Processing Chinese Compounds: A Survey of the Literature

James Myers

National Chung Cheng University, Taiwan

Introduction

Chinese is the poster child of compounding, the language to cite for an example of morphology without much affixation. This alone should make Chinese worthy of its own chapter in a book on compound processing, but another point in its favor is its notoriously unusual writing system; orthography, as we will see, plays a crucial role in how compounds are processed in Chinese. The literature on Chinese compound processing is rich and ever-growing (indeed, some of it is also described in the chapters in this book by Jarema, Nicoladis, and Levy, et al.). This chapter attempts to provide a general overview of this literature, the goal being not to argue for or review particular models, but rather to organize the key findings from a variety of sources in a way that highlights the universal and language-specific properties of Chinese compound processing (a very helpful earlier review by Taft, Y. Liu, & Zhu, 1999, describes only reading studies within a multilevel interactive-activation framework).

The chapter begins with overviews of Chinese, its orthography, and its morphology. The themes that organize the remainder of the chapter overlap with those addressed by other authors in this book. As with other languages, the Chinese literature consists primarily of reading experiments, but spoken word recognition and aphasiology have also been studied. The hope is that this review will not only introduce Chinese compounding to readers more familiar with other languages, but will also inspire them to build on paradigms or notions found in the Chinese literature but as yet unexplored elsewhere.

Linguistic Background

For all its fame, Chinese seems to be subject to some common misconceptions among psycholinguists who don't work on it. To begin with, the term 'Chinese' is itself ambiguous, since it can refer either to a particular language or to an entire language family (Sinitic). As a particular language, 'Chinese' refers to Mandarin (also known as Standard Chinese or Putonghua). Chinese orthography is essentially nonphonetic, making it usable (with some adjustments) for all Sinitic languages. For language teaching and dictionaries, Mandarin can also be written with various supplementary phonetic systems; Mandarin romanizations in this paper use Hanyu Pinyin, the official phonetic orthography of the PRC and Singapore. For typographical convenience, Chinese examples in the text are only given in their Pinyin transcriptions (suppressing the tone diacritics), but all examples are given in characters in the appendix.

As noted in the introduction, it is important to say a bit more about Chinese orthography. It is, of course, nonalphabetic, and it is sometimes implied that an alphabetic system would not work for Chinese due to the large number of homophones; there are, for example, about thirty distinct characters all pronounced in Mandarin as *yi* with a high level tone (K.-C. Ho, 1976). Such comments of course neglect the fact that spoken Chinese is understood without difficulty, a mundane but important observation to be returned to later.

The basic orthographic unit is the character, though often it is itself composed of elements that recur in other characters, often one part that gives some clue to the pronunciation and another part that roughly categorizes the meaning (the semantic radical); we will mostly ignore this 'subcharacter morphology' here (but see Taft, et al., 1999). A character virtually always represents one syllable and also almost always one morpheme, but belying the reputation of Chinese as an 'isolating' language, most words are not monomorphemic; Zhou and Marslen-Wilson (1995) estimate the proportion of disyllabic (two-character) words in Chinese as about 74% by type and 34% by token. As an example, the complex word *xiguazhi* (watermelon juice) is composed of the monosyllabic one-character morphemes *xi* (west), *gua* (melon), and *zhi* (jiuce); *xigua* itself means 'watermelon.' There are, however, some disyllabic morphemes, written with characters that share a semantic radical and are virtually never used independently of each other (called binding characters by Taft & Zhu, 1995). An example is *putao* (grape); both of its characters are topped by the semantic radical for plants and neither appears in any other word. Recent borrowings of polysyllabic morphemes tend to be written with existing characters, chosen for their pronunciation; for example, *shafa* (easy chair), adapted from English *sofa*, is written with the semantically irrelevant characters *sha* (sand) and *fa* (send out), making it apparently monomorphemic in Chinese as well. It should also be noted that even aside from these types of words, characters are not identical to morphemes, since they often shift arbitrarily in pronunciation, meaning, or both, from context to context. Thus the first morpheme in *xingren* (pedestrian, literally 'walk-person') and the second morpheme in *yinhang* (bank, literally 'silver-store') are written with the same character.

Familiarity with Chinese orthography has a profound influence on native speakers' understanding of the notion 'word,' which is often translated as *zi* (character). The more linguistically sophisticated prefer the term *ci*, but there is confusion about its precise meaning (historically it referred to a style of poetry). Spaces are never used to separate words in written text, making characters obvious but words much more abstract. Characters, even binding characters, are usually given separate entries in dictionaries and the distinction between word and phrase is much more unclear than it is in languages like English, even for linguists. Some, like

Chao (1968), have even declared that whether or not Chinese 'words' (in the English sense) exist at all is an empirical question, not one to be decided by universalist fiat. For example, Bates, S. Chen, P. Li, Opie, and Tzeng (1993) argue that *chifan* (eat, literally 'eat-rice') is a word, not a phrase, because it is semantically idiosyncratic, describing eating in general, even eating noodles. Yet its components can be separated syntactically (e.g., *chi-le-fan* 'ate,' literally 'eat-completive.aspect-rice'). Duanmu (1998) and Packard (2000) give somewhat more reliable wordhood tests, but such academic exercises have yet to influence the intuitions of ordinary Chinese readers. In experimental studies of the Chinese 'wordhood' question, both Hoosain (1992) and C.-H. Tsai, McConkie and Zheng (1998) asked readers to draw slashes between word boundaries; results showed much disagreement, suggesting that what for some readers was a word, was for others a phrase.

Assuming for the moment that Chinese does indeed have words, we turn now to word structure. It is often implied that Chinese morphology consists entirely of compounding, but this is not quite correct. First, modern Chinese has a number of productive affixational processes, including derivational affixation, inflectional affixation (cf. the misleading title of P. Li, Bates, & MacWhinney, 1993), and reduplication; overviews of Chinese morphology can be found in C. N. Li & Thompson (1981) and Packard (2000). An example of a productive derivational affix is *zhe*, roughly corresponding to agentive *-er* in English; thus 'author' is *zuozhe* (literally 'maker'). An example of inflection is the completive aspect marker *le*, as in *zuo-le* (made). Reduplication can be found, for example, in *ganganjingjing* (very clean), derived from *ganjing* (clean).

Second, by the common definition by which compounds must be composed of free words (see, e.g., Bloomfield, 1933, Fromkin & Rodman, 1998, and review in Dressler, this volume), a large proportion of Chinese compounds are not genuine compounds, since the morphemes that compose them are not free words; the proportion of bound characters in the Academia Sinica Balanced Corpus of Modern Chinese (K.-J. Chen, C.-R. Huang, Chang, & H.-L. Hsu, 1996) is around 36%. Compounds composed of bound morphemes even include many fully transparent ones, like *xiaozhang* (school president). This state of affairs has led to two opposite reactions among linguists, with some (e.g., Sproat & Shih, 1996; Packard, 2000) arguing that boundedness has no effect on the analysis of a word as a compound, while others (e.g., Dai, 1992; Starosta, Kuiper, Ng, & Z. Wu, 1998) disagree. In favor of the latter position, Starosta, et al. (1998) suggest that semantically opaque compounds like *dongxi* (thing, but literally 'east-west') should be considered monomorphemic, and that some bound morphemes should be reanalyzed as suffixes (e.g., *hao*, meaning 'good' in true compounds, merely marks successful completion in so-called resultative verbs like *xiehao* 'finished writing').

Third, compounding itself is a two-way operation, since consistent with word-based morphology (e.g., Anderson, 1992), a Chinese compound defines its component morphemes as much as they define it. As Libben (this volume) points out, the top-down element in compound semantics is a universal, but Chinese may take it to an unusual extreme, since new morphemes are readily created in Chinese by having individual characters inherit some aspect of compound meaning. Packard (2000) catalogs several examples of this, and it is easy to find more. For example, *feiji* (airplane) is literally 'fly-machine,' but in compounds such as *jichang* (airport, where *chang* is 'lot'), *ji* has inherited the meaning of *feiji* as a whole. This process can even affect apparently monomorphemic words; thus the *tai* (literally 'tower') of *Taiwan* (historically derived by metonymy from an Austronesian toponymic) has been reanalyzed as itself meaning 'Taiwan,' as in *Taibei* (Taipei, literally 'Taiwan-north') and *laitai* (come to Taiwan).

In spite of all this, however, the impression that compounding is the primary morphological operation in Chinese is basically correct. Nominal compounds include the canonical endocentric modifier-noun type, several examples of which have already been cited (e.g., *xiaozhang* 'school president' [NN], *feiji* 'airplane' [VN]). Exocentric nominal compounds include those of the usual semantically opaque (or translucent) type, such as *dongxi* (thing, cited above) or *huasheng* (peanut, literally 'flower-birth'), but also of the more interesting coordinative type, constructed productively from morphemes that contribute equally to the word meaning (e.g., *fumu* 'parents,' literally 'father-mother,' and *zici* '*zi* and *ci*,' seen in the title of Zhang, 1997). Chinese orthography also allows for the existence of oxymoronic 'monomorphemic compounds,' both binding words (e.g., *putao* 'grape') or phonetic borrowings composed of nonbinding characters (e.g., *shafa* 'easy chair'). Beyond these types, most psycholinguistic research has yet to explore systematically very far, so this chapter won't have much to say about verbal compounds (e.g., *qiuzheng* 'seek proof' [VN], *yanzheng* 'test and verify' [VV]), adjectival compounds like *ganjing* (clean, literally 'dry-clean'), and so on, though compounds of these types do appear occasionally in experimental materials.

This overview should suggest some predictions for processing. Since characters give readers a great deal of help in identifying morphemes (though not perfect, due to shifts in pronunciation and/or meaning), the most fundamental processing question is not the usual 'How are morphologically complex words decomposed?' but rather 'How are morphemes composed to form complex words?' Indeed, in the spirit of Chao (1968), we should first ask whether words play any role in Chinese reading at all. Nor is the word the only notion open to question; as we have seen, Chinese orthography also gives rise to similar vagueness in the concepts 'morpheme' (as opposed to character) and 'compound' (as opposed to morpheme and affixed word).

Listeners are faced with quite a different situation from readers, of course. As pointed out by Packard (1999), the rampant homophony of Chinese morphemes makes it implausible to suppose that listeners access word representations by first activating morpheme representations. However, they don't seem to simply access words as whole phonological forms either, given evidence that access is syllable-based; for example, in a phoneme-detection task Tseng, K.-Y. Huang, and Jeng (1996) found that it was the lexicality of syllables that made the difference, not that of disyllables. Nevertheless, morpheme identification is expected to be much more difficult for listeners than for readers, perhaps delaying the time morphemes become activated during compound processing.

In addition to its unusual characteristics, Chinese also shares many morphological properties with other languages. These include right-headed nominal compounds, a gradient distinction between transparent and opaque compounds, and compounds built hierarchically from other compounds (e.g., *xiguazhi* 'watermelon juice'). With regard to these, we expect Chinese to behave essentially the same as has been found with other languages.

The following review is thus organized into four major subsections. First we review evidence relating to the psychological reality of the word in Chinese compound processing. Next we describe studies investigating the activation of morphemes, looking at evidence from distributional patterns, semantic transparency, and aphasia. Finally we examine the role of compound structure in processing. These themes fit neatly into those discussed elsewhere in this book, but as we will see, there will be some surprising twists along the way.

Evidence for the Reality of a Word Level

The vagueness of the concept 'word' in Chinese has practical consequences for psycholinguistic methodology. As pointed out by Hung, Tzeng, & C.-Y. Ho (1999), the use of that old workhorse, the lexical decision task, is problematic in Chinese, since we can't expect naive participants to give consistent *ci* judgments when they aren't sure what a *ci* is. Nevertheless, results from lexical decision tasks converge with those from other tasks on the same conclusion: Although characters do have a quasi-wordlike status for Chinese readers, compounds in Chinese are indeed treated as lexical units at some level of processing, just as they are in other languages.

The first relevant piece of evidence comes from comparing *ci* judgments with *zi* judgments. Taft (2003) describes an experiment in which readers were asked to make lexical decisions about characters presented in isolation, some free and some bound. The participants who were asked to perform a character/pseudocharacter decision task (i.e., make *zi* judgments) showed no differences between free and bound characters, but those who had to make *ci* judgments on the same one-character items correctly gave more 'yes' responses to free than bound morphemes. Nevertheless, participants in the word-decision task still found it harder to reject bound but nonbinding characters than binding characters, although neither are free words. While these results confirm that Chinese readers do have a tendency to view characters as wordlike, the fact that boundedness affects *ci* judgments at all, while not affecting *zi* judgments, shows that the *ci* vs. *zi* distinction is a real factor in the judgment process.

The word superiority paradigm provides even stronger evidence for the reality of Chinese word units. In a word superiority effect, components are more readily recognized or identified in real lexical items than in nonwords (e.g., letters in English words; Reicher, 1969). Studies on Chinese have consistently found that lexical status at the *ci* level affects recognition at the *zi* level. Thus in separate studies, Cheng (1981) and Mattingly and Xu (1994) showed participants two-character strings that were either real words or pseudowords (two real characters in a nonsense string), and asked them to detect a given character. Both studies found a word superiority effect: participants performed better with real words. The effect appears to be rather automatic; J.-Y. Chen (1999) found a word superiority effect even when the characters had to be detected in sentences, regardless of sentential context. Hung, et al. (1999) were concerned that character detection is not identical to letter detection in English since characters themselves are meaningful, and thus they had participants identify character components (e.g., radicals) rather than whole characters. Even with this variation in methodology, a word superiority effect was found. I.-M. Liu (1988) also found a word superiority effect in the naming of characters in two-character strings, but no such effect for characters appearing in second and third position of three-character strings. This may possibly relate to the predominance of two-character words in Chinese; the default parsing strategy for three-character strings may be to break them up, and in fact, Yin, Derwing, and Libben (2004) have recently found evidence for just such a parsing strategy (discussed below).

Even when we turn to the *ci* decision studies that form the bulk of the research on morphological processing in Chinese, we find that they also provide evidence for the reality of the lexical status of the Chinese word: In all studies where word frequency is an experimental variable, it always has a facilitative effect on lexical decision time, separately from character, syllable, or morpheme frequency. We will note examples of this explicitly only when relevant, since word frequency effects are just as ubiquitous in Chinese as they are in other languages.

Character Frequency

Researchers interested in morphology can hardly be satisfied with the discovery that Chinese has words; they want to know about morphemes. In this section we consider one commonly recognized diagnostic of the activation of compound components arising from a distributional property: morpheme frequency effects. It is very difficult to calculate morpheme frequency in Chinese, however. Clearly *xing* in *xingren* (pedestrian) and *hang* in *yinhang* (bank) should be considered distinct morphemes despite being written with the same character, but it is less clear what to do with characters showing no change in pronunciation across contexts and subtler semantic shifts: the character *kuai* means 'happy' in *kuaile* (happy) but 'fast' in *kuaisu* (speed), *jia* means 'family' in *jiaren* (family member) but 'house' in *jiashi* (housework), and so on. In practice, therefore, in all studies manipulating the frequencies of compound components, what is actually manipulated is character frequency, and we will make this explicit in our summaries.

Taft, J.-T. Huang, and Zhu (1994) is a typical study. They gave participants a visual lexical decision task with two-character words varying in the frequency of the first and second characters, but with word frequency controlled. Similar to some studies in other languages (e.g., Taft & Forster, 1976), compounds with two high-frequency characters (HH, e.g., *chengguan* 'near the city gate,' literally 'city-close') were responded to more quickly than words containing a low-frequency character (HL, e.g., *jiajuan* 'wife and children,' literally 'family-concern'; LH, e.g., *anshu* 'eucalyptus,' literally 'eucalyptus-tree'). This facilitative (positive) effect of character frequency implies that word recognition does involve access of the component characters, as one would expect of Chinese reading. Interestingly, however, the effect was not completely consistent: LL words (e.g., *chouchang* 'disconsolate,' literally 'regretful-disappointed') were not accessed more slowly than HH words. Taft, et al. (1994) suspected that the binding words in their set of LL words were responsible (e.g., *meigui* 'rose'), because their orthographic cohesiveness may have given them an advantage in ease of access. We return to this suggestion, and how it may be understood more formally, in the section after this.

The role of character frequency in spoken word access is, as one would expect, quite different. Zhou and Marslen-Wilson (1994), also mentioned in Jarema (this volume), performed a complex set of nonprimed lexical decision experiments on spoken semantically transparent disyllabic Mandarin compounds, systematically varying syllable frequency, morpheme (character) frequency, and word frequency. Only one robust finding emerged: positive word frequency effects. There was no effect of character frequency, though negative syllable frequency effects were found for the first syllable of nonwords and for that of real words when word and morpheme frequency were controlled, suggesting cohort competition. As noted by Packard (1999) in commentary on this study, the results seem to imply that the lexical representations for spoken compounds are processed at some point as morphological wholes, though other paradigms, described below, have also provided evidence for morpheme activation in spoken compounds (e.g., Zhou & Marslen-Wilson, 1995).

Transition Probability and Family Size

Recall that Taft, et al. (1994) found, to their surprise, that written words composed of high-frequency characters were recognized no faster than words composed of low-frequency characters. While they specifically blamed binding words for these results, they also made a more general observation. Namely, for low-frequency initial characters, there are fewer character types that may follow it than for high-frequency initial characters, as they confirmed with an offline guess-the-next-character task. This higher transition probability for low-frequency characters means greater predictability of one character from the other. High character frequency thus exerts two conflicting forces on word recognition, both easing the activation of words and making it harder to narrow the choice down to just one of these words. Transition probability is related to family size (Schreuder & Baayen, 1997), which counts the number of word types containing a given morpheme; words containing morphemes from larger families will necessarily have lower morpheme transition probabilities, and vice versa. No matter how it is modeled, binding words will be predicted to be easier to access than nonbinding words; binding characters only appear in one word, making the character transition probability higher and character family size smaller. Indeed, like Taft, et al. (1994), experiments by Lü (1996), P. Tsai (1998), and H.-C. S. Hsiao (2004) have all found that binding words are recognized more quickly than nonbinding words, even fully opaque ones.

The cohesiveness of binding words can also be demonstrated in tasks involving isolated characters, as in Taft and Zhu (1995). A timed naming task found that first-position binding characters (e.g., *qiu* of *qiuyin* 'earthworm') were named faster than second-position binding characters (e.g., *yin*), suggesting that the pronunciation of binding words must be accessed as a whole. However, a timed real/pseudo-character decision task found no difference in reaction times due to binding character position, nor was there a positional effect in the naming of nonbinding but bound characters that always appear in the same position in different words (e.g., *xun* 'die for a cause' only appears in first position, *lü* 'companion' only appears in second position). Thus it appears that only some properties of binding characters, such as their pronunciation, are stored as part of a whole word, whereas all properties of nonbinding characters are stored independently. Mattingly and P. Hsiao (1997) confirmed this latter conclusion, finding no differences between bound (but not binding) characters and free characters in both an offline radical identification task and a timed real/pseudo-character decision task. However, Taft and Zhu (1997) failed to find a difference between two-character binding and nonbinding words in a primed naming task; for both types of words, first-character naming latency was sped up by prior presentation of the second character, which seems to imply that even the pronunciations of nonbinding words may be stored as wholes.

When the transition probability or family size are manipulated in written nonbinding words, the above effects are not found, and instead the more familiar positive character frequency effects return. In a study that investigated this systematically, Lü (1996) pitted words like *qiushi* (embarrassing thing, literally 'dry ration - thing') against words like *shanshi* (good deeds, literally 'good - thing'); both words are matched in frequency and share the same second character, but the first character in *qiushi* appears virtually only in this one word, while the first character in *shanshi* also appears in several other words. Lü found that words with larger families, like *shanshi*, were recognized more quickly, and the same effect (i.e., positive character frequency) also appeared when the manipulated character was in second position.

If negative character frequency effects occur under certain circumstances in reading, what do we expect with spoken compounds? Since I.-M. Liu, J.-T. Wu, and Chou (1996) have found that for readers, character frequency has its greatest effect on the early stages of processing, we don't expect a direct effect of character frequency itself in spoken word access. Thus for listeners the positive effects of high transition probability and small family size may be revealed more clearly. Myers and Gong (2002) examined this possibility in a small-scale partial replication of Zhou and Marslen-Wilson (1994). Transparent disyllabic Mandarin compounds were controlled for syllable frequency and cohort size; word frequency and character frequency were varied separately. Participants in a nonprimed auditory lexical decision experiment showed the usual positive word frequency effects, but also significant character frequency effects (the authors claim that the null results of Zhou & Marslen-Wilson, 1994, were due to their use of a smaller frequency range). In contrast to what is found with written transparent compounds, the character frequency effects were negative: Words with higher-frequency characters (e.g., *benneng* 'instinct') were responded to more slowly than words with lower-frequency characters (e.g., *banzou* 'accompaniment').

Surprisingly, precisely the same result is found even when the characters aren't morphemic at all. Reconfirming the earlier findings of P. Tsai (1998), Myers and Lai (2002) found in an auditory lexical decision task that (monomorphemic) binding words like *putao* (grape) were recognized more quickly than monomorphemic nonbinding words like *shafa* (easy chair), matched for word frequency, syllable frequency, syllable transition probability, first-syllable cohort size, and acoustic duration. The hypothesis that participants activated orthographic representations of the words and made their decisions on these was ruled out by the lack of an effect when the same items were presented visually (thus failing to replicate the findings of Lü, 1996). Thus again it appears that in processing spoken compounds, transition probability or family size plays a role without character frequency affecting responses directly. Clearly more work needs to be done to reconfirm this effect and explore its possible causes, but in the meantime it presents a cautionary message: Even when working with spoken words, be aware of the possible interference of orthography.

Semantic Transparency

Transparency is in principle independent of distributional properties, but it may covary inversely with them, since transparency is related to productivity and hence to lower transition probability and larger families. This leads to conflicting forces again. As we saw, compounds may be recognized faster if they are more 'cohesive' or have smaller families, but such compounds tend to be semantically more opaque (as is certainly the case for binding words). By contrast, if characters are automatically activated during reading, opacity should hurt word access (as pointed out by Libben, this volume), since activation of semantically irrelevant components will result in competition with the word. Perhaps given such conflicts, fully consistent effects of semantic transparency on compound access have proven difficult to demonstrate; for example, Jarema, Busson, Nikolova, Tsapkini, and Libben (1999) found that transparent compounds were recognized more quickly than opaque compounds in French, but more slowly in Bulgarian. The delicate nature of semantic effects can be seen in Chinese studies as well. Some (e.g., C.-H. Tsai, 1994; Myers, Derwing, & Libben, 2004) find that semantically transparent compounds like *huayuan* (garden, literally 'flower-yard') are responded to more quickly than semantically idiosyncratic or opaque compounds like *huasheng* (peanut, literally 'flower-birth'), while others find the reverse pattern (e.g., Su, 1998; Liang, 1992; C. Y. Lee, 1995), or no difference at all (S.-T. Chen, 1993; Myers, Libben, & Derwing, 2004).

In an attempt to clarify the role of semantic transparency, some studies also take component frequency into account. Peng, Y. Liu, and C. Wang (1999) first held semantic transparency constant and varied word and character frequency in a visual lexical decision experiment, and like Taft, et al. (1994), found a positive frequency effect both for words and for characters. When they held word frequency constant and crossed character frequency with semantic transparency, however, character frequency effects were found to depend on transparency: for transparent words the character frequency effect was positive, but for opaque words the effect was negative (i.e., higher character frequency meant slower word responses). Peng et al. (1999) suggest that activation at the compound level was inhibited for opaque compounds due to activation of the competing semantics of the component characters. This certainly makes sense, but when Liang (1992) and C. Y. Lee (1995) performed similar experiments, they found only positive character frequency effects for transparent compounds; there was no effect for opaque compounds.

A variation on the usual visual lexical decision task may help shed more light. S.-T. Chen (1993), C.-H. Tsai (1994), and C. Y. Lee (1995) varied stimulus-onset asynchrony (SOA) between the appearance of the first and second character of a two-character compound. While the details of their results varied, the general finding was that character frequency effects depending both on transparency (positive for transparent compounds, negative or absent for opaque compounds) and on SOA; at the longest SOAs (200-600 ms), S.-T. Chen (1993) found negative character frequency effects regardless of transparency. The effect of SOA implies that at least in this paradigm, negative character frequency effects can be caused by competition between the whole word and the component characters, consistent with Libben's proposal.

Myers, Libben, and Derwing (2004) addressed the nature of semantic transparency by using another novel variant on the visual lexical decision task. Opaque and transparent compounds were presented in separate blocks, with block order counterbalanced across participants. When transparent compounds were presented first, reaction times were the same for both types of compounds, but when opaque compounds were presented first, reaction times for the transparent compounds dropped off significantly over the course of the block. At the very least, this suggests that activation of character semantics is not obligatory, since otherwise such a context-dependent shift would not have been found: transparent compounds should always have been easier to recognize.

Priming

If the component morphemes of a complex word are activated during lexical access, the morphemes should be capable of priming or being primed. However, one must be cautious when designing and interpreting morphological studies (as often emphasized in the literature, e.g., Zhou & Marslen-Wilson, 2000), because the priming may actually be due to form similarity (e.g., a character in the prime may prime itself in the target) or word-level semantics (particularly if the prime and/or target are transparent). There are a number of techniques that can be used to deal with these problems, but overall the morphological priming literature is not altogether satisfying; too many poorly understood variables are in play at the same time. For Chinese, one of the most important is modality, and so this is how we arrange the review, looking first at visual-visual priming, then auditory-auditory priming, and finally cross-modal priming.

Visual-visual priming is the most commonly used in Chinese, and in almost all such studies, the primes and targets are both two-character strings. One of the experiments in Peng et al. (1999) employed a paradigm of this type, where the manipulation involved whether the first character of prime and target were identical or entirely unrelated. In order to ensure that any priming effects could not be due to whole-word semantic priming, the word meanings of primes and targets were unrelated. Note, however, that this requirement means that in the identical-character condition, the character could not contribute the same meaning to prime and target. This was the case both for transparent primes (e.g., prime *anning* 'quiet,' literally 'peace-peace,' begins with the same character as target *anzhuang* 'install') and opaque primes (e.g., prime *kuaihuo* 'happy,' literally 'happy/fast-life,' begins with the same character as target *kuaisu* 'speed'). Facilitative priming was found, but it was not a simple case of form priming, since the effect was found only with transparent primes. This is consistent with the hypothesis that readers activate the components of transparent compounds, but it doesn't seem quite right to say that what was activated here were morphemes, due to the meaning mismatch across prime and target. As Taft, et al. (1999) recognize in their analysis of these results, we must acknowledge a role for characters even in morphological processing.

One technique for distinguishing form priming from morpheme priming is to use a paradigm in which a prime component is related to the meaning of the whole target compound, so that prime and target do not share any characters. In an unpublished conference paper (Y. Liu & Peng, 1995) summarized in Taft, Y. Liu, and Zhu (1999), an experiment of this type is described, and again effects were only found for semantically transparent primes, not opaque primes. Again the experimenters were careful to make sure that their results couldn't be due to word-level semantic priming, since semantic primes, as wholes, were not related to their targets (e.g., transparent *beipan* 'betray,' literally 'back-rebel,' primed *fugai* 'cover,' but opaque *mashang* 'immediately,' literally 'horse-on,' did not prime *mianyang* 'sheep').

Y. Liu and Peng (1997) found another way to circumvent possible objections, using two-character primes that were opaque (e.g., *caoshuai* 'sloppy,' literally 'grass-command'), but paired with semantically two-character targets to which they were either related as wholes (e.g., *mahu* 'careless,' literally 'horse-tiger,' related to *caoshuai*), only via the prime's first character (e.g., *shumu* 'tree,' literally 'tree-wood,' related to *cao*), or only via the prime's second character (e.g., *lingdao* 'lead,' literally 'lead-guide,' related to *shuai*). At the shortest SOA (43 ms), priming only occurred in the whole-word condition, as might be expected of semantically opaque primes. However, at a longer SOA (143 ms), priming was found in all three conditions, revealing activation of whole words and both of their constituent morphemes. Another experiment compared transparent and opaque primes with an intermediate SOA of 86 ms, and morpheme priming was found only with transparent compounds. The morphemes in opaque compounds thus don't reveal their activation until late.

Zhou, Marslen-Wilson, Taft, and Shu (1999) further examined the time course of visual compound processing in a complex series of primed visual lexical decision experiments, using two-character primes and targets (apparently at least somewhat transparent) put into two SOA conditions (57 ms, 200 ms) or masked priming. For each target (e.g., *huagui* 'luxurious,' literally 'splendid-valuable'), the prime either shared the same morpheme written with the same character (e.g., *huali* 'magnificent,' literally 'splendid-beautiful'), shared a character used with a different meaning (e.g., *huaqiao* 'overseas Chinese,' literally 'China-bridge'), or shared a homophone (including same tone) written with different characters (e.g., *huaxiang* 'glide,' literally 'slide-soar'). Importantly, primes and targets were pretested to ensure that they were matched in degree of whole-word semantic relatedness across prime types. The positions of the key characters were also varied (both first character, both second character, or second in prime and first in target). The results showed that the morpheme priming effect (in ms) was consistently greater than character priming, and there was no homophone priming at all. The position of the key characters did not make much difference, except that masked morpheme priming was markedly reduced if the position differed across prime and target. These results are again consistent with previous evidence for component activation when reading Chinese compounds. However, the fact that spatial overlap was relevant for masked morpheme priming and that mere character matching also caused some priming, even at the late SOA of 200 ms (in the second-position and mixed-position conditions), shows again that the character also plays a crucial role in Chinese compound processing. This observation is strengthened by another experiment, which found facilitated lexical decisions for monomorphemic nonbinding words (e.g., *shafa* 'easy chair') from masked transparent primes beginning with the same character (e.g., shatan 'sandy beach').

Finally, Zhou et al. (1999) compared morpheme priming with word-level semantic priming head on, using for each target (e.g., *yisheng* 'doctor,' where *yi* means 'cure') either prime sharing the same first morpheme (e.g., *yiyuan* 'hospital') or a prime matched for word-level semantic relationship but without any morphological, orthographic, or phonological overlap at the character level (e.g., *hushi* 'nurse'). With masked primes, morphological priming added a facilitative effect on top of that for whole-word priming. However, this effect has proven difficult to replicate (S.-C. Chen, 2002).

Yet another way to deal with concerns over word-level semantic confounds is to use a long-term priming paradigm, in which earlier items in a list of single-stimulus trials act as primes for later items. Research on other languages has suggested that semantic priming disappears in such a paradigm, while morphological priming does not (see Feldman, 2003, for review). The only Chinese study employing this paradigm so far (Zhou & Marslen-Wilson, 1995) used spoken stimuli. Recall that the homophony of Chinese morphemes makes it unlikely that spoken compounds are activated via prior activation of their morphemes (Packard, 1999), and that Zhou and Marslen-Wilson (1994) had failed to find character frequency effects in auditory lexical decision. The negative character frequency effects found by Myers and Gong (2002) and Myers and Lai (2002), however, suggest that characters (or morphemes) are indeed activated at some later stage. Thus long-term priming should be a particularly useful way of revealing morpheme activation in spoken compounds.

Zhou and Marslen-Wilson (1995) ran a complex set of long-term priming experiments with lexical decision tasks, using auditory Mandarin disyllabic primes and targets. The design was the basis of Zhou, et al. (1999), just described, and so it involved the same four prime-target relations (identical, morpheme, character, homophone) and the same three positions of the relevant characters in primes and targets (both first, both second, second in prime and first in target). SOA was also varied, but across a larger time scale, with immediate priming, a short lag (one or two intervening items), a long lag (40 or more intervening items), and a medium lag somewhere between. Results showed similar patterns for both character and homophone primes, presumably because character primes were also homophones. These form priming effects were weak, variable, and sometimes inhibitory (in first position, implying cohort competition), but morpheme priming was always positive. Its effects also lasted longer than character and homophone priming, suggesting both that it differed from them and that it could not be due to semantic relatedness. Thus the morphemes in spoken compounds do appear to become activated in the course of lexical access.

One final technique for ruling out form-based explanations for morpheme priming effects is to use primes and targets of different modalities (e.g., Marslen-Wilson, Tyler, Waksler, & Older, 1994). For some reason, however, in Chinese this method only seems to have been used in two unpublished master's theses. In one of the experiments in C. Y. Lee (1995), auditory disyllabic Mandarin primes were used with two-character visual targets, revealing facilitatory priming when the prime was transparent and it shared an initial morpheme (same character, meaning, and pronunciation) with the target (e.g., prime *caiyi* 'suspicion,' target *caimi* 'solve riddles,' where *cai* means 'guess'). The cross-modal primed lexical decision experiment in P. Tsai (1998) used auditory disyllabic primes (e.g., *shudian* 'bookstore') and single character targets representing either the same first character (e.g., *shu* 'book') or a homophone of it (e.g., *shu* 'sparse'). Homophone priming was not found, but facilitatory character priming was found not only with transparent primes (e.g., *shudian*), but also with opaque (e.g., *huasheng* 'peanut,' literally 'flower-birth') and monomorphemic nonbinding primes (e.g., *jita* 'guitar,' literally 'lucky-third.person'). Further work seems to be needed to explore the cross-modal paradigm, or indeed priming paradigms generally, to accommodate the many conflicting findings within a single coherent model.

Aphasiological Evidence

We end this review of evidence for component activation in Chinese compound processing with intriguing, if somewhat controversial, evidence from aphasia. Morphologists are particularly interested in errors by aphasics that reveal knowledge of word-internal structure (see Semenza and Mondini, this volume). For example, it is found that Italian Broca's aphasics, who tend to have difficulty with verbs as free words, also have difficulty producing verbal morphemes in nominal compounds (Semenza, Luzzatti, & Carabelli, 1997). Precisely the same pattern has been reported in (Mandarin) Chinese by Bates, S. Chen, Tzeng, P. Li, and Opie (1991) and S. Chen and Bates (1998). Just as in Italian, picture-naming tasks reveal that Chinese Broca's aphasics have greater difficulty correctly producing verbs than nouns, and just as in Italian, this difficulty is also found with the verbal morpheme in VN compounds, such as *chifan* (eat, literally 'eat-rice'). Even more interesting, there appears to be a complementary pattern of errors for Wernicke's aphasics, who both make more errors when producing free nouns and when producing the nominal morpheme in such VN compounds.

The problem in interpreting such results, however, as Zhou, Ostrin, and Tyler (1993) point out in a critique of Bates, et al. (1991), is that it is hard to be certain that it is truly a sublexical phenomenon, since in a VN construction the N is the object of the V, just as in a standard Chinese verb phrase. Bates, et al. (1993) admit that they don't have a very good response given the stimuli that they used. S. Chen and Bates (1998) provide new data from a study that crossed syntactic category at both the word and morpheme level. Their test items thus included not only verbal VN compounds like *chifan,* but also nominal VN compounds like *feiji* (airplane, literally 'fly-machine'). The results for verbal VN compounds were replicated, but for the morphemes in nominal VN compounds, both Broca's and Wernicke's aphasics made more errors with the verbal morpheme. They ascribed this unexpected symmetry to the atypicality of nominal VN compounds, which unlike the situation in Italian, are marked in Chinese compared with nominal NN compounds. C.-L. Lee, et al. (in press) have recently provided additional evidence showing that aphasics do indeed retain awareness of the differences in typicality across compound types.

The Role of Compound Structure

The discussion of the differential processing of verbal and nominal VN compounds raises the more general question of the processing effect of compound structure. This question has been little investigated so far in Chinese, just as is unfortunately the case for other languages (see Libben, this volume), but relevant studies do exist.

The earliest study to look at compound structure in terms of headedness appears to be Zhang and Peng (1992). In a series of nonprimed visual lexical decision experiments, participants were presented with nominal coordinative compounds (e.g., *fuxiong* 'father and elder brothers') and nominal modifier-noun compounds (e.g., *muxiao* 'alma mater,' literally 'mother-school'). For coordinative compounds, positive character frequency effects were found for both positions, indicating equal importance, whereas for modifier-noun compounds, a positive character frequency effect was only found for the first position. In a later study, Zhang (1997) used a primed (visual-visual) lexical decision task to examine modifier-noun and coordinative compounds, with two-character primes and targets that either matched in the first or second character. Priming effects were reversed for the two structures: coordinative prime-target pairs resulted in faster responses if the second character matched, while modifier-noun prime-target pairs resulted in faster responses if the first character matched. This result partly fits the findings of Zhang and Peng (1992), in that there appeared to be activation of the first morpheme in modifier-noun compounds and the second morpheme in coordinative compounds.

However, the activation of the nonhead (first position morpheme) in modifier-noun compounds seems to conflict with the suggestions made by Libben (this volume), according to which the head (the second morpheme) should be more prominent. One possible way of accommodating the results is suggested by a nonprimed visual lexical decision task experiment described in Zhang (1997) that found that coordinative compounds were responded to more slowly than modifier-noun compounds of matching word frequency. This makes sense, given that coordinative compounds, though productive, appear to be marked in Chinese relative to modifier-noun compounds. Perhaps, then, access of the more typical type of nominal compound (modifier-noun) involves left-to-right processing, producing more prominent activation of the first component, while coordinative compounds require more active processing of both components. Before we speculate too far, however, it should be noted that speed-accuracy trade-offs in the Zhang and Peng (1992) study have made their results controversial, and they were not replicated by Taft, et al. (1994) in a reanalysis of their own results.

Recently the head/nonhead issue has been investigated in a series of visual-visual primed lexical decision experiments by Ji and Gagné (2004) that used modifier-noun compounds as both primes and targets and varied the semantic relationship between modifiers and between heads: modifier same (e.g., prime *shujia* 'bookcase' for target *shudian* 'bookstore'), modifier different (e.g., prime *shuhao* 'book number'), head same (e.g., prime *bingdian* 'cookie store'), or head different (e.g., prime *ridian* 'day store'). Faster responses were found for prime-target pairs with the same relation, both for modifiers and for heads. However, the advantage for having the same modifier relation was lost when target heads appeared 350 ms before the whole compound was shown, whereas the advantage for having the same head relation was not lost when target modifiers were delayed. These results can be interpreted as supporting a stronger role for heads in the processing of modifier-noun compounds.

Other types of morphological structures were investigated by J. C.-F. Hsu, Tzeng, and Hung (in press), who examined the role of morpheme syntactic category in compound processing in unimpaired adults. Asked to judge the syntactic category of nominal and verbal compounds containing nominal and verbal components, readers showed faster response times for nominal NN compounds than for nominal VV compounds, and the opposite was found with verbal compounds. Even if this pattern is due less to conflicts between the component and word level as they claim than to typicality effects (NN compounds are usually nominal, VV compounds usually verbal), it still suggests that morphological structure is available during compound processing.

The simplest aspect of compound structure is bare morpheme order. Thus we would expect compound components to be processed left to right, at least in the initial stages, and this is just what the evidence seems to show for a variety of European languages (see Jarema, this volume). The same seems to hold of Chinese. We have already seen the special status accorded the first morpheme in modifier-noun compounds (Zhang & Peng, 1992; Zhang, 1997), and other studies provide further support. For example, using a task in which participants were asked to freely associate to compound targets, J.-T. Huang (1979) found that participants tended to begin by listing words that shared the first character with the target, and only later moved to the second. Mattingly and Xu (1994) found that, that first-position characters showed a greater (positive) effect on character detection times in two-character pseudowords. However, there was no frequency effect for real words, which they interpreted as implying that characters in real words are processed in parallel (though of course it's always risky to make much of null results). Recently, Wong and H.-C. Chen (2002) employed eyetracking to study the influence of the semantic ambiguity of compound components on reading. They found longer eye fixations on the second character for compounds beginning with ambiguous characters used with a nondominant meaning (e.g., *bangzhu* 'gang master,' where *bang* usually means 'help,' not 'gang') than for those whose first characters were unambiguous (e.g., *wuzhu* 'house master').

The left-to-right nature of compound parsing is made particularly clear if one examines compounds longer those that have attracted virtually all of the field's attention. Yin, Derwing, and Libben (2004) recently did just this, using both a variety of tasks requiring readers to parse three-character compounds, both lexical and novel. It was found that even for monomorphemic nonbinding words (e.g., the phonetic loan *maikefeng* 'microphone'), which presumably have an entirely flat structure, the preferred parse was left-branching, that is [[XY]Z]. This preference is also found in Dutch and German (Baayen, Krott, Dressler, Jarema, & Libben, 2002), suggesting that it may be due to a universal parsing strategy related to the 'late closure' strategy of syntax (Frazier, 1987).

Myers, Derwing, and Libben (2004) and Myers, Libben, and Derwing (2004) tried to pin down the time course of position-in-the-string effects by using somewhat unusual variants on the visual lexical decision task. Following Libben, Gibson, Yoon, and Sandra (2003), they reasoned that position-dependent processing may be revealed by varying the location of semantic transparency in a compound. Thus their compound materials were pretested as being transparent in both positions (TT, e.g., *baise* 'white,' literally 'white-color'), opaque in both (OO, e.g., *shenjing* 'nerve', literally 'god-ways'), or a mixture (OT, e.g., *huoche* 'train,' literally 'fire-vehicle'; TO, e.g., *shiguang* 'time,' literally 'time-bright'). In both studies, OO and OT compounds were recognized the most slowly (precisely this same pattern has been found independently by H.-C. S. Hsiao, 2004, using a different paradigm and different materials). This not only provides more evidence for the special role of the first position, but also supports the explanation proposed by Libben (this volume) for transparency effects as being due to inhibition by opaque components.

Myers, et al. also reasoned that if position-in-the-string effects are due to early processing stages, the first-component effect should tend to play a more important role early on in access. In one experiment in Myers, Derwing, and Libben (2004) examining this hypothesis, isolated characters were used as primes and compounds containing these characters (in first or second positions) as targets, and in a complementary experiment, primes were compounds and targets were isolated characters. The expectation was that the compound processing stages probed in each experiment would relate to the length of time between display of the compound and onset of response (i.e., earlier stages for character-compound priming, later stages for compound-character priming). This predicted that responses to compound targets should be facilitated by first character primes, while responses to character targets wouldn't be affected by character position in the compound primes. This is precisely what was found. Semantic transparency of the first component was also found to play an early role: in character-compound priming, TO compounds showed the greatest sensitivity to character position, being responded to much faster if its first character was prime, while for compound-character priming there were no prime type by target type interactions. Myers, Libben, and Derwing (2004) probed time course in a different way, by placing the prime within the compound targets themselves. That is, the first or second character, normally black, briefly (50 ms) flashed bright red, either early (0 ms SOA) or late (200 ms SOA) during presentation of the four types of compounds. Again, when the flashing occurred early, the effect was greatest for first-position characters, and there was an interaction with transparency: opaque-initial compounds (OO, OT) were slowed, while others (TO, TT) were relatively unaffected.

Conclusions

A number of general conclusions emerge from this survey of Chinese compound processing. The most fundamental of these is the difficulty of defining the Chinese word, a problem closely related to the nature of Chinese orthography. Researchers who come to Chinese merely to fit data from yet another language into a universal theory of morphological processing should take this issue seriously. In fact, the experiments with visual stimuli make it quite clear that characters have a great influence on the reading of words, and there is some evidence that they influence the access of spoken words as well. In spite of this, however, we can also safely conclude that words are also treated as units at some point during lexical access in Chinese.

Another safe conclusion is that the degree to which the components of a compound are activated, and when, depends greatly on properties such as semantic transparency, modality (spoken versus written), and distributional properties of the component characters (e.g., their mutual predictability and whether or not they are binding characters). The precise way in which these factors influence word access is not always clear, but in general one can say that activation of characters is essentially obligatory in the reading of compounds, and that activation of these components occurs from left to right. When accessing spoken compounds, however, rampant homophony greatly affects the role of compound components. None of these findings are particularly unusual when seen in a cross-linguistic perspective, giving us increased confidence in universal principles of compound processing.

Nevertheless, the firmest conclusion that can be drawn from this survey is that much work remains to be done. In particular, the various possible explanations for negative component frequency effects (e.g., those based on transition probability vs. those based on family size) need to be sorted out, priming effects require further study, and the special challenges posed by spoken Chinese deserve far more attention than they have hitherto received. More research is also needed on compound processing by Chinese aphasics (see Packard, 1993), the acquisition of compounds (see Hsieh, 1989, for Mandarin, and Tsay, Myers, & X.-J. Chen, 1999, for another Sinitic language), and the production of compounds by nonimpaired adults (e.g., T.-M. Chen & J.-Y. Chen, 2003, make the surprising claim that morphemes play at best a minimal role in spoken compound production).

It is also unknown how the processing of compounds fits into a more general picture of morphological processing. In particular, are there separate processing modes for compounding and affixation? The first study to look at this question in Chinese (W. Wang & Myers, 2004) employed two diagnostics: first-morpheme frequency effects should be less context sensitive for compounds than for suffixed words (Andrews, 1986), and compounds sharing the first morpheme should prime each other cross-modally while suffixed words should not (Marslen-Wilson, et al., 1994). Results did point in the direction of there being two distinct processing modes for compounding and affixation, even in Chinese, with its fuzzy boundaries between morphological operations. However, as there has been little attention to this question in the morphological processing literature in general, we can't be sure that the right diagnostics were chosen, or even that reliable diagnostics truly exist.

This chapter, though long, is thus far from complete; countless new studies on Chinese compound processing still wait to be done.

References

Anderson, S. R. (1992). *A-morphous morphology*. Cambridge: Cambridge University Press.

Andrews, S. (1986). Morphological influences on lexical access: Lexical or nonlexical effects? *Journal of Memory and Language, 25*, 726-740.

Baayen, H., Krott, A., Dressler, W.U., Jarema, G. & Libben, G. (2002, October). *Interfixes as boundary markers in compounds*. Paper presented at the Third International Conference on the Mental Lexicon, Banff, Alberta, Canada.

Bates, E., Chen, S., Li, P., Opie, M., & Tzeng, O. (1993). Where is the boundary between compounds and phrases in Chinese? A reply to Zhou et al. *Brain and Language, 45*, 94-107.

Bates, E., Chen, S., Tzeng, O., Li, P., & Opie, M. (1991). The noun-verb problem in Chinese aphasia. *Brain and Language, 41*, 203-233.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart and Winston.

Chao, Y.-R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.

Chen, J.-Y. (1999). Word recognition during the reading of Chinese sentences: evidence from studying the word superiority effect. In J. Wang, A. W. Inhoff & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 239-256). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chen, K.-J., Huang, C.-R., Chang, L.-P., & Hsu, H.-L. (1996). SINICA CORPUS: Design methodology for balanced corpora. *Language, Information and Computation, 11*, 167-176.

Chen, S., & Bates, E. (1998). The dissociation between nouns and verbs in Broca's and Wernicke's aphasia: Findings from Chinese. *Aphasiology, 12,* 5-36.

Chen, S.-C. (2002, December). *Morphological processing and semantic processing in reading Chinese compound words*. Paper presented at the Tenth International Conference on the Cognitive Processing of Chinese and Other Related Asian Languages, Taipei.

Chen, S.-T. (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, T.-M., & Chen, J.-Y. (2003, August). Morphological processing in the production of Chinese compound words. The 9th Annual Conference on the Architectures and Mechanisms for Language Processing, Glasgow, Scotland.

Cheng, C.-M. (1981). Hanzi renzhi de licheng. [Perception of Chinese characters.] *Acta Psychologica Taiwanica, 23*, 137-153.

Dai, J. X.-L. (1992). Chinese morphology and its interface with the syntax (Doctoral dissertation, Ohio State University, 1992). *Dissertation Abstracts International, 53* (08), 2792.

Duanmu, S. (1998). Wordhood in Chinese. In J. Packard (Ed.), *New approaches to Chinese word formation: Morphology, phonology and the lexicon in Modern and Ancient Chinese* (pp. 135-196). Berlin: Mouton de Gruyter.

Feldman, L. B. (2003). What the repetition priming methodology reveals about morphological aspects of word recognition. In J. S. Bowers & C. J. Marsolek (Eds.), *Rethinking implicit memory* (pp. 124-138). Oxford: Oxford University Press.

Frazier, L. (1987). Sentence processing: A tutorial review. In M. Coltheart (Ed.) *Attention and performance, Vol. XII: The psychology of reading* (pp. 559-586). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Fromkin, V., & Rodman, R. (1998). *An introduction to language* (6th ed.). New York: Harcourt Brace College Publishers.

Ho, K.-C. (1976). A study of the relative frequency distribution of syllabic components in Mandarin Chinese. *Journal of the Institute of Chinese Studies, 8,* 275-352.

Hoosain, R. (1992). Psychological reality of the word in Chinese. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 111-130). Amsterdam: North-Holland.

Hsiao, H.-C. S. (2004, June-July). The role of semantic transparency in the processnig of Mandarin compounds. Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Hsieh, M.-L. (1989). The acquisition and comprehension of lexical entries by Mandarin-speaking children. Unpublished master's thesis, Fu Jen Catholic University, Taipei, Taiwan.

Hsu, J. C.-F., Tzeng, O. J.-L., & Hung, D. L. (in press). Syntactic effects on constituent components in word recognition: Evidence from compound nouns and compound verbs in Chinese. *Maryland Working Papers in Linguistics*.

Huang, J.-T. (1979). Zhongwen ci lianxiang de yishixing ke fenli jiashe [Time-dependent separability hypothesis of Chinese word association]. *Acta Psychologica Taiwanica, 21*, 41-48.

Hung, D. L., Tzeng, O. J. L. & Ho, C.-Y. (1999). Word superiority effect in the visual processing of Chinese. In O. J. L. Tzeng (Ed.) *Journal of Chinese Linguistics Monograph Series No. 13: The biological bases of language*, 61-95.

Jarema, G., Busson, C., Nikolova, R., Tsapkini, K., & Libben, G. (1999). Processing compounds: A cross-linguistic study. *Brain & Language, 68*, 362-369.

Ji, H., & Gagné, C. L. (2004, June-July). Lexical and relational influences on the processing of Chinese modifier-noun compounds. Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Lee, C. Y. (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.

Lee, C.-L., Hung, D. L., Tse, J. K. P., Lee, C.-Y., Tsai, J.-L. & Tzeng, O. J.-L. (in press). Processing of disyllabic compound words in Chinese aphasia: Evidence for the processing limitations account. *Brain and Language*.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.

Li, P., Bates, E., & MacWhinney, B. (1993). Processing a language without inflections: A reaction time study of sentence interpretation in Chinese. *Journal of Memory and Language*, *32*, 169-192.

Liang, M. Y. (1992). Recognition processing in reading compositional and idiomatic words. Unpublished master's thesis, National Tsinghua University, Hsinchu, Taiwan.

Libben, G., Gibson, M., Yoon, Y. B., & Sandra, D. (2003). Compound fracture: The role of semantic transparency and morphological headedness. *Brain and Language, 84*, 50-64.

Liu, I.-M. (1988). Context effects on word/character naming: Alphabetic versus logographic languages. In: I.-M. Liu, H.-C. Chen, & M. J. Chen (Eds.), *Cognitive aspects of the Chinese language, vol. 1* (pp. 81-92). Hong Kong: Asian Research Service.

Liu, I.-M., Wu, J.-T., & Chou, T.-L. (1996). Encoding operation and transcoding as the major loci of the frequency effect. *Cognition*, *59*, 149-168.

Liu, Y., & Peng, D. (1995). *Time-course of Chinese compound word recognition*. Poster presented at the Seventh International Conference on the Cognitive Processing of Chinese and Other Asian Languages, Hong Kong.

Liu, Y. & Peng, D. (1997). Meaning access of Chinese compounds and its time course. In H.-C. Chen (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 219-232). Hong Kong: The Chinese University Press.

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

Marslen-Wilson, W., Tyler, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101* (1), 3-33.

Mattingly, I. G. & Hsiao, P. (1997). Constituent superiority in Chinese. In H.-C. Chen (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 207-218). Hong Kong: The Chinese University Press.

Mattingly, I., G. & Xu, Y. (1994). Word superiority in Chinese. In H.-W. Chang, J.-T. Huang, C.-W. Hue & O. J. L. Tzeng (Eds.), *Advances in the study of Chinese language processing, vol. 1* (pp. 101-111). Taipei: Department of Psychology, National Taiwan University.

Myers, J., & Gong, S. (2002). Cross-morphemic predictability and the lexical access of compounds in Mandarin Chinese. *Folia Linguistica*, *26* (1-2), 65-96.

Myers, J., & Lai, Y. (2002, December). *The auditory lexical access of monomorphemic compounds*. Paper presented at the Tenth International Conference on the Cognitive Processing of Chinese and Other Related Asian Languages, Taipei.

Myers, J., Derwing, B., & Libben, G. (2004). The effect of priming direction on reading Chinese compounds. *Mental Lexicon Working Papers, 1*, 69-86. University of Alberta, Canada.

Myers, J., Libben, G., & Derwing, B. (2004, June-July). The nature of transparency effects in Chinese compound processing. Poster presented at Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Packard, J. L. (1993). *A linguistic investigation of aphasic Chinese speech*. Dordrecht: Kluwer.

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

Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge: Cambridge University Press.

Peng, D., Liu, Y. & Wang, C. (1999). How is access representation organized? The relation of polymorphemic words and their morphemes in Chinese. In J. Wang, A. W. Inhoff, & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 65-89). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology, 81* (2), 275-280.

Schreuder, R., & Baayen, H. (1997). How complex simplex words can be. *Journal of Memory and Language, 37*, 118-139.

Semenza, C., Luzzatti, C., & Carabelli, S. (1997). Morphological representation of a nouns: A study on Italian aphasic patients. *Journal of Neurolinguistics, 10* (1), 33-43.

Sproat, R., & Shih, C. (1996). A corpus-based analysis of Mandarin root compounds. *Journal of East Asian Linguistics, 5*, 49-71.

Starosta, S., Kuiper, K., Ng, S., & Wu, Z. (1998). On defining the Chinese compound word: Headedness in Chinese compounding and Chinese VR compounds. In J. L. Packard (Ed.), *New approaches to Chinese word formation* (pp. 347-370). Berlin: Mouton de Gruyter.

Su, Y.-C. (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, M. (2003). Morphological representation as a correlation between form and meaning. In E. Assink & D. Sandra (Eds.), *Reading complex words* (pp. 113-138). Amsterdam: Kluwer.

Taft, M., & Forster, K. I. (1976). Lexical storage and retrieval of polymorphemic and polysyllabic words. *Journal of Verbal Learning and Verbal Behavior, 15* (6), 607-620.

Taft, M., Huang, J.-T. & Zhu, X. (1994). The influence of character frequency on word recognition responses in Chinese. In H.-W. Chang, J.-T. Huang, C.-W. Hue, & O. J. L. Tzeng (Eds.), *Advances in the study of Chinese language processing, vol. 1* (pp. 59-73). Taipei: Department of Psychology, National Taiwan University.

Taft, M., Liu, Y. & Zhu, X. (1999). Morphemic processing in reading Chinese. In J. Wang, A. W. Inhoff, & H.-C.Chen (Eds.) *Reading Chinese script: A cognitive analysis* (pp. 91-113). Mahwah, NJ: Lawrence Erlbaum Associates.

Taft, M., & Zhu, X. (1995). The representation of bound morphemes in the lexicon: A Chinese study. In L. B. Feldman (Ed.), *Morphological aspects of language processing* (pp. 293-316). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Taft, M., & Zhu, X. (1997). Using masked priming to examine lexical storage of Chinese compound words. In H.-C. Chen (Ed.), *Cognitive processing of Chinese and related Asian languages* (pp. 233-241). Hong Kong: The Chinese University Press.

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

Tsai, C.-H., McConkie, G. W. & Zheng, X. (1998). *Lexical parsing by Chinese readers*. Poster presented at the Advanced Study Institute on Advances in Theoretical Issues and Cognitive Neuroscience Research of the Chinese Language, University of Hong Kong.

Tsai, P. (1998). *Auditory access of disyllabic words in the Chinese mental lexicon*. Unpublished master's thesis, National Chung Cheng University, Chia-Yi, Taiwan.

Tsay, J., Myers, J. & Chen, X.-J. (1999). Tone sandhi as evidence for segmentation in Taiwanese. In E. V. Clark (Ed.), *Proceedings of the Thirtieth Annual Child Language Research Forum* (pp. 211-218). Stanford, CA: Center for the Study of Language and Information.

Tseng, C.-H., Huang, K.-Y. & Jeng, J.-Y. (1996). The role of the syllable in perceiving spoken Chinese. *Proceedings of the National Science Council, Part C: Humanities and Social Sciences, 6* (1), 71-86.

Wang, W., & Myers, J. (2004, June-July). The processing of affixation and compounding in Chinese. Paper presented at Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Wong, K. F. E., & Chen, H.-C. (2002, December). *Morphemic ambiguity resolution in reading Chinese two-character words: An eye-monitoring study*. Paper presented at the Tenth International Conference on the Cognitive Processing of Chinese and Other Related Asian Languages, Taipei.

Yin, H., Derwing, B. L., & Libben, G. (2004, June-July). Branching preferences for large lexical structures in Chinese. Poster presented at Poster presented at the Fourth International Conference on the Mental Lexicon, Windsor, Canada.

Zhang, B. & Peng, D. (1992). Decomposed storage in the Chinese lexicon. In H.-C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 131-149). Amsterdam: North-Holland.

Zhang, B. (1997). Zhongwen shuangzici zai xinli cidian zhong de chucun moshi [Storage model of Chinese two-character words in the mental lexicon]. In R. Peng, H. Shu & H. Chen (Eds.), *Hanyu renzhi yanjiu* [Chinese cognitive research] (pp. 217-230). Jinan, China: Shandong Jiaoyu Publishers.

Zhou, X., & Marslen-Wilson, W. (1994). Words, morphemes, and syllables in the Chinese Mental lexicon. *Language and Cognitive Processes, 9* (3), 393-422.

Zhou, X., & Marslen-Wilson, W. (1995). Morphological structure in the Chinese mental lexicon. *Language and Cognitive Processes, 10* (6), 545-600.

Zhou, X., & Marslen-Wilson, W. (2000). Lexical representation of compound words: Cross-linguistic evidence. *Psychologia, 43* (1), 47-66.

Zhou, X., Marslen-Wilson, W., Taft, M. & Shu, H. (1999). Morphology, orthography, and phonology in reading Chinese compound words. *Language and Cognitive Processes, 14* (5/6), 525-565.

Zhou, X., Ostrin, R. K. & Tyler, L. K. (1993). The noun-verb problem and Chinese aphasia: comments on Bates et al. (1991). *Brain and Language, 45*, 86-93.

Author Note

For help in finding references and for helpful suggestions, my thanks to Gary Libben and Gonia Jarema, the three Oxford reviewers, Marcus Taft, Hsuan-Chih Chen, Shuping Gong, Tingyu Huang, Hongbo Ji, Ching-yuan Lin, Huichuan Liu, Cui-xia Weng, Peiwen Tsai, and Jane Tsay. Naturally, all errors in description or interpretation are my own responsibility: caveat emptor! The author may be contacted at the Graduate Institute of Linguistics, National Chung Cheng University, Min-Hsiung, Chia-Yi 621, Taiwan, or by email at Lngmyers@ccu.edu.tw.

Appendix

Chinese examples cited in the text.*

Pinyin	Characters	Pinyin	Characters	Pinyin	Characters
anning	安寧	huaqiao	華僑	qiuzheng	求證
anshu	桉樹	huasheng	花生	ridian	日店
anzhuang	安裝	huaxiang	滑翔	shafa	沙發
baise	白色	huayuan	花園	shanshi	善事
bangzhu	幫主	huoche	火車	shenjing	神經
banzou	伴奏	hushi	護士	shiguang	時光
beipan	背叛	jiajuan	家眷	shudian	書店
benneng	本能	jiaren	家人	shuhao	書號
bingdian	餅店	jiashi	家事	shujia	書架
caimi	猜謎	jichang	機場	shumu	樹木
caiyi	猜疑	jita	吉他	Taibei	台北
caoshuai	草率	kuaihuo	快活	Taiwan	台灣
chengguan	城關	kuaile	快樂	wuzhu	屋主
chifan	吃飯	kuaisu	快速	xiaozhang	校長
chi-le-fan	吃了飯	laitai	來台	xiehao	寫好
chouchang	惆悵	lingdao	領導	xiguazhi	西瓜汁
ci	詞	lü	侶	xingren	行人
dongxi	東西	mahu	馬虎	xun	殉
feiji	飛機	maikefeng	麥克風	yanzheng	驗證
fugai	覆蓋	mashang	馬上	yinhang	銀行
fumu	父母	meigui	玫瑰	yisheng	醫生
fuxiong	父兄	mianyang	綿羊	yiyuan	醫院
ganganjingjing	乾浄	muxiao	母校	zi	字
ganjing	乾乾浄浄	putao	葡萄	zici	字詞
huagui	華貴	qiushi	糗事	zuo-le	做了
huali	華麗	qiuyin	蚯蚓	zuozhe	作者

* Tone marks are suppressed in the Pinyin transcriptions. Characters are in the traditional form used in Taiwan, even though some of the experiments cited were conducted using the simplified characters standard in the PRC; these differences don't affect the points made in the text.