Chinese Compound Processing in Sentences

with Rapid Serial Visual Presentation

by

Guangting Wang

A thesis submitted in partial fulfillment of the requirements for the degree of'

Doctor of Philosophy

Department of Linguistics
University of Alberta

Abstract

Due to the uniqueness of Chinese orthographic features and the pervasiveness of compounding in modern Chinese, psycholinguistic research in the past decades has shown great interest in the visual recognition of Chinese compound words. Models of compound processing make different predictions about whether compounds have whole-word representations, whether compound words and even characters are initially decomposed and recognized on the basis of their morphemic subunits, and at what point the meanings associated with these units come into play. Clearly, the debate is unresolved. The research presented in this thesis aims at contributing to this area of inquiry through a series of experiments addressing the reading of Chinese compounds.

The present dissertation reports four naming experiments (with in all 12 sub-experiments) , three of which manipulated different situations in which Chinese compound words are read: one in which two constituent characters are presented on the same line, one in which they are split across two lines, and one in which the order of the constituents is reversed to form a semantically different word, and the fourth one extended compounding from compound words to compound characters. For each of these situations, the exposure duration in Rapid Serial Visual Presentation (RSVP) of the compound was either 100ms, 200ms, or 500ms. Subsequently, a stimulus potentially related to the compound was presented on the computer screen. Participants were asked to read this target word out loud. The naming latencies were recorded, and entered as response variable into a mixed-effects regression model with the lexical properties (such as frequency, character complexity and character family size) and experimental factors (exposure duration, presentation type) as predictors.

In the first three experiments on compound word processing, significant effects of

the frequency of the compound prime word were observed for the naming latencies to the target word for the shortest exposure duration (100 ms). Interestingly, the sign of the frequency effect depended on whether the target word was morphologically related or unrelated to the compound prime. Facilitation was present in the related condition, but inhibition in the unrelated condition, indicating that this frequency effect is semantic in nature. This pattern persisted even when the two constituent characters were split over two lines.

Comparison of the results between short and longer RSVP presentation rates (i.e. 100ms versus 200ms and 500ms) showed that the compound frequency effect was subject to fast decay: it was present for a 100ms exposure duration, but absent for 200ms and 500ms exposure durations. This suggests the semantic priming effect is subject to fast decay in short-term memory (STM).

Finally, we did not obtain any evidence that naming a component of a compound character or an unrelated character would be interpreted with reference to the meaning of the preceding prime compound character (Experiment 4), a finding that is very different from what emerged for two-character compound primes and single-character targets (Experiments 1, 2 and 3). This suggests that the components of single characters, read out loud after presentation of the prime sentence, are processed as semantically void, purely orthographic parts of characters, comparable to letters in English words.

Taken together, my findings provide evidence for rapid access to the meanings of compounds read in sentential context, for fast decay of these meanings, and for the importance of contextual integration in short-term memory.

Preface

This thesis is an original work by Guangting Wang. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Research Ethics Board, Project Name “Word Recognition in Reading Chinese: a Psycholinguistic Point of View”, No. Pro00038499, April 17, 2013.

Acknowledgements

First of all, I would like to thank my co-supervisors Dr. Harald Baayen, Dr. Antti Arppe and Dr. Patrick Bolger. I express my deepest gratitude to Dr. Harald Baayen, who has guided me throughout the journey. His constant support, constructive suggestions and patient guidance were fundamental for the development and the completion of my Ph.D. program. He changed my thinking pattern towards science. I look forward to future collaboration and publications with Harald. My sincere thanks also goes to Dr. Antti Arppe who helped me a lot in the last two years, having academic meetings with me every week, giving professional suggestions on my writing, and cordially organizing my candidacy and defense. I would like to thank Dr. Patrick Bolger, from whom I learned many useful lab techniques.

I would also like to thank my supervisory committee member, Dr. Benjamin Tucker, for his comments and suggestions on my thesis writing. I would like to thank Dr. James Myers for being the external examiner of my defense, giving insightful comments and questions. My gratitude also goes to Dr. Xiaoting Li for her constructive suggestions in both my candidacy and defense.

I've been fortunate to have the opportunity to study in the Department of Linguistics, University of Alberta. I am grateful to Dr. John Newman, who was the Chair of the Department when I started my program, for helping me come to Canada. Thanks to Oliver Rossier, who gave a lot of solid suggestions on my program, and impressed me by his efficiency and friendly smile. Thanks to Diane McKen, Elizabeth French, Gail Mathew, as the wonderful administration staff, for your constant support and assistance.

Last, but by no means least, I would like to thank my wife, Wei Tang. She was always there cheering me up and stood by me through the good times and bad. Without her I would not be who I am.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

The Chinese script is the oldest writing system still in use. It enables speakers of various Sinitic languages (Mandarin, Cantonese, Wu, etc.) to share a written language based on Mandarin grammar and vocabulary, bringing a degree of cohesion to the vast and heterogeneous Han Chinese community (Hanley et al., 1999; McNaughton and Ying, 1999). Chinese in this thesis, without special indication, refers to Mandarin Chinese. Over the last 30 years, many studies have investigated the processing of Chinese compound words and characters (e.g. Hoosain, 1992; Zhang et al., 1992; Taft et al., 1994; Zhou and Marslen-Wilson, 1994, 1995, 1999; Liu, et al.,1997; Peng, et al., 1999; Packard, 1999, 2000; Tsai et al., 2006; Chen and Chen, 2006; Myers, 2006, 2010, Mok, 2009) . Of special interest in this research is how readers respond to the challenges posed by the Chinese writing system as compared to alphabetic scripts.

1.1 **Modern Chinese orthographic features**

Chinese orthography differs substantially from that of alphabetic languages (in what follows, the Chinese orthography in this thesis refers to the modern simplified Chinese writing system). There are two basic features that are of central importance for visual word processing in Chinese. First, up to the word level, there are four orthographic layers: strokes, components, characters, and words. A Chinese word contains one or more characters (e.g. 手/ʃəʊ(214)/ *hand*; 知道 /dʒi(55)daʊ/ *know*; 电视机/diɛn(51)ʃi(51)tɕi(55)/ *TV set*). These characters in turn consist of one or more components. For example, the single-character word 手 has only one component. The left character 知 in the two-character word 知道 comprises two components: 矢 and 口. A component is composed of a single stroke or a cluster of strokes. For instance, the component 矢 is made up of 5 strokes "丿", "一", "一", "丿" and "丶".

The second feature is the lack of word boundary markers in writing. In conventionally written or printed Chinese scripts, characters are evenly spaced. There are no extra space boundaries for words like in some alphabetic languages such as English. For instance, the common Chinese sentence ”我们实验室每周三下午开会。” will most probably be segmented and literally translated as:

我们 / 实验室 / 每 / 周三 / 下午 / 开 / 会。
Our / lab /every / Wed. / afternoon/ holds /a meeting.

The fact of having no word boundary in written Chinese may make it problematic to isolate the words in the text. For example, if a sentence consists of 8 characters, there will be 7 character boundaries among them. Since each character boundary could also be a potential location of a word boundary, the 8-character sentence would theoretically generate dozens of possible words, and these words could further create a complex problem for word segmentation. Among the many possible segmented word strings, most of time there is only one correct parse, but ambiguity about the proper segmentation can persist.

Since the Chinese word is not explicitly marked or segmented in writing, there are ongoing debates as to whether actually there are words in written Chinese, about whether the character is the basic processing unit, or whether sequences of characters should or should not be analyzed as “compound words"'. All these questions have consequences for the conception of the mental lexicon in Chinese. Therefore, it is necessary to have a brief look at the status of Chinese words and compound words.

1.2 Chinese words and compound words

Unlike Classical Chinese, which was morphologically essentially an isolating language (Branner, 2001) with a strong tendency towards monosyllabicity (Arcodia, 2007), present-day Mandarin features a number of different morphological phenomena and prefers multi-syllabic words (Lin, 2001). The commonly held view that Mandarin has almost no morphology is a misconception. Truth be told, Mandarin morphology is certainly simple compared to the morphological systems of Finnish, Arabic or Turkish. However, there do exist words, which are different from either characters or phrases, which function as syntactically independent meaningful units in modern Chinese (Ge, 1985; Wei, 2004; Peng and Pan, 2010; Huang et. al., 2012), for example, 蚂蚁 (/ma(214).i(214)/, *ant*), 苹果 (/piŋ(35)guo(214)/, *apple*), and 东道主 (/duŋ(55)dɑu(51)dʒu(214)/, *host*). Although words are not visually salient in Chinese writing, the existence of a word level in Mandarin is demonstrable both through linguistic argumentation and experimentation. Packard's book on Chinese morphology (2000) presents an overview of the linguistic arguments about words (the most important being that Mandarin syntactic rules operate on words, not characters/morphemes). Based on a careful examination and comparison of many definitions of the word in Chinese, Packard (2000) proposed a definition that incorporates both syntactic and semantic criteria. In what follows, I accept as a working hypothesis Packard's definition of a word for modern Chinese: a word is a smallest meaningful syntactically independent unit (p. 12). As a semantic unit, a word expresses one meaning. As a syntactic unit, it may freely play different syntactic roles (for instance, a noun could be a subject or an object).

Then, what is a Chinese compound word? The definition of Chinese compound words is similar to that of English: Chinese compounds are words containing two or more root morphemes, i.e., they are combinations of words. For example, the two-character compound word 鞋带(/ɕiɛ(35)dai(51)/, sh*oelace*) consists of 鞋 (/ɕiɛ(35)/, *shoe*) and 带(/dai(51)/, *belt*); the compound word 打印 (/da(214).in(51)/, *print*) consists of 打 (/da(214)/, *hit*) and 印 (/in(51)/, *print*). Following Packard, a compound word is an onomasiological unit expressing one meaning. Even transparent, endocentric compounds (like 鞋带) cannot be fully reduced to a compositional function of the meanings of their components. Therefore, compounds must have a whole-word conceptual representation in the mental lexicon.

There are two important facts about modern Chinese words that are worth mentioning. First, in modern Chinese, most words contain more than one character. For a direct comparison between the proportions of single-character and multi-character words, let us take an example of Yuan and Huang's statistics based on the Chinese Morpheme Data Bank at Tsinghua University (Yuan and Huang, 1998). From Table 1.1, one can see that in modern Chinese, 86.32% of all Chinese words contain more than one character. A more recent statistic on two-character words from the State Language Work Committee (2008) shows that 72.05% of the most commonly used Chinese words (40352 out of 56008) are two-character words. Secondly, most single characters function as onomasiological units themselves, or function as root morphemes (Taft, 1995; Zhou, 1999; Libben, 2006).

Table 1.1

Statistics of words containing different number of characters (Yuan & Huang, 1998)

status	**number**	**percentage**
Single-character words	15951	13.68
2-character words	78230	67.09
3-character words	6700	5.75
4-character words	14200	12.18
5-character words	1163	1
other	356	0.3
total	116,600	100

This thesis focuses on two-character compounds in Chines, as most words in Chinese contain two characters (about 70 % of Mandarin words according to the Institute of Language Teaching and Research, 1986). Compounds also predominate among Mandarin neologisms (95 %) which suggests compounding is productive in modern Chinese, whereas derivatives (words created from other words, usually through the addition of prefixes or suffixes, for instance by adding the suffix 家 (-er): 画 (to paint) to 画家(painter) and 歌唱 (to sing) to 歌唱家 (singer), account for only a little more than 2 % of all newly coined words (Ceccagno and Basciano, 2007).

1.3 **Context and motivation of the study**

Given that most words in Chinese are two-character compounds, it is not surprising that one of the hotly debated issues in the literature on visual word recognition in Chinese concerns the representation and processing of compound words and characters. What then is the experimental evidence supporting the linguistic notion of “word”? The Word Superiority paradigm has been used to provide evidence for the mental reality of Chinese words. The word superiority effect concerns the finding that letters are more readily recognized and identified in real words than in nonwords (see Reicher, 1969, for English). Studies on Chinese have consistently found that lexical status at the word level affects recognition at the character level (Cheng, 1981; Mattingly and Xu, 1994; Chen 1999; Hung, et al., 1999). Furthermore, Taft (2003) reported a lexical decision task in which readers had to decide whether the characters presented (some free and some bound) were a word. Readers gave more correct “yes” responses to free-morphemic characters than to bound characters. This finding also supports the notion of a word as an independent onomasiological unit.

A second type of evidence for the reality of the word, that we can find in almost all studies, is the facilitatory effect of whole word frequency, independent of character frequency and syllable frequency.

However, researchers interested in morphology can hardly be satisfied with the (obvious) claim that Chinese has words. Some researchers have argued that compound words are stored in decomposed form in the Chinese lexicon (Zhang and Peng, 1992; Taft and Zhu, 1994,1999; Wu et al.,1999). They point to character frequency effects, which they take to imply that word recognition proceeds through initial access of a compound's constituent characters. The multilevel interactive activation model of Taft and Zhu (1994) in fact has several decompositional layers, including submorphemic units such as radicals.

A decompositional view in which the character is a central unit is shared by several Chinese linguists, who take characters to be at the core of both Chinese grammar and of Chinese language processing （字本位 in Chinese) (Xu, 1991, 1994, 1997, 2005; Pan, 2002; Wang, 2009; Wang, 2008).

So far, the debate on the mental representation and processing of Chinese compounds in reading is still ongoing. Although some researchers believe that Chinese words may be treated as units at some point during lexical access and that characters have a great influence on the reading of words, the relationship and time course of holistic whole-word units and decomposed sub-unit representations is still unclear and controversial. Furthermore, relatively little is known about the time course of lexical access when words are read in sentences.

There are some striking similarities and differences between compound words and complex characters. Compounding seems not to be restricted to the juxtaposition of characters in the sentence. Characters themselves can be also composed of other characters. Most characters contain more than one component. Here, it is important to distinguish between a "character component" (部件 in Chinese) and a "radical" (部首 in Chinese). Radicals are a subset of the set of components, namely, they are those components that are used in Chinese dictionaries for looking up characters (Wang, 1981). Components, by contrast, comprise all internally dividable stroke patterns. According to *Xinhua Dictionary (*10th revised edition, Chinese Academy of Social Sciences, 2004), there are 264 radicals. But there are around 648 constituent components in the 7,786 standard modern Chinese character set (*A Dictionary of Chinese Character Information,* Science Press, 1988). According to the statistics from the 7,786 characters in *A Dictionary of Chinese Character Information*, the characters containing no more than four components cumulatively account for 94.9%. On the other hand, less than 5% of characters have only one component (details shown in Table 1.2).

Given that many single-character words can be decomposed into characters that are themselves words, and that single-character words can be combined to form two-character (or multiple-character) compounds, the question arises what the status is of characters in two-character compounds on one hand, and of characters embedded in characters, on the other. How are these different kinds of words, with very different kinds of internal complexity, represented and accessed in the mental lexicon? Are these characters and compounds accessed as wholes, are they decomposed, or are they both stored and decomposed? Furthermore, when compounds are put into a sentential comprehension environment, what would be the pattern and relationship between the compound's holistic and sub-units representations? These questions bring us to the central questions of this thesis.

Table 1.2

Percentage of characters containing different numbers of components based on 7786 standard Chinese characters from *A Dictionary of Chinese Character Information*

Number of components	Percentage	Cumulative percentage
1	4.149	4.149
2	34.04	38.189
3	40.321	78.51
4	16.39	94.9
5	4.149	99.049
> 5	0.951	100

1.4 Evolution and organization of this dissertation

The initial goal of this study was to collect evidence about the role of and the relationship between holistic and sub-units of Chinese compounds and characters during lexical access, using a regression design including lexical variables such as frequency and stroke complexity for both larger and smaller units (two-character compounds, single-character constituents of these compounds, and the character components embedded in single characters). The idea was that a rapid presentation of the prime compound, followed equally rapidly by further sentential context, would provide a window on when and under what conditions whole words and constituents are activated. If constituents are activated first, and whole words last, then a short exposure time should reveal effects of constituent frequency, whereas a long exposure time would show effects of whole-word properties reflecting access to the compound's meaning. The results of the experiments, however, yielded a pattern of results that did not match any of the scenarios envisioned before the experiments were carried out. These results fit better with a theory in which semantics are activated rapidly, within a 100 ms time window, and in which semantic activation also decays rapidly (within a 200 ms time window post stimulus onset). Instead of tapping into the process of lexical access, the Rapid Serial Visual Presentation (RSVP) paradigm that I used appears to tap into post-access processes of contextual integration and the time-course of the availability of activated concepts in short-term memory (STM).

The organization of this dissertation is as follows. This first chapter is a brief introduction to the background and motivation of the study. Chapter two presents a literature review, zooming in on the debate about the mental representations involved in compound processing and a few related studies about the time course of the activation of the associated meanings in reading Chinese compound words. Chapters three to six are respective reports and discussions of the four experiments' results. The final chapter is the general discussion, in which the results are summarized and pitted against the different models in the literature on lexical processing.

Chapter 2 Literature Review

Due to the uniqueness of Chinese orthographic features and the pervasiveness of compounding in modern Chinese, psycholinguistic research over the last decades showed great interest in the visual recognition of compound words. Models of compound processing make different predictions about whether the compound words or characters are initially decomposed and recognized on the basis of their morphemic subunits or whether they can directly be accessed as whole words, and at what point the compound meaning begins to play a role.

2.1 Decompositional models

One prominent view in psycholinguistics holds that complex words are decomposed into their constituent parts. For Chinese, the decomposition view holds that compound words are first decomposed into their sub-lexical or internal morphemic units (the constituent characters or even character components), which subsequently are organized to form and activate the mental lexical representation of the whole word and its meaning (Taft, 1994, 2003). In other words, this obligatory decomposition approach proposes that the recognition of morphologically complex words always begins with a rapid morphemic segmentation (Feldman,& Andjelković,1992; Feldman, 1995, 2003; Rastle & Davis, 2008; McCormick, Rastle & Davis, 2008). This parsing process is posited to take place without exception, and to be sensitive to surface cues to morphological complexity in its orthography (e.g., corn in corner). According to this morpho-orthographic view, a Chinese 2-character word such as "父亲"(/fu(51)tɕʰin(55)/ *father*) is always first decomposed into "父"(*father*) and "亲(*kin*)" in lexical access.

Traditional research of sub-unit processing in Chinese started from compound character recognition (due to the traditional view that characters are the basic processing units), focusing on a special type of characters containing a semantic component (providing information about a character's meaning) and a phonetic component (encoding phonological information). This type of compound characters, hereafter called SP (semantic-phonetic) characters, account for about 81% of the 7,000 most frequent characters in modern Chinese (Li & Kang, 1993). The sub-lexical processing of these components plays an important role in the recognition of whole characters. Much of the sub-lexical evidence was collected in studies on both phonological regularity (i.e., whether the whole character is pronounced the same as its phonetic radical) and consistency (i.e., whether a set of characters having the same phonetic radical are pronounced the same) through behavioral tests (Fang, Horng, & Tzeng, 1986; Zhou & Marslen-Wilson, 1999) and recent ERP experiments (Hsu, Tsai, Lee, & Tzeng, 2009). Moreover, consistent with previous behavioral studies, it has been reported in ERP studies that the semantic information of the embedded phonetic radical is activated, and that at least for low-frequency words, the phonological representation of the phonetic radical can be automatically activated. Experimental evidence for the activation of the semantic radical has also been reported (Feldman & Siok, 1997; Leck, Weekes, & Chen, 1995; Ju and Jackson, 1995). These authors conclude that semantic radicals play a role in character categorization and that the effect cannot be attributed to graphic similarity alone.

A second line of research focused on multi-character compound words. The first documented study of Chinese compound word processing is Zhang and Peng (1992), who found that the constituent characters' frequency was predictive for lexical decisions to compound word targets. This was taken to indicate that constituent characters are activated in Chinese compound word processing. In the following decade, a great amount of research on Chinese compound words was conducted and several theoretical frameworks were developed based on the research fruit of alphabetical languages. Influential models include the Multi-level Interactive Activation Model (Taft & Zhu, 1995; Taft, Liu & Liu, 1999), Dual-route models (Baayen et al., 1997; Coltheart et al.,2001; Pinker, 2011), the Intra/Inter Connection Model (Peng, Liu & Wang, 1999) and the Non-Hierachical Semantically Based Model (Zhou, Marslen-Wilson & Shu, 1999). All models of Chinese compound word processing acknowledge the co-existence of morpheme and word-level representations. The Multi-level Interactive Activation Model assumes that morpheme-level representation is set at a lower level than whole word level representation. The other two models contend that the two types of representations are set at the same level. However, a review of the literature shows that few developments have been made since these early studies.

There are reasons to doubt that Chinese characters are the optimal unit through which Chinese words are necessarily accessed. First of all, because of its high productivity, a character may have many very different meanings when it appears as a constituent in different words. Such characters are quite similar to homographs in English. For example, the meanings of the character 处(/tʃu214/) in the words 处女 (means *virgin* in English) and 处罚 (means *punish* in English) are entirely different although the pronunciation is the same. Moreover, the pronunciation of some characters is also undetermined with the visual form of a single character. For example, 会 can be pronounced as /huei(51)/ in 会议 (/huei(51).i(51)/, means *conference* in English) and /kuai(51)/ in 会计 (/kuai(51)tɕi(51)/, means *accountant* in English). For such characters, the context is the essential guide to the proper phonological form. Thus, it is difficult to determine the exact meaning and pronunciation of a character based on its orthographic form alone. A word context is necessary for final determination. Considering the above two aspects, if component characters are the basic representational units, the processing of Chinese compound words would not be efficient. However, if Chinese compound words are represented as unitary wholes, the match between visual input and internal mental representation would be as fast as a dictionary look-up process. Hence, there are processing efficiency advantages for representing Chinese compound words as orthographic wholes in the mental lexicon.

2.2 Holistic representation

Various studies have reported experiments that suggest that indeed compound words may be processed as wholes and accessed through holistic representations. One finding supporting holistic processing is the word superiority effect (WSE, i.e., a letter is more quickly and accurately accessed in a real meaningful word than in a meaningless letter string). This effect was also found in Chinese word recognition. Liu and Peng (1997) observed that when target compounds for lexical decisions appeared immediately after prime compounds, opaque primes (e.g. 草率 /tsau(214)ʃuai(51)/, *careless*, literally 'grass-lead') facilitated responses to targets semantically related to the prime's whole-word meaning (e.g. 马虎 /ma(214)hu/, *careless*) but not to targets related to component morphemes (e.g. 树木 /ʃu(51)mu(51)/, *tree*, related to 草/tsau(214)/, *grass*). Mok (2009) found a stronger word superiority effect for character detection in compounds that contained at least one semantically opaque morpheme/character as compared with fully transparent compounds. This suggests that opaque compounds have word-level representations. Zheng (1981) also found a word superiority effect in reading Chinese 2-character words. His single-character target words had shorter decision latencies when the prime 2-character word which contained the target was a real word rather than a non-word.

Holistic processing of Chinese two-character words has also been observed in adult skilled readers on silent reading of reversible words (Peng, Ding, Wang, & Zhu, 1999; Zhang et al., 2004). In Chinese, there is a group of two-character words for which reversing their morphemes results in another real word. For example, the word 领带 (/liŋ(214)dai(51) /, tie) has a real reverse-order word 带领 (/dai(51)liŋ(214)/, *to lead*). This pair of two-character words differs only in the order of their constituent characters. They are of interest to linguists and psychologists because they are particularly meaningful and convenient for investigating the respective roles of morphemes and whole words in word recognition. Because sub-word properties remain the same within pairs, they offer excellent opportunities to pit whole word properties such as compound frequency against constituent properties (such as constituent frequency), while keeping other predictors such as stroke, component, and radical properties matched. Peng et al.(1999), through semantic priming experiments, found that, with a brief priming time, i.e. SOA (stimulus-onset asynchrony) =57ms, there is a semantic priming effect on the original words, but no significant semantic priming effect on its reversed word. If the word's meaning were accessed through its morphemic units, access to the meaning of the reversed word that contains exactly the same morphemes should also have been facilitated. Therefore, they argued for the holistic processing in recognizing Chinese words.

Li and Logan (2008) found an object-based attention effect in reading Chinese words. In their research, subjects were shown four characters in each trial, appearing in two rows and two columns centered at the fixation point. Their results showed that subjects' visual attention shifted more quickly when 2 characters could form a valid word. This result was in line with results on object-based visual attention (Egly, Driver, & Rafal, 1994; Li, Liu, & Ma, 2011), and indicated that processing of these Chinese words, similar to objects, affected the subjects' deployment of attention. Again, this was taken as evidence that Chinese compound words are processed as wholes.

Many studies on reading English conclude that words' attributes affect readers' eye-movement patterns (Kuperman, Bertram & Baayen, 2008; Rayner, 1998, 2009; White, 2008). Similarly, in reading Chinese, word-level factors such as frequency, predictability, and complexity also impact fixation duration and fixation position and word skipping (Rayner, Li, Juhasz, & Yan, 2005; Yan, Tian, Bai, & Rayner, 2006). Yan et al.,(2006) found that frequencies of word and character affect subjects' fixation duration, but word frequency modulated the effect of character frequency. The effect of character frequency was attenuated with high frequency target words while it was quite apparent with low frequency target words. Other researchers (Rayner, Li, Juhasz, & Yan, 2005) have shown that Chinese readers fixated for less time on high- and medium-predictable target words than on low-predictable target words. They were also more likely to fixate on low-predictable target words than on high- or medium-predictable target words. These previous studies have shown that, in visual word recognition, word-related variables affect readers' eye-movement control, which then reasonably supports the holistic processing of Chinese words.

Recent studies (Bai et al., 2008), investigating the consequences of spaces in the reading of Chinese, have also supported whole word processing. Bai et al. (2008) added spaces in various ways, separating real words, or splitting two-character compounds, or placed between every two characters. They found that sentences with real-word parsing were as easy to read as unspaced text. The other two ways of inserting spaces produced longer reading times. As adding spaces within compounds turned out to be disruptive, their data support the possibility that words, not individual characters, are the primary representational units in Chinese reading.

An important consideration that fits well with the evidence of the importance of word-level properties (such as compound frequency) in lexical processing is that many, if not all, compounds in Chinese have varying degrees of non-compositionality. As semantically opaque meanings cannot be computed by rule (by definition), theories that first decompose can only work by assuming table look-up that is blind to semantics. But the importance of semantic relations between components and compounds shows that a simply form-matching approach does not do justice to the full complexity of reading in Chinese.

2.3 The time course of semantic activation

Understanding a word's meaning is the ultimate goal of lexical access. If Chinese compounds have word level representations in the mental lexicon, as many have argued, then the question arises of when and how the associated meaning is activated. In the past decades, a few studies have been carried out on the time course of semantic activation of Chinese compound words. The lexical decision experiments of Zhou and Marslen-Wilson (2000) showed strong semantic priming effects at both short (57 ms) and long (200 ms) stimulus onset asynchrony (SOA), without phonological effects, when Chinese compound words were read. Zhou et al. (1999) also found that, at both 57ms and 200ms SOA, semantic activation of constituent morphemes of Chinese compound words was constrained by both orthographic and phonological activation, even though phonology by itself had no significant influences on whole-word semantic activation. According to the constituency model of Chinese word recognition (Perfetti and Tan, 1999), semantic activation only occurs later, after mandatory phonological activation. However, the empirical findings in support of this view have proven difficult to replicate (see, for example, Chen and Shu, 2001; Xie and Zhou, 2003). In a more recent functional magnetic resonance imaging study (Zhan et al, 2013), by applying a lexical decision task to pseudo-homophones of Chinese compound words, compound semantic activation was also detected with rapid presentation of the stimuli.

The relationship between the phonological and semantic activation in reading Chinese compound words is still controversial. One strong phonological view was that access to the semantic representations of the constituent morphemes and of the whole word could be assumed to be mainly mediated by the phonological activation of constituent syllables (and information about the co-occurrence of these syllables, see Zhou & Marslen-Wilson, 2000). Direct access from orthography to semantics could play at most a minor role. However, one problem with this claim is that most of the supporting experimental evidence, which is mainly based on the comparison of the relative time course of phonological and semantic activation in reading words, proves to be difficult to replicate (see, for example, Chen & Shu, 2001; Zhou & Marslen-Wilson, 2000). Based on evidence from studies using different experimental paradigms tapping directly into semantic activation, including semantic categorization (Chen et al., 1995; Leck et al., 1995; Sakuma et al, 1998; Wydell, Patterson, & Humphreys, 1993), semantic judgment (Xu, Pollatsek, & Potter, 1999; Zhou, Pollatsek, & Marslen-Wilson, 1999), and phonologically mediated semantic priming (Zhou, 1997; Zhou & Marslen-Wilson, in press b), Zhou and his colleagues (1999) further argued that it is orthography, rather than phonology, that plays a relatively more important role in determining semantic activation in reading Chinese. Therefore, the relationship between semantic and phonological elements of the Chinese mental lexicon is still under debate.

Compared to semantic activation, the potential phonological activation of a compound word in Chinese was studied less. Among the few studies, questions, like whether there is obligatory phonological activation in the natural reading of Chinese (Zhou et al., 1999) and whether reading Chinese is solely meaning-based (Weekes et al. 1998), are still controversial. What makes Chinese come to its own is its written form (Perfetti and Tan 1998). Therefore, this thesis mainly focused on the orthographic impact, rather than the uncertain phonological effects, on the Chinese compound processing.

2.4 Summary and research questions

The role of holistic representations of compounds is still a controversial issue in research on the lexical processing, both for languages such as English, and for Chinese. If there is a word-level orthographic representation, the point in time at which it is engaged in lexical processing is hotly debated. In Chinese psycholinguistics, these questions are complicated by the additional layers of complexity that come with the orthography.

The research questions originally motivating this study are: Do holistic orthographic representations for Chinese two-character compound words play a functional role in reading? If so, when are such holistic representations engaged in lexical processing: from the start, or only late, after the constituent characters have been processed? Does it matter whether the compound word is read in a sentence rather than in isolation? What are the differences and similarities between the processing of two-character compound words and compound characters that are themselves constructed from components that are themselves characters?

In order to answer these questions, four word naming experiments, each with three sub-experiments, were carried out. In each of these experiments, a “prime word” was presented in a sentence, followed by a “target word” that the participant was asked to read out loud.

The first set of experiments presented two-character compounds with both constituent characters on the same line in the sentence. The second set of experiments presented two-character compounds, the constituent characters of which were presented on two different lines. The third set of experiments worked with two-character reversible compounds, and manipulated the order of the constituents. The final set of experiments focused on compound characters. All words were presented in sentences, which were displayed, using Rapid Serial Visual Presentation (RSVP), in three-character groups, with three exposure durations: 100ms, 200ms and 500ms. The naming latencies were analyzed with mixed-effect regression models, using, as predictors, exposure durations and covariates for the lexical distributional properties of the “prime word” in the sentence, and the “target” character/word to be read out loud.

My original prediction was that how Chinese compounds are read is time-sensitive and fluid with respect to the different exposure durations of the word or character in the sentence. In fast and efficient reading, I expected holistic orthographic representations to be involved, rather than the decomposed orthographic representations of the constituent characters or components. With the passing of time, i.e., at longer exposure durations, I expected morphemic sub-units to come into play. In other words, my hypothesis was that when reading is fast and efficient, the whole takes precedence over the parts. But when attention is focused on the whole for longer periods of time, the parts would become more prominent. This prediction was in line with a supra-lexical account proposed by Giraudo & Grainger (2001, 2003).

Chapter 3 Experiment 1: Two Constituent Characters in the Same Line

In Experiment 1, participants read a 9-character sentence, divided into three 3-character groups and containing as a prime a two-character word. The sentence was presented using Rapid Serial Visual Presentation (RSVP) of the successive three groups of three characters. Immediately after having read the sentence, subjects were asked to read out loud a single character that either had occurred as part of the "prime" compound (the related condition), or that had not occurred in and not semantically related to the preceding sentence (the unrelated condition). The experiment was conducted three times with three different groups of participants. In Experiment 1a, each 3-character group was presented for 100ms at the center of the screen. In Experiment 1b, the exposure duration was 200ms, and in Experiment 1c it was 500ms (see Figure 3.1 on page 24 for the details of this procedure). I predicted that orthographic whole-word representations would be essential for compound processing, and would be most prominently detectable at the fastest representation rate (100ms), whereas at longer exposure durations, the orthographic representations of the constituent characters would be more important.

3.1 Experiment 1a: 100ms per 3-character group

3.1.1 Method

Participants

Twenty Chinese speakers (12 female and 8 male) participated in the Experiment. All of them were native Mandarin Chinese, and had normal or corrected-to normal vision. Some received course credit for participation; others participated on a voluntary basis. Their average age was 23.75 (SD= 7.4). Participants had lived in

Canada for an average of 2.25 years (SD=2.21).

Stimuli and Apparatus

The stimuli consisted of 60 sentences, each with 9 characters that were presented rapidly in groups of three. All the stimuli used in Experiment 1 (with their English translations) are shown in Appendix 2. In every sentence, a two-character prime word was embedded in the second 3-character group as the last two characters. For the related condition, the target character that was displayed after the sentence, and that was to be read out loud, was the first character of the two-character word. Thus, this character was always the middle-character position in the second group.

All stimuli were presented and all the responses were collected with DMDX using an HP Windows 7 computer with a microphone (PNICAM PLS-350) attached to record the participants' naming responses.

For the experiment, 60 Chinese two-character compound words were selected from CNCORPUS (20,000,000 characters; Ministry of Education Institute of Applied Linguistics, 2009). See Appendix 2 for all details.

Why these presentation rates?

The presentation rate was selected based on results from two previous studies (Liao and Zhang, 1996; Wang, 2005) and my own pilot on Chinese reading speed. The statistical results from Liao and Zhang's Chinese reading experiments on 40 university students showed that, at fast efficient sentence reading, the speed ranged from 70 ms/character (hereafter, mspc) to 134 mspc with an average of 103 mspc; at normal reading, from 156 mspc to 290 mspc with an average of 196 mspc. Wang (2005) found the reading rate of 30 adult Chinese readers (mean age= 27.2) was

175-324 mspc (averagely 235 mspc) in normal reading of sentences. In addition to these studies, in my own pre-tests on seven Chinese readers, I found the minimum presentation rate for presenting 3-character group while maintaining understanding of the meaning of the sentence varied between 80ms to 120 ms. Therefore, to make sure that comprehension by the subjects is guaranteed, I set 100ms as the fastest presentation rate for each 3-character group.

Why Rapid Serial Visual Presentation?

Rapid Serial Visual Presentation is a method of displaying the text word-by-word (or by word groups) in a fixed focal position. It was first proposed for screen reading by Foster in 1970 though the term had been employed more than 10 years before (Gilbert, 1959). Since then, RSVP has become a tool of studying reading behavior. With RSVP, one screen can contain one word, multiple words, one sentence, or even one paragraph, and quite a lot research with RSVP has been done on reading Chinese. For example, Xu and Zhu (1997) did experiments on reading Chinese with RSVP by presenting different numbers of Chinese characters (1, 3, 5, 7 characters to be exact) per unit, and found the comprehension was not affected much by the number of characters, but that the 3-character condition afforded the highest accuracy. Shen et al. (2001, 2002) tested the readability of RSVP in Chinese texts and found 1) the display rate of 300 characters per minute (i.e 200ms/character) was more appropriate to reading comprehension than 600 characters per minute; 2) a fixed computer-controlled display rate worked better than a self-paced display rate; 3) there was a trade-off between reading speed and comprehension, unsurprisingly.

There are three reasons for my adoption of an RSVP of sentences with three successive characters presented at a time. One is to prevent the participants from using task strategies when reading the two-character compound primes. By providing further meaningful context, attention is drawn away from the prime. The second reason is the possibility of manipulating exposure duration. The last reason is practical. When using digital hi-tech products such as PDA, electronic readers, smart phones, vehicle GPS, etc., all of which have limitations on their interfaces due to small monitor displays, the actual way in which Chinese is read is actually very close to reading with RSVP.

The rationale for grouping and presenting three characters as a group with RSVP can be found in studies on the visual field and perceptual span of Chinese readers (Inhoff & Liu, 1997, 1998; see also Chen & Tang, 1998). The average perceptual span in Chinese natural reading is about three characters (to be exact, extending 1 character to the left of fixation and 2 characters to the right when reading from left to right). Given the small overlap between the perceptual spans of successive fixations during the reading of Chinese text (Inhoff & Liu, 1998), the three-character grouping presents itself as the most appropriate way to present Chinese sentences with Rapid Serial Visual Presentation. Furthermore, Xu and Zhu's study (1997) on the readability of Chinese scripts with RSVP, presenting different number of Chinese characters per unit, also supports three-character units as a good choice.

Procedure

Subjects were seated in front of a PC screen at a distance of approximately 60 cm. They were instructed to silently read the sentence presented with RSVP and then read aloud the target character presented after comprehension of the sentence, as quickly and accurately as possible. At the beginning of the experiment, the participants were instructed about the procedure. Then there was a practice phase to make sure they understood the task. The practice phase consisted of 10 examples of a sentence followed by a constituent character as target, and 10 examples of a sentence followed by a non-related target character. All 60 experimental stimuli (sentences and characters) were presented on the screen in Chinese Song Font with a size of 48 in a mixed random order. Each trial began with a cross as a fixation point at the center of the screen for 300ms, then, the sentence was quickly shown in 3-character groups with RSVP at speed of 100ms per group on the screen. Immediately after the sentence, the constituent character of the two-character prime word, or a character which did not appear in the sentence, was presented on the screen for naming. Finally, the participants were required to repeat the main point of the sentences that they had read to the researcher, to make sure that they had understood the sentences correctly. If there was no response within 2,000 milliseconds for naming the character, the experiment proceeded to the check for comprehension of the sentence's meaning. Both no-response trials and trials with incorrect answers to the comprehension question were coded as errors. The participants' pronunciation of each character was recorded from stimulus onset of the target character by DMDX and saved as WAV files for further checking of pronunciation errors. Each set of 10 experimental trials was followed by a short break.

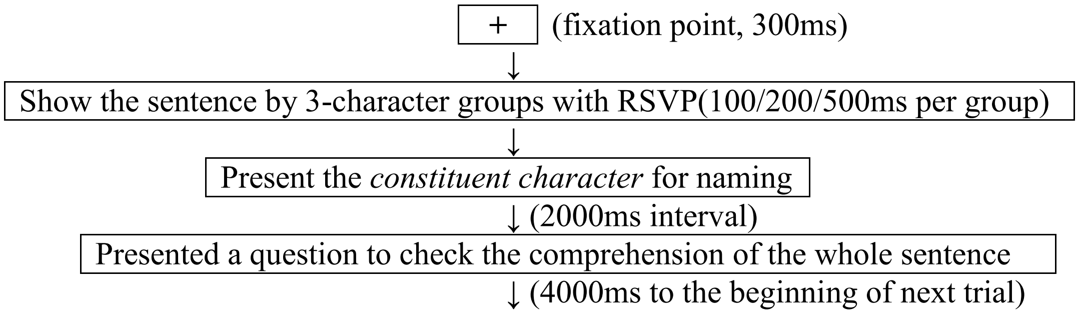

Figure 3.1. The procedure of Experiment 1

All the following experiments in this chapter and the other chapters were designed in the same way. The difference with the following two sub-experiments is the time manipulation on the presentation of the prime RSVP sentence, with exposure duration extended to 200ms and 500ms respectively. The differences with the experiments in subsequent chapters also concerns the sentences read, and the manipulation of the type of the prime and its relation with the target.

Linguistic variables included and collinearity

In experiment 1, the linguistic variables related to either the compound words or their constituent characters. They can be categorized into three major types which are standardly reported in studies on visual word recognition: frequency, complexity, and family size. The family size in this experiment refers to the number of the words that the named character could form as the initial character. Most of these predictors were obtained from the national corpus "CNCORPUS" (20,000,000 characters; Ministry of Education Institute of Applied Linguistics, 2009). I extracted the following frequency counts: the frequency of the related target character (*Freqchar*), the frequency of the unrelated target character (*Freqnr*), the frequency of the compound word (*FreqWord*), and the frequency of the syllable of the target character *(PYFreq,* taken from a Chinese electronic text corpus containing in total 258,852,642 characters, Da, 2004). Measures of orthographic complexity are the number of strokes of the related character (*Nstroke*), the number of strokes of the unrelated character (*Nstrokenr*), and the number of strokes of the compound word (*NstrokeWord*). Finally, the family sizes measures that I included are the family size of the related target character (*FamSize*) and the family size of the unrelated target character (*FamSizenr*).

In order to address the problem of multicollinearity (Baayen, 2008; Belsley et al., 1980; Wurm & Fisicaro, 2014) in regression modeling, I checked the correlation between each pair of covariates. The correlation matrix (Table 3.1) shows a wide range of correlations, with a minimal absolute correlation equal to 0.0002 (shown as zero when rounded to two digits after decimal), and as maximal correlation r = 0.57. This relatively high correlation concerns the stroke complexities (*NstrokeWord*, *Nstroke* and *Nstrokenr*). As none of the correlations exceeded 0.7, no further measures for reducing collinearity were necessary. As will become clear below, across all analyses, only a handful of predictors reached significance, which makes it even more unlikely that collinearity would be distorting my analyses.

Table 3.1

Correlation matrix between pairs of variables

	1	2	3	4	5	6	7	8	9
1. Freqchar	1	0.33	0.5	0.18	-0.56	-0.29	-0.23	0.4	0.01
2.FreqWord	0.33	1	0.08	0.21	-0.3	-0.46	-0.25	0.42	0.06
3.Freqnr	0.5	0.08	1	0.06	-0.23	-0.08	-0.49	0	0.31
4.PYFreq	0.18	0.21	0.06	1	0	-0.12	0.15	0.16	0.02
5.Nstroke	-0.56	-0.3	-0.23	0	1	0.53	0.57	-0.25	0.1
6.NstrokeWord	-0.29	-0.46	-0.08	-0.12	0.53	1	0.3	-0.32	0.05
7.Nstrokenr	-0.23	-0.25	-0.49	0.15	0.57	0.3	1	0.02	-0.17
8.FamSize	0.4	0.42	0	0.16	-0.25	-0.32	0.02	1	0.08
9.FamSizenr	0.01	0.06	0.31	0.02	0.1	0.05	-0.17	0.08	1

All failures to respond and the stimuli that elicited a naming latency shorter than 300 ms were eliminated from the analysis (34 data points, 2.83 %). A total of 1,167 valid data points remained for the analysis of the naming latencies (henceforth RTs) for the target characters.

3.1.2 Results and Discussion

All data were analyzed using R 3.0.2 (R Development Core Team, 2013) and the

R packages: lme4 (Bates and Sarkar 2007), lmerTest (Alexandra Kuznetsova, 2013) and languageR (Baayen, 2009; cf. Baayen, 2008). I analyzed the data by using linear mixed effects models, entering Subjects, Items and Manner of the onset phonemes as random-effect factors into the model equation (see Baayen et al., 2008).

The manner of the onset phonemes (i.e. stops, fricatives, approximants, glidses, etc.) was included as a predictor to control for potential voice key artifacts. However, the standard deviation estimated for this random-effect factor was invariably very close to zero across all experiments. This indicates that this predictor did not capture relevant variance. This might be part of the task effect. When participants' tasks involved comprehending the meaning of the whole sentence with rapid reading, the phonological traits might be attenuated. I therefore removed it from the model specification in all analyses reported below. Or, this might be due to the insensitiveness of the microphone.

The fixed-effect factor condition was examined in interaction with the covariates for the properties of prime and target. Table 3.2 summarizes key results after removing insignificant variables. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 1a in Appendix 1.

Table 3.2

Comparison of the significant variables between two conditions in experiment 1a:

Condition	Significant variables	Mean RT (ms)
Related (constituent) characters	the frequency of the character (t= -3.92) the frequency of the prime word (t=-2.09) the number of strokes of the character (t=-2.68)	853
Unrelated characters	the frequency of the character (t= -4.47) the frequency of the prime word (t=2.34)	919

For the related condition, the frequency of the target character (t=-3.92, p<0.01), the frequency of the prime two-character compound (t=-2.09, p<0.01), and the number of strokes of the characters (t=-2.68, p<0.01) were statistically significant. These results suggest that in very rapid reading of Chinese sentences, not only the sub-units, but also holistic units were playing a role. For the unrelated condition, the frequency of the target character (t=-4.47, p<0.01) was again significant. The frequency of the prime two-character compound (t=2.34, p<0.01) was also significant, but, surprisingly, with opposite sign. Whereas in the related condition, the frequency of the prime afforded facilitation, it gave rise to inhibition in the unrelated condition. Figure 3.2 presents the two regression lines for prime two-character compound frequency. The slopes for the frequency of the target, however, did not differ across conditions (t=0.36, p=0.72). In both conditions, a greater frequency of the target character predicted shorter naming latencies, as expected. Furthermore, the response time for unrelated characters was much greater than that for related characters (by more than 60 ms, see table 3.2.

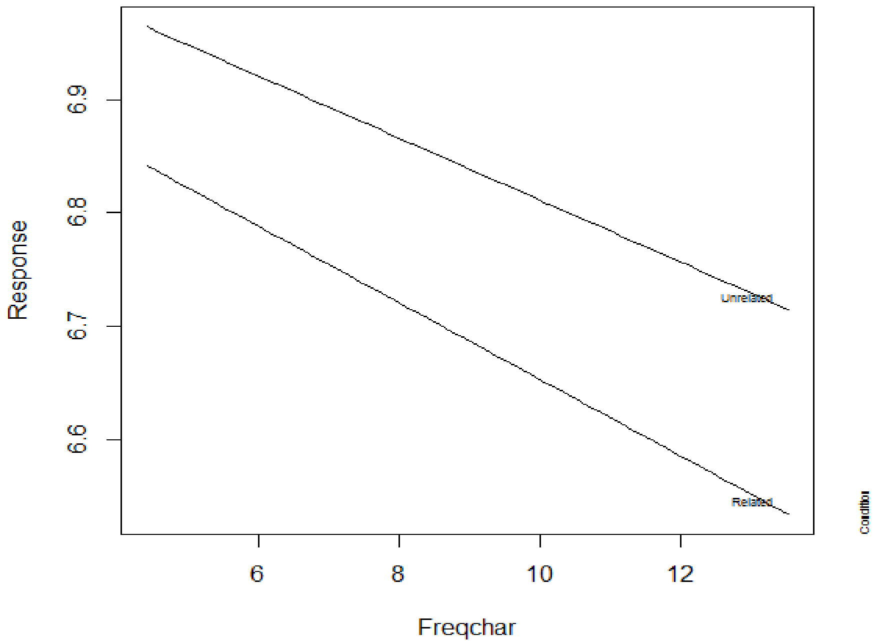

Figure 3.2 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 1a (100ms).

3.2 Experiment1b: 200ms per 3-character group

3.2.1 Method

Participants

Twenty Chinese speakers (11 female and 9 male) participated in the Experiment. All of them were native Mandarin Chinese, and had normal or corrected-to normal vision. Some received course credit for participation; others participated on a voluntary basis. The average age was 22.45 (SD= 5.3). Participants had lived in Canada for an average of 2 years (SD=2.27).

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 1a.

Procedure

The procedure was the same as Experiment 1a, but the presentation duration of each 3-character group has been changed from 100ms in Experiment 1a to 200ms in Experiment 1b.

3.2.2 Results and Discussion

All trials with failures to respond and trials that elicited a naming latency shorter than 300 ms were eliminated from the analysis (35 data points, 2.92 % of the total number of data points). The remaining 1,165 valid data points for correct responses were entered into a mixed-effects regression analysis of the response latencies to the target characters.

Table 3.3 lists the covariates that reached significance in the two conditions, as well as the mean naming latencies for the two conditions. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 1b in Appendix 1. For the related condition, a greater frequency of the primed character predicted shorter response latencies ($t= -4.36$, $p<0.01$). For the unrelated condition, the target's frequency also predicted facilitations ($t=-6.39$, $p<0.01$), with a significantly more negative slope (-0.04, compared with -0.03 for the related condition, see Figure 3.3). The number of strokes of the character ($t=-2.15$, $p<0.01$) also reached significance, but only in the unrelated condition. Here, characters with more strokes elicited shorter RTs. Apparently, with a longer exposure duration of 200ms, only properties of the target character came into play. There was no evidence that would support the possibility that the two-character compound prime influenced the naming latencies. As in Experiment 1a, responses in the unrelated condition were substantially longer than those in the related condition.

Table 3.3

Comparison of the significant variables between two conditions in experiment 1b

Condition	Significant variables	Mean RT (ms)
Related (constituent) characters	the frequency of the character (t= -4.36)	774
Unrelated characters	the frequency of the character (t= -6.39) the number of strokes of the character (t=-2.15)	829

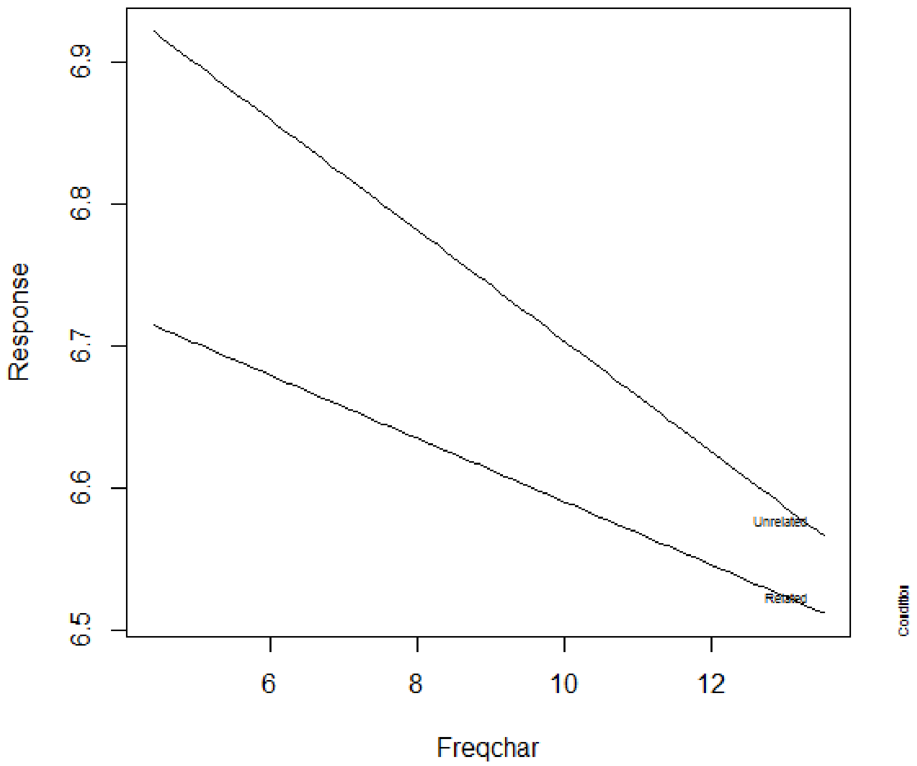

Figure 3.3 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 1b (200ms)

3.3 Experiment1c: 500ms per 3-character group

3.3.1 Method

Participants

Twenty native Mandarin Chinese (16 female and 4 male) in Edmonton took part in the present sub-experiment. The average age was 21.8 (SD= 4.7). Participants had lived in Canada for an average of 3.56 years (SD=2.84)

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 1a.

Procedure

The procedure was the same as Experiment 1a, but the presentation duration of each 3-character group has been changed from 100ms in Experiment 1a to 500ms in Experiment 1c.

3.3.2 Results and Discussion

All trials in which no response was registered, and all trials with a naming latency shorter than 300 ms were eliminated from the analysis (53 data points, as 4.42% of the total). A total of 1,147 valid data points remained.

Table 3.4 summarizes the main statistical findings. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 1c in Appendix 1. The only significant covariate was the frequency of the target character, in interaction with condition. As in Experiment 1b, the facilititatory effect of target character frequency was reduced in the related condition (see Figure 3.4).

As in the preceding experiments, responses in the unrelated condition were substantially delayed compared to the related condition.

Table 3.4

Comparison of the significant variables between two conditions in experiment 1c

Condition	Significant variables	Mean RT (ms)
Related (constituent) characters	the frequency of the character (t= -4.66)	825
Unrelated characters	the frequency of the character (t= -9.80)	870

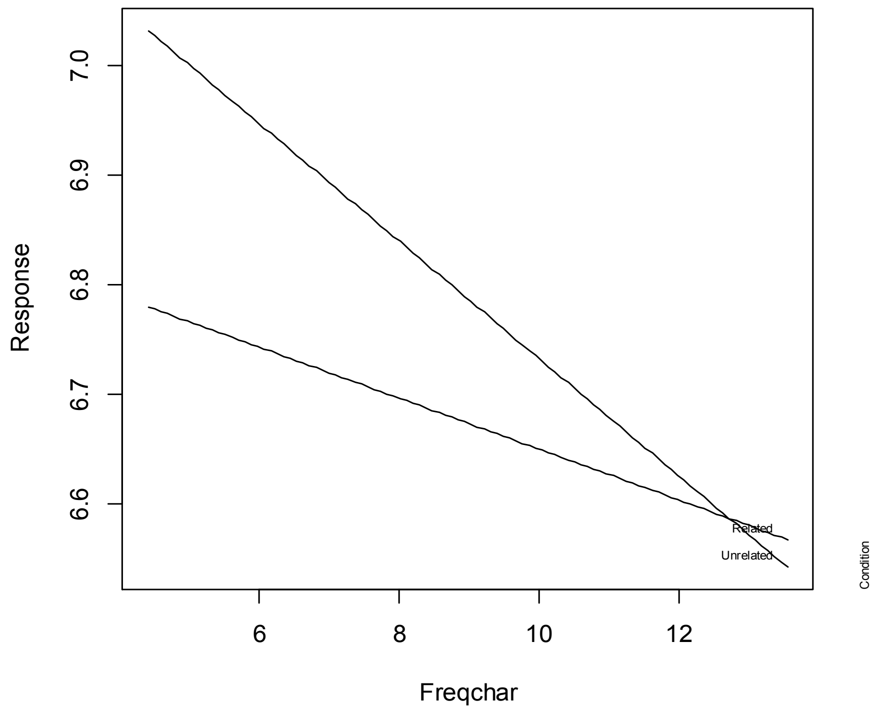

Figure 3.4 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 1c (500ms).

3.4 General Discussion of Experiment 1

Table 3.5 presents a summary of the three sub-experiments. The coefficients in parenthesis represent the slopes of the corresponding covariates. Figure 3.5 plots the slopes of target character frequency at the three exposure rates (100ms, 200ms and 500ms).

From Table 3.5, we can see that the pattern in the naming latencies changed substantially when the speed with which the RSVP sentence was presented was manipulated. At the most rapid presentation rate, an effect of compound frequency was present. Interestingly, in the related condition, the effect of compound frequency was facilitatory, whereas in the unrelated condition, its effect was inhibitory. At longer presentation rates, these compound frequency effects disappeared. Furthermore, at longer presentation rates, the magnitude of the facilitatory effect of the frequency of the character to be read aloud increased for the unrelated condition.

Table 3.5
Summary of Experiment 1

Condition	**100ms**	**200ms**	**500ms**
Mean RT for related characters	853ms	774ms	825ms
Mean RT for unrelated characters	919ms	829ms	870ms
on the related (constituent) characters	the frequency of the character (-0.03); the frequency of the prime word; the number of strokes of the character	the frequency of the character (-0.03)	the frequency of the character (-0.03)
on the unrelated characters	the frequency of the character (-0.03); the frequency of the prime word	the frequency of the character (-0.04); the number of strokes of the character (-0.007)	the frequency of the character (-0.05)

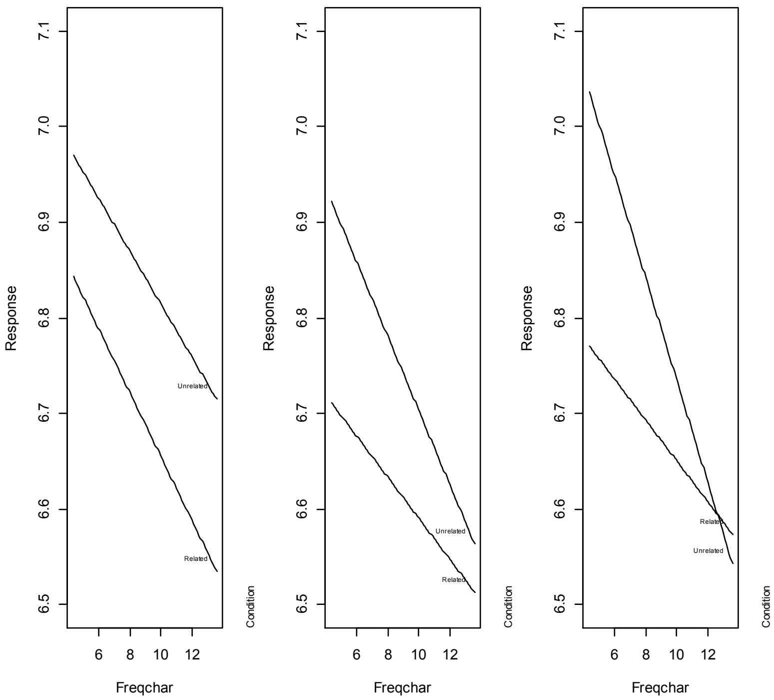

Figure 3.5 Comparison of the regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) across three sub-experiments in Experiment 1

First of all, we need to understand whether the whole-word frequency effect for the prime that we observe here an orthographic or a semantic effect. As shown in Table 3.5, the opposite signs of the slopes of the regression line for the two conditions at 100ms suggest that it may be a semantic frequency effect. If it were an effect reflecting familiarity with a compound' s orthographic form, it should either have been inhibitory across both conditions (reflecting competition between orthographic forms), or it should have been facilitatory for the related condition, and non-significant for the unrelated condition. This was not what we observed. From a semantic perspective, however, the facilitation in the related condition can be understood as being due to shared meanings facilitating a response, whereas in the unrelated condition, naming a character that does not make any sense given the preceding sentence is surprising, and requires more processing time before a response can be initiated. The substantially delayed responses in the unrelated condition fits well with this semantic interpretation of the frequency effect of the prime. In the related condition, the target is semantically related to the preceding sentence, and hence more expected, and easier to read out loud. But the unrelated character is truly unrelated, hence unexpected and unpredictable, leading to longer responses.

Secondly, why is the slope of the frequency effect for the target character less negative in the related as compared to the unrelated condition? It is known that word frequency effects may be attenuated or even disappear in context (Hyona, Vainio & Laine, 2002; Petten and Kutas, 1991). These findings suggest that the more predictable a word is, the more likely it is that a frequency effect disappears. Since word frequency estimates an a-priori probability, without taking context into account, it makes sense that when context makes a character more predictable, as in the related condition, that the a-priori probability of the character becomes less predictive. These considerations lead to the conclusion that in Experiment 1b, as in Experiment 1a, naming latencies hinge on how well the character is predictable from the preceding sentence. Since the frequency effect of the two-character prime is no longer significant, it seems likely that with the longer exposure duration, the meaning of the prime has decayed.

The above pattern of results is consistent with the hypothesis that, in text reading, the meaning of the compound is subject to fast decay. When the presentation rate is 100ms, the time that elapses between reading the compound and reading the target character is a mere 100ms. Even though two or three other words were read during this time, the meaning of the target compound appears to be still available, giving rise to facilitation in the related condition, but inhibition in the unrelated condition. This inhibitory effect in the unrelated condition may reflect a problem with integrating the target character into the current discourse representation generated from the sentence. This problem does not occur in the related condition. For instance, having read 鞋带 (/ɕiɛ(35)dai(51)/, *shoelace*) in the sentence, reading out loud 鞋 (/ɕiɛ(35)/, *shoe*) is easier the more frequently one has encountered the word 鞋带. When the presentation rates were 200ms or more, readers apparently were no longer able to connect the target character to the compound word in the preceding discourse. All that remains is a frequency effect of the target itself, which in the unrelated condition becomes stronger (as shown in Figure 3.5, the slopes became steeper with coefficients from -0.03 to -0.04 to 0.05) with the temporal distance to the sentence. This suggests that, in the related condition, where the effect of the target's frequency does not change, some effect of having been read in the preceding text remains present.

If decomposition is obligatory, as models like the Interactive Activation Model predict, the effect of target character frequency should increase in strength for longer exposure of the prime, i.e., with slower presentation of the sentence. This is not what I found. Furthermore, for the unrelated condition, the target character frequency should have been similar across the three experiments, but an increase was observed. In addition, under obligatory decomposition, the prime compound word should not have shown a significant inhibitory effect on naming the unrelated target character in the 100ms condition.

When reading compound words in RSVP sentences, semantic processes dominate in the present task, leaving little to observe with respect to lexical properties, such as syllable frequency, character family size, and character complexity.

Chapter 4 Experiment 2: Constituent Characters Split Across Two Lines

Experiment 2 examines the impact of the prime exposure duration and relevant linguistic predictors on naming the constituent character in a special case — split-up two-character compound primes. It is conceivable that when a two-character Chinese word is split up and shown on two consecutive lines, compositional processes would be favored over holistic processes, especially since Chinese does not make use of continuation markers like the English hyphen. I originally predicted that splitting a compound word over two lines would give rise to stronger effects of the compound's constituent characters.

4.1 Experiment2a: 100ms per 3-character group

4.1.1 Method

Participants

Twenty students and Chinese speakers (12 female and 8 male) participated in the present sub-experiment. All of them were native Mandarin Chinese, and had normal or corrected-to normal vision. Some received course credit for participation; others participated on a voluntary basis. The average age was 25 (SD= 5.5). Participants had lived in Canada for an average of 2.15 years (SD=1.75).

Stimuli and Apparatus

Both sentences and words were different from those used in the Experiments 1. The stimuli consisted of 60 sentences, 12 characters long (therefore four groups in each RSVP sentence), and in every sentence the two-character prime word was embedded in the second and third 3-character groups. Both the sentences and prime compound words are different from Experiment 1. All the stimuli used in Experiment 1 (with their English translations) are shown in Appendix 3. The reason for creating 12-character (four groups) sentences instead of 9-character ones is to maintain the neutral position of the compound words at the end of the second 3-character group and the beginning of the third group. Because normal average length of modern Chinese sentences ranges from 8-12 characters (Liao and Zhang, 1996; Wang, 2005), the 12-character sentences remain a common sentence length. The character target named was the first character of the compound word. For the experiment, 60 Chinese two-character compound words were selected from CNCORPUS (20,000,000 characters; Ministry of Education Institute of Applied Linguistics, 2009). For full details, see Appendix 3.

Apparatus was the same as Experiment 1.

Procedure

The general procedure frame was the same as for Experiment 1.

Linguistic variables included and collinearity

The linguistic variables included in the analysis were the same as those of Experiment 1. For the convenience of reading the following correlation matrix table, I repeat them as a reminder: frequencies include the frequency of the related target character (*Freqchar*), the frequency of the unrelated target character (*Freqnr*), the frequency of the compound word (*FreqWord*), and the frequency of the syllable of target character *(PYFreq*); as complexity measures, I included the number of strokes of the related character (*Nstroke*), the number of strokes of the unrelated character (*Nstrokenr*), and the number of strokes of the compound word (*NstrokeWord*); the family size measures comprised the family size of the related target character (*FamSize*) and the family size of the unrelated target character (*FamSizenr*). Table 4.1 presents all pairwise correlations among the linguistic predictors. Since no correlations have absolute values exceeding 0.7, the issue of multicollinearity does not arise.

Table 4.1

Correlation matrix between pairs of variables

	1	2	3	4	5	6	7	8	9
1.Freqchar	1	0.31	0.51	0.18	-0.46	-0.09	-0.24	0.42	0.1
2.FreqWord	0.31	1	0.04	0.2	-0.09	-0.35	0.13	0.35	0.02
3.Freqnr	0.51	0.04	1	0.1	-0.25	0.05	-0.53	0.04	0.32
4.PYFreq	0.18	0.2	0.1	1	-0.04	-0.05	-0.04	0.1	0.01
5.Nstroke	-0.46	-0.09	-0.25	-0.04	1	0.34	0.54	-0.09	0.12
6.NstrokeWord	-0.09	-0.35	0.05	-0.05	0.34	1	-0.01	-0.2	0.11
7.Nstrokenr	-0.24	0.13	-0.53	-0.04	0.54	-0.01	1	0.07	-0.16
8.FamSize	0.42	0.35	0.04	0.1	-0.09	-0.2	0.07	1	0.25
9.FamSizenr	0.1	0.02	0.32	0.01	0.12	0.11	-0.16	0.25	1

4.1.2 Results and Discussion

As for the preceding experiments, trials for which no response was registered, and trials for which a naming latency was recorded that was less than 300 ms were removed from the data set (27 data points, 2.25% percent of all trials). A total of 1,173 valid data points remained for statistical analysis.

Similar to Experiment 1, in Experiment 2, there were also two conditions for the target characters to be read out loud, namely a related condition (the target is a constituent character) and the unrelated condition (the target is a character that did not appear in the sentence). Table 4.2 presents an overview of the results of the statistical analysis.

In the related condition, the frequency of the target character facilitated responding ($t=-4.27, p<0.01$), as did the frequency of the two-character compound prime ($t=-6.32, p<0.01$). In the unrelated condition, only the frequency of the target character was predictive, with similar slope ($t=-0.34$, $p=0.73$), see Figure 4.1 and Table 2a in Appendix 1 for detailed slope/coefficients. Responses in the unrelated condition were, on average, slightly shorter than in the related condition, but this difference did not reach significance ($t=-0.54$, $p=0.59$). A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 2a in Appendix 1.

Table 4.2

Comparison of the significant variables between two conditions in experiment 2a

Condition	Significant variables	Mean RT (ms)
Related (constituent) characters	the frequency of the character (t= -4.27) the frequency of the prime word (t=-6.32)	858
Unrelated characters	the frequency of the character (t= -7.38) the frequency of the prime word (t=1.15)	840

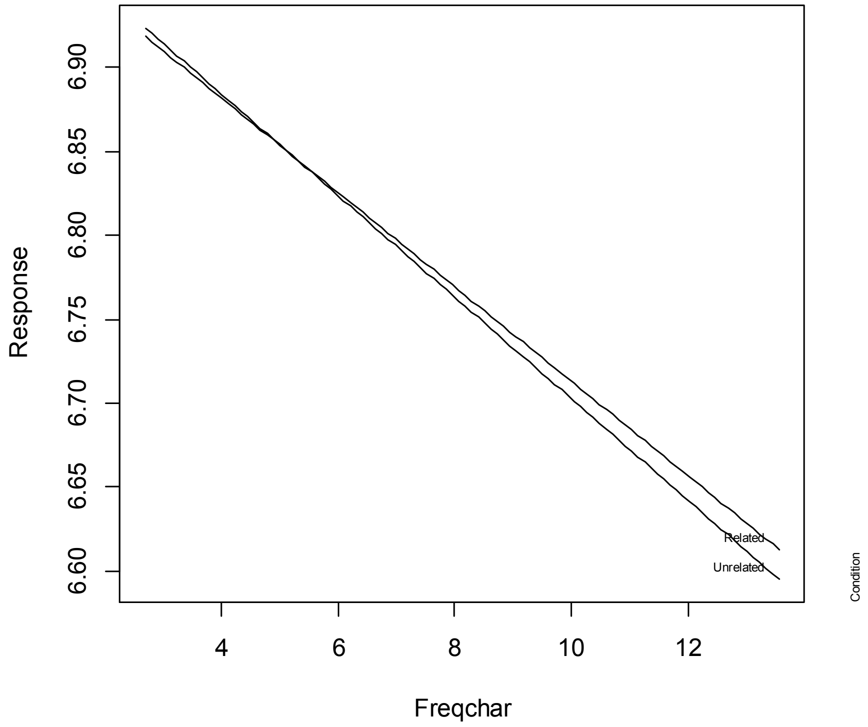

Figure 4.1 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 2a (100ms)

4.2 Experiment2b: 200ms per 3-character group

4.2.1 Method

Participants

Twenty native Mandarin Chinese readers (8 female and 12 male) from around the University of Alberta took part in the present sub-experiment. The average age was 25.7 (SD= 7.8). Participants had lived in Canada for an average of 3.18 years (SD=2.18).

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 2a.

Procedure

The procedure was the same as the one used in Experiment 2a, but the presentation duration of each 3-character group was changed from 100ms in Experiment 2a to 200ms in this experiment.

4.2.2 Results and Discussion

Trials with response failures, erroneous responses, or RTs less than 300 ms were removed from the data set (21 data points, 1.75%), leaving 1,179 valid data points for analysis.

Table 4.3
Comparison of the significant variables between two conditions in experiment 2b

Condition	**Significant variables**	**Mean RT (ms)**
Related (constituent) characters	the frequency of the character (t= -4.52) the family size of the character (t=3.54)	774
Unrelated characters	the frequency of the character (t= -7.98)	817

Table 4.3 presents an overview of the central results of the statistical analysis. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 2b in Appendix 1. Only properties of the target character reached significance. In the related condition, the frequency of the target character was facilitatory (t=-4.52, p<0.01), with a slope that was significantly attenuated (t=-4.28, p<0.01) when compared to the more negative slope characterizing the effect in the unrelated condition. In the related condition, but not in the unrelated condition, the family size of the target character was inhibitory (t=3.54, p<0.01). Responses to targets in the unrelated condition were elongated compared to those in the related condition (t=6.63, p<0.01). To summarize, with longer exposure durations, the

meaning of the two-character compound prime has decayed. All that remains as evidence from prior semantic processing is the reduced effect of the frequency of the target character, and the inhibitory effect of its family size. This latter effect suggests that due to the decay of the meaning of the actual compound that was read, uncertainty has increased about which compound sharing the character as left constituent was at issue. The more potential alternative compounds exist, the greater this uncertainty, and the longer the naming latency became.

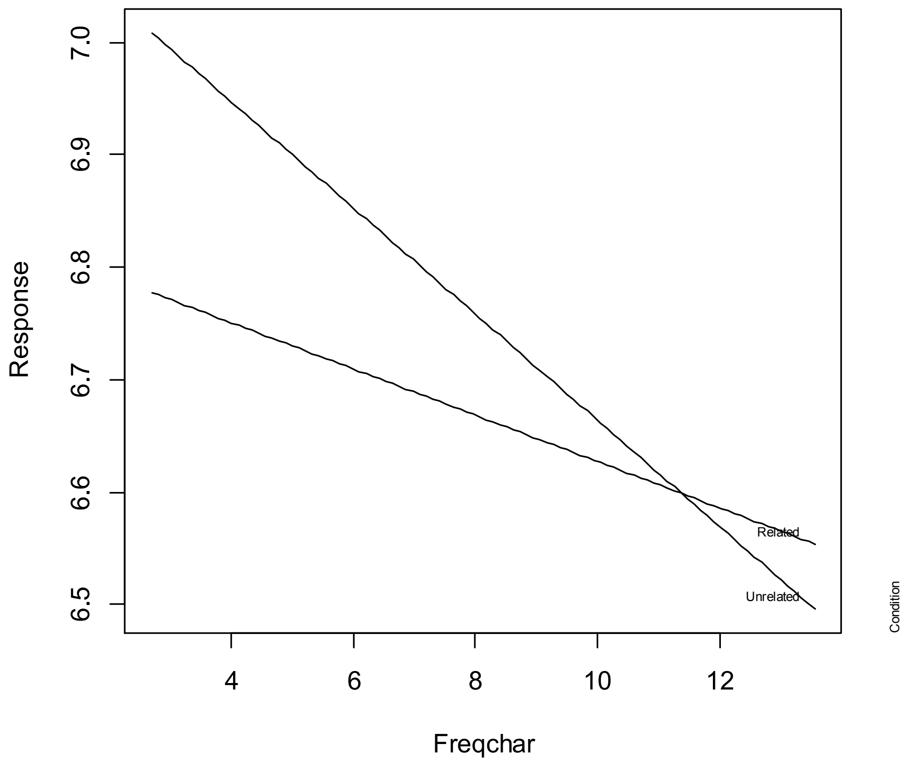

Figure 4.2 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 2b (200ms)

4.3 Experiment2c: 500ms per 3-character group

4.3.1 Method

Participants

Twenty native Mandarin Chinese readers (17 female and 3 male) from the University of Alberta took part in the present sub-experiment. The average age was 21 (SD= 4.12). Participants had lived in Canada for an average of 2.74 years (SD=2.51).

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 2a.

Procedure

The procedure was the same as Experiment 2a, but the presentation duration of each 3-character group was changed from 100ms in Experiment 2a to 500ms.

4.3.2 Results and Discussion

After removal of response failures, errors, and trials with extremely short response latencies less than 300 ms (35 trials, 2.92%), 1,165 valid data points remained for statistical analysis.

Table 4.4
Comparison of the significant variables between two conditions in experiment 2c

Condition	**Significant variables**	**Mean RT (ms)**
Related (constituent) characters	the frequency of the character (t= - 5.14) the family size of the character (t=3.88)	815
Unrelated characters	the frequency of the character (t= -13.66) the number of strokes of the character (t=4.36)	866

Table 4.4 presents an overview of the results of the statistical analysis. A more detailed table of coefficients and associated statistics on the significant variables can

be found in Table 2c in Appendix 1.The pattern of results is very similar to that of Experiment 2b, with a frequency effect for the target character that offers more facilitation in the unrelated condition (t=-7.62, p<0.01), an inhibitory effect of family size only in the related condition(t=3.88, p<0.01), and significant slower responses in the unrelated condition (t=7.27,p<0.01). Additionally I found that in the unrelated condition, an inhibitory effect emerged of the stroke complexity of the target character. No such effect was present for the related condition.

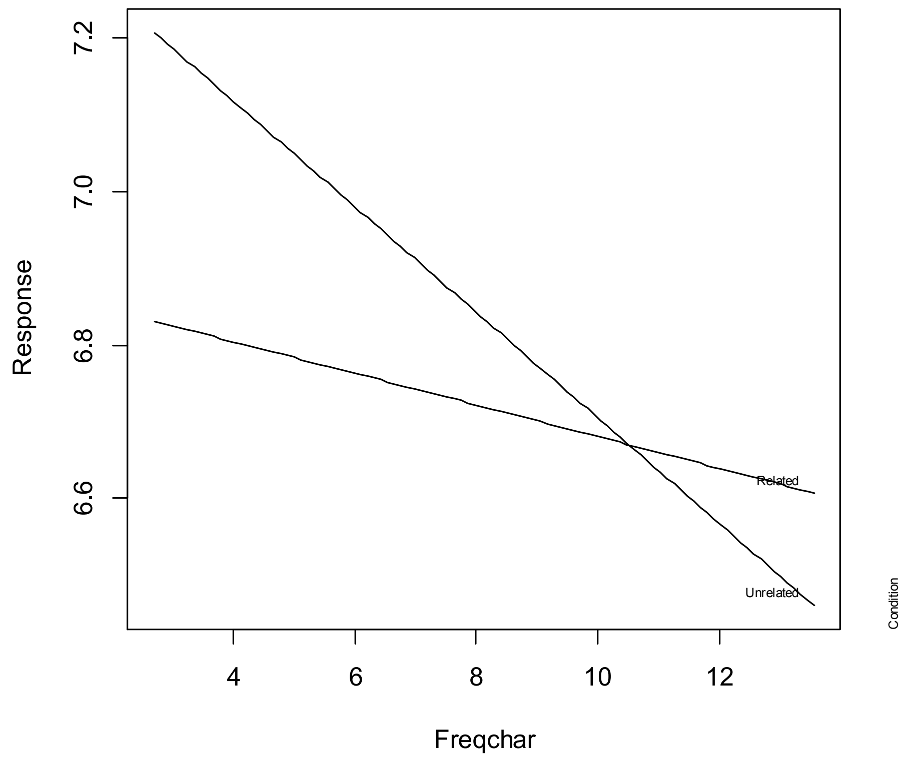

Figure 4.3 Regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) in experiment 2c (500ms).

4.4 General Discussion of Experiment 2

Table 4.5 brings together the main findings of the three sub-experiments in this chapter. The numbers in parenthesis are the slopes of the respective covariates. Figure 4.4 presents the regression lines for target character frequency at the three presentation rates of 100ms, 200ms and 500ms.

As in Experiment 1, the pattern in the naming latencies changed substantially with the speed with which the RSVP sentence was presented. At the most rapid presentation rate, an effect of the frequency of the two-character compound prime was again present. In the related condition, it was facilitatory, whereas in the unrelated condition, it was inhibitory (albeit failing to reach significance). As before, at longer presentation rates, effects of prime compound frequency disappeared. Furthermore, the family size of the target character showed significant inhibition in the related condition at the two longer presentation rates (200ms & 500ms).

Table 4.5
Summary of Experiment 2.

Condition	100ms	200ms	500ms
Mean RT of Related Characters	858 ms	774 ms	815 ms
Mean RT of Unrelated Characters.	840 ms	817 ms	866 ms
On the related (constituent) characters	the frequency of the character (-0.03); the frequency of the prime word (-0.02)	the frequency of the character(-0.03); the family size of the character (0.06)	the frequency of the character(-0.04); the family size of the character (0.06)
On the unrelated characters	the frequency of the character (-0.03)	the frequency of the character (-0.04)	the frequency of the character(-0.05) ; the number of strokes of the character (0.008)

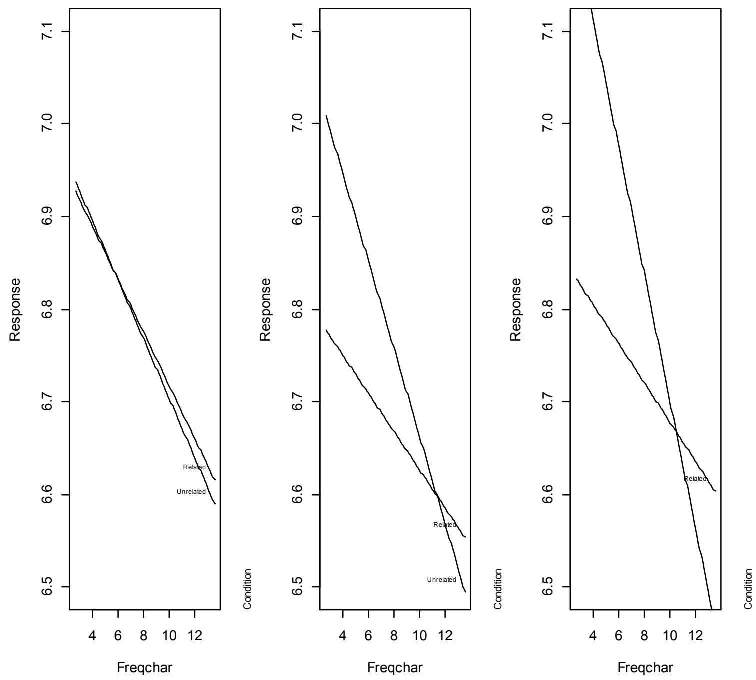

Figure 4.4 Comparison of the regression lines for character frequency (*Freqchar*) as predictor of naming latencies (*Response*) under two conditions across three sub-experiments in Experiment 2

These results suggest, first of all, that the amount of priming was significantly reduced when the target two-character compound prime was split over two lines. Nevertheless, in the related condition, albeit only in this condition, the prime's frequency of occurrence was predictive of the naming latency for the target, affording shorter latencies for more frequent primes. It is important to notice that the time that elapsed between the presentation of the first character of the compound prime (in the second set of three characters) and the presentation of the target character is 100 ms longer than in Experiment 1a. Thus, there are two factors here that may help explain why the results of Experiment 2a differ in the way they do from those of Experiment 1a. First, the prime compound was split across two lines, which necessitated a saccade to the next line. This may have rendered access to the compound's meaning more time-consuming. Second, the time that elapsed between the initial appearance of the prime in the sentence and its second presentation for naming was 100 ms longer. This may help explain why in the unrelated condition, only a non-significant trend towards inhibition remained.

The replication of the compound frequency effect in Experiment 2 provides further support for the hypothesis that the meaning of the compound word is subject to fast decay. Interestingly, this short-lived compound frequency effect emerges even when the compound was split over the two successive lines of text. This fits well with the compound effect being semantic in nature. If the compound frequency effect had been a form effect, due to an orthographic representation for the whole compound being activated while the compound was fixated on, then the split over two lines should have eliminated its effect, contrary to fact.

One remarkable difference from Experiment 1 was the inhibition effect from family size at longer presentation rates (200ms and 500ms). This must be due to the compound now being presented split over two lines. When participants were looking at the first constituent character, which was the last character the second line, the second character, which was the first character of the third line, was not visible: It would only appear 200 or 500ms later. Due to this delay, which was 100-400 ms longer than in the fastest presentation condition, participants apparently began to consider possible compound continuations. The number of possible continuations is estimated by the left character's family size. Upon reading the second character, only one of these family members was consistent with the visual input. The remaining family members had to be suppressed. This suppression may have caused the inhibitory effect of the first character's family size when this character had to be read out loud. The set of potential compounds considered during the reading of the sentence may have had some residual activation, but since they had to be eliminated as viable continuations upon reading the second character of the prime, this residual activation functioned as a source of uncertainty, leading to longer response times.

The reduction in the slope of the regression line for target character frequency in the related condition as compared to the unrelated condition, replicating Experiment 1, is consistent with the hypothesis that in the related condition, the target character is less surprising, thereby rendering a context-less probability estimate less effective.

Chapter 5 Experiment 3: Transposed Constituent Characters

Experiment 3 makes use of a different experimental manipulation. Instead of splitting a compound across two lines, the order of the constituents is reversed. The two-character compound primes were carefully selected so that this manipulation would result in another, existing, compound (compare *houseboat* and *boathouse* in English). A pair of examples of such reversible two-character compounds in Chinese are the words 带领 (/dai(51)liŋ(214)/, *to lead*) which has as its reversed counterpart 领带(/liŋ(214)dai(51) / *tie*). Reversible compounds are potentially revealing about lexical processing because the lexical-distributional properties of the constituents remain unaffected by the reversal. Only the order of the constituents changes, and the meaning of the compound itself. The materials of Experiment 3 were selected such that the meaning of the reversed word was semantically unrelated to the meaning of the original word. In this experiment, the target words to be read out loud are the compound itself, un-reversed in the original condition, or its reversed counterpart in a reversed condition. Thus, the compounds in the reversed condition are semantically unrelated, but orthographically related.

5.1 Experiment3a: 100ms per 3-character group

5.1.1 Method

Participants

Twenty Chinese speakers (10 female and 10 male) participated in the present sub-experiment. All of them were native Mandarin Chinese, and had normal or corrected-to normal vision. They were compensated for their time with course credit, except for those subjects who took part on a voluntary basis. The average age was 23 (SD=4.98). Participants had lived in Canada for an average of 1.88 years (SD=1.61).

Stimuli and Apparatus

Both sentences and words were different from those used in the Experiments 1 and 2. The stimuli consisted of 60 sentences, each 9 characters long, therefore requiring three 3-character groups for rapid serial presentation. For the experiment, 60 Chinese two-character compound words and their reversed words were selected from CNCORPUS. All the stimuli used in Experiment 3 (with their English translations) are shown in Appendix 4. In every sentence, the two-character prime word was embedded in the second 3-character group as the last two characters. The target word named was the semantically-unrelated reversed word of the two-character compound word (the unrelated condition) or the unchanged original two-character compound (the related condition).

The apparatus was the same as Experiment 1.

Procedure

The procedure was identical to that of Experiment 1. The only difference was that the target was a two-character word instead of a single character, and that compound targets in the unrelated condition shared their characters with the two-character compound prime, albeit in a different order.

Linguistic variables included and collinearity

The linguistic variables included in Experiment 3 fall into the three groups. As frequency measures, I included the frequency of the first character of the original word (*Freq1st*), the frequency of the second character of the original word (*Freq2nd*), the frequency of the original compound word (*Freqword*), the frequency of the reversed word (*FreqRev)*, the frequency of the syllable of the first character of the original word *(PY1stFreq*), and the frequency of the syllable of the second character of the original word *(PY2ndFreq*). As measures of orthographic complexity, I considered the number of strokes of the first character of the original word (*Nstroke1st*), the number of strokes of the second character of the original word (*Nstroke2nd*), the number of strokes of the compound word (*NstrokeWord*). As measures of family size, I took into account the family size of the first character of the original word (*Fam1st)* and the family size of the second character of the original word (*Fam2nd)*. Figure 5.1 shows the pairwise correlation matrix for this set of predictors. High correlations (greater than 0.7) were present only for the orthographic complexity measures *Nstroke1st* and *NstrokeWord, Nstroke2nd* and *NstrokeWord.* In the statistical analyses, only one of them was included at a time, to avoid risks of distortion due to collinearity.

Table 5.1

Correlation matrix between pairs of variables

	1	2	3	4	5	6	7	8	9	10	11
1.Freq1st	1	0.16	0.3	0.43	0.57	-0.02	-0.4	-0.19	-0.38	0.34	0.15
2.Freq2nd	0.16	1	0.26	0.23	-0.11	0.77	-0.08	-0.36	-0.26	0.32	0.32
3.FreqWord	0.3	0.26	1	0.66	0.06	0.11	-0.07	0.03	-0.03	0.06	0.06
4.FreqRev	0.43	0.23	0.66	1	0.16	0.05	-0.02	0.01	-0.01	0.09	0.15
5.Fam1st	0.57	-0.11	0.06	0.16	1	-0.1	-0.44	-0.24	-0.43	0.09	-0.08
6.Fam2nd	-0.02	0.77	0.11	0.05	-0.1	1	-0.06	-0.56	-0.38	0.18	0.27
7.Nstroke1st	-0.4	-0.08	-0.07	-0.02	-0.44	-0.06	1	0.27	0.82	-0.26	0.01
8.Nstroke2nd	-0.19	-0.36	0.03	0.01	-0.24	-0.56	0.27	1	0.77	-0.07	-0.15
9.NstrokeWord	-0.38	-0.26	-0.03	-0.01	-0.43	-0.38	0.82	0.77	1	-0.21	-0.09
10.PY1stFreq	0.34	0.32	0.06	0.09	0.09	0.18	-0.26	-0.07	-0.21	1	0.09
11.PY2ndFreq	0.15	0.32	0.06	0.15	-0.08	0.27	0.01	-0.15	-0.09	0.09	1

5.1.2 Results and Discussion

After removing trials eliciting very short response latencies (less than 300 ms), trials that elicited errors, and trials for which no response was obtained (16 data points, as 1.33 percent of the total), 1,184 data points remained for statistical investigation.

An overview of the pattern of results is given in Table 5.2. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 3a in Appendix 1. In the (semantically) unrelated condition, only the frequency of the reversed two-character compound was predictive ($t=-4.00$, $p<0.01$). None of the predictors related to the original prime word or to the constituent characters were informative about the naming latencies. In this unrelated condition, responses were delayed by approximately 40 ms ($t=3.94$, $p<0.01$) compared to the related condition.

In the related condition, in which the original prime compound was simply repeated, the frequency of the compound itself ($t=-3.27$, $p<0.01$, with a slope with similar magnitude as in the unrelated condition, $t=-0.79$, $p=0.43$, see Figure 5.1 and Table 3a in Appendix 1 for detailed slope/coefficients) as well as the frequency of its second character ($t=-2.35$, $p<0.05$) were predictive, both with negative slopes. Inspection of these compounds revealed that in most cases their second character is the head of the (endocentric) compound. This suggests an effect of headedness, with faster responses when the compound and its head refer to the same kind of objects, states, events, or properties.

Table 5.2

Comparison of the significant variables between two conditions in experiment 3a

Condition	Significant variables	Mean RT (ms)
on the reversed words	the frequency of the reversed word ($t= -4.00$)	935
on the original words	the frequency of the original word ($t= -3.27$) the frequency of the last character ($t=-2.35$)	896

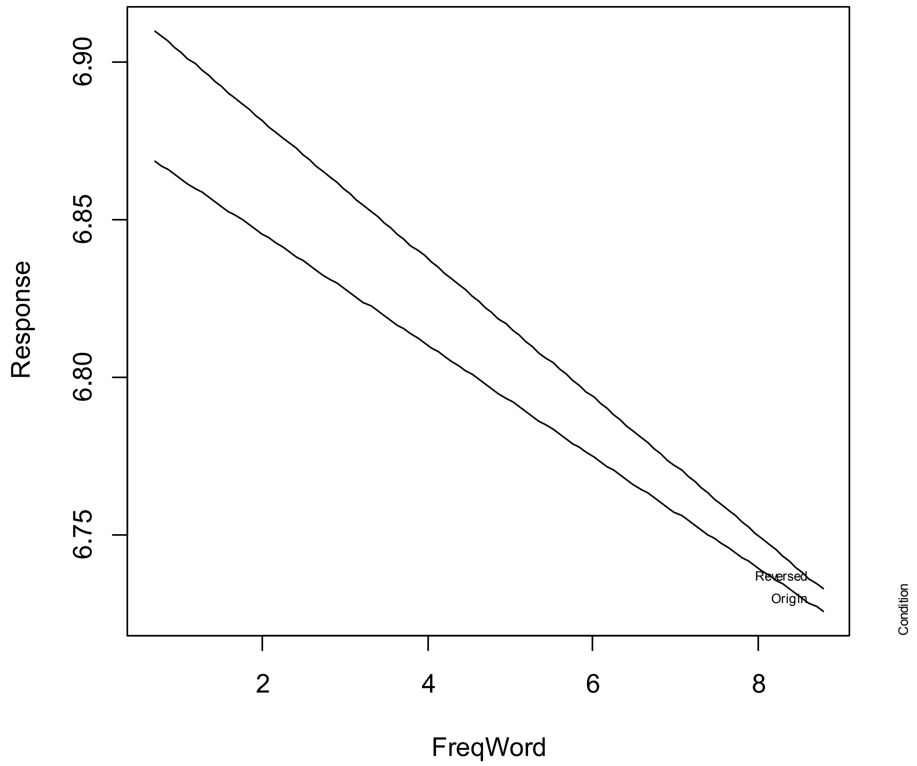

Figure 5.1 Regression lines for the target word frequency (*FreqWord*) as predictor of naming latencies (*Response*) in experiment 3a (100ms).

5.2 Experiment3b: 200ms per 3-character group

5.2.1 Method

Participants

Twenty native speakers of Mandarin Chinese (13 female and 7 male) participated in this sub-experiment. All had normal or corrected-to normal vision. They received course credit for participation, or took part in the study on a voluntary basis. Their age was on average 26 (SD= 7.76). Participants had lived in Canada for an average of 2.98 years (SD=2.59).

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 3a.

Procedure

The procedure was the same as that of Experiment 3a, but now the presentation duration of each 3-character group was changed from 100ms to 200ms.

5.2.2 Results and Discussion

As in the preceding experiments, no-response trials, trials with incorrect responses, and trial with extremely short RTs, were discarded (40 data points, 3.33%), leaving 1,160 data points for statistical evaluation.

Table 5.3

Comparison of the significant variables between two conditions in experiment 3b:

Condition	Significant variables	Mean RT (ms)
on the reversed words	the frequency of the reversed word (t= -3.39) the frequency of the initial character (t=-2.79) the number of strokes of the initial character(t=2.66)	951
on the original words	the frequency of the initial character (t=-5.61)	883

Table 5.3 outlines the pattern of results in Experiment 3b. A more detailed table of coefficients and associated statistics on the significant variables can be found in Table 3b in Appendix 1.For the unrelated condition (with reversed compounds as targets), two frequency effects emerged. More frequent target compounds elicited shorter response latencies ($t=-3.39$, $p<0.01$). Furthermore, when these compounds had higher frequency first characters, response latencies were also facilitated ($t=-2.79$, $p<0.01$). Initial characters, however, were responded to more slowly in proportion to

their number of strokes. As to the related condition, in which the prime compound was simply repeated, only the frequency of the initial character was predictive, with the expected negative slope. There was no effect of the (unreversed) compound's own frequency of occurrence (see Figure 5.2). Words in the unrelated condition required an additional 68 ms on average compared to words in the related condition.

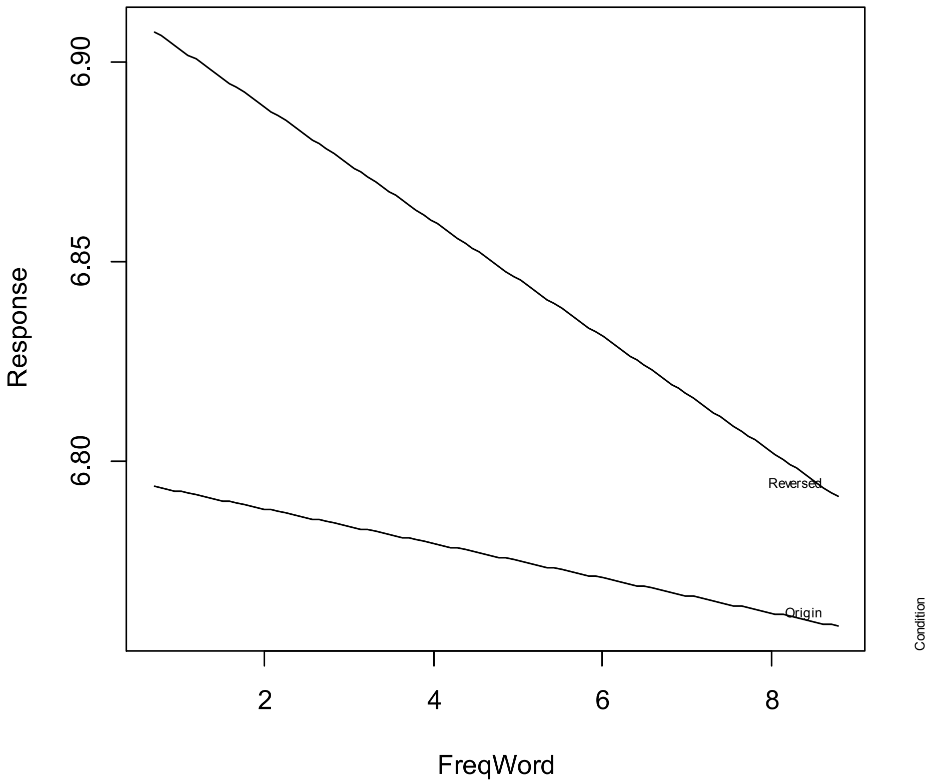

Figure 5.2 Regression lines for the target word frequency (*FreqWord*) as predictor of naming latencies (*Response*) in experiment 3b (200ms).

5.3 Experiment3c: 500ms per 3-character group

5.3.1 Method

Participants

Twenty Madarin Chinese speakers (13 female and 7 male) participated in the present sub-experiment. All of them were native Mandarin Chinese, and had normal or corrected-to normal vision. They got course credits for participation or voluntarily took part in the study. The average age was 21.85 (SD= 4.66). Participants had lived in Canada for an average of 2.85 years (SD=2.72)

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 3a.

Procedure

The procedure was the same as Experiment 3a, but the presentation duration of each 3-character group has been changed from 100ms in Experiment 3a to 500ms.

5.3.2 Results and Discussion

After removal of unusable trials (using the same criteria as for the preceding experiments, 18 data points, 1.5%), 1,182 data points remained available for data analysis. Table 5.4 provides an overview of the pattern of results obtained, which was similar to that of Experiment 3b, except for the lack of statistical support for an effect of stroke complexity. Similar to Experiment 3b, the frequency of the target word was not significant for naming the repeated original word (see Figure 5.3). In the unrelated condition, the frequency of the initial character (t=-2.46, p=0.02) was facilitatory as well. There was a hint of an inhibitory effect of the frequency of the unreversed compound in the unrelated condition (t= 1.8, p=0.08), suggesting the possibility of semantic competition between prime and target. Responses in the related condition were faster than those in the unrelated condition, albeit by only some 20 ms on average (t=3.1, p<0.01).

Table 5.4

Comparison of the significant variables between two conditions in experiment 3c:

Condition	Significant variables	Mean RT (ms)
on the reversed words	the frequency of the reversed word (t=-2.26) the frequency of the initial character (t=-2.46)	803
on the original words	the frequency of the initial character (t=-7.68)	782

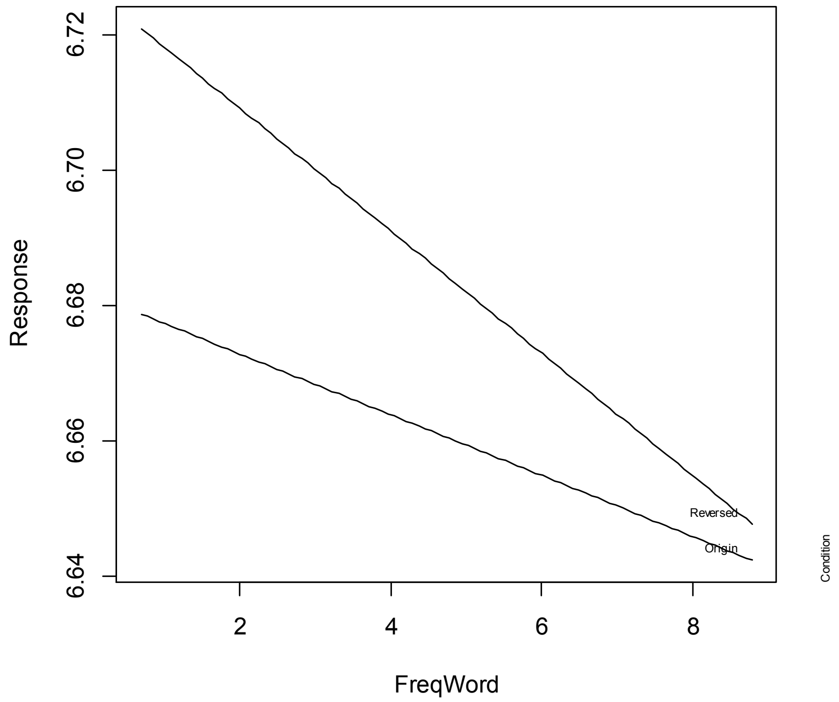

Figure 5.3 Regression lines for the target word frequency (*FreqWord*) as predictor of naming latencies (*Response*) in experiment 3c (500ms).

5.4 General Discussion of Experiment 3

An overview of the results of Experiments 3a-c is given in Table 5.5, with the slopes of the covariates in parentheses. Figure 5.4 illustrates how the slope of target compound frequency changes with exposure duration.

Table 5.5

Summary of Experiment 3

Condition	**100ms**	**200ms**	**500ms**
Mean RT for the reversed words	935 ms	951 ms	803 ms
Mean RT for the original words	896 ms	883 ms	782 ms
on the reversed words	the frequency of the reversed word (-0.02)	the frequency of the reversed word (-0.016); the frequency of the initial character (-0.02); the number of strokes of the initial character (0.006)	the frequency of the reversed word (-0.01); the frequency of the initial character (-0.03)
on the original words	the frequency of the original word (-0.02); the frequency of the last character(-0.02)	the frequency of the initial character (-0.03)	the frequency of the initial character (-0.03)

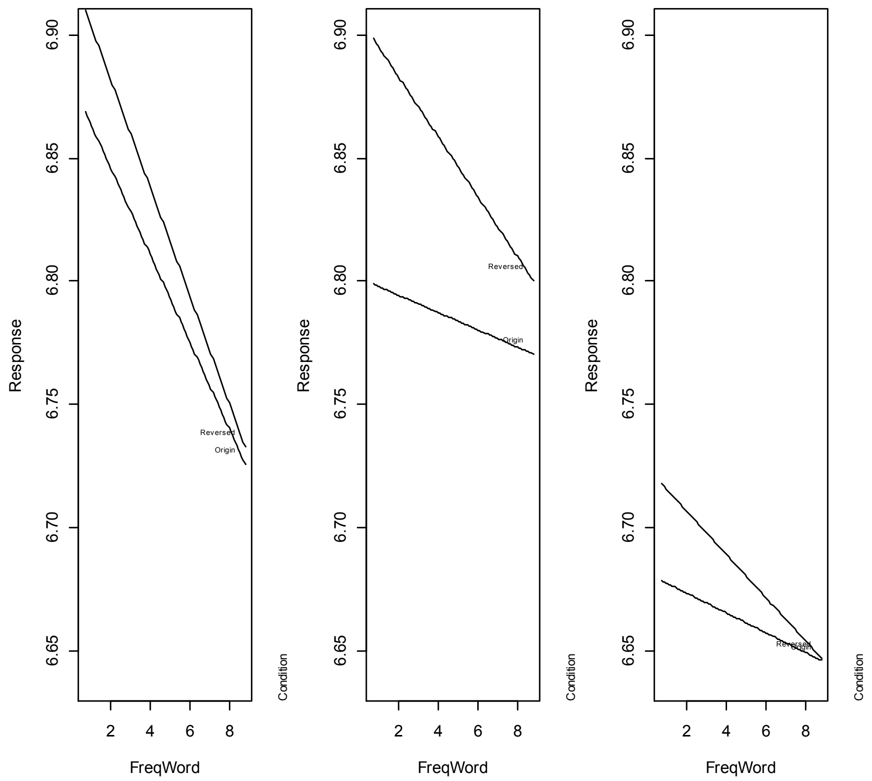

Figure 5.4 The comparison of the regression lines for the target word frequency (*FreqWord*) as predictor of naming latencies (*Response*) across three sub-experiments in Experiment 3

As in Experiments 1 and 2, manipulation of the speed of presentation led to different patterns of results. At the most rapid presentation rate (100ms), naming latency decreased with increasing frequency of the target word, irrespective of whether the target word was identical to the prime or reversed. The difference between the two experimental conditions was that only in the related condition (where prime and target are identical), the frequency of the head of the compound word showed significant facilitation. Presentation of the compound in reversed order eliminated this effect.

At longer presentation rates, the first character of the target word revealed significant facilitation for both conditions. At these longer presentation rates, the target word's frequency was predictive only when naming the reversed words, which have different meanings from the un-reversed prime words (cf. 领带*, a tie;* 带领*, to lead*).

These results are consistent with the hypothesis that the compound's meaning is subject to rapid decay when read in sentential context. Two pieces of evidence are important. First, in reading the repeated original word (the related condition), the compound frequency effect was only found at the fastest presentation rate (i.e. 100ms), but not at longer presentation durations. Second, for the related condition, when exposure duration was 100ms, the frequency of the second character of the prime compound word, which in about 70% of the cases was the head of that compound, was facilitatory. Since, the frequency of the second character was never predictive at longer presentation rates, and since the same holds for the frequency of the compound, it seems likely that the frequency effect of the second character also reflects short-lived semantic activation, but now not of the compound itself, but of its super-ordinate semantic class. The more well-established the super-ordinate class is, the faster a response can be initiated. Considered jointly, these two results – which are characteristic of reading a word read previously in the sentence – support the hypothesis that the meanings of the words in the sentence have been accessed by the time the response to the target is executed. The disappearance of these effects at longer exposure durations fits well with the hypothesis that, once activated, words' semantics are subject to fast decay.

Table 5.5 further clarifies that the slope of the word frequency effect is attenuated for increasing exposure duration in the unrelated condition, slopes from -0.02 to -0.016 to -0.01. This suggests that at longer exposure durations, participants begin to anticipate the appearance of reversed compounds. This anticipation makes a context-free, a priori estimate of the (reversed) compound increasingly less useful, which in turn results in an increasingly weak effect of frequency (probability) of occurrence.

When the presentation rates were longer (200ms and 500ms), the frequency of the initial character (and at 200ms also its number of strokes) was predictive irrespective of whether the compound was reversed. Apparently, at these longer presentation rates, access to the first character’s phonology becomes important. The more frequent the first character is, the faster its phonology can be accessed, and the faster the naming response can be initiated.

Chapter 6 Experiment 4: Character in Character

After having examined the consequences of reading two-character compounds in a sentence for the reading aloud of a character (Experiments 1 and 2) or a two-character compound (Experiment 3), Experiment 4 investigates whether characters that are embedded in other characters give rise to a similar pattern of results as observed in Experiment 1. In what follows, I use the term "component character" to refer to characters in other characters, and the term "matrix character" to refer to the character in which other characters are embedded. If component characters are functional units in the same way as the character constituents of two-character compounds when reading sentences, a facilitatory effect of the matrix character is expected in the related condition (in which a component character is to be read out loud), and there should be an inhibitory effect of prime matrix character frequency for the response time in the unrelated condition (in which a new character that did not appear in the preceding sentence, nor in any of the characters in the preceding sentence).

6.1 Experiment 4a: 100ms per 3-character group

6.1.1 Method

Participants

Twenty native speakers of Mandarin Chinese (10 female and 10 male) participated in Experiment 4a. They had normal or corrected-to-normal vision. Those who did not participate on a voluntary basis received course credits for participation. Their average age was 22.8 (SD= 5.21). Participants had lived in Canada for an average of 1.92 years (SD=2.11).

Stimuli and Apparatus

The stimuli consisted of 60 sentences, each with a length of 9 characters, which were read three at a time in rapid serial succession. All the stimuli used in Experiment 4 (with their English translations) are shown in Appendix 5. In every sentence, the matrix character prime was embedded in the second 3-character group at the central position, in the most neutral position of the total sentence. Each of the matrix characters had the common left-right component structure. The component character selected as target for naming was always the left component of the matrix character. For this experiment, 60 Chinese two-component matrix characters were selected from the CNCORPUS. The frequencies of matrix characters and the component characters were also extracted from this corpus. The family size of components (i.e., the number of matrix characters that a component character is a part of) was determined by consulting the GF 0014-2009 Specification of Common Modern Chinese Character Components (Ministry of Education, 2009).

Apparatus was the same as Experiment 1.

Procedure

The procedure was identical to that of Experiment 1.

Linguistic predictors

The linguistic covariates for Experiment 4 included the frequency of the two-component matrix character (*FreqChar*), the frequency of the component character target (*FreqComp*), the frequency of the target character presented in the unrelated condition (*Freqnr*), the number of strokes of the two-component matrix character (*NstrokeChar*), the number of strokes of the component target character

(*NstrokeComp*), the number of strokes of the target character for the unrelated condition (*Nstrokenr*), and the family size of the component character (*CompFam*). Table 6.1 presents the pairwise correlations of the predictors. As all correlations are low or moderate (<0.7), no issues of multicollinearity arose.

Table 6.1

Correlation matrix between pairs of variables

	1	2	3	4	5	6	7
1.NstrokeChar	1	0.57	0.05	-0.13	-0.05	-0.29	-0.36
2.NstrokeComp	0.57	1	-0.07	-0.11	-0.01	-0.14	-0.52
3.Nstrokenr	0.05	-0.07	1	0.27	0.11	-0.14	0.24
4.FreqChar	-0.13	-0.11	0.27	1	0.1	-0.2	0.23
5.FreqComp	-0.05	-0.01	0.11	0.1	1	0.09	0.1
6.Freqnr	-0.29	-0.14	-0.14	-0.2	0.09	1	0.11
7.CompFam	-0.36	-0.52	0.24	0.23	0.1	0.11	1

6.1.2 Results and Discussion

Using the same criteria as for the preceding experiments, 26 data points were lost (2.17%). The remaining 1,174 data points were subjected to a mixed-effects analysis of covariance. Table 6.2 provides a summary of the main pattern of results, which is straightforward: the only significant predictor was the frequency of the character to be read out loud, with slopes that did not differ across the related and unrelated conditions (t=-1.48, p=0.14). On average, the response latencies in the related condition were shorter than those in the unrelated condition, but this difference did not reach full significance (t=1.7, p=0.08).

Table 6.2

Comparison of the significant variables between two conditions in experiment 4a

Condition	Significant variables	Mean RT (ms)
Related (component) characters	the frequency of the component character (t= -4.39)	888
Unrelated (previously unseen) characters	the frequency of the unrelated character (t= -4.54)	900

I conclude that there is no solid evidence that, when reading is rapid, with 100ms exposure durations, naming latencies are affected by whether or not a character occurred as component character earlier in the sentence.

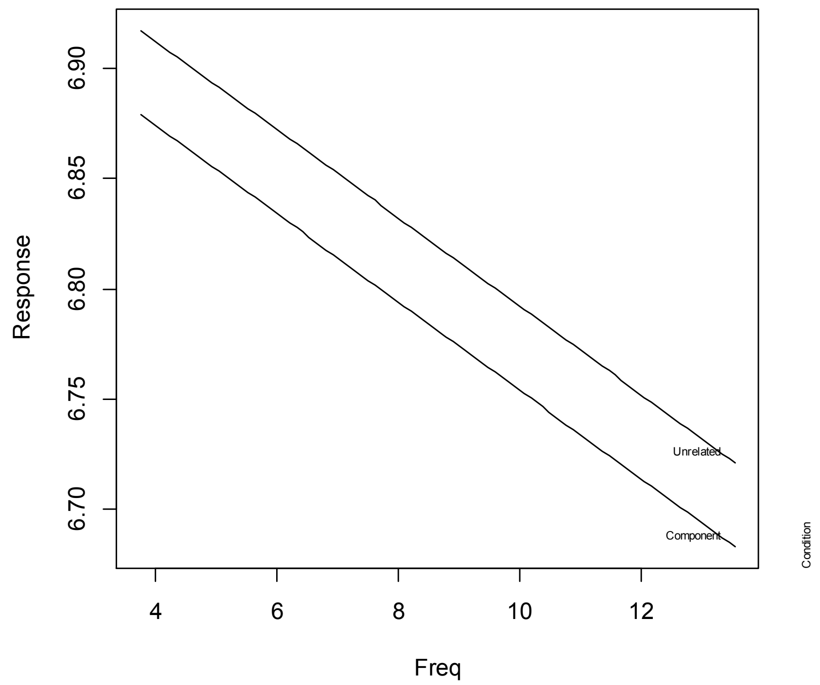

Figure 6.1 Regression lines for character frequency (*Freq*) as predictor of naming latencies (*Response*) in experiment 4a (100ms).

6.2 Experiment 4b: 200ms per 3-character group

6.2.1 Method

Participants

A total of twenty native speakers of Mandarin Chinese took part in this sub-experiment (11 female and 9 male), all with normal or corrected-to-normal vision. They received course credit for their participation, or voluntarily took part in the study.

The average age was 26.7 (SD= 8). Participants had lived in Canada for an average of 2.37 years (SD=2.04).

Stimuli and apparatus

The stimuli and the apparatus were the same as Experiment 4a.

Procedure

The procedure was the same as that of Experiment 4a, but the presentation duration of each 3-character group was changed from 100ms in Experiment 4a to 200ms.

6.2.2 Results and Discussion

A total of 1,175 valid data points remained after removal of errors, missing responses, and overly short responses (RT less than 300 ms, 25 data points, 2.1%). A summary of the results is provided by Table 6.3. For both the related and unrelated conditions, the only covariate that reached significance was the frequency of the target character. Unsurprisingly, more frequent characters elicited shorter response latencies. There was no evidence of any effect of the matrix character prime on the naming latency for the target character.

Table 6.3

Comparison of the significant variables between two conditions in experiment 4b

Condition	Significant variables	Mean RT (ms)
Related (component) characters	the frequency of the component character (t= -5.70)	819
Unrelated (previously unseen) characters	the frequency of the unrelated character (t= -5.89)	845

Figure 6.2 presents the regression lines for target character frequency for the two conditions. The slopes of the two lines (-0.02 for unrelated and -0.03 for component) are not significantly different from each other (t=0.65, p=0.52). In this sub-experiment, the average response time in the unrelated condition was significantly longer (on average, by 26ms, t=2.35, p=0.02) than that in the related condition.

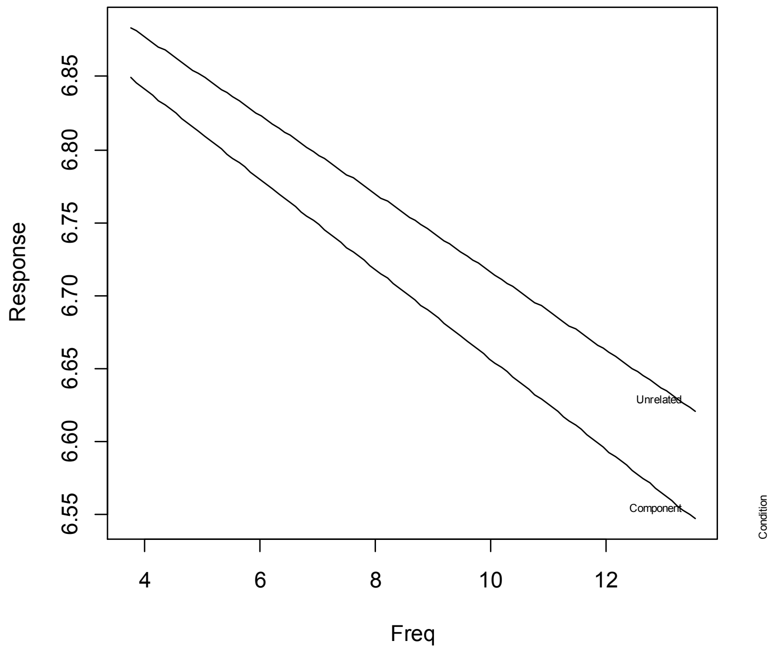

Figure 6.2 Regression lines for character frequency (*Freq*) as predictor of naming latencies (*Response*) in experiment 4b (200ms).

6.3 Experiment 4c: 500ms per 3-character group

6.3.1 Method

Participants

Twenty native Mandarin Chinese (16 female and 4 male) from Edmonton took part in the present sub-experiment. The average age was 21.1 (SD= 3.46). Participants had lived in Canada for an average of 2.85 years (SD=2.52).

Stimuli and apparatus

The stimuli and the apparatus were the same as those used for Experiment 4a

Procedure

The procedure was the same as that implemented for Experiment 4a, but the presentation duration of each 3-character group was changed from 100ms in Experiment 4a to 500ms.

6.3.2 Results and Discussion

After removal of 19 data points (1.58%, due to errors, missing responses, or extremely short response latencies less than 300 ms), 1,181 data points were available for statistical analysis. Table 6.4 compares the significant linguistic predictors in Experiment 1c.

Table 6.4
Comparison of the significant variables between two conditions in experiment 4c

Condition	Significant variables	Mean RT (ms)
Related (component characters)	the frequency of the component character (t= -7.83) the number of strokes of the matrix character (t=-2.26)	834
Unrelated (previously unseen) characters	the frequency of the unrelated character (t= -7.94)	852

Results mirrored those obtained for the 200ms exposure duration: responses were slower in the unrelated condition, and target character frequency provided facilitation with statistically indistinguishable slopes (t=0.78, p=0.44, see Figure 6.3). A new finding is that at this long presentation duration, the number of strokes of the matrix character reached significance, but only in the related condition. Target characters that are part of matrix compounds with higher stroke complexity were named more quickly. This is likely to be a task effect, with readers having time to ponder the characters and their components while reading the sentence, after having noticed (thanks to the slow pacing of reading) that character components were later presented for reading out loud.

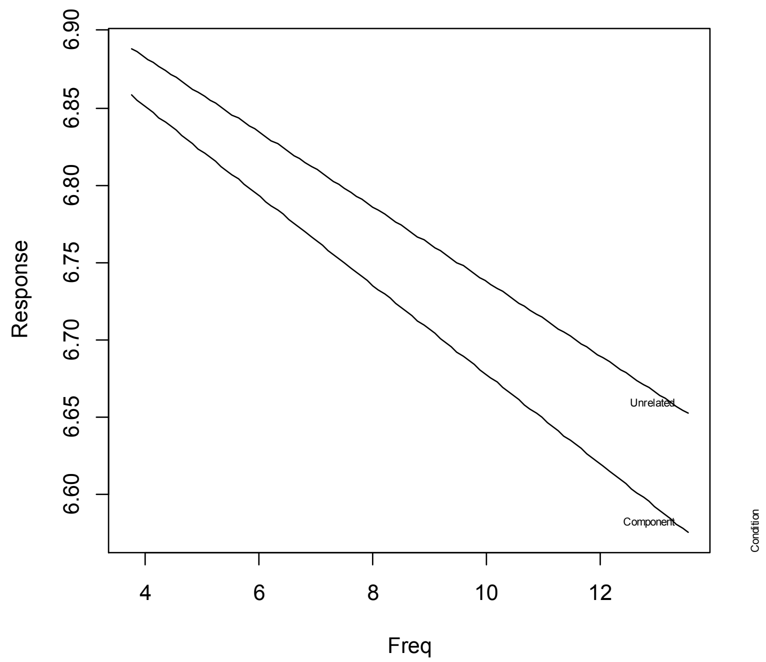

Figure 6.3 the comparison of target character frequency (*Freq*) effect on naming latencies (*Response*) in experiment 4c (500ms)

6.4 General Discussion of Experiment 4

Table 6.5 provides a summary overview of the results of the three sub-experiments with matrix characters and their components. The numbers in parentheses are the coefficients for the slopes of the covariates. Figure 6.4 shows how slopes change with exposure durations.

Table 6.5

Summary of Experiment 4

Condition	**100ms**	**200ms**	**500ms**
Mean RT for related characters	888 ms	819 ms	834 ms
Mean RT for unrelated characters	900 ms	835 ms	852 ms
on the related (constituent) characters	the frequency of the component character (-0.02)	the frequency of the component character (-0.03)	the frequency of the component character (-0.03); the number of strokes of the matrix character (-0.01)
on the unrelated characters	the frequency of the unrelated character (-0.02)	the frequency of the unrelated character (-0.02)	the frequency of the unrelated character (-0.03)

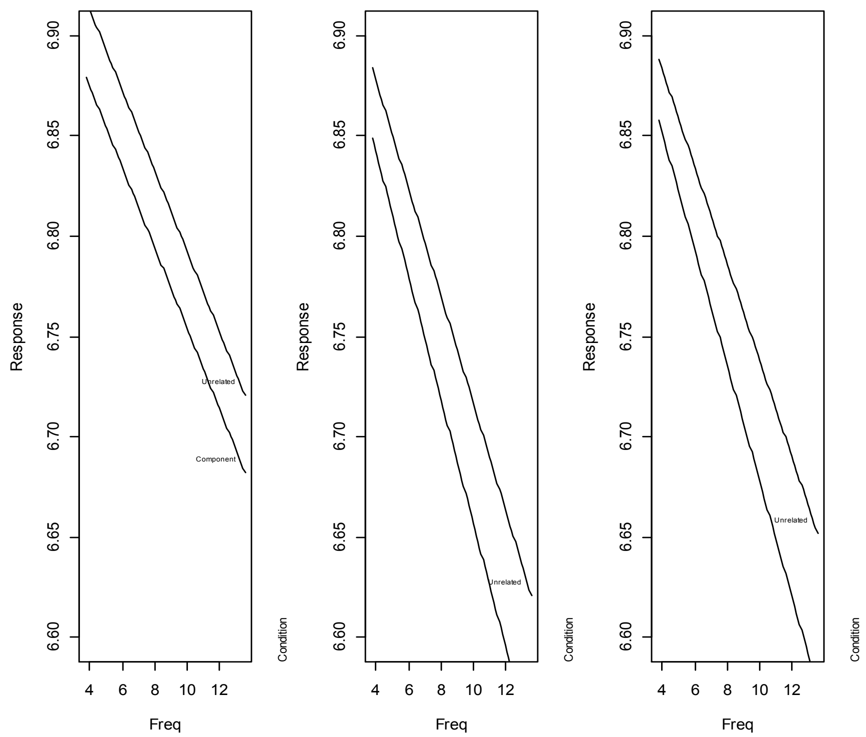

Figure 6.4 Comparison of the regression lines for character frequency (*Freq*) as predictor of naming latencies (*Response*) across three sub-experiments in Experiment 4

Compared to the previous three experiments with two-character compound primes, the effects of the linguistic variables on the naming latencies in Experiment 4 did not change substantially when the reading speed of the RSVP was manipulated. Across all the three sub-experiments, the frequency of the target character was significant for naming, both for component characters in the related condition and for previously unseen characters in the unrelated condition. The slope of target character frequency did not vary significantly across the three sub-experiments and also did not vary between conditions. This suggests that the target characters were read independently of their primes in the sentences. The frequency of the matrix character prime never reached significance.

Two findings suggest that there was some awareness on the part of the participants of the prior occurrence of component characters in matrix characters read in the sentence. First, response latencies tended to be longer in the unrelated condition, not significantly so at the shortest exposure duration, but well supported at the longer exposure durations. This suggests some kind of priming is occurring. The emergence of a matrix character stroke complexity effect at the longest exposure duration in the related condition sheds further light on what might be going on. Subjects may have become aware of the experimental manipulation in the related condition. This may have led to attention being directed to character components during reading with longer exposure durations. Furthermore, backward priming (Koriat, 1981; Peterson and Simpson, 1989; Seidenberg et al.,1984; Shelton and Martin, 1992) may have taken place, with subjects realizing that they had seen the character before, which may have given them confidence to proceed more quickly with naming the character.

In summary, there is little evidence in this experiment that in normal reading of sentences, the characters embedded in other characters would function as units with semantic consequences for subsequent processing. Instead, it seems likely that the component characters in Experiment 4 are processed as meaningless orthographic units comparable to letters in English words. The pattern of results of the present experiment differs fundamentally from that of Experiment 1 (and 2). Etymologically, the target character may be semantically related to its matrix character, but we see no evidence that, during normal reading of sentences for comprehension, these kinds of relations would play a role.

Chapter 7 General Discussion

This final chapter of my thesis begins with a discussion of the results of Experiments 1-4 against the background of current models of lexical access. I first argue in section 7.1 that none of these models succeeds in explaining the full pattern of results obtained. In section 7.2, I further develop the hypothesis proposed in the preceding chapters that my Experiments are tapping into post-access processes, rather than into processes of lexical access. The final section places this line of research in a broader context.

7.1 **Inadequacy of the current lexical access models**

As discussed in the literature review (Chapter 2), over the past decades of psycholinguistic research, different models of visual lexical processing have been proposed. Central issues concern whether morphologically complex written words are initially and obligatorily decomposed and recognized on the basis of their morphemic subunits or whether they are directly accessed as whole words. A further issue concerns how semantics is involved in word recognition and processing. Current influential models of lexical processing include the sub-lexical model (which posits obligatory initial decomposition followed by whole-word look-up), the supra-lexical model (which posits whole-word look-up followed by later decomposition), dual route models (which posits involvement of both sub-units and whole word at the initial stages of lexical processes) and the triangle model (a parallel-distributed-processing model without morphemes or constituents). In what follows, I first examine whether these morphological models help explain my findings.

7.1.1 Sub-lexical model (the obligatory decomposition model)

This class of models assumes that morphologically complex words are first automatically decomposed into their morphemic subunits, which then in turn activate the lexical representation of the whole word (Taft, 1994, 2003; Taft and Zhu, 1995). A classic paper applying this approach to reading Chinese compounds is the Multilevel Interactive Activation Model by Taft et al. (1999). Figure 7.1 illustrates the architecture of this model, which contains several layers of orthographic representations: a layer of stroke units, which connects to a layer of radical units, which are linked up to morphemic units (constituent character), which in turn are linked to compound word units. The units in each pair of successive layers are connected to each other, allowing activation to flow between them. Morphemes and words are linked to both semantic and phonological representations. Decomposition is obligatory in this theory. The processing of compound words can only be realized through pre-lexical parsing of the two-character compound word into morphemes (here, constituent characters). Such a pre-lexical mechanism is regarded as being semantically "blind", in that the parsing process uniquely relies on the orthographic characteristics of the morphemes and characters (morpho-orthographic decomposition).

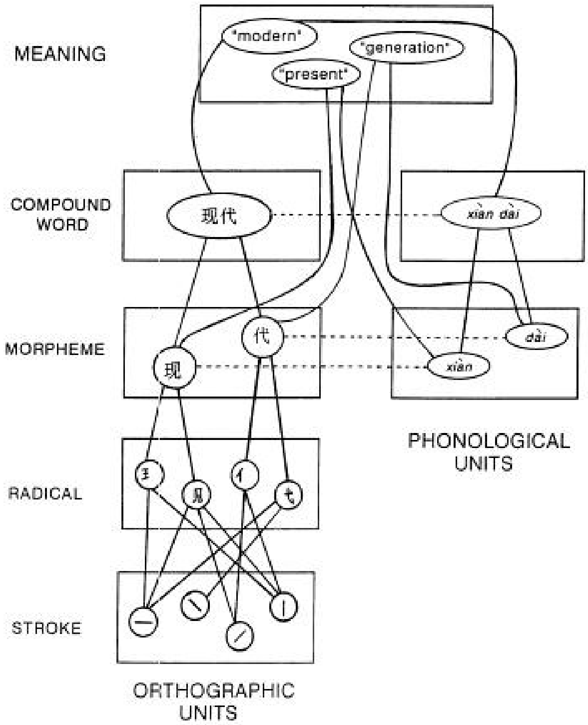

Figure 7.1 The Multilevel Interactive Activation Model (Taft et al., 1999，Fig.5.1, p95)

However, the Multilevel Interactive Activation Model does not explain the full pattern of results of the present study.

First, recall that both Experiments 1 and 2 showed a significant inhibitory effect from the prime word onto the target character in the unrelated condition. This inhibitory effect is not predicted by the model. It might be argued that the compound, once activated, sends activation back to its constituent morphemes, and that subsequently these morphemes inhibit unrelated characters thanks to the within-layer inhibitory links between morphemes. However, such an explanation seems post-hoc and it remains to be shown that in an actual computational implementation this would really work.

Second, if pre-lexical decomposition is obligatory, the effect of target character frequency should increase in strength for longer exposure of the prime, i.e., with slower presentation of the sentence. The assumption here is that for very rapid serial presentation, there is little time for activation to spread from the strokes to the radicals, and from the radicals to the characters. At slower presentation rates, therefore, the effects of constituent frequency should dominate. This is not what I found in Experiments 1 and 2. In the related condition, the slope of the character frequency remained constant. In the unrelated condition, where no change was predicted because there was no priming, the magnitude of the slope of the character frequency effect increased with presentation duration, unexpectedly. Perhaps, in the unrelated condition, the target characters were inhibited more strongly for faster presentation rates. Furthermore, this does not explain why in the related condition the slope of the frequency of the character was not modulated by the presentation rates.

Third, Taft's model predicts a late effect of compound frequency. This prediction would have been more consistent with the compound frequency effect at the slower representation rates instead of the observed effect at the fast representation rate only.

Fourth, since Taft (1997) argued that constituent characters and compounds should facilitate each other as units located at different levels, the compound words should always facilitate their constituent characters. The more compound words are stored in the mental lexicon, the more facilitation should be obtained, hence, the family size should be facilitatory. However, this was opposite to what we found in Experiment 2, where the family size of the constituent character was significantly inhibitory. The multilevel interactive activation model also implies that, between any particular two units at the same level, there should be inhibition. The relationship between this within-level inhibition and the facilitatory role of family size from the word level to character level seems to be internally inconsistent.

7.1.2 Supra-lexical model

The second type of model provides a supra-lexical account of morphological decomposition. It is the exact opposite to the sub-lexical models in terms of the processing order of the word and its sub-morphemic units. According to the supra-lexical hypothesis, the decomposition of a word occurs only after the whole word has been accessed in the lexicon (Giraudo & Grainger, 2001, 2003). The morphemic representations then in turn activate higher level semantic representations, which send back activation to corresponding form representations. Figure 7.2 presents the basic structure of the model. This account thus differs from the sub-lexical models in that morphological decomposition involves a semantically based search for morphemes (morpho-semantic decomposition)

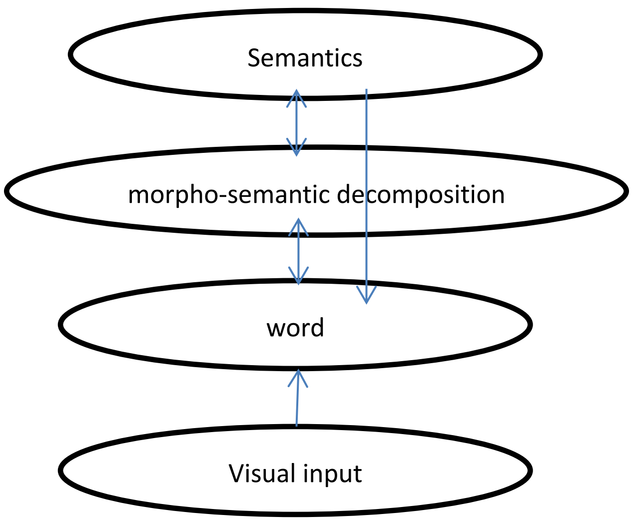

Figure 7.2 the supra-lexical model (Giraudo & Grainger, 2001, 2003)

The supra-lexical model fits well with the fast decay of the frequency effect of the two-character compound prime. The compound representation is the first to be accessed and hence the first to decay. However, two findings are not well explained.

First, the inhibitory effect from the two-character compound prime in the unrelated condition of Experiment 1 (100 ms exposure duration) for the target character and the facilitation for the related target character remains unexplained. If the compound word is accessed first, and subsequently decomposed into constituent meanings, these constituent meanings should make it easier to read out aloud the character presented in the related condition. However, one would not expect any inhibition as observed in the unrelated condition.

Secondly, in Experiments 1 and 2, the slope of the target character's frequency was larger than that of the prime compound's frequency at the shortest exposure duration (100ms). If the word representation is activated first, one might have expected the slope for the compound to be the larger one.

7.1.3 Dual route models

The term "dual route model" refers to the three different kinds of models. For speech production, Pinker's *words and rules* model (2011) has two routes, a route using rules for regular complex words and a route using lexical look-up for all other words. As the primes in Experiments 1,2 and 3 are all morphologically complex, they would all be processed in the same way, namely, by rule. Transposed into a comprehension model, it becomes indistinguishable from an obligatory decomposition model. As discussed above, such a model does not offer useful insight into the variegated pattern of results that we obtained.

A second kind of dual route model is the model for reading aloud developed by Coltheart et al. (2001). Their Dual Route Cascaded Model (DRC) contains a lexical route and a non-lexical Grapheme-Phoneme Conversion (GPC) route. An attempt to adopt the DRC to reading Chinese words, strictly following Coltheart's model for English, would result in a model for the reading aloud of Chinese compound words as shown in Figure 7.4.

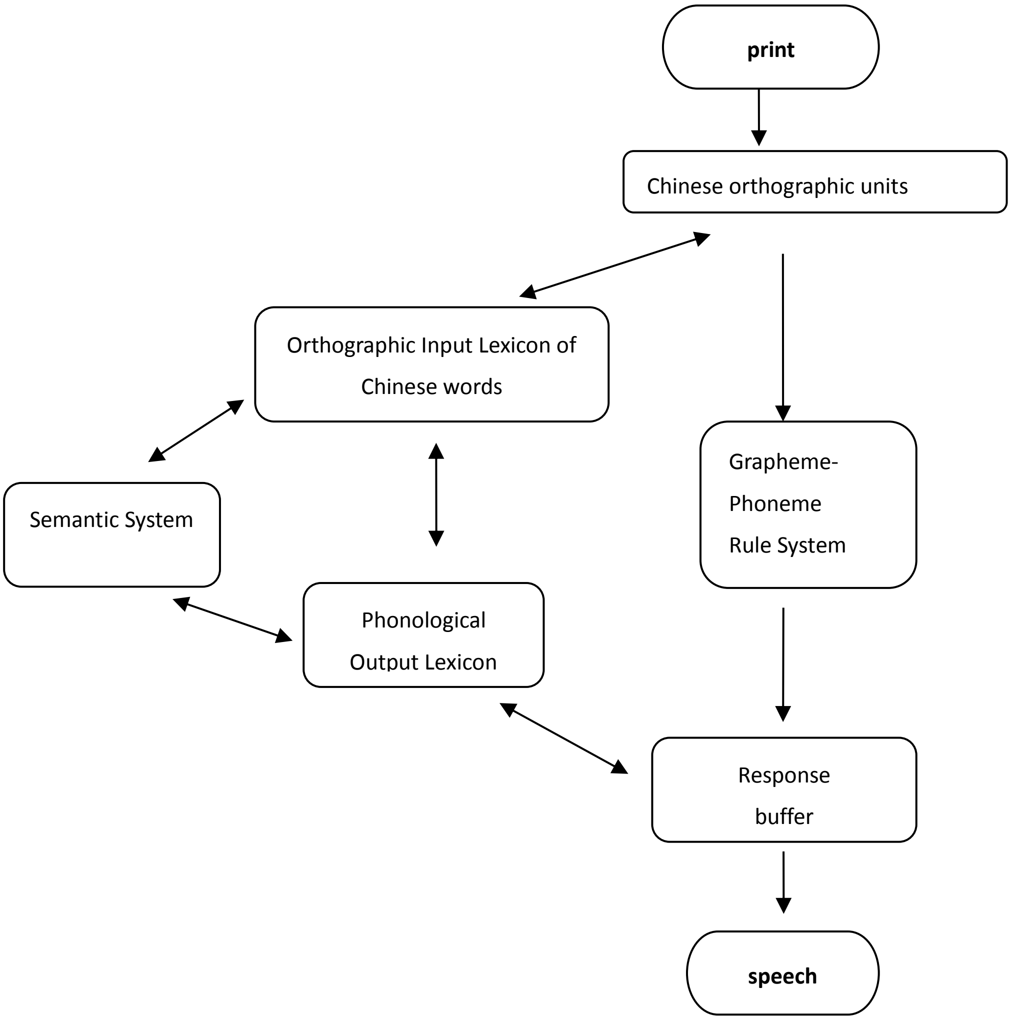

Figure 7.3 An attempted application of DRC to reading Chinese compound words.

However, Grapheme-Phoneme Conversion (GPC) is not typically suitable for Chinese due to the lack of a GPC route. As Coltheart (2001) himself mentioned in his article, “The Chinese, Japanese, and Korean writing system are structurally so different from English writing system that a model like the DRC model would simply not be applicable: For example, monosyllabic non-words cannot be even written in the Chinese script or in Japanese kanji, so the distinction between a lexical and non-lexical route for reading aloud cannot even arise.” (p.236) For Chinese, grapheme-to-phoneme conversion would have to apply to characters, but here rules are more difficult to formulate than for English by several orders of magnitude.

Therefore, Chinese word recognition through dual routes remains questionable.

Furthermore, the DRC model is concerned mainly with mono-morphemic words, and does not further our understanding of the processing of morphologically complex words. More specifically, it does not clarify how the silent reading of a two-character compound prime might affect subsequent reading aloud of a constituent character as tested in Experiments 1 and 2. It also does not shed light on how the silent reading of a two-character compound word might subsequently determine the processing of another two-character (non-reversed, or reversed) compound as tested in Experiment 3.

Third, the DRC, just as all other models of lexical access that I am aware of, remains silent about a possible role of contextual knowledge during lexical access. In order to explain the whole word frequency effect on naming latencies at the fastest reading speed (i.e. 100ms), for instance, the DRC model would need to be extended with some short-term memory for the preceding discourse.

The third kind of dual route model is the parallel dual-route model proposed by Baayen et al., (1997). According to this model, the access of a word's meaning would proceed through either that word's constituents or through its whole-word representation. Which route completes lexical access first would depend on the frequencies of the constituents and of the compound. Although the model might predict facilitation from previous reading of a two-character compound prime, it cannot help us understand why in Experiment 1 the reading of a target character in the unrelated condition is adversely affected. Furthermore, the inhibitory effect of family size observed in Experiment 2 is incompatible with this model's predictions.

7.1.4 Triangle Model (parallel-distributed-processing)

Seidenberg and McClelland (1989) developed a connectionist model, with three main banks of units for orthography, phonology and semantics, in a triangular arrangement as shown in Figure 7.5. All these units are connected to sets of hidden units which are part of the system, but that do not have connections to other external systems. The weights in the connections between the orthographic input units, hidden units, semantic units and phonological output units are adjusted through training. During training, the model is presented with an orthographic letter string which leads to activation of, for instance, the phonological units. The back-propagation algorithm ensures that the weights on the connections in the network are adjusted, gradually towards the desired output.

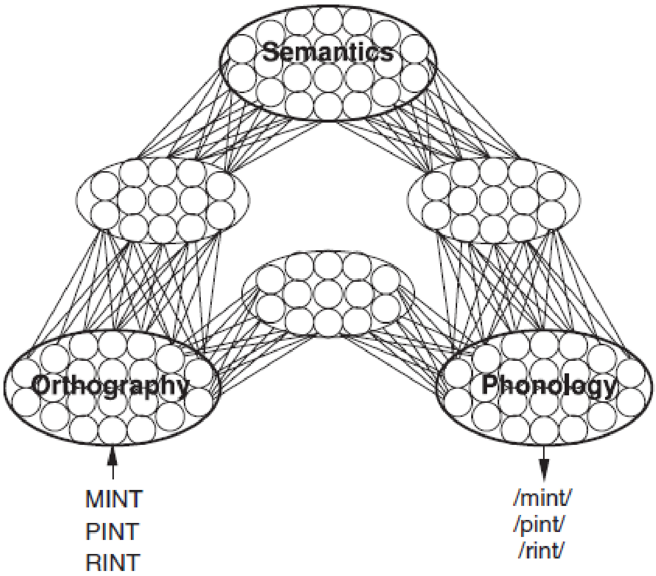

Figure 7.4 The triangle connectionist framework for lexical processing (Seidenberg and McClelland, 1989)

The triangle model has been shown to account for a wide range of experimental data (Seidenberg, and McClelland 1989; Makague et al., 2001; Seidenberg, 2005). However, just as the above mentioned models, it cannot provide an explicit reason for the opposing effects of compound frequency at the fastest reading rate, the family size effect and the change of target character frequency slope.

7.2 Compound processing in sentential environment

Summing up the preceding section, current models of lexical access do not offer a satisfactory account for the main findings in my experiments. This may be due to these models being incorrect, but it may also be due to models for lexical access being applied to data that are primarily revealing about post-access processes. The original predictions for my experiments were framed from the perspective that lexical access would be relatively slow, with semantic access being completed around 400ms post stimulus onset (compare the N400 effect in electrophysiological studies). However, since the present experiments testify that whole sentences can be read with good comprehension within 300ms, lexical access must take place more rapidly --- otherwise, a stack of semantically unprocessed words would accumulate in the course of rapid reading. Importantly, since reading for comprehension can proceed this rapidly, with an exposure duration of 100ms for three characters, the moment in time at which in Experiment 1 a stimulus is presented for naming is likely to occur after lexical access to the two-character compound prime has been already fully completed. Thus, at the shortest exposure duration in rapid reading, the short-lived activation of the two-character compound prime is still felt, but at longer exposure durations, due to activation decay, the consequences of prior understanding of the compound are no longer detectable.

The hypothesis that the present experiments primarily tap into post-access processes sheds light on why so many of the lexical predictors that have been observed in experiments using isolated word presentation failed to be significant for my data - to my surprise.

Word recognition is typically described as the ability of a reader to recognize written words correctly and virtually effortlessly. It is sometimes referred to as isolated word recognition because it involves a reader's ability to recognize words individually from a list without needing similar words for contextual help (Kruidenier, 2002). In the past 30 years, research on isolated word recognition paradigms has led to models in which the rich information characterizing individual words is dissected and assigned to many different levels of representation, with orthographic features being distinguished from morphological units and from phonological information. According to these models, readers need to access the information at these different levels, and to re-synthesize the resulting streams in order to access a word's meaning. However, outside the lab, the reading of isolated words is seldom necessary. Studies on isolated word recognition may give us hints for understanding lexical processing, but their ecological validity is questionable, especially for a language such as Chinese, for which the interpretation of characters often crucially depends on the neighboring characters.

The present study sought to understand lexical processing by presenting words in a more ecologically valid sentential environment. With sentential context facilitating anticipation and interpretation of often highly ambiguous characters, it is perhaps not surprising that, across all the 12 sub-experiments, many of the predictors typically reported to be predictive for isolated word reading in Chinese failed to reach significance (e.g. syllable frequency, family size of the constituent characters/components, the complexity of constituents characters/components, see Duanmu, 1998; Ge, 1985; Hsu et al.,2009; Inhoff & Liu,1998; Myers, 2006, 2010; Packard, 2000; Perfetti & Tan, 1999; Peng et al.,1999; Taft, et al., 1999; Wu et al., 1999; Zhan et al., 2013; Zhou et al., 1999; Zhou and Marslen-Wilson, 1995,2000). When sentential context helps readers anticipate upcoming words, and helps resolve ambiguity of interpretation, many of the lexical-distributional properties of radicals, components, strokes, and characters become much less important, if not irrelevant.

What then are the post-access processes involved in the present experiments, and what kinds of representations are involved in these processes? First consider Experiment 4. Matrix characters with two component characters appear to have been read in the rapid exposure condition of 100ms without readers realizing that the character presented to them for naming had been presented before as a character component. Given the very different meanings that components and matrix characters can have, it seems likely that readers retrieved a matrix character's meaning straight from its full character, in all likelihood while taking into account the flanking characters and the preceding sentential context. For rapid reading, I do not have evidence for or against the activation of the component characters, but in the absence of any evidence pro or contra, the most parsimonious solution is to assume that multi-component matrix characters are not understood by an initial dissection into components, followed by a hash-table like component-driven lookup of the compound's meaning. The absence of robust priming effects at the 100ms exposure duration in Experiment 4 argues against this scenario. I conclude that the character is a likely candidate for a linguistic sign, with a holistic form representation (the written character) and a holistic meaning. (This is, of course, a simplified set-up, as many characters are multiply ambiguous.)

Next, consider the reading of two-character compounds. The main pattern of results is as follows. First, at the shortest exposure duration (100ms), the frequency of the two-character compound prime is predictive; at longer exposure durations, it is not.

The recurrent predictor at longer fixation durations is the frequency of the target character. Second, splitting the two-character compound prime across two lines of text still allows an effect of the frequency of the two-character compound to emerge. Third, in the related condition, the slope of the frequency effect of the two-character compound prime is negative, whereas in the unrelated condition, it is positive (albeit not significantly so in Experiment 2). I maintain that this pattern of results is best understood as follows. At the initial stage of processing, lexical access takes place, resulting in the activation of the meaning of the two-character compound prime. Whether this activation process was mediated by decomposition or not is an issue that my data do not speak to, and that I will therefore remain silent about. As word meanings become available, they have to be integrated in the reader's understanding of the sentence. Once integrated into the sentence/discourse, these meanings must decay to make place for new meanings. Without decay, working memory would become cluttered with ever increasing numbers of meanings as we read through sentences and texts.

My experiments suggest that the meaning of the two-character compound prime remains available within a time window of 200 to 400 ms post stimulus onset. This time window is deduced from the data as follows, under the assumption that 100ms suffices for understanding the characters presented in groups of three in the experiment (since people can actually understand the sentences in the fastest presentation condition). Therefore, the meaning of the target character (presented for reading aloud) must also be available after 100 ms. It follows that, in the fastest presentation condition, the meaning of the prime is available 100ms after prime stimulus onset, and that the meaning of the target is available 200 ms later. This is the lower boundary for the time period during which the meaning of the prime is available. Table 7.1 lays out this time-course.

Table 7.1

Time course of Experiment 1a (100ms) between the prime word and target character.

Time point	Event
0ms	visual presentation of the two-character compound prime in the second three-character group
100ms	meaning of the prime is available, last three characters of sentence shown
200ms	visual presentation of the target character for naming
300ms	meaning of the target is available

When the exposure duration is 200ms, a period of 400ms elapses between the activation of the prime's meaning and the activation of the meaning of the target, as shown in Table 7.2

Table 7.2

Time course of Experiment 1b (200ms) between the prime word and target character.

Time point	Event
0ms	visual presentation of the two-character compound prime in the second three-character group
100ms	meaning of the prime is available
200ms	visual presentation of the final three characters in the sentence
400ms	presentation of the target character for naming
500ms	meaning of the target character is available

Since the prime's frequency is not predictive when the exposure duration is 200ms, the upper boundary for the availability of the prime's meaning must be 400ms. Therefore, the time window for semantic priming ranges from 200 ms to 400ms. When the prime is semantically related to the target, word naming is facilitated, but when it is not semantically unrelated, the meaning of the prime word is in the way, and delays the onset of articulation of the prime.

The small time window for the activation of the two-character compound prime's meaning and its interaction with the target's meaning suggests this is an effect arising in short-term memory (STM). Short-term memory is a memory system involved in storing small amounts of information for a brief period of time (Baddeley et al., 2009). Thus, whatever one has just read is in his or her short-term memory. Most of this information is eventually lost. STM has not only a limited duration but also a limited capacity of information (Goldstein, 2008), which fits well with the hypothesis that it is impossible to stack large numbers of meanings in STM, and that hence decay must take place.

When participants were naming an unrelated character in the fastest presentation rate in Experiment 1, inhibition was observed from the two-constituent character prime's frequency. This can perhaps be understood as an instantiation of anti-priming. Anti-priming (Marsolek, 2008) is the phenomenon that one’s ability to identify something is impaired by his or her identification of a previous object. For this explanation to hold, it would be necessary that the prime and target meanings share semantic features. Exploring this possibility is left to future research.

7.3 Task effects

Experiment 3 adds to the preceding set of finding by showing that reversing the order of the constituents gives rise to an effect of category congruency between prime and target in the related condition. This semantic effect is as short-lived as the prime frequency effect, and likely to be constrained by the abovementioned constraints on short-term memory. This headedness effect is likely to be a task effect, induced by a list composition in which a compound and its reverse twin occur in closer proximity to each other over the course of the experiment than is usual in normal text. By directing attention to whether the target denotes a category a subcategory of which was mentioned in the preceding sentence, subjects were able to optimize their naming latencies.

The inhibitory effect of character family size in Experiment 2 is also a candidate for a task effect. It arose only at longer exposure durations, with an additional 200ms (or 500ms) of processing time before presentation of the target compared to Experiment 1, due to the sentences in Experiment 2 having four sets of three characters, instead of three sets of three characters. As the target character was always the first character of the two-character compound prime, which in Experiment 2 appeared at the end of the second line of the sentence, subjects may have been anticipating (while waiting for the next line to appear) what compound might be at issue, using the left character's family as search space. After presentation of the head (on the third line of the sentence), the alternative continuations had to be discarded. The reappearance of the first character, now in isolation for reading out loud, must have been surprising, and rendered difficult by its family members being incongruent with the preceding sentence.

Finally, recall that in Experiments 1 and 2, at longer presentation rates, the magnitude of the facilitatory effect of the frequency of the target character to be read aloud increased for the unrelated condition. This indicates that at longer exposure durations, the preceding sentence offers reduced constraints on interpretation. As a consequence, the target character's a-priori probability could become a more robust estimator of naming difficulty.

7.4 Relevance of this study

This study has both theoretical and practical relevance.

The main highlights, from a theoretical perspective, are (I) the insight it offers on activation decay after lexical access has been completed, (II) the evidence it provides that whole-word frequency effects can be understood as semantic effects (which occur even when a compound is split out over two lines in the sentence), and (III) the rapidity with which Chinese sentences can be read with comprehension using rapid serial visual presentation. From a methodological perspective, studying the effects of priming in a sentence context is innovative, and, as I have shown, productive.

The practical significance of this study is that it provides evidence that Chinese can be read at a high speed without significant loss of comprehension. This information may be useful for designing software or apps intended to increase the efficiency with which Chinese is read. The RSVP technology is particularly good for smaller devices like smartphones and smart watches, and may help render superfluous actions such as scrolling and zooming. Our results suggest a possible Optimal Recognition Rate (ORS) of around 100ms for three-character groups for fast-reading apps on Chinese smart-phones, GPS and other small-screen readers. For English, several apps are available for rapid reading using RSVP. For example, the Spritz app, which is coming to the Samsung Galaxy S5 and Samsung Gear 2 watch, English words appear one at a time in rapid succession which amazingly allows the typical college-level readers to finish all 309 pages of "*Harry Potter and the Sorcerer's Stone*" in under 77 minutes (Kleinman, 2014). It seems worth exploring whether displaying text in RSVP mode will help Chinese people to read more efficiently in the modern demanding fast-pace world.

7.5 Limitation and further research

This experimental study is, of course, limited by the technique it uses, which turned out to be helpful for understanding post-access processing, but that does not shed light the process of lexical access itself. As a Chinese proverb says: “one’s strong point always reflects one's weak point at the same time”(人之所长，即其所短). The evidence for fast decay of compounds' meanings fits well with the short range at which spillover effects have been found in eye-tracking studies (usually to the next word, possibly to the word after that), but whether two-character compounds are read decompositionally is still an open question. I regard this as a limitation of the method. Therefore, in future research work, I hope to use other techniques, such as neuro-physiological methods (EEG, ERP, fMRI, etc.) and eye-tracking. From traces of oxygenation or patterns of eye fixations, more evidence can be collected to improve our understanding of how Chinese words are read.

Similar studies could be carried out to Chinese, Japanese, Korean, which constitute the main East Asian languages (also termed as CJK in the field of software and communications internationalization). These languages all have a shared characteristic: their writing systems all completely or partly use Chinese characters — Hànzì in Chinese, kanji in Japanese, hanja in Korean.

References

Arcodia , G. F. (2007). Chinese: a language of compound words? In Montermini, F. , Boyé , G. , and Hathout , N. (eds), *Selected proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville, MA: Cascadilla Proceedings Project, 79 – 90.

Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language, 37(1),* 94–117.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R.* Cambridge: Cambridge University Press.

Baddeley, A. D., Eysenck, M., & Anderson M. C. (2009). *Memory.* New York: Psychology Press.

Bai, X., Yan, G., Liversedge, S. P., Zang, C., & Rayner, K. (2008). Reading spaced and unspaced Chinese text: evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 34,*1277–1287.

Bates D, & Sarkar D (2007). *lme4: Linear Mixed-Effects Models Using S4 Classes*. R package version 0.9975 -12, URL http://CRAN.R-project.org/.

Beyersmann, E., Coltheart,M., and Castles, A. (2012). Parallel processing of whole words and morphemes in visual word recognition. *The Quarterly Journal OF Experimental Psychology,65(9),* 1798–1819.

Bowers, Jeffrey S.; Colin J. Davis; Derek A. Hanley. (2004).Automatic semantic activation of embedded words: Is there a “hat” in “that”? *Journal of Memory and Language. 52(1)*:131-143.

Branner , D. P. (2001). Chinese languages: classical Chinese. In Garry , J. and Galvez Rubino , C. R. (eds), *Facts about the World's Languages*. New York : H.

W. Wilson, 134 – 8

Ceccagno , A. and Basciano , B. (2007). Compound headedness in Chinese: an analysis of neologisms. *Morphology, 17:* 207 – 31.

Chen, H.-C., Cheung, S.L., & Flores d'Arcais, G.B. (1995). Orthographic and phonological activation in recognizing Chinese characters. *Psychological Research, 58,* 144–153.

Chen, H.-C., & Tang, C.-K. (1998). The effective visual field in Chinese. *Reading & Writing, 10*, 245-254.

Chen, H. C., and Shu, H. (2001). Lexical activation during the recognition of Chinese characters: evidence against early phonological activation. *Psychon. Bull. Rev. 8,* 511–518.

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, Train-Min and Jenn-Yeu Chen (2006) "Morphological encoding in the production of compound words in Mandarin Chinese", *Journal of Memory and Language 54*, 491-514.

Cheng C M. (1981). Perception of Chinese characters. *Acta Psychological Sinica, 23:* 137–153.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review, 108,* 204-256.

Da, Jun. (2004). A corpus-based study of character and bigram frequencies in Chinese e-texts and its implications for Chinese language instruction. In Zhang, Pu, Tianwei Xie and Juan Xu (Eds.), *The studies on the theory and methodology of the digitized Chinese teaching to foreigners: Proceedings of the 4th International Conference on New Technologies in Teaching and Learning Chinese,* 501-511. Beijing: The Tsinghua University Press,

Da, Jun. (2004). *Syllable frequencies irrespective of tones*.
Extracted from: http://lingua.mtsu.edu/chinese-computing/phonology/syllable.php.

Diependaele, K., Sandra, D., & Grainger, J. (2009). Semantic transparency and masked morphological priming: The case of prefixed words. *Memory & Cognition, 37(6),* 895–908.

Dictionary of Chinese Character Information.(1988). Beijing, China: Science Press.

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.

Egly, R., Driver, J., &Rafal, R. D. (1994). Shifting visual-attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161–177.

Elisabeth Beyersmann , Max Coltheart & Anne Castles (2012) Parallel processing of whole words and morphemes in visual word recognition, *The Quarterly Journal of Experimental Psychology, 65:9*, 1798-1819.

Feldman, L.B. (1995). *Morphological aspects of language processing.* Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Feldman, L.B. (2003). Morphological processing as revealed through the repetition priming task. In J. Bowers and C. Marsolek (Eds.). *Rethinking Implicit Memory*. Oxford, UK: Oxford University Press.

Feldman, L.B. & Andjelkovi, D. (1992). morphological analysis in word recognition.

In L. Kats & R. Frost (Eds.), *Orthography, Phonology, Morphology, and Meaning, Advances in Psychology.* (pp.343-360). Amsterdam: Elsevier.

Feldman, L. B., O'Connor, P. A., & Moscoso del Prado Martin, F. (2009). Early morphological processing is morpho-semantic and not simply morpho-orthographic: A violation of form-then-meaning accounts of word recognition. *Psychonomic Bulletin & Review, 16(4),* 684–691.

Feldman, L. and Pastizzo, M.(2003). Morphological facilitation: The role of semantic transparency and family size. In R. H. Baayen and R. Schreuder (Eds.), *Morphological structure in language processing.* pp.233-258 .Berlin: Mouton de Gruyter.

Fiorentino, R. D. Poeppel (2007). Compound words and structure in the lexicon. *Language and Cognitive Processes, 22,* 953–1000

Ge, Benyi (1985). *Study of Chinese Words* (in Chinese). Jinan: Shandong Education Press.

Gilbert, Luther Calvin. (1959). Speed of processing visual stimuli and its relation to reading. *Journal of Educational Psychology ,Vol.50*, 8-14.

Giraudo, H., & Grainger, J. (2001). Priming complex words: Evidence for supralexical representation ofmorphology. Psychonomic *Bulletin & Review, 8,* 127–131.

Giraudo, H., &Grainger, J. (2003). A supralexical model for French derivational morphology. In D. Sandra & H. Assink (Eds.), *Reading complex words* (pp. 139–157). Amsterdam, The Netherlands: Kluwer.

Goldstein, E. Bruce. (2008).*Cognitive psychology: connecting mind, research and everyday experience (3rd Edition).* CA: Thomson/Wadsworth.

Hoosain, R. (1992). Psychological Reality of the word in Chinese. In H.-C. Chen &

O.J.-L. (Eds.), *Language processing in Chinese* (pp. 111-130). Amsterdam, Netherlands: North-Holland.

Hsu, C.H.. Tsai. J.L. Lee, C.Y., &Tzeng. O.L. (2009). Orthographic combinability and phonological consistency effects in reading Chinese phonograms: An event-related potential study. *Brain and Language, 108*, 56-66.

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.

Hyona,J., Vainio, S. & Laine, M.(2002). morphological effect obtains for isolated words but not for words in sentence context. *European Journal of Cognitive Psychology, 14 (4),* 417–433.

Inhoff, A. W., & Liu, W.　(1998). The perceptual span and oculomotor activity during the reading of Chinese sentences.　*Journal of Experimental Psychology: Human Perception &Performance , 24,* 20-34.

Institute of Language Teaching and Research. (1986). 现代汉语频率词典 [*Modern Chinese frequency dictionary*]. Beijing: Beijing Language Institute Press.

Janssen, N., Bi, Y., & Caramazza, A. (2008). A tale of two frequencies: Determining the speed of lexical access for mandarin Chinese and English compounds. *Language and Cognitive Processes, 23(7–8),* 1191–1223.

Juhasz, Barbara J.; Berkowitz, Rachel N.(2011). Effects of morphological families on English compound word recognition: A multitask investigation. *Language & Cognitive Processes,26:* 4-6.

Koriat, A. (1981). Semantic facilitation in lexical decision as a funtion in lexical decision as a function of prime-target association. *Memorv & Cognition, 9,* 53-65.

Kruidenier,John (2002). *Researchbased Principles for Adult Basic Education Reading*

Instruction. Extracted from http://lincs.ed.gov/publications/pdf/adult_ed_02.pdf

Kuperman, V., Bertram, R., and Baayen, R. H. (2008). Morphological dynamics in compound processing. *Language and Cognitive Processes, 23:*1089-1132.

Kuperman V., R. Schreuder, R. Bertram, R.H. Baayen (2009). Reading poly-morphemic Dutch compounds: Toward a multiple route model of lexical processing. *Journal of Experimental Psychology: Human Perception and Performance, 35,*876–895.

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

Lin , H. (2001). *A Grammar of Mandarin Chinese*. Munich: Lincom Europa.

Li, X. S., & Logan, G. (2008). Object–based attention in Chinese readers of Chinese words: Beyond Gestalt principles. *Psychonomic Bulletin & Review, 15*, 945–949.

Liu, Ying and Dan-ling Peng, Meaning access of Chinese compounds and its time course, in *Hsuan-Chih Chen, ed., Cognitive Processing of Chinese and Related Asian Languages*, Hong Kong: The Chinese University Press, 1997, 219-232.

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

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

Marsolek, Chad J..(2008). What antipriming reveals about priming. *Trends in Cognitive Sciences. 12(5):*176-181.

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.

McCormick, S. F., Rastle, K. & Davis, M. H. (2008). Is there a 'fete' in 'fetish'? Effects of orthographic opacity on morpho-orthographic segmentation in visual word recognition. *Journal of Memory and Language. 58, 2,* 307-326.

McKague, M.; Pratt, C.; Johnston, M. B..(2001). The effect of oral vocabulary on reading visually novel words: a comparison of the dual-route-cascaded and triangle frameworks. *Cognition, 80(3):*239-270

Myers, James. (2006). "Processing Chinese compounds: A survey of the literature", in: Gary Libben and Gonia Jarema, eds., *The Representation and Processing of Compound Words*, Oxford: Oxford University Press, 169-196.

Myers, James. (2010). "Chinese as a natural experiment", *The Mental Lexicon 5/3,* 423-437.

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

McNaughton , W. and Ying , L. (1999). *Reading and Writing Chinese, 2nd ed.* Boston, MA : Tuttle Publishing

Mok, LehWoon,(2009).Word-superiority effect as a function of semantic transparency of Chinese bimorphemic compound words. *.Language and Cognitive Processes 24/7-8*, 1039-1081.

Packard, J. L. (2000). *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge : Cambridge University Press.

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

Pan Wenguo. (2002). *Character-Based View and the Study of Chinese. (in Chinese)* Shanghai: East-China Normal University Press.

Perfetti, C.A. & Tan, L.H.(1999). The constituency model of Chinese word recognition. In J. Wang, A. W. Inhoff & H.-C. Chen (Eds.), *Reading Chinese script: A cognitive analysis* (pp. 115-134). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Peng, D., Ding, G., Wang, C. Taft, M. & Zhu, X. (1999). The processing of Chinese reversible words: The role of morphemes in lexical access. *Acta Psychological Sinica, 31* , 36–46.

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

Peng, Zerun & Pan Wenguo (2010). Which is better for Chinese Linguistics: character-centered or word-centered?(in Chinese) *Journal of Tonghua Normal Universit*y, 09, 6-11.

Perfetti, C.A., & Tan, L.H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identication. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 101–118.

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

Peterson, R. R.,& Simpson, G. B. (1989). The effect ofbackward priming on word recognition in single-word and sentence contexts. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 15,* 1020- I032.

Petten, C.V. and Kutas, M. (1991). Kutas M: Influences of semantic and syntactic context on open and closed class words. *Memory and Cognition, 19(1),* 95--112

Pinker, Steven. (2011). *Words and rules: the ingredients of language*. New York:

Harper Perennial.

Rastle, K. & Davis, M. H. (2008). Morphological decomposition based on the analysis of orthography. *Language and Cognitive Processes. 23, 7-8,* 942-971.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Rayner, K., Li, X. S., Juhasz, B. J. & Yan, G. (2005). The effect of word predictability on the eye movements of Chinese readers. *Psychonomic Bulletin & Review, 12*, 1089–1093.

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

Seidenberg, Mark S. (2005). Connectionist models of word reading. *Current Directions in Psychological Science (Wiley-Blackwell), 14(5),* 238-242.

Seidenberg, Mark S. and McClelland, James L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review, Vol 96(4)*: 523-568.

Seidenberg, M. S., Waters, G. S., Sander, M., & Langer, P.(1984). Pre- and postlexical loci of contextual effects on word recognition. *Memory & Cognition, 12,* 315-328.

Shelton,J. R., & Martin, R. C. (1992). How semantic is automatic semantic priming? *Journal of Experimental Psychology: Learning, Memory, & Cognition, 18,* 1191-1210.

Shen Mowei, Chen Xin & Tao Rong (2001).The readability of gliding leading format and RSVP in Chinese texts. *Psychological Science, 04,* 393-395.

Shen Mowei, Fu Dejiang, Zhang Guangqiang & Chen Xin (2002).The readabilty of Chinese text presented in gliding leading format and RSVP with self-paced speed. *Acta Psychologica Sinica,34(1)* ,43-48.

State Language Work Committee. (2008). *Lexicon of Common Words in Contemporary Chinese*. Beijing: The Commercial Press.

Sakuma, N., Sasanuma, S., Tatsumi, I.F., & Masaki, S. (1998). Orthography and phonology in reading Japanese Kanji words: Evidence from the semantic decision task with homophones. *Memory and Cognition, 26,* 75–87.

Taft, M. (2004). Morphological decomposition and the reverse base frequency effect. *Quarterly Journal of Experimental Psychology A, 57,*745–765

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

Taft, M., & Zhu, X. (1995). The representation of bound morphemes in the lexicon: A Chinese study. In L. Feldman (Ed.), *Morphological aspects of language processing* (pp293-316). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Taft, M..& Zhu, X. (1997). Using masked priming to examine lexical storage of Chinese compound words. In H.-C., Chen (Eds). T*he cognitive processing of Chinese and related Asian languages.* HK: Chinese University Press.

Taft, M., Liu, Y., & Zhu, X. (1999). Morphemic processing in reading Chinese. In A.

Inhoff; J. Wang,& H-C. Chen (Eds.) *Reading Chinese script: A cognitive analysis*. New Jersey: Lawrence Erlbaum Associates.

Taft, M. (1985). The decoding of words in lexical access: A review of the morpho-graphic approach. In D. Besner, T. G Waller, & G E. MacKinnon (Eds.) *Reading research Advances in theory and practice, Vol. V,* New York: Academic Press.

Taft, M. (1994). Interactive-activation as a framework for understanding processing. *Language and Cognitive Processes, 9*, 271-294.

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.

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

Tsai, Jie-Li, Chia-Ying Lee, Ying-Chun Lin, Ovid J. L. Tzeng and Daisy L. Hung. (2006). "Neighborhood size effects of Chinese words in lexical decision and reading". *Language and Linguistics 7/3,* 659-675.

Wang, Hongjun.(2008).Linguistic Levels and the Levels of the Theory of Taking Zi as the Chinese Basic Unit(in Chinese). *Language Teaching and Linguistic Studies, 03*, 1-11.

Wang, Jun.(2009). A Study on the Relative Factors of Foreign Students' Chinese Character Learning. Language (in Chinese). *Teaching and Linguistic Studies, 01*, 9-16.

Wang, S. Y. (1981). Language structure and optimal orthography. In O.J.L. Tzeng&

H. Singer (Eds.), *Perception of print: Reading research in experimental psychology*(pp. 223-236). Hillsdale, NJ: Erlbaum

Weekes B, Chen MJ & Lin YB. (1998). Differential effects of phonological priming on Chinese character recognition. *Reading Writing, 10,* 201–222.

Wei, Huiping (2004). Reflections on the Relationship Between Chinese Characters and Words (in Chinese). *Journal of Nanjing Normal University (Social Science Edition): 01,* 135- 140.

White, S. J. (2008). Eye movement control during reading: effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 205–223.

Wu ,Ningning, Xiaolin Zhou & Hua Shu (1999): Sublexical Processing in Reading Chinese: A Development Study, *Language and Cognitive Processes, 14:5-6,* 503-524

Wurm, L. H. & Fisicaro,S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language, 72,* 37–48.

Wydell, T.N., Patterson, K.E., & Humphreys, G.W. (1993). Phonologically mediated access to meaning for Kanji: Is a ROWS still a ROSE in Japanese Kanji? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 491–514.

Xie, D., and Zhou, X. (2003). Earlier phonological activation in semantic activation in reading Chinese? A replication study. *Acta Psychol. Sin.* 14–22.

Xing , J. Z. (2006). *Teaching and Learning Chinese as a Foreign Language: A Pedagogical Grammar.* Hong Kong : Hong Kong University Press.

Xing, H. (2008). The over 800 pairs of reversed Chinese words in the modern Chinese lexicon (in Chinese). Extracted from http://www.blog.sina.com.cn/s/blog_5517e35b01009zmb.html

Xu Baihua & Zhu Hai (1997). Factors affecting the readability of Rapid Serial Visual Presentation of Chinese texts (in Chinese), *Psychological Science,01,* 11-15.

Xu, Tongqiang.(1991). On Semantic Syntax -- An Exploration of the Linguistic Structural Basis and the Methodology of Grammar Study (in Chinese). *Language Teaching and Linguistic Studies, 03,* 38-62.

Xu, Tongqiang. (1994)."Character" and the Methodology of Chinese Language Research (in Chinese). *Chinese Teaching in the World, 03,* 1-14.

Xu, Tongqiang (2005). Zi as the Basic Structural Unit and Linguistic Studies. *Language Teaching and Linguistic Studies, 06,* 1-11.

Yan, G., Tian, H., Bai, X., & Rayner, K. (2006). The effect of word and character frequency on the eye movements of Chinese readers. *British Journal of Psychology, 97,* 259–268.

Yuan, C. & Huang, C (1998). Word-formation study based on the Chinese Morpheme Data Bank (in Chinese). *Chinese Teaching in the World,2*: 7-12.

Zhan,Jiayu, Yu,Hongbo, & Zhou,Xiaolin. (2013). fMRI evidence for the interaction between orthography and phonology in reading Chinese compound words. *Frontiers in Human Neuroscience, vol 7:* 1-11

Zhang, B., &Peng, D. (1992). Decomposed storage in the Chinese lexicon. In H.-C. Chen & 0. J. Zheng (Eds.), *Language processing in Chinese*. Amsterdam: North-Holland.

Zhang, Biyin and Danling Peng. (1992). "Decomposed storage in the Chinese lexicon". in:Hsuan-Chih Chen and Ovid J. L. Tzeng , eds., *Language Processing in Chinese*, Amsterdam: North-Holland, 131-149.

Zhang,J.X. J. Zhuang, L. Ma, W. Yu, D. Peng, G. Ding et al.(2004). Semantic processing of Chinese in left inferior prefrontal cortex studied with reversible words .*NeuroImage, 23*, 975–982

Zhang, B., &Peng, D. (1992). Decomposed storage in the Chinese lexicon. In H.-C. Chen & O. J.-L. Tzeng (Eds), *Language processing in Chinese. Advances in psychology, 90* (pp. 131-149). Amsterdam, Netherlands: North-Holland.

Zhou, X. (1997). The limited effect of phonology on access to meaning in reading Chinese. In D.-L. Peng, H. Shu, & H.-C. Chen (Eds), *The cognitive research of Chinese* (pp.159–184). Shandong: Shandong Educational Publisher.

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

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

Zhou, X., W. Marslen-Wilson (1995). Morphological structure in the Chinese mental Lexicon.*Language and Cognitive Process, 10,* 545–600.

Zhou, X., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition. Vol. 26 (5),* 1245-65.

Appendix 1

The mixed-effects modeling results of 12 sub-experiments

(after removing the insignificant predictors)

Experiment 1a:

Random effects					
Groups	Name	Variance	Std.Dev.		
Item	(Intercept)	0.001936	0.04400		
subject	(Intercept)	0.006720	0.08198		
Residual		0.047648	0.21829		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.119027	0.055904	218.200000	127.345	< 2e-16 ***
Freqchar:ConditionRelated	-0.031822	0.005618	318.200000	-3.922	4.69e-07 ***
Freqchar:ConditionRelated	-0.033833	0.004377	398.900000	-4.475	1.24e-05 ***
FreqWord:ConditionRelated	-0.011149	0.003993	151.400000	-4.475	0.00763 **
FreqWord:ConditionUnrelated	0.009808	0.003420	105.900000	2.344	9.75e-05 ***
Nstroke:ConditionRelated	-0.009754	0.002219	372.600000	-2.684	3.70e-05 ***
Nstroke:ConditionUnrelated	-0.002345	0.002438	316.200000	-0.962	0.33690

Experiment 1b:

Random effects					
Groups	Name	Variance	Std.Dev.	Corr	
Item	(Intercept)	0.0017022	0.04126		
subject	(Intercept)	0.0302338	0.17388		
	Freqchar	0.0001261	0.01123	-0.96	
Residual		0.0336708	0.18350		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.016186	0.057799	55.400000	121.389	< 2e-16 ***
Freqchar:ConditionRelated	-0.039218	0.004634	64.300000	-4.367	2.31e-05 ***
Freqchar:ConditionUnrelated	-0.02992	0.0045	58.000000	-6.398	1.14e-08 ***
Nstroke:ConditionRelated	-0.001234	0.00204	444.500000	-0.605	0.545647
Nstroke:ConditionUnrelated	-0.006924	0.001896	425.000000	-2.154	0.00783 **

Experiment 1c:

Random effects					
Groups	Name	Variance	Std.Dev.	Corr	
Item	(Intercept)	0.0025447	0.05045		
subject	(Intercept)	0.0361709	0.19019		
	Freqchar	0.0001579	0.01256	-0.93	
Residual		0.0294242	0.17153		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.030588	0.050065	25.190000	140.429	< 2e-16 ***
Freqchar:ConditionRelated	-0.031075	0.004189	36.530000	-4.663	1.83e-05 ***
Freqchar:ConditionUnrelated	-0.048186	0.003945	28.860000	-9.806	8.93e-12 ***

Experiment 2a:

Random effects					
Groups	Name	Variance	Std.Dev.		
Item	(Intercept)	0.0002472	0.01572		
subject	(Intercept)	0.0166680	0.12910		
Residual	0.0467291	0.21617			
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.973030	0.040025	61.600000	174.218	< 2e-16 ***
Freqchar:ConditionRelated	0.030955	0.004045	323.100000	-4.277	2.37e-05 ***
Freqchar:ConditionUnrelated	-0.030054	0.003063	354.500000	-7.382	8.18e-14 ***
FreqWord:ConditionRelated	-0.021339	0.003559	232.900000	-6.321	8.05e-09 ***
FreqWord:ConditionUnrelated	0.006173	0.002855	155.300000	1.159	0.2516

Experiment 2b:

Random effects					
Groups	Name	Variance	Std.Dev.		
Item	(Intercept)	0.001811	0.04255		
subject	(Intercept)	0.017511	0.13233		
Residual		0.038074	0.19513		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.959e+00	4.590e-02	8.320e+01	151.613	< 2e-16 ***
Freqchar:ConditionRelated	-3.261e-02	6.406e-03	8.443e+02	-4.520	2.87e-05 ***
Freqchar:ConditionUnrelated	-3.876e-02	3.245e-03	1.039e+03	-7.981	7.88e-14 ***
FamSize:ConditionRelated	6.506e-02	1.040e-02	2.902e+02	3.542	0.000493 ***
FamSize:ConditionUnrelated	1.460e-02	6.590e-03	9.080e+01	-0.317	0.7517

Experiment 2c:

Random effects					
Groups	Name	Variance	Std.Dev.		
Item	(Intercept)	0.001631	0.04039		
subject	(Intercept)	0.017316	0.13159		
Residual		0.038424	0.19602		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.048483	0.052522	126.800000	134.200	< 2e-16 ***
Freqchar:ConditionRelated	-0.037927	0.006804	675.200000	-5.143	1.05e-08 ***
Freqchar:ConditionUnrelated	-0.049248	0.003791	688.900000	-13.662	< 2e-16 ***
Nstroke:ConditionRelated	-0.002504	0.001962	528.300000	-1.277	0.2023
Nstroke:ConditionUnrelated	0.008212	0.001930	636.500000	4.361	2.02e-05 ***
FamSize:ConditionRelated	0.059564	0.010394	273.000000	3.884	1.08e-05 ***
FamSize:ConditionUnrelated	0.004887	0.006612	100.400000	-0.038	0.969999

Experiment 3a

Random effects			
Groups	Name	Variance	Std.Dev.
Item	(Intercept)	0.001932	0.04396
subject	(Intercept)	0.021120	0.14533
Residual		0.044448	0.21083

Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.012008	0.056719	144.800000	123.626	< 2e-16 ***
FreqWord:ConditionOrigin	-0.015606	0.004340	565.300000	-3.27 6	0.000352 ***
FreqWord:ConditionReversed	-0.019731	0.004354	608.000000	-4.002	7.05e-06 ***
FreqLast:ConditionOrigin	-0.019731	0.005088	578.900000	-2.358	0.004842 **
FreqLast:ConditionReversed	-0.009969	0.005257	620.900000	-1.896	0.058410

Experiment 3b

Random effects				
Groups	Name	Variance	Std.Dev.	Corr
Item	(Intercept)	1.552e-03	0.039393	
subject	(Intercept)	1.701e-02	0.130433	
	Freqchar	8.083e-05	0.008991	-0.95
Residual		4.637e-02	0.215345	

Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.093e+00	6.378e-02	1.937e+02	111.208	< 2e-16 ***
FreqWord:ConditionOrigin	-5.196e-03	4.710e-03	8.620e+01	-1.11	0.27012
FreqWord:ConditionReversed	-1.378e-02	4.621e-03	8.150e+01	-3.393	0.00117 **
FreqInitial:ConditionOrigin	-3.460e-02	5.376e-03	5.018e+02	-5.616	2.87e-9 ***
FreqInitial:ConditionReversed	-2.254e-02	5.331e-03	4.556e+02	-2.791	0.00552 **
NstrokeInitial:ConditionOrigin	3.562e-03	2.265e-03	4.223e+02	1.573	0.11658
NstrokeInitial:ConditionReversed	6.251e-03	2.485e-03	4.824e+02	2.663	0.00802 *

Experiment 3c

Random effects			
Groups	Name	Variance	Std.Dev.
Item	(Intercept)	0.0004804	0.02192
subject	(Intercept)	0.0195173	0.13970
Residual		0.0458899	0.21422

Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.988948	0.051958	105.940000	134.512	< 2e-16 ***
FreqWord:ConditionOrigin	0.001817	0.003822	269.050000	0.475	0.63490
FreqWord:ConditionReversed	-0.009936	0.003710	264.410000	-2.268	0.02417 *
FreqInitial:ConditionOrigin	-0.035137	0.004508	273.540000	-7.685	1.35e-13 ***
FreqInitial:ConditionReversed	-0.026593	0.004390	282.020000	-2.465	0.01433 *

Experiment 4a

Random effects					
Groups	Name	Variance	Std.Dev.		
Item	(Intercept)	0.002564	0.05064		
subject	(Intercept)	0.011309	0.10634		
Residual		0.033587	0.18327		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.964970	0.037055	101.100000	187.962	< 2e-16 ***
Freq:ConditionComponent	-0.021195	0.003316	958.600000	-4.392	4.02e-05 ***
Freq:ConditionUnrelated	-0.021259	0.002917	989.700000	-4.541	3.52e-05 ***

Experiment 4b

Random effects					
Groups	Name	Variance	Std.Dev.	Corr	
Item	(Intercept)	0.0008568	0.02927		
subject	(Intercept)	0.0306817	0.17516		
	Freq	0.0001252	0.01119	-0.94	
Residual		0.0393638	0.19840		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	6.972771	0.047706	23.480000	146.160	< 2e-16 ***
Freq:ConditionComponent	-0.031752	0.004083	33.200000	-5.707	3.59e-09 ***
Freq:ConditionUnrelated	-0.025603	0.003791	25.010000	-5.893	2.07e-09 ***

Experiment 4c

Random effects					
Groups	Name	Variance	Std.Dev.	Corr	
Item	(Intercept)	0.0015994	0.03999		
subject	(Intercept)	0.0438109	0.20931		
	Freqchar	0.0001848	0.01359	-0.93	
Residual		0.0332898	0.18245		
Fixed effects:					
	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	7.046948	0.067130	50.110000	104.975	< 2e-16 ***
Freq:ConditionComponent	-0.027667	0.004646	41.630000	-7.836	1.29e-12 ***
Freq:ConditionUnrelated	-0.026939	0.004531	38.030000	-7.946	6.73e-13 ***
NstrokeChar:ConditionComponent	-0.009569	0.004216	84.830000	-2.265	0.0258 *
NstrokeChar:ConditionUnrelated	-0.004285	0.004057	72.870000	-1.056	0.2943

Appendix 2. Stimuli used in Exp. 1				
No.	prime sentence	compound embedded	related target character	unrelated target character
1	他父亲/是厂长/兼书记 His father is the factory director and secretary.	厂长 factory director	厂 factory	天 sky
2	老师说/讲卫生/不得病 The teacher said that, if careful about hygiene, you won't get an infection.	卫生 hygiene	卫 protect	服 clothes
3	他查不/出失火/的原因 He can't find the cause of fire.	失火 catch fire	失 lose	港 port
4	结婚可/是自己/的大事 Marriage is one's own big event.	自己 self	自 self	去 go
5	她不认/识其他/男同学 She doesn't know the other boy classmates.	其他 other	其 it	家 home
6	祖父没/给父亲/留遗产 My grandfather didn't leave any wealth to my father.	父亲 father	父 father	捡 pick
7	总督再/次发布/垦荒令 The governor issued the order of land reclamation.	发布 issue	发 distribute	翻 overturn
8	我采取/的方法/很简单 The method I adopted was very simple.	方法 method	方 square	墙 wall
9	这得换/个角度/去思考 It requires us to think from a different perspective.	角度 angle	角 horn	吐 vomit
10	哥哥刚/把事情/告诉我 My brother told me the thing just now.	事情 thing	事 thing	量 amount
11	稳准狠/是目前/的方针 Steadiness, accuracy and toughness is the present principle.	目前 at present	目 eye	居 reside
12	他喜欢/听电影/的声效 He likes to listen to the sound effects of films.	电影 film	电 electricity	世 world
13	我又梦/见女魔/追杀我 I dreamed about being chased by a female devil.	女魔 female devil	女 female	界 boarder
14	他总带/着电脑/去听课 He always carried his computer to the class.	电脑 computer	电 electricity	港 port
15	反叛军/被击毙/两百人 Two hundreds of rebels were shot to death.	击毙 shoot to death	击 shoot	内 inside
16	这是生/产面膜/的厂家 This is a facial-mask factory.	面膜 facial mask	面 face	如 as
17	美国官/兵昨天/撤走了 American troops withdrew yesterday.	昨天 yesterday	昨 yesterday	四 four
18	孕妇产/前征兆/很明显 The woman's antenatal symptom is very obvious.	征兆 omen	征 journey	位 position
19	新雇员/的要求/比较高 The new employees demand more.	要求 demand	要 need	吃 eat
20	我对中/国功夫/有兴趣 I'm interested in Chinese Kung Fu.	功夫 Kung Fu	功 work	穿 wear
21	这是降/低运费/的目的 This aims at reduce the transportation cost.	运费 transportation cost	运 transport	计 plan
22	芦苇草/能改良/盐碱地 Reeds could improve the saline-alkali soil.	改良 improve	改 change	百 hundred
23	我国政/府采取/新措施 Our government adopted some new measures.	采取 adopt	采 pick	分 minute
24	这是来/自宇宙/的射线 These rays come from outer space.	宇宙 universe	宇 universe	之 of
25	背诵基/本句型/很有用 It's very useful to recite the basic sentence structures.	句型 sentence structure	句 sentence	床 bed
26	他发表/的评论/影响大 His published comments exerted great influence.	评论 comment	评 comment	上 up
27	那是硫/化物质/的气味 It's the smell of sulfide.	物质 substance	物 object	书 book
28	这是全/球畅销/的教材 This is the world's bestselling textbook.	畅销 sell well	畅 smooth	口 mouth
29	煤焦油/有防腐/的作用 The coal tar has an effect of corrosion-protection.	防腐 corrosion prevention	防 protect	岸 bank
30	精神备/受折磨/和煎熬 There is intense mental torment and suffering.	折磨 torment	折 break	进 enter
31	那是十/分愚蠢/的行为 That's very stupid action.	愚蠢 stupid	愚 stupid	出 out
32	淋巴体/的功能/最活跃 The lymph system is most active.	功能 function	功 work	货 goods

33	林志炫/的唱法/太妙了 Lin Zhixuan's singing style was so wonderful.	唱法 singing style	唱 sing	值 value
34	这句话/的意思/是什么 What's the meaning of this sentence.	意思 meaning	意 idea	均 average
35	还有几/处谜团/未解开 There are still several mysteries unsolved.	谜团 mystery	谜 riddle	来 come
36	她忘不/了鞋匠/的身影 She can't forget the figure of the shoemaker	鞋匠 shoemaker	鞋 shoe	自 self
37	他不需/要鼠标/和键盘 He doesn't need a mouse or a keyboard.	鼠标 mouse	鼠 mouse	外 out
38	我要树/立远大/的理想 I want to establish lofty ideals.	远大 great	远 far	省 province
39	初测试/的结果/很成功 The result of the preliminary test was very successful.	结果 result	结 bear (fruit)	区 district
40	他征求/各部门/的意见 He seeks advice from every department.	部门 department	部 part	中 middle
41	这就是/他梦想/的地方 This is the place that he dreamed of.	梦想 dream	梦 dream	国 country
42	政府封/锁沉船/的水域 The government blocked off the wreck waters.	沉船 sunken ship	沉 sink	数 number
43	老教授/很满意/这成果 The old professor is very satisfied with the achievements.	满意 satisfied	满 full	字 character
44	我热情/地招待/客人们 I warmly received all the guests.	招待 serve	招 beckon	出 out
45	现在无/法摧毁/癌细胞 It's impossible to destroy cancer cells now.	摧毁 destroy	摧 break	版 edition
46	这完全/是背离/事实的 This is completely against the fact.	背离 deviate	背 back	博 extensive
47	她五天/才摆脱/了危险 It took five days for her to get out of the danger.	摆脱 get out of	摆 swing	览 look
48	要保护/有智障/的孩童 Protect the mentally retarded children.	智障 retarded	智 wisdom	会 meeting
49	部长享/有豁免/的待遇 The minister enjoys the diplomatic immunity.	豁免 immunity	豁 exempt	明 bright
50	新政策/的机遇/已出现 The opportunity for the new policy has emerged.	机遇 opportunity	机 chance	月 moon
51	我佩服/他敬业/的精神 I admire his devotion to work.	敬业 devotion to work	敬 respect	日 sun
52	此处铁/矿资源/很丰富 There is rich iron mineral resource here	资源 resource	资 capital	在 at
53	新激素/能调节/血糖量 The new hormone can adjust blood sugar.	调节 adjust	调 tune	蒜 garlic
54	冲积岛/是淤沙/积成的 The alluvial island was made of the accretion of sand.	淤沙 accretion of sand	淤 accretion	桌 table
55	郑成功/的墓碑/很高大 The gravestone of Zheng Chenggong is very tall.	墓碑 gravestone	墓 grave	歌 song
56	她演奏/着激昂/的乐曲 She is playing an exciting tune.	激昂 exciting	激 arouse	器 device
57	弟弟露/出憨傻/的笑容 The younger brother gave a silly smile.	憨傻 silly	憨 foolish	摇 shake
58	这纯粹/是瞎编/出来的 This was totally fabricated.	瞎编 fabricate	瞎 blind	镇 town
59	他没学/过打造/或焊接 He never learned forging and welding.	打造 forge	打 hit	慢 slow
60	手机长/途漫游/很花钱 The mobile roaming is very expensive.	漫游 roam	漫 unrestrained	慰 comfort

Note: English translation is provided beneath every stimulus. In the experiments, only Chinese sentences or characters were presented.

Appendix 3. Stimuli used in Exp 2

No.	prime sentence	compound embedded	related target character	unrelated target character
1	艾米丽/告别丈/夫之后/回家了 Amily went home after saying farewell to her husband.	丈夫 husband	丈 ten feet	香 fragrant
2	副校长/遭到严/肃批评/和处理 The vice president is seriously criticized and punished.	严肃 serious	严 strict	晨 morning
3	她已为/还原本/来面目/并不难 She thought it's not difficult to restore the original looks.	本来 original	本 root	特 special
4	电路板/标准尺/寸缩小/近一半 The standard size of the circuit board shrank to half.	尺寸 size	尺 ruler	点 dot
5	探险队/前赴南/极考察/冰川层 The expedition set off for exploring the South Pole glacier.	南极 south pole	南 south	门 door
6	哨兵在/前方发/现敌人/的坦克 The sentry found the enemy's tank in the front.	发现 find	发 send out	新 new
7	这里的/各种文/件资料/很齐全 There are all kinds of documents and materials here.	文件 document	文 writing	体 body
8	对数据/进行分/析整理/较困难 It is relatively difficult to sort and analyze the data.	分析 analyze	分 divide	布 cloth
9	这一类/心脏血/管疾病/很难治 It is difficult to cure this type of cardiovascular disease.	血管 blood vessel	血 blood	高 tall
10	她成为/北京电/影学院/特招生 She became a student with special admission of Beijing Film Academy.	电影 film	电 electricity	紫 purple
11	他刚刚/才把事/情经过/告诉我 He told me the whole story just now.	事情 thing	事 thing	制 produce
12	阳光是/可以直/接利用/的能量 Sunlight can be used directly.	直接 directly	直 straight	婴 baby
13	一男子/手持弓/箭站立/在墙角 A man holding the bow and arrow is standing at the corner.	弓箭 bow and arrow	弓 bow	年 year
14	外围的/观众干/脆站着/看比赛 The audience on the outer rim just stood up to watch the game.	干脆 just	干 dry	儿 son
15	此处的/洞窟开/凿至今/上千年 This cave was dug more than a thousand years ago	开凿 dig	开 open	奶 milk
16	将军是/带着口/罩前去/参观的 The general went to visit it wearing a mask.	口罩 mask	口 mouth	区 district
17	开头的/大写字/母表示/出生地 The initial capitalized letters stand for the birth place.	字母 letter	字 character	出 exit
18	对工人/加强安/全教育/很必要 It is necessary to strengthen safety education among the staffs.	安全 safety	安 quiet	心 heart
19	她爸在/影视公/司担任/总经理 Her father is the general manager of the film company.	公司 company	公 public	蒸 steam
20	我妻子/爱听历/史故事/和传说 My wife like listening to the historical stories and legends.	历史 history	历 calendar	式 style
21	气候的/恶劣破/坏人们/的生活 The poor weather has ruined people's living.	破坏 destroy	破 break	市 city
22	毕业生/这段时/间忙着/找工作 The graduates are busy on jib-hunting recently.	时间 time	时 time	粒 particle
23	这是有/广泛好/评的重/要原因 This is an important reason for the extensive favorable comments.	好评 favorable comments	好 good	泼 splash
24	这等于/唐山地/震释放/的能量 This equals to the amount of energy released by the Tangshan earthquake.	地震 earthquake	地 place	人 people
25	据估计/企业利/润总额/将增加 It's estimated that the total profit of the enterprise will increase.	利润 profit	利 profit	吃 eat
26	秦始皇/建造规/模宏大/的宫殿 Qin Shihuang built a large-scale palace.	规模 scale	规 compass	仅 only
27	飞机有/记录关/键信息/的装置 There is the device to record key information	关键 key	关 close	能 can
28	看球赛/必须选/择合适/的位置 It's necessary to choose an appropriate position to watch the football game.	选择 choose	选 choose	蟹 crab
29	吉祥物/需要可/爱健康/的造型 The mascot needs to have a lovely and healthy appearance.	可爱 lovely	可 can	戴 wear
30	公安局/开始追/查非法/的药品 The police starts to found out the illegal drugs	追查 investigate	追 chase	两 two
31	成龙在/大赛闭/幕式上/有表演 Cheng Long will perform at the closing ceremony.	闭幕 closing	闭 close	古 ancient
32	网络上/炒股秘/籍大多/不可信 Most of the online stocks-trading tips are not reliable.	秘籍 tip	秘 secret	且 and

33	屋檐上/挂满腊/肉香肠/等年货 There was preserved meat and sausages hanging the eaves for New Year.	腊肉 preserved meat	腊 preserved	虽 although
34	他只卖/画有鸵/鸟图案/的陶器 He only sells the pottery with the patterns of ostriches.	鸵鸟 ostrich	鸵 ostrich	医 doctor
35	副经理/负责新/年庆典/的组织 The deputy manager is responsible for the New Year celebrations.	新年 New Year	新 new	下 below
36	这是他/父亲最/后一次/坐飞机 This is the last time his father took the plane.	最后 last	最 most	金 gold
37	他们是/终日奔/波各地/的记者 They are journalists travelling among different places all days.	奔波 travelling	奔 run	指 point
38	杂志上/刊登征/婚广告/不便宜 It's not cheap to place a marriage-seeking ad in a magazine.	征婚 marriage-seeking	征 ask for	仁 kindness
39	他相信/国外赌/博公司/的预测 He believes in the prediction from the foreign gambling company	赌博 gambling	赌 gamble	何 way
40	本访谈/具有综/艺表演/的特点 This interview has the features of an entertainment show.	综艺 entertainment	综 synthesize	若 if
41	这一次/顺水航/程长达/三千米 This downstream voyage has covered a distance of more than 3000 miters	航程 voyage	航 sail	违 violate
42	近几年/银行存/款利息/十分低 The interest on deposit in recent years was very low.	存款 deposit money	存 deposit	反 against
43	为植物/生长创/造有利/的条件 Create favorable conditions for the plant's growth.	创造 create	创 initiate	有 have
44	公司招/秘书要/求写得/很清楚 The company's requirements for the secretary's position was written clear.	要求 requirement	要 demand	鹅 goose
45	近五年/法国经/济状况/好转了 There is an upturn in the economy in France in the recent five years	经济 economy	经 sacred book	规 regulation
46	想发展/需要突/破常规/的思路 Development demands the breakthrough of the conventional thinking.	突破 breakthrough	突 dash	例 e example
47	你先要/了解熟/悉人体/的结构 You should first get familiar with the structure of human body.	熟悉 familiar	熟 cooked	即 namely
48	他知道/隐瞒事/实真相/的后果 He knows the consequence of hiding the facts	事实 fact	事 thing	属 belong to
49	她已为/兑现诺/言奠定/了基础 She has laid a solid foundation for realizing her promise.	诺言 promise	诺 promise	西 west
50	新一代/大型激/光技术/已应用 The new generation large-scale laser tech has been put into use.	激光 laser	激 arouse	罪 crime
51	原来的/噪音标/准一定/不能变 The original noise standards must be retained.	标准 standard	标 sign	网 net
52	污染对/身心健/康危害/很严重 Pollution does much harm to the mental and physical health.	健康 health	健 healthy	网 net
53	这里是/每日新/鲜蔬菜/供应点 This is the daily supply point of the fresh vegetable.	新鲜 fresh	新 new	定 must
54	葡萄牙/培训警/察的反/恐能力 Portuguese government trained their police forces for against terrorism.	警察 police	警 police	星 star
55	我认为/这场比/赛成绩/不算好 I think this match's result was not vey good.	比赛 match	比 compare	杯 cup
56	要想想/怎么帮/助农民/富起来 Consider how to help farmers to get rich.	帮助 help	帮 help	款 fund
57	我只要/跟他随/便聊聊/就行了 I just want to casually chat with him.	随便 casually	随 follow	豆 bean
58	瓶子里/可以清/楚看到/残留物 The remaining substance in the bottle can be clearly seen	清楚 clear	清 clear	元 Yuan
59	北朝鲜/表明愿/意放弃/核计划 North Korea declared that it’s willing to give up its nuclear program.	愿意 willing	愿 willing	监 prison
60	沙尘暴/最新消/息传遍/了全国 The latest news about the dust storm has spread throughout the country.	消息 news	消 disappear	对 correct

Appendix 4. Stimuli used in Exp 3

No.	prime sentence	compound embedded	the reversed compound
1	商人们/想巴结/新市长 The merchants want to flatter the new mayor.	巴结 flatter	结巴 stutter
2	父亲要/用白灰/粉刷墙 The father is going to paint the wall with white ash.	白灰 white ash	灰白 greyish white
3	我必须/想办法/挣点钱 I have to find a way to earn some money.	办法 method	法办 punish by law
4	各行业/都办公/自动化 Office automation has been applied across industries	办公 office work	公办 run by government
5	红色台/风警报/已发布 Red typhoon warning has been issued.	警报 warning	报警 report to the police
6	这个字/的笔画/很复杂 The stroke pattern of this character is complicated.	笔画 stroke (pattern)	画笔 paintbrush
7	他们十/分虚心/地求教 They ask for advice very modestly	虚心 modest	心虚 feel guilty
8	总经理/的领带/天天换 The general manager changed his tie every day.	领带 tie	带领 lead
9	这药品/的成分/很简单 The composition of the medicine is very simple.	成分 composition	分成 dividend
10	终于到/了出发/的时间 It's finally the time to set off.	出发 set off	发出 send out
11	她们分/发传单/和图片 They distribute the leaflets and pictures.	传单 leaflet	单传 one son for each generation
12	我指出/有过错/的一方 I pointed out the party at fault.	过错 fault	错过 miss
13	他要上/告人大/常委会 He plans to report to the standing committee of NPC.	人大 the National People's Congress	大人 adult
14	我打开/了台灯/写日记 I turned on the desk lamp and start writing dairy.	台灯 desk lamp	灯台 lamp stand
15	小明的/话感动/了老师 Xiaoming's words touched the teacher.	感动 touch	动感 dynamic
16	他倾诉/对女儿/的思念 He poured out the thought of his daughter.	女儿 daughter	儿女 child
17	她一句/不顺耳/就闹气 She gets angry once listening to anything she doesn't want to hear.	顺耳 pleasant to the ear	耳顺 sixty years of age
18	每个人/都负重/十公斤 Everyone carries ten kilos.	负重 carry	重负 heavy burden
19	她不习/惯海外/的生活 She is not used to living abroad	海外 abroad	外海 open seas
20	他讲述/的故事/很感人 The story he told was very touching.	故事 story	事故 accident
21	士兵们/向国王/表忠心 The soldiers express their loyalty to the king.	国王 king	王国 kingdom
22	她扔了/不好看/的衣服 She threw the ugly clothes	好看 beautiful	看好 be bullish
23	此人有/个外号/小诸葛 This person has a nickname "Little Zhuge".	外号 nickname	号外 the newspaper extra
24	我父亲/的后事/已办完 My father's funeral has been done.	后事 funeral	事后 afterwards
25	后花园/的茶花/要开了 The camellia in the back garden is going to blossom.	茶花 camellia	花茶 flower tea
26	农场加/大黄牛/饲养量 The amount of cattle in the farm is increased.	黄牛 cattle	牛黄 bezoar
27	我决定/在会议/上发言 I decide to give a speech at the meeting.	会议 meeting	议会 parliament
28	这是什/么动机/和目的 What are the motivations and the goals.	动机 motivation	机动 locomotive
29	含金量/的计算/很准确 The calculation of gold contained is very accurate.	计算 calculation	算计 scheme against
30	这体现/了道家/的思想 This reflects the spirit of Taoism.	道家 Taoism	家道 family financial situation
31	我没什/么家世/和传闻 I have no family background or stories.	家世 family background	世家 old and well-known family
32	这个月/的奖金/减少了 The bonus of this month decreased.	奖金 bonus	金奖 gold award
33	二十四/个节气/很科学 The twenty-four solar terms are scientific.	节气 solar terms	气节 integrity

34	山洞口/的直径/约一米 The diameter of the mouth of the cave is about one meter.	直径 diameter	径直 directly
35	水污染/对居民/影响大 Water pollution has big effect on the residents.	居民 resident	民居 folk house
36	教授给/出科学/的解释 The professor gave a scientific explanation.	科学 science	学科 subject
37	解放之/前文盲/相当多 There were quite many illiterate people before the liberation.	文盲 an illiterate person	盲文 braille
38	乒乓球/赛马上/就结束 The table tennis will end immediately.	马上 immediately	上马 start (a project)
39	他是见/过世面/的孩子 He is a child with a lot of experience of the world.	世面 world	面世 come out
40	母亲只/对名著/感兴趣 The mother likes classic books only.	名著 classic book	著名 famous
41	他爱编/写年历/的工作 He enjoys the job of compiling calendars.	年历 calendar	历年 over the years
42	没有哪/个女子/不爱美 There is no lady who doesn't want to be beautiful.	女子 lady	子女 child
43	他一直/在盘算/怎么说 He was contemplating how to explain.	盘算 contemplate	算盘 abacus
44	郭德纲/的评书/很搞笑 Guo Degang's storytelling is very amusing.	评书 storytelling	书评 book review
45	校门口/的汽水/不干净 The soft drinks sold at the school gate is unhygienic.	汽水 soft drink	水汽 vapor
46	据传说/他前生/是花匠 It's said he had been a gardener in the previous existence.	前生 previous existence	生前 before death
47	要成功/的前提/是努力 The precondition of success is the efforts.	前提 precondition	提前 beforehand
48	一整天/他强压/着怒火 He suppressed his anger for the whole day.	强压 suppress	压强 pressure
49	老局长/的情人/有两个 The old director has two lovers.	情人 lover	人情 favor
50	她已确/定人生/的目标 She has decided her life goal.	人生 life	生人 stranger
51	此发明/为世人/所瞩目 This invention is the focus of people's attention.	世人 people	人世 world
52	中国古/代文人/很高傲 Chinese ancient intellectuals were arrogant.	文人 intellectual	人文 humanity
53	出拳之/际上身/应挺直 Straighten the upper body when punching.	上身 upper body	身上 on the body
54	这是一/篇养生/的食谱 This is recipe of health preserving.	养生 health preserving	生养 give birth to and bring up
55	为理想/的实现/而奋斗 Strive to realize your dream.	实现 realize	现实 real world
56	他新收/的手下/很机灵 His new understrapper is very clever.	手下 understrapper	下手 start
57	他有什/么长处/和短处 What are his merits and demerits?	长处 merit	处长 director
58	我儿子/很用功/地学习 My son is hard-working.	用功 hard-working	功用 utility
59	陈奕迅/不愿意/拍电影 Chen Yixun is not willing to cast a film.	愿意 be willing to	意愿 wish
60	改变阵/型发挥/了作用 Changing the team formation has worked.	发挥 exert	挥发 volatilization

Appendix 5. Stimuli used in Exp 4				
No.	prime sentence	compound character embedded	target component character	unrelated target character
1	他倒骑/着驴翻/山越岭 He went across the mountains riding a donkey.	驴 donkey	马 horse	特 special
2	他听得/到林子/里鸟鸣 She could hear the birds' singing in the woods	林 woods	木 wood	立 stand
3	李老师/的邮箱/塞满了 Professor Li's mailbox has been filled	邮 mail	由 from	长 long
4	邻居总/是吵得/没个完 The neighbors always quarrel.	吵 quarrel	口 mouth	水 water
5	注意你/的财产/和安全 Pay attention to your possessions and safety.	财 wealth	贝 shell	关 close
6	他说话/措辞相/当强硬 His wording was very uncompromising.	辞 diction	舌 tongue	东 east
7	队长的/右肘部/受伤了 The captain's right elbow was hurt.	肘 elbow	月 moon	间 between
8	我发现/微妙的/小变化 I found the subtle change.	妙 wonderful	女 female	是 be
9	这个是/祭灶神/的仪式 This is to offer a sacrifice to kitchen god.	灶 kitchen	火 fire	会 meeting
10	我需要/被删除/的文件 I need the deleted files	删 delete	册 booklet	齐 together
11	这报社/出版了/教科书 This newspaper company published textbooks.	版 printing plate	片 sheet	不 not
12	她穿着/一双白/色皮鞋 She is wearing a pair of white leather shoes.	双 double	又 again	面 noodle
13	使用的/材料是/优质品 The materials used were of the high quality.	料 material	米 rice	信 trust
14	驻伊美/军殴打/在押者 US troops in Iraq hit the prisoner.	殴 hit	区 district	体 body
15	这是第/一轮的/淘汰赛 This is the first round of the knockout.	轮 wheel	车 car	保 protect
16	这是客/流畅通/的时段 This is the period when the passenger flow is smooth.	畅 smooth	申 express	想 think
17	他讲的/是明朝/的事情 What he told was the story about the Ming Dynasty.	明 bright	日 sun	期 period
18	他们发/请帖宴/请亲友 They send out invitations to friends and relatives to dinner.	帖 post	巾 towel	骗 cheat
19	华山是/秦岭的/一部分 Mount Hua is part of Qinling Mountains.	岭 hill	山 mountain	心 heart
20	草原是/畜牧业/的产区 The grassland is the livestock concentrated zone	牧 herd	牛 cattle	限 limit
21	她带两/个孩童/摆地摊 She sets up a street stall with her two children.	孩 child	子 son	表 watch
22	花朵呈/鲜艳的/紫红色 The flower was bright purple red.	艳 colorful	丰 full of	物 subject
23	电流只/能顺着/导线走 The electric current can only flow along the wire.	顺 along	川 river	年 year
24	选一个/十秒的/小广告 Select a 10-second advertisement.	秒 second	禾 cereal	烧 burn
25	他按家/乡规矩/摆宴席 He set a feast according to his hometown rules.	矩 rule	矢 arrow	学 learn
26	葡萄果/肉甜而/多汁水 The flesh of the grape was sweet and juicy	甜 sweet	舌 tongue	难 difficult
27	北京人/期盼着/奥运会 People in Beijing are expecting the Olympic Games.	盼 expect	目 eye	澡 bath
28	她正在/搽粉准/备演出 She is putting on powder, ready to perform.	粉 powder	米 rice	权 power
29	卧底被/凶残地/杀害了 The undercover was killed brutally.	残 incomplete	歹 bad	因 because
30	核电站/放射污/染很强 There is strong radio contamination from the nuclear station.	射 shoot	身 body	真 real
31	驯鹿有/顽强的/耐寒力 The reindeer is very cold-tolerant.	顽 obstinate	元 first	时 time
32	他还是/没赶上/末班车 He finally missed the last bus.	赶 drive	走 walk	每 every

33	具体的/谁都不/太了解 Nobody knows the details.	都 all	者 person	变 change
34	每人都/有功绩/和过错 Each man has his merits and demerits.	功 achievement	工 work	像 similar
35	金属箭/离弦飞/出很远 The metal arrow flew far away.	弦 string	弓 bow	伞 umbrella
36	她发现/特殊的/矿物质 She found a special mineral substance.	殊 special	歹 bad	组 team
37	森林减/少致使/沙漠化 Deforestation causes desertification	致 send	至 to	向 towards
38	全聚德/烤鸭店/不外卖 Quanjude Roast Duck Restaurant doesn't deliver food.	鸭 duck	甲 first	笨 stupid
39	选民们/很积极/地投票 The electors actively participated in the voting.	积 accumulation	禾 cereal	信 letter
40	这病是/由蚊子/传播的 This disease was mainly spread by mosquitoes.	蚊 mosquito	虫 insect	机 machine
41	恐龙在/地球上/绝迹了 Dinosaurs have been extinct from the earth.	球 ball	王 king	猜 guess
42	桌子上/一堆统/计报表 There are a pile of statistical reports on the desk.	堆 pile	土 earth	查 check
43	同事们/闲聊着/走进来 The colleagues walked in, chatting with each other.	聊 chat	耳 ear	读 read
44	战士们/解救被/拐儿童 The soldiers rescued the abducted child.	救 rescue	求 beg	北 north
45	那儿有/一碑纪/念逝者 There is a tablet in memory of those who died.	碑 tablet	石 stone	己 self
46	此病需/要辅助/治疗法 This disease requires the adjutant therapy.	辅 supplement	车 car	初 first
47	他们有/颗炽热/爱国心 They were passionately patriotic.	炽 flaming	火 fire	用 use
48	风筝从/视野中/消失了 The kite was no longer in the vision.	野 wild	里 inside	报 report
49	这是最/活跃的/活火山 This is the most active volcano.	跃 jump	足 foot	古 ancient
50	他鄙视/轮船上/的工人 He looked down upon the workers on the ship.	船 ship	舟 ship	对 correct
51	那是一/道靓丽/的风景 That's a beautiful scenery.	靓 beautiful	青 blue	耳 ear
52	那几个/流氓行/骗为生 The rascals live on the credulity of the people.	氓 vagrant	亡 death	错 wrong
53	他在手/上划一/小口子 He snicked his hand.	划 gash	戈 dagger-axe	爱 love
54	孩子们/把档案/弄乱了 The children messed the files around.	档 records	木 wood	骨 bone
55	她站在/宽敞的/院子里 She is standing in the spacious yard.	敞 uncovered	尚 still	漏 leak
56	请你们/把帽子/摘下来 Please take off your hats.	帽 hat	巾 towel	信 letter
57	两只白/天鹅远/远飞来 Two white swans flew close.	鹅 goose	我 I	竹 bamboo
58	包厢里/只剩下/三个人 There are only three people in the box room.	剩 leave	乘 multiply	行 walk
59	他露出/很舒心/的笑容 He smiled happily.	舒 comfortable	舍 house	棚 hut
60	这是为/人妇职/责所在 This is the responsibility of a wife.	妇 woman	女 female	异 different