds) between the onset of the auditory stimulus and the onset of the pitch tracking the second syllable. Regarding the semantic transparency test, we generated two questionnaires, one for each of the component characters of the target words such that the participants would not have to rate the semantic relatedness of the character in both positions to the meaning of the whole word in the same questionnaire. For each questionnaire, there were a total of 120 words (42 target items were mixed with 78 filler items that were already known to have either high or low scores of semantic relatedness). All of the items were presented in randomized orders, with half of them being rated for the semantic relatedness to their first character and the other half to the second component character. Two groups of students at National Chung Cheng University, 20 in each, were thus recruited respectively for each questionnaire to give their ratings by answering questions like "*On a scale from 1 to 6, how semantically related is the given component character to the target word? (1 = least related, 6 = most related)*." The statistical information of the UPs and the mean semantic transparency scores for the experimental items is given in Appendix III.

Separate regressions were then run for each participant, with acoustic durations, UPs, mean semantic relatedness scores, and other factors considered in materials preparation discussed above as predictors (e.g., compound type, word frequency, syllable frequency, character frequency, syllable MI, and character MI), and RTs as the dependent variable. As Table 2 summarizes, the pattern of the processing of the differences between binding words and opaque non-binding words could still be observed ($B = -39.59$, $SE = 12.52$, $p < 0.05$) and was not due to the variation in the UP of the auditory stimuli or semantic conflict since there in fact was no significant effect of the durations of the stimuli, their UPs, or mean semantic transparency scores (duration: $B = -0.112$, $SE = 0.092$, *ns*; UP: $B = -0.061$, $SE = 0.085$, *ns*; average of semantic transparency scores: $B = 10.436$, $SE = 6.153$, *ns*).

Table 2.   Experiment 1: Summary of regression analysis for variables predicting reaction times (using UP and Dur as independent variables)

Variable	*B*	SE	df	t
Intercept	2615.955	239.4352	18	10.92552 *
CompType	-39.5906	12.52347	18	-3.16131 *
WordFreq	8180.669	4826.364	18	1.694996
AveSem	10.4364	6.153293	18	1.696068
Dur	-0.11201	0.091797	18	-1.22024
UP	-0.06086	0.084561	18	-0.71968
Syl1LogFreq	-107.134	19.16454	18	-5.59024 *
Syl2 LogFreq	-72.893	21.66884	18	-3.36396 *
SylMI	-30.146	4.802931	18	-6.27657 *
Char1LogFreq	-0.00109	0.000477	18	-2.27548 *
Char2LogFreq	-0.00037	0.00059	18	-0.63174
CharMI	3.014343	2.254266	18	1.337173

Note. CompType=compound type, WordFreq=word frequency, AveSem=average of semantic transparency scores for each syllable, Dur=duration of acoustic stimuli, UP=uniqueness point, SylLogFreq=syllable log frequency, SylMI=syllable mutual information, CharLogFreq=character log frequency, CharMI= character mutual information, *B* = mean by-participant raw regression coefficients, *SE* = standard errors conducted across by-participant regression coefficients.
* *p* < .05

This analysis suggests that even when acoustic durations, UPs, and the effects of the nuisance factors of semantics were factored out, the bindingness effect still existed for listeners. Since listeners cannot begin to make a lexical decision before reaching the UP, we re-analysed the results measuring RTs from the UP, using the same predictors. The second regression analysis assumes that the participants could not make any decision until they heard the onset of the pitch of the second syllable. As Table 3 summarizes, the difference between binding and non-binding words was still found (B = 2-30.54, SE = 11.785, $p < 0.05$), which further confirmed that the processing difference between binding words and opaque non-binding words was not due to the variation in UP or to the semantic factor resulting from the conflict between the character semantics and the whole word meanings (average of semantic transparency scores: B = -2.052, SE = 5.387, *ns*).

Table 3.   Experiment 1: Summary of regression analysis for variables predicting reaction time (using RT measured from UP as the dependent variable)

Variable	*B*	SE	df	t
Intercept	2468.502	246.0055	18	10.03433 *
CompType	-30.5443	11.78547	18	-2.59169 *
WordFreq	46198.05	5094.17	18	9.068808 *
AveSem	-2.05241	5.386766	18	-0.38101
Syl1LogFreq	-142.614	21.51366	18	-6.62901 *
Syl2 LogFreq	-161.952	27.12877	18	-5.96974 *
SylMI	-47.0156	5.952171	18	-7.89889 *
Char1LogFreq	-0.00039	0.000794	18	-0.48621
Char2LogFreq	-0.00229	0.000678	18	-3.37974 *
CharMI	-3.65125	2.096684	18	-1.74144

Note. CompType=compound type, WordFreq=word frequency, AveSem=average of semantic transparency scores for each syllable, SylLogFreq=syllable log frequency, SylMI=syllable mutual information, CharLogFreq=character log frequency, CharMI= character mutual information, *B* = mean by-participant raw regression coefficients, *SE* = standard errors conducted across by-participant regression coefficients.
* $p < .05$

Interestingly, as one may have noted, the word frequency effect or the character MI effect was not consistently observed. While no word frequency effect was found in the earlier model (Table 2), it re-emerged in the latter model (Table 3). We speculate that it could be that the measurement of the RTs from the start of the word in the first model

resulted in too much RT noise over the duration of the whole word to detect a whole-word effect like word frequency, whereas when measured from the UP, the shorter duration from there to the response resulted in less variance (noise), allowing the word frequency effect to stand out. Therefore, the latter model (Table 3) that took RTs measured from the UP as the dependent variable seems to more approximate the processing of spoken words than the prior model (Table 2). The lack of a character MI effect, though, is a bit more mysterious. But if we look at the regression model in Table 3, character MI is approaching significance ($p = 0.09$). Thus, the insignificance of character MI might simply come from a lack of enough observations in the regression analyses. And since experiments that find strong word frequency effects in the processing of Chinese compounds usually do not also include character, syllable frequency, and/or MI as continuous variables, it could be that these variables may have effects that overlap with, or conflict with each other, even though both regressions models still unanimously found a robust bindingness effect.

The results thus replicated those of Tsai (1998): spoken binding words were recognized faster than spoken opaque non-binding words, even after controlling for a variety of other possible confounds, both lexical (word frequency) and phonological (syllable frequency). The key factor in faster decisions for spoken binding words thus seemed to be the difference in the orthography of the words, i.e., the degree of being as an orthographically whole unit that is expressed by cross-character predictability. The higher the degree is, the faster responses it may lead to. Moreover, the orthographic effect emerged in both fast and slow responders.

Two questions arise about these results, however. The first concerns the effectiveness of the phonological controls. These were based on the largest corpora (i.e., the Academia Sinica Balanced Corpus) at the time the current research was conducted; however, the corpora are derived from written Chinese, not spoken Chinese, and the frequency counts for a given syllable that included all its homophones of different tones might be inaccurate. It is thus conceivable that binding words and opaque non-binding words differ systematically in syllable frequency, cross-syllable predictability, cohort size, or some other factor that we have yet to have

considered, in such a way that binding words are easier to access, for purely phonological reasons. Although the syllable frequencies of the target items have been recalculated by summing the occurrences of the homophones with only the same tone in the _Mandarin Conversational Corpus Wordlist_[2] (Tseng 2004) and showed no significant differences across the two conditions (binding: log frequency $M = 1.89$ ($SD = .78$); non-binding: log frequency $M = 1.85$ ($SD = .94$); $t(40) = -0.151$, ns), the possibility is also indicated by the robust facilitative effects of syllable frequency and syllable MI, in both the first regression analysis in Table 2 (syllable 1 frequency: $B = -107.13$, $SE = 19.165$, $p < 0.05$; syllable 2 frequency: $B = -72.893$, $SE = 21.669$, $p < 0.05$; syllable MI: $B = -30.146$, $SE = 4.803$, $p < 0.05$) and in the second regression analysis in Table 3 (syllable 1 frequency: $B = -142.614$, $SE = 21.514$, $p < 0.05$; syllable 2 frequency: $B = -161.952$, $SE = 21.129$, $p < 0.05$; syllable MI: $B = -47.016$, $SE = 5.952$, $p < 0.05$).

Even if the orthographic effect is real, a second question arises: What causes it? Does it follow from the way phonology is represented in the lexicon, or from some post‑lexical process? On the face of it, it does not seem plausible that our particular orthographic effect reflects lexical phonological representations, due to the large number of homophones in Chinese. After hearing the first syllable, listeners would have no way of knowing that *han4* is the beginning of the non-binding word *han4bao3* 漢堡 (hamburger) while *zhang1* is the beginning of the binding word *zhang1lang2* 蟑螂 (cockroach). If they immediately activate all of the characters representing the first syllable, and/or the cohort of all words beginning with this syllable, we would expect that cross-character frequency would have an inhibitory effect, rather than the facilitatory effect that we found. This is because the lower character frequency of binding words relative to that of non-binding words means that binding words will have more competitors and/or competitors with a higher

---

[2] The _Mandarin Conversational Corpus Wordlist_ (Tseng 2004) is grounded on the three corpora, the Mandarin Topi-oriented Conversation Corpus, the Mandarin Conversational Dialogue Corpus and the Mandarin Map Task Corpus, completed under the Mandarin Spoken Corpora Project conducted by Dr. Shu-Chuan Tseng at the Academia Sinica in Taiwan. These corpora consist of a total of 60 conversations (42 hours) and 16,746 lexical items and 405,435 tokens.

syllable frequency than non-binding words. In order to be more confident that a pre-lexical account cannot work for our results, however, we need to collect data of a different sort.

We thus decided to address both of these questions by performing a second experiment, using the same stimuli as in Experiment 1, but now used in an auditory repetition task, which involves phonological processing, but no post-lexical decision component. If there are any hidden phonological differences between conditions, or if the orthographic effect is pre-lexical, a difference in RTs between conditions should still be found.

3. EXPERIMENT 2: SPEEDED REPETITION

3.1 Method

3.1.1 Participants

Twenty students, ten males and ten females, at National Chung Cheng University in southern Taiwan, were paid for their participation in this experiment. None had participated in Experiment 1.

3.1.2 Designs and materials

An auditory speeded repetition task was used. The real-word auditory stimuli from Experiment 1 were utilized.

3.1.3 Procedure

Participants were tested individually before PC compatible computers in a sound-attenuated room. The stimuli were presented over headphones in random order (different for each participant). In each trial, participants were first shown the English word "Attention" in the center of the screen, followed 1000 ms later by the auditory stimulus. Upon hearing the stimulus, participants had to repeat the stimulus back into a microphone as quickly as possible. The response time limit was set to

3000 ms and the repetition latencies (RTs) were measured from the onset of the stimuli to the onset of participants' responses. The experimental procedure was co-ordinated using the E-prime software (Schneider, Eschman, and Zucolotto 2002). Before the main experiment, each participant was given ten practice trials. The whole experiment took approximately ten minutes per participant.

3.2 Results and Discussion

Errors included the incorrect repetition of the given stimulus and RTs above the time limit of 3000 ms. About 3.8% of all responses were excluded using these criterion. The mean by-participant RTs and error rates, computed for the remaining data, are reported in Table 4. From this table, we can see that the mean RTs for both sets are quite close. The mean RT for binding words is now 20 ms slower than that for non-binding words, although this difference was non-significant both by participant and by item ($t_1(19) = 1.68$, *ns*; $t_2(40) = 0.90$, *ns*).

Table 4. Experiment 2: Repetition latencies (ms) and error rates (%)

	Mean RT (*SD*)	Percentage error	Example
Binding	1120 (105)	2.4%	蟑螂 "cockroach"
Opaque non-binding	1099 (108.16)	1.4%	漢堡 "hamburger"

The lack of a significant difference in repetition latencies in Experiment 2 suggests that the processing difference in Experiment 1 could not be due to purely phonological factors. Indeed, the 20 ms trend in the reverse direction may reflect a purely acoustic factor, since the mean duration for binding words (873 ms) was 32 ms longer than that for non-binding words (841 ms), though this difference was not significant either. By contrast, recall that Experiment 1 found a 67 ms effect in the opposite direction, that is, faster responses to acoustically longer stimuli.

62

The lack of a significant difference in repetition latencies for binding and non-binding words in Experiment 2 is also consistent with the hypothesis that the orthographic effect found in Experiment 1 was post-lexical. This is because the auditory repetition task is known to have a lexical component, sensitive to lexical properties such as word frequency and semantics (see review in Bates and Liu 1996). Like the visual character naming task analysed by Liu, Wu, and Chou (1996), however, auditory repetition does not contain an explicit post-lexical decision stage. Thus the pre-lexical stages for auditory lexical decision and auditory repetition are quite similar, but differ in post-lexical processing. Therefore, since an auditory LDT (Experiment 1) found a difference between binding and non-binding words while an auditory repetition task (Experiment 2) did not, it appears that the effect found in Experiment 1 was post-lexical.

If the orthographic effect seen in Experiment 1 was post-lexical, what exactly happened? The simplest possibility is that after the participants had found the words in their phonological lexicon, they still had to access the orthographic forms of the words before they could finalize their lexical decisions. This leads to a straightforward prediction: the response patterns seen with written stimuli should be identical to those found with spoken stimuli. This is equivalent to saying that written mono-morphemic compounds should show a negative frequency effect, with faster responses for words of low-frequency characters (binding words) than for words of higher-frequency characters (opaque non-binding words). As noted in the introduction, such a pattern has indeed been found for written opaque compounds.

However, other results suggest that the post-lexical process in the effect seen in Experiment 1 may be more complex than simply a "consultation" of mental orthographic representations. First, written opaque compounds do not always show negative character frequency effects, nor is the response to written opaque compounds always faster than to written transparent compounds. By contrast, the binding vs. non-binding effect found in Experiment 1, in addition to replicating the only other study (Tsai 1998) with different materials, was also robust enough to be found even when the responses from the slower and faster participants were analysed separately. Second, the automatic activation

of orthographic units in auditory lexical decision should cause positive character frequency effects in the case of transparent compounds just as they are found for that of visual lexical decisions, but Zhou and Marslen-Wilson (1994) failed to find any character frequency effects. Myers and Gong (2002) even found negative frequency effects in auditory lexical decision using transparent compounds, just like what we found using opaque compounds in Experiment 1.

We therefore propose that while participants in the auditory LDT are making post‑lexical use of orthographic information, only an abstract subpart of this information rather than an orthographic image is used. For a reader, information about specific characters may be useful in deciding word-level lexicality (the whole word is, after all, orthographically represented in terms of its component characters), which is why character frequency affects visual lexical decisions. For a listener, however, information about characters is not useful, especially since there are so many homophones. What may be beneficial to listeners, however, is cross-character predictability, which correlates with word-level lexicality. If listeners can access this information in an abstract form, independent of the character frequencies from which it is derived, it would result in their being sensitive to cross-character predictability without their being sensitive to character frequencies directly: higher-predictability words would be easier to recognize, even though they are composed of lower-frequency characters. This would explain the negative character frequency effects of Myers and Gong (2002) for spoken transparent compounds, if what were really reflected were positive cross-character predictability effects. By contrast, other than the information of cross-character predictability available, readers also have access to character frequency information, thus resulting in positive character frequency effects for transparent words (generally low cross-character predictability) and null or negative character frequency effects for opaque words (character frequency effects being cancelled out by higher cross-character predictability).

In order to test this alternative explanation for the orthographic effect, we conducted a third experiment, using the same materials as in Experiment 1, but now in written form. We expect the binding vs. non‑binding contrast to disappear.

64

4. EXPERIMENT 3: VISUAL LEXICAL DECISION

4.1 Method

4.1.1 Participants

Twenty three students, nine males and 14 females, at National Chung Cheng University in southern Taiwan, were paid for their participation in this experiment. None had participated in Experiments 1 or 2.

4.1.2 Designs and materials

A visual LDT was used. The same stimuli from Experiment 1 were presented visually. The non-word foils from that experiment were designed to be transcribable with commonly-used characters (also see Appendix I).

4.1.3 Procedure

Participants were tested individually before PC compatible computers in a sound-attenuated room. The stimuli were visually presented in pseudo-random order (different for each participant), where no more than four consecutively presented targets were pairs of real words and their corresponding non-word foils. In each trial, participants first saw the symbol "+ +" in the center of the screen, which lasted for 1000 ms, after which it was replaced by the visual stimulus. Participants then had to give as rapid and accurate a response as possible as to whether the stimulus was a real word, by pressing either the *zhen1ci2* 真詞 (real word) button on the right side or *fei1ci2* 非詞 (non-word) button on the left side of the response box. The time limit for response was set to 3000 ms and the RT was measured from the onset of the stimuli to the onset of participants' responses. The experimental procedure was co-ordinated using the E-prime software (Schneider, Eschman, and Zucolotto 2002). There were 14 practice trials before the main

experiment began. The whole experiment took approximately 10 minutes per participant.

4.2 Results and Discussion

The data from five participants were discarded because the error rates of their responses to real-word items were higher than 15%. Errors included both incorrect categorization of words and non-words and RTs above the time limit of 3000 ms. About 18% of all responses were excluded using these criterion. The mean by-participant RTs and error rates, computed for the data from the remaining participants, are reported in Table 5 (also see Appendix IV for mean RTs and error rates for each item).

Table 5.   Experiment 3: Mean lexical decision times (ms) and error rates (%)

	Mean RT (*SD*)	Percentage error	Example
Binding	552 (63.15)	7.93%	蟑螂 "cockroach"
Opaque non-binding	547 (55.15)	10.1%	漢堡 "hamburger"

As the table shows, the mean RT for binding words (552 ms) was only 6 ms longer than that for opaque non-binding words, and this slight difference was not significant by participant or by item ($t_1(17) = 0.56$, *ns*; $t_2(40) = 0.42$, *ns*). Although a few character repetitions were found among the non-binding words (e.g., *da2* 達 in *ma3da2* 馬達 and *lei2da2* 雷達, *ke4* 克 in *jia2ke4* 夾克 and *tan3ke4* 坦克, and *lei2* 雷 in *lei2she4* 雷射 and *lei2da2* 雷達), the null results cannot simply be owing to the counteracting force between the repetition priming for non-binding words and the higher cross-character predictability for binding words. This is because the processing pattern remained the same even after the above suspicious items were screened out (binding: *M* = 555.98; non-binding: *M* = 549.46; $t(17) = -0.659$, *ns*). In addition, if the repetition

priming effect in Experiment 3 were present and strong enough to facilitate non-binding word recognition, the same effect should have been observed in Experiment 1.

The error rates also did not reach statistical difference ($t_1(17) = 1.05$, *ns*; $t_2(40) = 0.87$, *ns*) although readers tended to make more mistakes in the case of opaque non-binding words (10.1%) than for binding words (7.93%). Note that, however, the overall error rate in the visual LDT (Experiment 3) was higher than that in the auditory LDT (Experiment 1), as one may have noticed. We speculate that the difference could be owing to the fact that, given the presentation of a mix of binding and regular (non-binding) compounds, the readers might feel it harder to identify regular compounds as real words since the inconsistency in the radicals may force them to realize that the recognition of the character combinations of non-binding compounds could not reliably depend on any visual help, unlike as in the case for binding words, where the characters shared the same radical. Therefore, the readers might develop a strategy of assuming that all of the binding items were real and think that all of the regular compounds were fake, thereby making more errors in identifying non-binding words. Moreover, as noted earlier in Experiment 1, it could also be that the presence of the auditory foils called for more careful attention to be paid to the task of spoken word recognition, therefore further sharpening the difference in error rates.

We also analysed the results using the Lorch-Myers repeated-measures regressions with the relevant factors as predictors, i.e., compound type, word frequency, syllable frequency, character frequency, syllable MI, and character MI, and treating RTs as the dependent variable. Table 6 summarizes the results of the regressions, which reveal the hidden effect of bindingness that the simple *t*-test did not manifest ($B = -21.60$, *SE* = 8.45, $p < 0.05$) and the facilitative effect of syllable information (syllable 1 frequency: $B = -0.0009$, *SE* = 0.0003, $p < 0.05$; syllable MI: $B = -6.185$, *SE* = 2.827, $p < 0.05$) but no word frequency effect ($B = -3338.83$, *SE* = 2790.21, *ns*).

Table 6.   Experiment 3: Summary of regression analysis for variables predicting reaction times

Variable	*B*	SE	df	t
Intercept	711.4735	40.46704	17	17.58156 *
CompType	-21.6003	8.452307	17	-2.55555 *
WordFreq	-3338.83	2790.209	17	-1.19662
AveSem	1.993291	6.994778	17	0.284968
Char1LogFreq	9.81255	9.465974	17	1.036613
Char2LogFreq	-18.197	9.612009	17	-1.89316
CharMI	-2.94336	1.097149	17	-2.68274 *
Syl1LogFreq	-0.00089	0.000305	17	-2.91171 *
Syl2 LogFreq	-0.00042	0.000241	17	-1.75054
SylMI	-6.18529	2.827449	17	-2.18759 *

Note. CompType=compound type, WordFreq=word frequency, AveSem=average of semantic transparency scores for each syllable, CharLogFreq=character log frequency, CharMI= character mutual information, SylLogFreq=syllable log frequency, SylMI=syllable mutual information, *B* = mean by-participant raw regression coefficients, *SE* = standard errors conducted across by-participant regression coefficients.
* $p < .05$

These results seem to rule out the possibility that the participants in Experiment 1 were simply making lexical decisions based on mental orthographic forms, since readers in the present experiment should have also shown faster responses to binding words. Instead, the results are more consistent with our alternative explanation. That is, at the post-lexical stage, listeners had access to an abstract subpart of orthographic

information, i.e., cross-character predictability, without also having access to the actual character frequencies. Readers have both types of information available, and since predictability and character frequency are inversely correlated, they may conflict, each cancelling out the effect of the other. This hypothesis seems to receive more support from the regression analysis where the robust effect of bindingness was seen, whereas character frequency itself had no effect.

A somewhat similar "cancelling out" account has also been used by Taft (2004). In defending the claim that polymorphemic words are obligatorily decomposed during lexical access (in reading), Taft argued that the finding of a whole-word (surface) frequency effect without a morpheme (base) frequency effect occasionally reported in the literature does not undermine the morphological decomposition view. Instead, he argued that decomposition occurs obligatory in an early stage of processing, while the components are recombined into words in a later stage. When matched in surface frequency, words with higher base frequency would take readers more time than those with lower base frequency to recombine the stem and affix at the later stage. This counteracts the benefit of easier access to the higher frequency stem at the earlier decomposition stage, resulting in no overall base frequency effect. This hypothesis applies here because previous research supports the intuitively obvious assumption that Chinese readers obligatorily decompose words into characters, so if recombination is affected by cross-character predictability, it is in principle possible for the lower cross-character predictability of non-binding words to cancel out the benefit of easier access to their higher-frequency characters.

Yet, as plausible as the "cancelling out" account is, one may still argue that the listeners in Experiment 1 recognized binding words faster than non-binding words simply because the presence of more binding items in the sub-semantic "animal" category might have primed the lexical decision process (see material discussion in Experiment 1). However, such a scenario for semantic priming has to be grounded on a less probable assumption that the (lack of) RT difference during auditory (Exp. 1) and visual perception (Exp. 3) might have come from the (absence of) activation of semantic representations, not the orthographic information per se. That is, readers only had to judge whether the

presented items matched any item physically (or orthographically) in the mental lexicon without activating semantic information, and the difference in the processing time between visual word recognition (552 ms) and spoken word recognition (1473ms) may reflect the absence of semantic activation. However, the problem of this assumption is that it contradicts the view of robust whole-word semantic activation among Chinese reading research (e.g., Liu and Peng 1997, among others). Moreover, the difference in processing time does not necessarily reflect the absence of semantic activation, but maybe simply express a fundamental difference in difficulty in processing between visual and auditory processing.[3]

As a final note before ending the discussion, we address the reliability of the cross-character predictability account by analysing the results of both Experiment 1 and Experiment 3, utilising "modality" (i.e., Experiment 1 vs. Experiment 3) as an additional predictor in another repeated-measures regression model. The regression results reveal a significant interaction between bindingness effect (i.e., compound type) and modality ($B = 5.204$, $SE = 2.55$, $p < 0.05$). The results from the different regression models taken together imply that "modality" does indeed significantly affect whether bindingness affects RTs, and the lack of an effect (based on *t*-test analysis) in Experiment 3 is arguably not due to any uncontrolled nuisance but is justified according to the operation of the "cancelling out" effect resulting from the counteracting information of character frequency and cross-character predictability.

To gain more confidence in the ascription of the above "cancelling out" account and to rule out the possibility that listeners in Experiment 1 made their decisions simply based on "orthographic representations", a follow-up experiment using a cross-modal interference LDT paradigm was thus conducted.

---

[3] Since visual processing can always be done with eyes moving back and forth, which is not possible for auditory processing, our cognitive system has evolved to adapt to the perceptual difference, hence leading to longer information persistence in the auditory sensory store (Darwin, Turvey, and Crowder 1972). Such a difference in the length of information persistence may reflect the fundamental difference in the processing difficulty of the two modalities.

5. EXPERIMENT 4: CROSS-MODAL LEXICAL DECISION

The rationale for the experimental design is that if it were the written (image) forms that were accessed by listeners in Experiment 1, the types of visual interference (i.e., character-like fake words vs. star symbols "※※") should influence spoken word recognition. That is, character-like fake words should cause the orthographic interference at the stage, where participants "consult" orthographic representations, and hence may possibly reduce the advantage of faster processing for binding words. By contrast, the star symbols "※※" without any orthographic or semantic information should have no influence during auditory word recognition.

5.1 Method

5.1.1 Participants

A total of twenty Taiwanese college students, ten males and ten females, were recruited and equally distributed into two groups, each of which was assigned one version of the stimulus sets. None of them had participated in any of the previous experiments.

5.1.2 Designs and materials

The task involved auditory lexical decision accompanied with visual interference. The same set of real-word auditory stimuli from Experiment 1 was adopted, and the corresponding visual interference items contained two types: fake-word items like "十亻臥" created by exchanging radicals of real Chinese characters with True Type (a character-forming program by Microsoft Corporation) and star symbols "※※". Two versions of the presentation sets were generated with each containing two lists. Each list contained half the target items paired with fake-word visual interference and the other half with star interference "※※" so that participants would not see the same item paired with both interference types within the same experimental session.

5.1.3 Procedure

The procedure is identical to that used in Experiment 1 except that a visual interference item was also presented in the center of the monitor simultaneously with the auditory stimulus. The duration of each visual interference item lasted for 300 msec. Therefore, participants were instructed to pay full attention to both spoken and visual targets, and then to decide whether the auditory target they heard was a real word or not.

5.2 Results and Discussion

No data were discarded, but about 5% of all responses were excluded due to errors of incorrect categorization of (non)words and RTs above the time limit of 3000 ms. The mean RTs and error rates, computed for the data of remaining participants, are reported in Table 1.

Table 7. Experiment 4: Mean lexical decision times (ms) and error rates (%)

	Star	Fake-word
Binding	1400.2 (3.3%)	1332.2 (2.4%)
Opaque non-binding	1526.8 (5.7%)	1419.1 (2.9%)

A two-way ANOVA model with repeated measures was conducted, taking Word Type (binding vs. non-binding) as the within-subject factor and Interference Type ("※※" and fake-word) as the between-subject factor. The statistical analysis showed a main effect of Word Type [$F(1, 18) = 82.097$, $p < .05$] but not of Interference Type [$F(1, 18) = 0.078$, *ns.*] or interaction [$F(1, 18) = 0.123$, *ns.*], suggesting that listeners processed binding words faster than non-binding words no matter the kind of visual interference that was presented.

72

The slower RTs for non‑binding words across the two interference conditions indicate that the RT difference during auditory perception (Experiment 1) cannot be due to the activation of true orthography. Because if it were the case, the processing advantage for binding words should have been reduced by the interference of the fake-word items, which should have been revealed in an interaction between Word Type and Interference Type. Instead, the results seem to support the view that listeners make use of an abstract subpart of orthographic information, i.e., cross-character predictability, during spoken word recognition.

6. GENERAL DISCUSSION

Even with better controlled stimuli, the results for Experiment 1 replicated the orthographic effect in auditory lexical decision first reported in Tsai (1998): spoken binding words are recognized faster than spoken opaque non-binding words. The results for Experiment 2 confirmed that phonological factors were not the cause of the pattern seen in Experiment 1, and also implied that the effect is post-lexical. The results for Experiment 3 showed that the benefit of bindingness was not observed during reading, suggesting that the effect in Experiment 1 was due to an abstract form of cross-character predictability unhindered by the influence of character frequency. Below we summarize the arguments for our two main claims about the effect: that it is post-lexical and that it comes into effect due to cross-character predictability.

We believe that the orthographic effect is post-lexical for at least four reasons. First, since we controlled syllable information (syllable frequency and cross-syllable predictability) and cohort size, as in the design for Experiment 1 and confirmed by means of another task in Experiment 2, there seem to be no phonological cues in our materials that may influence processing at an early (pre-lexical) stage. It is not possible for a listener to know that *han4* is the beginning of a non‑binding word while *zhang1* is the beginning of a binding word; realization of the difference must occur after the lexicon has been contacted. Second, since the processes involved in lexical decision (Experiment 1) and repetition (Experiment 2) share a pre-lexical stage

but not the post-lexical stage, the lack of an orthographic effect in repetition implies that this effect is localized to the post-lexical stage. Third, a pre-lexical explanation of the orthographic effect would require characters (or phonological units somehow modified by experience with characters) to be activated before the whole word was accessed in the lexicon. Since the characters in binding words are of much lower frequency relative to those in non-binding words, binding words will have more competitors and/or competitors with higher syllable frequency than non-binding words, predicting slower binding word responses, opposite to what we in fact found. Finally, the pre-lexical activation of characters during listening would seem to predict that responses to spoken and written words should be exactly the same, but Experiments 1 and 3 gave quite different results.

We also claim that the orthographic effect results from the benefits of cross-character predictability being available to listeners without being counteracted by the information of character frequency that cancels out these benefits for readers. More precisely, cross-character frequency may be a measure of a more abstract quality of "bindingness" available to both listeners and readers, derived from orthography but not expressed in terms of orthography. Years of experience reading binding and non-binding words would train the brain to treat binding words as more word-like than non-binding words. Since we claim that the cross-character predictability is not derived online from the component characters, this information would have to be stored with the words in some abstract form, and would not become available until after the words had been initially accessed. Note, however, that one may wonder why the timing of the "orthographic effect", i.e., post-lexical, in our experiments contradicts the "orthographic consistency effect" reported to occur quite early in recent studies of ERPs on alphabetic languages (e.g., Pattamadilok et al. 2009). As was argued earlier, the "orthographic effect" does not result from the activation of real orthographic representation but from the abstract information extracted from orthography. This claim was further supported by the results from follow-up Experiment 4 since the robust orthographic effect still appeared during the recognition of the spoken mono-morphemic compounds no matter under which visual interference type, "※※" or

fake words consisting of pseudo-characters. Therefore, it is not necessary to put both processes on a par and time-lock the two effects equally during the course of spoken word recognition.

This proposal is supported by the sharp contrast in the results from auditory lexical decision (a strong bindingness effect) and from visual lexical decision (no bindingness effect), and it seems to account for the facts better than several alternatives to be considered below:

According to one of these alternatives, the lack of character effects in Experiment 3 could be interpreted as evidence that reading is morpheme-based, rather than truly character-based. In transparent compounds, characters are morphemic, thus giving rise to character-based effects, but binding and opaque non-binding words are both mono-morphemic, so both types are accessed as wholes, making word frequency the only relevant factor, but since binding and opaque non-binding words were matched in word frequency, there were no processing differences. Proposals like this have been made about the processing of opaque compounds in Chinese by Chen (1993) and Tsai (1998), among others. However, in addition to being in conflict with the occurrence of negative character frequency effects sometimes found with opaque written compounds, this view would also have no explanation for the results for Experiment 1, where the responses for the two types of mono-morphemic compounds behaved quite differently.

Another alternative could be that it is character semantics and not cross-character predictability that is crucial in the orthographic effect. Non-binding characters mean something, and presumably their meanings conflict with the meanings of the opaque compounds they compose, while binding characters mean nothing, and thus cannot conflict with the meanings of the compounds they compose. The activation of characters during the recognition of spoken words would thus lead to inhibitory semantic conflicts in the case of non-binding words, which their responses relative to binding words, just as found in Experiment 1. However, not only does this hypothesis require the immediate activation of the correct characters in spite of rampant homophony, but it also has no explanation for the lack of an orthographic effect in Experiment 3, where character semantics were presumably even more immediately available. In addition, even after we entered the mean semantic

transparency scores in the regression analysis, they did not have a significant effect on the RTs observed in Experiment 1.

A more interesting version of the above alternatives would be some version of the morphology autonomy hypothesis, which claims that the encoding of morphological structure is autonomous of semantics. Roelofs and Baayen (2002) tested the autonomy hypothesis in spoken word production by examining the preparation effect for transparent and opaque compounds in Dutch. They found that the size of implicit morpheme priming was almost the same for both types of compound, implying that morphemes exist in the memory representations even for opaque complex words. Applied to the bindingness effect in Chinese, such a morphology autonomy hypothesis might mean that while binding words are truly mono-morphemic, non-binding words are composed of quasi-morphemes of some sort which have some influence on spoken word recognition. The lack of an effect in Experiment 3 would still be unexplained, however.

Finally, the most sophisticated alternative would be to adopt the multi-level interactive-activation framework for the reading of Chinese described in Taft and Zhu (1995, 1997) and Taft et al. (1999). This framework could be applied to the results for Experiment 1 as follows. First, the phonological information provided by the first syllable *han4* activates the lemma for 漢 (Chinese) as well as the lemma for the mono-morphemic compound *han4bao3* 漢堡 (hamburger). The same would apply with the first syllable *zhang1*, which would activate the lemmas both for more common homophones such as 章 (chapter) and for the actual target *zhang1lang2* 蟑螂 (cockroach). At this early stage in the processing, no distinction can be made between the two words. At a later stage, however, the activated lemmas send feedback to the associated orthographic character units. It is at this later stage where *han4bao3* and *zhang1lang2* differ: activation of the lemmas "Chinese" and "hamburger" send feedback to the same character unit at the character level while the lemmas for "chapter" and "cockroach" send feedback to different character units. Then, when the phonological information for the second syllable becomes available, the lemma of the compound "cockroach" wins out. The feedback from the compound lemma helps to increase activation of the correct character for *zhang1* 蟑, and to

differentiate it from the wrong one 章 (chapter). However, because the characters for 漢 (Chinese) and 漢堡 (hamburger) are the same, the feedback from the character level does not help distinguish the lemmas. It is the later stage in which the feedback from the lemmas has increased the activation threshold for the correct characters of binding words, hence making responses for binding words faster, as found in Experiment 1. The framework does not make a sharp distinction between pre- and post‑lexical processing, but it does imply that the orthographic effect will emerge relatively late in the process, only after activation has spread from the lemma level downward to the character level.

The problem with this model is that, once again, it incorrectly predicts that the same orthographic effect during spoken word recognition may also emerge in reading. Again the character 漢 should activate the lemmas for both "Chinese" and "hamburger," and again the character 蟑 should activate only the lemma for "cockroach," ultimately leading to faster response times to binding words, contrary to what was found. At this point we have to conclude that a model based on the storage of cross-character predictability works better than any of the other alternatives that we have considered in accounting for the orthographic effect present only during the auditory lexical access.

This study demonstrates the unexpected complexity of spoken word processing in Chinese, an area that is solely neglected, but as we hope to have shown, it is an area that may provide surprising information about word recognition and the nature of morphological processing. Moreover, the orthographic effect demonstrates the need for caution in interpreting studies that claim to find morphological units in the processing of spoken Chinese (e.g., Zhou and Marslen-Wilson 1995), since it may be the case that the access to some sort of orthographic information plays a role in how lexical decisions are being made. However, since a model relying solely on cross-character predictability is insufficient to address all of the issues relating to spoken compound processing, clearly more work needs to be done to reconfirm this effect and explore its possible causes and interaction with other variables during spoken word recognition. This situation could likely be alleviated by shifting some of the resources currently spent on Chinese reading research over to the study of the processing of the spoken word in Chinese. In particular, there are two

obvious directions for further research: One is the examination of whether the orthographic effect is modulated by word frequency since word-internal-structure effects like semantic transparency or morpheme frequency have sometimes been reported to be modulated by word frequency in Chinese word-reading studies (e.g., Liang 1992; Chen 1993; Lee 1995, among others). The other direction is an investigation of the depth of semantic processing and the time course of the process of recognition between the visual and auditory lexical access. Given the great difference in the processing speed for the same set of materials in the two modalities (Visual vs. Auditory: 552 ms vs. 1473 ms), the linguistic variables involved and the timings of activation during the course may not be the same (cf. Holcomb and Neville 1990). The validity of our post-lexical account grounded on the parallel drawn between the behavioural data of listening and reading may have been inflated. Thus, electro-physiological data that show the time course and scalp distribution would help to clarify features of the process that may be yet unknown. It is expected that studies taken even just one step further would help to integrate research on spoken compound processing and word processing into a more general study of the processing of language in real time.

REFERENCES

Bates, Elizabeth, and Hua Liu. 1996. Cued shadowing. *Language and Cognitive Processes* 11.6:577-581.

Booth, James R., Douglas D. Burman, Joel R. Meyer, Darren R. Gitelman, Todd B. Parrish, and Marek Marsel Mesulam. 2004. Development of brain mechanisms for processing orthographic and phonologic representations. *Journal of Cognitive Neuroscience* 16:1234-1249.

Chao, Yuen-Ren. 1968. *A Grammar of Spoken Chinese.* Berkeley: University of California Press.

Chen, Shao-Tsu. 1993. *Hanyu gouci zai yuedu licheng zhong dui yuyi chufa xiaogai de yingxiang* [The influence of semantic activation effects on Chinese morphology during reading.] Unpublished master's thesis, National Chung Cheng University, Chia-Yi, Taiwan.

Chen, Keh-Jiann, Chu-Ren Huang, Li-Ping Chang, and Hui-Li Hsu. 1996. SINICA CORPUS: Design methodology for balanced corpora. *Language, Information and Computation* 11:167-176.

Chinese Knowledge Information Processing Group. 1994. *Zhongwen shumianyu pinlü cidian* [A word frequency dictionary for written Chinese] (CKIP Technical Report No. 94-01). Taipei, Taiwan: Academia Sinica.

Chinese Knowledge Information Processing Group. 1995. *Mandarin Chinese character frequency list based on national phonetic alphabets*. (CKIP Technical Report No. 95-01). Taipei: Academia Sinica.

Church, Kenneth W., and Patrick Hanks. 1990. Word association norms, mutual information, and lexicography. *Computational Linguistics* 16.1:22-29.

Darwin, Christopher J., Michael T. Turvey, and Robert G. Crowder. 1972. An auditory analogue of the Sperling partial report procedure: Evidence for brief auditory storage. *Cognitive Psychology* 3:255-267.

DeFrancis, John. 1984. *The Chinese language: Fact and fantasy*. Honolulu, HI: University of Hawaii Press.

Dijkstra, Ton, Ardi Roelofs, and Steffen Fieuws. 1995. Orthographic effects on phoneme monitoring. *Canadian Journal of Experimental Psychology* 49:264-271.

Frauenfelder, Uli H., Juan Segui, and Ton Dijkstra. 1990. Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance* 16.1:77-91.

Fu, Ding-Yi. 1986. *Lianmian zidian*. [Dictionary of Chinese binding words.] Taipei, Taiwan.

Hallé, Pierre A., Céline Chéreau, and Juan Segui. 2000. Where is the /b/ in "absurde" [apsyrd]? It is in French listener's minds. *Journal of Memory and Language* 43:618-639.

Henderson, Leslie. 1985. Toward a psychology of morphemes. *Progress in the psychology of language*, ed. by Andrew W. Ellis, 15-72. London: LEA.

Holcomb, Phillip J., and Helen J. Neville. 1990. Auditory and visual semantic priming in lexical decision: A comparison using event-related brain potentials. *Language and Cognitive Processes* 5.4:281-312.

Hoosain, Rumjahn. 1992. Psychological reality of the word in Chinese. *Language processing in Chinese: Advances in psychology*, ed. by Hsuan-Chih Cheng, and Ovid J.-L. Tzeng, 111-130. Amsterdam: North-Holland.

Jakimik, Jola, Ronald A. Cole, and Alexander I. Rudnicky. 1985. Sound and spelling in spoken word recognition. *Journal of Memory and Language* 24:165-178.

Lai, Ming, and Tai-Yi Lin. 1987. *Zui xin Lin Yutang han-ying cidian*. [The new Lin Yutang Chinese-English dictionary.] Hong Kong: Dasheng Publishing.

Lee, Chia-Ying. 1995. *Hanyu zuheci yu chengyuci zai xinli cidian zhong de biaozheng fangshi*. [The representation of compounds and idioms in the Chinese mental lexicon.] Unpublished master's thesis, National Chung Cheng University, Chia-Yi, Taiwan.

Liang, Mei-Ya. 1992. *Recognition processing in reading compositional and idiomatic words*. Unpublished master's thesis, National Tsinghua University, Hsinchu, Taiwan.

Liu, In-Mao, Jei-Tun Wu, and Tai-Li Chou. 1996. Encoding operation and transcoding as the major loci of the frequency effect. *Cognition* 59:149-168.

Liu, Ying, and Dan-Ling Peng. 1997. Meaning access of Chinese compounds and its time course. *Cognitive processing of Chinese and related Asian languages*, ed. by Hsuan-Chih Chen, 219-232. Hong Kong: The Chinese University Press.

Lorch, Robert F., and Jerome L. Myers. 1990. Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16.1:149-157.

Lü, Ching-Ching. 1996. *Zhongwen yuci renzhi licheng*. [Chinese word recognition.] Unpublished doctoral thesis, National Tsinghua University, Hsinchu, Taiwan.

Mandarin Daily News Association. 1981. *Guoyu ribao wailaiyu cidian*. [Mandarin daily news dictionary of loan words.] Taiwan: Taipei.

McCandliss, Bruce D., Laurent Cohen, and Stanislas Dehaene. 2003. The visual word form area: expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences* 7:293-299.

Miller, Kimberly M., and Diane Swick. 2003. Orthography influences the perception of speech in alexic patients. *Journal of Cognitive Neuroscience* 15:981-990.

Myers, James. 2006. Processing Chinese morphology: A survey of the literature. *The representation and processing of compound words*, ed. by Gary Libben and Gonia Jarema, 169-196. Oxford: Oxford University Press.

Myers, James, and Shu-Ping Gong. 2002. Cross-morphemic predictability and the lexical access of compounds in Mandarin Chinese. *Folia Linguistica* 26.1.2:65-96.

Orfanidou, Eleni, William D. Marslen-Wilson, and Matthew H. Davis. 2006. Neural Response Suppression Predicts Repetition Priming of Spoken Words and Pseudowords. *Journal of Cognitive Neuroscience* 18:1237-1252.

Packard, Jerome L. 1999. Lexical access in Chinese speech comprehension and production. *Brain and Language* 68:89-94.

Pattamadilok, Chotiga, José Morais, Paulo Ventura, and Régine Kolinsky. 2007. The locus of the orthographic consistency effect in auditory word recognition: Further evidence from French. *Language and Cognitive Processes* 22:1-27.

Pattamadilok, Chotiga, Laetitia Perre, Stéphane Dufau, and Johannes C. Ziegler. 2009. On-line orthographic influences on spoken language in a semantic task. *Cognitive Neurosciences* 21:169-179.

Peng, Dang-Ling, Ying Liu, and Chun-Mao Wang. 1999. How is access representation organized? The relation of polymorphemic words and their morphemes in Chinese. *Reading Chinese script: A cognitive analysis*, ed. by Jian Wang, Hsuan-Chih Chen, Ralph Radach, and Albrecht Inhoff, 65-89. Lawrence Erlbaum Associates: Mahwah, New Jersey.

Perre, Laetitia, and Johannes C. Ziegler. 2007. On-line activation of orthography in spoken word recognition. *Brain Research* 1188:132-138.

Radeau, Monique, Philippe Mousty, and Paul Bertelson. 1989. The effect of the uniqueness point in spoken-word recognition. *Psychological Research* 51:123-128.

Roelofs, Ardi, and Harald Baayen. 2002. Morphology by itself in planning the production of spoken words. *Psychonomic Bulletin and Review* 9.1:132-138.

Schneider, Walter, Amy Eschman, and Anthony Zuccolotto. 2002. *E-Prime reference guide.* Pittsburgh: Psychology Software Tools Inc.

Seidenberg, Mark S., and Michael K. Tanenhaus. 1979. Orthographic effects on rhyming monitoring. *Journal of Experimental Psychology: Human Learning and Memory* 5:531-545.

Starosta, Stanley, Koenraad Kuiper, Siew-Ai Ng, and Zhi-Quian Wu. 1998. On defining the Chinese compound word: Headedness in Chinese compounding and Chinese VR compounds. *New approaches to Chinese word formation*, ed. by Jerome Lee Packard, 347-370. Berlin: Mouton de Gruyter.

Su, Yi-Ching. 1998. The representation of compounds and phrases in the mental lexicon: evidence from Chinese. *University of Maryland Working Papers in Linguistics* 6:179-199.

Taft, Marcus. 2003. Morphological representation as a correlation between form and meaning. *Reading complex words*, ed. by Egbert Assink and Dominiek Sandra, 113-138. Amsterdam: Kluwer.

Taft, Marcus. 2004. Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology* 57A.4:745-765.

Taft, Marcus, and Gail Hambly. 1985. The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language* 24.3:320-335.

Taft, Marcus, Jong-Tsun Huang, and Xiao-Ping Zhu. 1994. The influence of character frequency on word recognition responses in Chinese. *Advances in the study of Chinese language processing vol. 1*, ed. by Hsing-Wu Chang, Chih-Wei Hue, and Ovid J.-L. Tzeng, 59-73. Taipei: Department of Psychology, National Taiwan University.

Taft, Marcus, Yu Liu, and Xiao-Ping Zhu. 1999. Morphemic processing in reading Chinese. *Reading Chinese script: A cognitive analysis*, ed. by Jian, Wang, Albrecht W. Inhoff, and Hsuan-Chih Chen, 91-113. Mahwah, NJ: Lawrence Erlbaum Associates.

Taft, Marcus, and Xiao-Ping Zhu. 1995. The representation of bound morphemes in the lexicon: A Chinese study. *Morphological aspect of language processing*, ed. by Laurie B. Feldman, 293-361. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Taft, Marcus, and Xiao-Ping Zhu. 1997. Using masked priming to examine lexical storage of Chinese compound words. *Cognitive processing of Chinese and related Asian languages*, ed. by Hsuan-Chih Chen, 233-241. Hong Kong: The Chinese University Press.

Tanenhaus, Michael K., Helen P. Flanigan, and Mark S. Seidenberg. 1980. Orthographic and phonological priming in auditory and visual word recognition. *Memory and Cognition* 8:513-520.

Tsai, Chih-Hao. 1994. *Effects of semantic transparency on the recognition of Chinese two-character words: Evidence for a dual-process model.* Unpublished master's thesis. Graduate Institute of Psychology, National Chung Cheng University, Chia-Yi, Taiwan.

Tsai, Pei-Wen. 1998. *Auditory access of disyllabic words in the Chinese mental lexicon.* Unpublished master's thesis, Graduate Institute of Linguistics, National Chung Cheng University, Chia-Yi, Taiwan.

Tseng, Shu-Chuan. 2004. Processing Spoken Mandarin Corpora. *Traitement Automatique des Langues. Special Issue: Spoken Corpus Processing* 45.2:89-108.

Ventura, Paulo, José Morais, Chotiga Pattamadilok, and Régine Kolinsky. 2004. The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes* 19:57-95.

Yip, Chi-Wing M. 2001. Phonological priming in Cantonese spoken-word processing. *Psychologia* 44.3:223-229.

Zecker, Steven G. 1991. The orthographic code: Developmental trends in reading-disabled and normally-achieving children. *Annals of Dyslexia* 41:178-192.

Zhang, Bi-Ying, and Dan-Ling Peng. 1992. Decomposed storage in the Chinese lexicon. *Language processing in Chinese: Advances in psychology*, ed. by Hsuan-Chih Cheng and Ovid J.-L. Tzeng, 131-149. Amsterdam, Netherlands: North-Holland.

Zhou, Xiao-Lin 2000. Phonology in lexical processing of Chinese: Priming tone neighbours. *Psychological Science* 23.2:133-140.

Zhou, Xiao-Lin, and William D. Marslen-Wilson. 1994. Words, morphemes and syllables in the Chinese mental lexicon. *Language and Cognitive Processes* 9.3:393-422.

Zhou, Xiao-Lin, and William D. Marslen-Wilson. 1995. Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes*, 10.6:545-600.

Zhou, Xiao-Lin, William D. Marslen-Wilson, Marcus Taft, and Hua Shu. 1999. Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes* 14.5/6:525-565.

Ziegler, Johannes C., and Ludovic Ferrand. 1998. Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review* 5:683-689.

Yu-da Lai
Department of English Language, Literature, and Linguistics
Providence University
Taichung, Taiwan 433, ROC
yudalai@pu.edu.tw

James Myers
Graduate Institute of Linguistics
National Chung Cheng University
Chiayi, Taiwan 621, ROC

APPENDIX I: LIST OF EXPERIMENTAL STIMULI

	Binding words					Opaque non-binding words			
Target	WF	SMI	CMI	Foil	Target	WF	SMI	CMI	Foil
檳榔	0.00519	16.57	21.78	檳 浪	瓦斯	0.00762	14.61	17.28	瓦 四
玻璃	0.00438	13.34	21.42	玻 利	咖啡	0.00217	14.55	20.99	咖 肥
葡萄	0.00154	14.91	21.95	葡 套	雷射	0.00186	11.41	17.13	雷 奢
蝴蝶	0.00075	12.65	22.81	蝴 爹	馬達	0.00122	10.44	11.44	馬 搭
玫瑰	0.00065	9.01	23.10	玫 鬼	坦克	0.00114	10.84	14.22	坦 殼
蟑螂	0.00063	11.17	24.83	蟑 浪	卡通	0.00111	11.80	11.92	卡 同
葫蘆	0.00057	12.82	21.94	葫 露	雷達	0.00102	12.08	13.47	雷 搭
蘿蔔	0.00047	10.92	23.82	蘿 播	寶貝	0.00073	7.46	15.62	寶 北
珊瑚	0.00041	9.83	20.08	珊 互	夾克	0.00067	10.93	14.83	夾 苛
琥珀	0.0003	8.97	25.20	琥 波	沙發	0.00057	8.17	10.06	沙 法
螞蟻	0.00029	5.68	25.22	螞 依	漢堡	0.00042	6.95	17.33	漢 抱
駱駝	0.00023	14.04	23.07	駱 托	菩薩	0.00037	13.04	21.54	菩 撒
鴛鴦	0.00023	12.84	26.32	鴛 陽	納粹	0.00035	12.95	17.06	納 催
琵琶	0.00022	13.91	26.61	琵 怕	芭蕾	0.00028	9.27	21.08	芭 勒
蝙蝠	0.00015	7.04	26.85	蝙 副	尼龍	0.00027	9.83	12.27	尼 攏
蜘蛛	0.00013	5.07	24.58	蜘 主	派對	0.00021	6.34	6.73	派 堆
膀胱	0.0001	10.40	24.39	膀 逛	幾何	0.00037	5.13	10.08	幾 喝
鷺鷥	0.00048	4.17	26.35	鷺 寺	賓果	0.0003	9.92	11.05	賓 鍋
橄欖	0.00019	10.53	23.21	橄 爛	沙拉	0.00033	9.79	11.73	沙 喇
茉莉	0.00016	3.65	20.46	茉 哩	分貝	0.00028	4.98	10.34	分 北
骷髏	0.00017	12.46	27.36	骷 漏	可樂	0.00058	7.61	8.42	可 勒
Average	0.000821	9.91	29.87			0.00104141	0.48	14.02	

Note. WF = word frequency (proportion of the number of word tokens out of 9,529,233 words), SMI = syllable mutual information, CMI = character mutual information.

APPENDIX II: MEAN RTS (MS) AND ERROR RATES (%) IN EXPERIMENT 1

	Binding words					Opaque non-binding words				
Target	RT	Error	Foil	RT	Error	Target	RT	Error	Foil	RT	Error
檳榔	1460	5	檳浪	1529	11	瓦斯	1433	0	瓦四	1480	0
玻璃	1403	5	玻利	1464	0	咖啡	1619	0	咖肥	1408	11
葡萄	1532	0	葡套	1348	5	雷射	1524	0	雷奢	1667	0
蝴蝶	1403	11	蝴爹	1231	0	馬達	1609	0	馬搭	1560	0
玫瑰	1417	0	玫鬼	1375	0	坦克	1576	16	坦咳	1624	11
蟑螂	1468	0	蟑浪	1492	0	卡通	1674	5	卡同	1624	0
葫蘆	1523	0	葫露	1283	11	雷達	1567	5	雷搭	1422	5
蘿蔔	1454	0	蘿播	1268	16	寶貝	1521	11	寶北	1553	0
珊瑚	1566	5	珊互	1213	0	夾克	1472	0	夾苛	1360	16
琥珀	1518	0	琥波	1426	5	沙發	1617	0	沙法	1324	0
螞蟻	1642	21	螞依	1598	0	漢堡	1614	5	漢抱	1367	5
駱駝	1462	5	駱托	1579	0	菩薩	1428	5	菩撒	1701	5
鴛鴦	1528	5	鴛陽	1492	0	納粹	1626	0	納催	1368	5
琵琶	1519	11	琵怕	1558	5	芭蕾	1429	0	芭勒	1217	16
蝙蝠	1477	0	蝙副	1433	5	尼龍	1611	5	尼攏	1612	5
蜘蛛	1392	0	蜘主	1409	5	派對	1478	5	派堆	1583	11
膀胱	1362	0	膀逛	1301	16	幾何	1446	0	幾喝	1445	0
鷥鷥	1392	0	鷥寺	1517	5	賓果	1483	5	賓鍋	1474	16
橄欖	1566	0	橄爛	1444	5	沙拉	1472	0	沙喇	1518	21
茉莉	1363	0	茉哩	1370	11	分貝	1507	0	分北	1407	0
骷髏	1486	5	骷漏	1571	5	可樂	1635	0	可勒	1422	11
Average	1473	3.5		1424	5		1540	3		1483	7

Note. RTs = reaction times.

APPENDIX III: SEMANTIC RELATEDNESS SCORES & UP LENGTHS

	Binding words							Opaque **non-**binding words					
Target	Char.1	Score	Char.2	Score	Avg.	UP	Target	Char.1	Score	Char.2	Score	Avg.	UP
檳榔	檳	3.5	榔	4.3	3.9	652	瓦斯	瓦	1.55	斯	3.1	2.32	657
玻璃	玻	2.4	璃	3.95	3.17	764	咖啡	咖	4.3	啡	1.95	3.12	654
葡萄	葡	4.7	萄	3.7	4.2	650	雷射	雷	2.75	射	3.2	2.97	724
蝴蝶	蝴	2.35	蝶	5.05	3.7	661	馬達	馬	1.85	達	2.5	2.17	832
玫瑰	玫	2.55	瑰	2.15	2.35	818	坦克	坦	2.15	克	2.55	2.35	1050
蟑螂	蟑	2.4	螂	2.4	2.4	639	卡通	卡	3.2	通	3.2	3.2	891
葫蘆	葫	4.65	蘆	3.65	4.15	766	雷達	雷	1.85	達	2.84	2.34	1025
蘿蔔	蘿	3.25	蔔	4.1	3.67	737	寶貝	寶	3.3	貝	4.05	3.67	795
珊瑚	珊	4	瑚	3.95	3.97	806	夾克	夾	2	克	2.47	2.23	867
琥珀	琥	4.45	珀	3.2	3.82	828	沙發	沙	1.75	發	1.15	1.45	740
螞蟻	螞	3.4	蟻	3.6	3.5	806	漢堡	漢	3.35	堡	4.05	3.7	842
駱駝	駱	3.7	駝	5.2	4.45	773	菩薩	菩	2.25	薩	1.95	2.1	769
鴛鴦	鴛	4.17	鴦	3.95	4.05	768	納粹	納	2.8	粹	2.55	2.67	700
琵琶	琵	3.25	琶	3.4	3.32	790	芭蕾	芭	2.75	蕾	3.15	2.95	788
蝙蝠	蝙	3.75	蝠	3.1	3.42	675	尼龍	尼	3.6	龍	3.35	3.47	839
蜘蛛	蜘	4.45	蛛	5.4	4.92	706	派對	派	2.15	對	1.9	2.02	967
膀胱	膀	3.95	胱	4.4	4.17	762	幾何	幾	2.25	何	2.5	2.37	760
鷺鷥	鷺	4.55	鷥	4	4.27	629	賓果	賓	1.95	果	2.55	2.25	802
橄欖	橄	2.55	欖	2.65	2.6	629	沙拉	沙	2.6	拉	2.25	2.42	936
茉莉	茉	3.9	莉	2.5	3.2	689	分貝	分	3.65	貝	2.8	3.22	818
骷髏	骷	3.2	髏	2.7	2.95	674	可樂	可	3.15	樂	2.5	2.82	767

Note. Char. = character, Score = average semantic relatedness rating scores to the target compound (1 = least related, 6 = most related), Avg. = average semantic relatedness rating scores of component characters to their target compound word, UP = uniqueness point (ms).

APPENDIX IV: MEAN RTS (MS) AND ERROR RATES (%) IN EXPERIMENT 3

						Opaque **non**-binding words				
		Binding words								
Target	RT	Error	Foil	RT	Error	Target	RT	Error	Foil	RT	Error
檳榔	522	0	檳浪	522	6	瓦斯	527	0	瓦四	525	6
玻璃	582	22	玻利	572	22	咖啡	549	6	咖肥	571	6
葡萄	527	6	葡套	560	6	雷射	513	0	雷奢	535	0
蝴蝶	521	0	蝴爹	534	0	馬達	544	0	馬搭	549	11
玫瑰	573	11	玫鬼	573	11	坦克	650	22	坦咳	650	17
蟑螂	542	0	蟑浪	551	0	卡通	607	17	卡同	607	17
葫蘆	521	0	葫露	540	0	雷達	563	28	雷搭	567	11
蘿蔔	516	0	蘿播	527	11	寶貝	550	6	寶北	559	6
珊瑚	626	11	珊互	643	17	夾克	558	33	夾苛	568	22
琥珀	485	0	琥波	509	0	沙發	561	11	沙法	552	11
螞蟻	660	39	螞依	620	17	漢堡	571	17	漢抱	544	17
駱駝	549	0	駱托	598	11	菩薩	512	0	菩撒	528	0
鴛鴦	565	0	鴛陽	589	11	納粹	646	28	納催	536	17
琵琶	595	0	琵怕	595	0	芭蕾	488	11	芭勒	530	11
蝙蝠	504	0	蝙副	522	0	尼龍	517	6	尼攏	517	6
蜘蛛	516	0	蜘主	541	0	派對	545	11	派堆	559	11
膀胱	510	0	膀逛	528	0	幾何	527	6	幾喝	516	6
鷺鷥	515	11	鷺寺	545	11	賓果	478	0	賓鍋	505	0
橄欖	580	22	橄爛	617	22	沙拉	552	0	沙喇	561	17
茉莉	600	28	茉哩	631	28	分貝	472	11	分北	493	11
骷髏	582	6	骷漏	605	6	可樂	536	6	可勒	646	6
Average	552	7.9		568	8.5		547	10		553	9.7

Note. RTs = reaction times.

Yu-da Lai and James Myers

單語素複合詞於口語中文之辨識歷程

賴昱達　　　麥傑
靜宜大學
國立中正大學

本研究旨在探討單語素複合詞於口語中文中的詞彙存取歷程，並進一步了解文字構型於口語字詞辨識歷程中所扮演的角色。傳統漢語語言學定義複合詞為包含二個以上的國字字串，其國字是否為詞素乃非必要條件，依其定義，中文單語素雙字詞即屬於複合詞的一種；而單語素複合詞又可依其文字構型上的完整性，亦即內部國字間的緊密度，再分為連綿詞與音譯詞兩類，前者之內部國字只可與出現在該詞的相鄰國字共同出現，例如"蟑螂"，後者之內部國字還可與其它國字出現於其它字詞當中，例如"漢堡"。為了解此文字構型完整性的差異是否影響口語字詞辨識之歷程，本研究控制了目標詞的詞頻、音節頻率、跨音節預測度、群組個數並操弄其組成國字間的緊密度，分別進行了口語詞彙判斷作業（實驗一）、複誦作業（實驗二）、閱讀詞彙判斷作業（實驗三）與跨模組詞彙判斷作業（實驗四）。實驗一結果發現連綿詞所需的判斷時間較音譯詞來得短，證實了於口語字詞辨識過程裡文字構型完整性效果的存在。實驗二結果發現連綿詞與音譯詞的複誦反應時間沒有顯著差異，由於複誦過程所反應的是音韻處理過程，因此實驗一所發現的效果並非音韻上的差異所造成，同時也可推測該效果應出現於後詞彙加工階段。實驗三結果發現連綿詞與音譯詞的詞彙判斷時間沒有顯著差異，說明了實驗一所發現的效果並非來自於文字構型視覺上的激發，而是跨字預測度的資訊所造成，因為在實驗四不同類型的視覺干擾下，仍可發現與實驗一相同的效果。本研究並討論此效果的成因與其在中文口語詞彙實驗裡的意涵。

關鍵詞：口語字詞辨識歷程、單語素複合詞、文字構型