Examination of the Homophone Density Effect in Spoken Chinese

WANG, Wenna

A Thesis Submitted in Partial Fulfillment
of Requirements for the Degree of
Doctor of Philosophy
in
Psychology

The Chinese University of Hong Kong
August 2011

Thesis/Assessment Committee

Professor Alan Chun-nang Wong (Chair)
Professor John Xuexin Zhang (Thesis Supervisor)
Professor Liqiang Huang (Committee Member)
Professor Xiaojian Li (External Examiner)

Abstract of thesis entitled:

Study of the homophone density effect in spoken Chinese

Submitted by WANG, Wenna

for the ~~degree~~ of Doctor of Philosophy in Psychology

at the Chinese University of Hong Kong in August 2011

This study examined the homophone family size effect in spoken Chinese using monosyllabic and disyllabic homophones, and disyllabic words with a homophonic syllable. For each of these three types of words, the family size effect was measured in Exp. 1 to 3 using the auditory lexical decision task. In Exp. 1, participants showed slower response to monosyllabic homophones with a larger family size than those with a smaller family size. In Exp. 2, participants showed an opposite family size effect for disyllabic homophones (e.g. /mi4 feng1/密封 and 蜜蜂), that is, higher accuracy to disyllabic homophones than disyllabic non-homophones. In Exp. 3, the family size effect for disyllabic words with a homophonic syllable was found to be dependent on word frequency, an inhibitory effect was present for low frequency words, regardless of whether the first or the second syllable was the homophonic syllable. Using the same three types of words and the same task, Exp. 4 to 6 examined the neural basis of the family size effect with event-related potentials (ERPs). Results showed that the family size effect mainly resulted from access to the multiple semantic meanings associated with a spoken word. In addition, there was evidence that the multiple word-forms of a homophone were activated in a relatively early state (300-500 ms)

during spoken word recognition when the first syllable of a disyllabic word was homophonic.

The study provides the first systematic set of data for the homophone family size effect, demonstrating its presence in spoken Chinese word recognition.

The results indicate that such effects could be mainly attributed to activation of multiple semantic representations, in line with the primary goal of meaning access for spoken word comprehension. However, it seems that listeners sometimes do activate orthographic information for fine discrimination when the meaning was uncertain.

摘要

本研究旨在探讨同音字族因素（homophone family size ）在中文听觉词汇识别中的作用。在各实验中，实验统一采用了听觉词汇判断任务；实验 1 的结果发现，同音字族因素在中文单音词的识别中起抑制作用，即被试对同音字族越大的词的反应越慢。实验 2 则考察同音字族在双音节词词汇识别中的效应，结果发现字族因素起促进作用，即被试对同音词（例如，/mi4feng1/蜜蜂、密封）的反应正确率高于非同音词。实验 3a 则变化双音词的首字音节的同音字族（大、小）和整词频率（高、低）两个因素；结果发现字族因素与整词频率有交互作用，即字族因素仅在低频整词条件中体现出抑制作用。实验 3b，变化双音词的尾字音节的同音字族（大、小）和整词频率（高、低）两个因素；结果与实验 3a 一致。

然而，上述结果的本质原因是——基于同音词的多个同音字形还是多个语义的共同激活导致？对应于实验 1-3，采用相同的实验任务和实验材料，实验 4 -6 应用脑电技术进一步探讨该问题。结果发现：上诉三类材料中，不论是哪种类型的词，同音字族的效应均可在语义加工的相关脑电电极上发现差异；然而，对于同音字族位于首音节的情况，还发现了字形加工相关的脑电电极上发现同音字族效应。

上述结果表明，虽然同音字族因素的不用在各类词中所体现的效应并不统一，但该效应稳健的存在于中文听觉词汇识别过程中。其次，字族因素的本质属性归结为同音词多个语义的共同激活；然而，个别情况下相应的字形信息仍会被激活，从而辅助听觉词汇的语义更好通达。

Acknowledgements

I would like to first express my deepest gratitude to my thesis supervisor, Prof. John X. Zhang and for the continuous help and support of my Ph.D study and research. Very special thanks are also given to prof. Xiaojian Li and his students for the help to prepare experiment materials and carry out the behavior studies. My sincere thanks and appreciation also goes to the rest of my thesis committee: Prof. Liqiang Huang, and Prof. Alan C.N. Wong, for their encouragement, insightful comments, and hard questions. Appreciation is also given to my classmates and friends Ning Ning, Cao Hui, Shiwei Jia, Li Chunxia, Jiang Biyun, Fang Hailun, Fu Ruijuan, Su lijuan, Sun Nana and XiaoYi who gave me a lot of help in conducting the experiments.

TABLE OF CONTENTS

List of Tables

List of Figures

Chapter 1

Literature review

1.1 Homophone as a critical feature of spoken Chinese

The Chinese language differs greatly from many other languages in the world (Zhang, in press). According to the Chinese Frequency Dictionary (Liu, 1990), Chinese words include monosyllabic and multi-syllabic words composed of syllables ranging from 1 to 7, e.g., one-syllable words */shui3/* (水), seven-syllable words */da3 po4 sha1 guo1 wen4 dao4 di4/* (打破沙锅问到底). Monosyllable is the basic unit of spoken Chinese. Study of the processing of monosyllables and how they as components are related to whole words is essential to understanding the cognitive processes underlying spoken word recognition in Chinese. Below, I will first review some fundamental features of spoken Chinese.

Compared with spoken languages using an alphabetic script, Chinese monosyllabic words have a distinctive feature in that more than 75% of its syllables have homophones. This is because there are in total only a small number of around 1200 different syllables in Chinese. The homophonic mates for each syllable range from 2 to 48 or more (e.g. /jie2/ 节、结、洁、杰、婕……). Therefore, homophone family size is much larger than that in other languages. For majority of the multi-syllabic words, their constituent morphemes are homophonic syllables, e.g., 容忍，繁荣, and 融化 sharing */rong2/*, 按部就班 and 黯然失色 sharing the first syllable */an4/*. As Chinese spoken words, the abundance of homophone is a prominent issue, it is critical to understand how homophones or homophonic morphemes are processed.

Generally speaking, there are two important properties to homophone syllables that can be illustrated with the following example. All the following 15 characters ‘节、结、洁、截 、杰、竭、捷、劫、羯、桀、颉、碣、诘、睫、婕’ pronounced as /jie2/ in Chinese and they are called homophone mate with each other. The first property is that each word in this homophonic set has its individual usage frequency. The second property is that different syllables are associated with different numbers of homophone mates, referred to as homophone density (Ziegler et al., 2000). The central issues in homophone research are to understand how homophone frequency and homophone density affect the processing of monosyllable words. When a monosyllable is presented alone, e.g., when we only listen to a single syllable /jie2/, we would not be able to know which “jie2” it is. Such an absence of one-to-one correspondence between sound and meaning and orthography generates great ambiguity for spoken word recognition. However, listeners seldom experience comprehension difficulty of homophones in daily conversation. One line of studies suggests that context can help listeners to resolve lexical ambiguity. Early research in Chinese often examines sentence context to see how such dis-ambiguity process occurs.

1.2 Homophone processing in spoken language

With or without context, there are two questions that can be asked about homophone processing. First, are all the homophone mates of a particular syllable activated when one hears this syllable or only a subset of more appropriate ones?

Second, is there a competition process where syllables with a greater number of homophone mates will suffer stronger competition (Zhou, 1994) during spoken word recognition? If not, will the number of homophone mates influence processing of Chinese spoken word and how?

Li and Yip (1998) carried out several experiments in Cantonese by asking subjects to listen to a sentence and to name a visual character probe presented at the end of the sentence. For example, they may present the following two sentences followed by one of the three visual probe words.

1) 我唸喺動物園裏面最大嗰隻係象.

Probe: (a) 鼠 (biased) (b) 棍 (unbiased) (c) 质 (unrelated control)

2) 婆婆話佢行路唔方便，想揸支杖.

Probe: (a) 棍 (biased) (b) 鼠 (unbiased) (c) 质 (unrelated control)

The character '象' and '杖' are the two critical priming words that are homophonic to each other in Cantonese, both sounding as /chang/. As 'elephant' is the most frequently used word meaning among all homophones of /chang/, the authors considered it to be the dominant homophone, also referred to as the high frequency homophone. As 'stick' is the least frequently used meaning among all homophones of /chang/, the authors considered it to be the subordinate homophone, also referred to as the low frequency homophone. In sentence 1), the context biases the meaning of the sound /chang/ to be 'elephant' and in sentence 2), the context biases the meaning of the sound /chang/ to be ‘stick’. In addition to homophone frequency and context manipulation, they also included two more factors of homophone density and SOA. For homophone density, a given homophone has more (4 or more) or less (2 or 3) homophone mates. SOA is the onset time difference between the auditory prime word and the visual target word. For one level called Onset SOA, the target appeared the same time as the prime. For the other level called Offset SOA, the target appeared after the prime was turned off.

The results showed no significant response time differences across the different types of probes for the Onset SOA condition. There was some effect for the Offset SOA condition. Specifically, when the probe did not fit with the context meaning, response to the target ‘鼠’ in sentence (2) was faster than responding to the target ‘棍’ in sentence (1), indicating that the dominant meaning of ‘elephant’ was activated faster than that of the subordinate meaning of ‘stick’. However, when the probe fitted with the context meaning, response to the target “鼠’ in sentence (1) and to the ‘棍’ in sentence (2) were not different, indicating that no significant difference between dominant (high frequency) homophone and subordinate (low frequency) homophone. The message seems to be that context mediates the homophone frequency effect of homophonic monosyllable spoken word. However, in this study, there was no effect of homophone density observed, even in a further experiment using the gating paradigm (Yip, 2000). An additional piece of information from this study was that spoken words can be recognized before their acoustic ending, and different words have different isolate point (IP) for recognition, as reported in Li and Yip (1996).

Though the above studies support the hypothesis that contexts contribute in eliminating ambiguity during spoken homophone recognition, the effect of homophone frequency was not very clear. In other words, when listening to a homophonic syllable, are all its homophones activated as per the exhaustive access hypothesis or only a given homophone with context-consistent meaning per the context-dependency hypothesis?

To address this question, Yip (2007) conducted an experiment by using the same cross-modal naming task as above. He manipulated similar experimental variables to contrast high vs. low frequency homophone, homophone density, context type, along with SOA, and probe type. One set of example sentences and probes are shown below.

1) 間屋咁焗，你快啲去開晒啲/cheung,窗/. ---dominant

2) 軍火專家話呢啲全部都係真/cheung,槍/. ---subordinate

3) 我要你哋而家即刻走去開/cheung,窗,槍/. ---ambiguous

The probe was one of the following, 門 (moon4), 彈 (daan6), 麵(min6), 豆(dou6). Note that 麵 is a high frequency control for 門; and 豆 is a low frequency control for 彈.

One novel aspect of this study was that Yip added an ambiguous context as in sentence 3) where both the dominant meaning and the subordinate meaning were consistent with the context. As before, there were contexts biasing the dominant meaning and the subordinate meaning of a homophone conditions, as in sentence 1) and 2). The ambiguous context condition provided a baseline for testing whether homophones of different frequency were activated when they were equally likely primed by the context.

Another improvement was that separate controls were provided for the dominant probes and subordinate probes. For the SOA factor, there were three levels, (a) IP (isolation point) meaning the least presentation duration of an acoustic word when the word can be correctly recognized, (b) OS, the probe appeared at the offset time of the acoustic homophone presentation, (c) the probe appeared 300 ms after the offset of the homophone.

The results corroborate outcomes of previous studies. First, there was no homophone density effect. Second, the context interacted with word frequency. For the ambiguous context condition, the dominant words showed shorter response time as the SOA was longer (IP > OS > 300ms). Clearly different from Li and Yip (1998), there was a homophonic word frequency effect and both high and low frequency homophones were activated in the non-predictive context condition, i.e., when the context did not prime the target probe.

Although most of studies showed no homophone density effect, this effect was once demonstrated in one study (Yip, 2002). Similar as other studies, Yip used auditory sentence materials with the critical change that the ending homophone was presented in a gating paradigm. Participants were asked to write down what they believe the word was and also their confidence level under a particular gating time. From time course data analysis, participants' correct recognition time was found to be influenced by both context type and homophone density, and their interaction.

The mean recognition time was not different between high and low density conditions in the high constraint context (both dominant and subordinate context type). But under the ambiguous context, the recognition time was longer in high homophone density condition than did in low homophone density condition. This inhibitory effect of homophone density may indicate that the more competitive the homophones the more time was needed to settle on the correct answer. The most surprising contribution from this study is that it first documented an inhibitory homophone density effect in spoken word recognition. It also indicates that Chinese spoken word can be recognized sooner before the ending of the auditory input with context than when presented in isolation (mean IP is 23.4%, 27.6% and 55% for dominant, subordinate and ambiguous context, respectively).

To summarize, except for the homophone density effect, the above studies all showed a clear picture that both dominant and subordinate homophones are activated when participants listen a spoken homophone and that context plays an important role in selecting the correct meaning. Further, in an ambiguous context, homophone frequency has a positive effect.

1.3 Processing of isolated spoken words

In speech with enough context information, there seems to be not much difficulty in sound-meaning matching for spoken monosyllabic word. However, it remains unclear to what extent spoken homophonic word can be identified without sentence context. Critically, do homophone frequency (e.g., dominant meaning) and homophone density have any effects in processing isolated spoken words?

To address these questions, researchers have examined the role of the above related variables by using several different paradigms, including the cross-modal matching task, the cross-modal phonological matching task, the cross-modal priming task, and the dictation task.

In one study (Chen & Ning, Exp. 2, 2005), the data demonstrated a homophone density effect indirectly. Actually, this study aimed to investigate whether there is a homophone effect in Chinese language, defined as the longer response time to homophones than to non-homophonic control words in lexical decision tasks (Pexman et al, 1999; 2001). Chen and Ning adopted a cross-modal priming task and asked participants to make a lexical decision for a visual homophonic character after hearing its own syllable. Their data showed that individual homophonic frequency has a facilitatory effect, and that characters with many homophones had longer response time than did characters with less homophones, under the low frequency condition. It means that the homophone density inhibitory effect only occurs in the low frequency condition.

In a set of dictation studies (Zhang et al, 1999; Yi et al, 2001), children were asked to write down the second morpheme of disyllabic spoken word, e.g. listen to /lao3shi1/ and write down "师". These earlier data found that children showed replacement errors by writing out the homophone mates of the targets when the target words had many homophone mates, which hinted that the homophone density may possibly interfere with homophone processing. To further examine this possibility, Zhou and Shu's (2008) employed a cross-modal matching task by manipulating (1) the homophone density of ending syllable in disyllabic spoken word, (2) the whole words semantic transparency and (3) the ISI within auditory and visual stimuli by. The participants' task was to decide whether the visual character was the same as the ending morpheme of the auditory word. The results from 3 different ISI conditions revealed a robust homophone density effect with low homophone density stimuli showing faster response time. The results suggest that homophone indeed competed more strongly with other members of the family when there are more homophone mates, relative to when there are fewer mates.

Briefly, these two studies seem to indicate that homophone density effect is indeed present for isolate spoken word processing.

1.4 Processing of disyllabic homophone in spoken Chinese

Other than monosyllable words, disyllable word in Chinese may also have homophones. Although compared with monosyllable words, the homophonic rate is sharply reduced in disyllabic words, there is still about 15% of disyllables (totally 4033 out of 31180 disyllables) with homophonic mates ranging from 2 to 6 (e.g. /bao3jian4/ 保健、宝剑；/shi4li4/ 势利、示例、视力、势力、事例). Therefore, the study of the disyllabic homophone word may also help to reveal spoken word processing. The effect of homophone density effect has not been addressed in any research so far, partly because there are only 2 homophone mates for most (about 92%) disyllabic homophone spoken words.

Similar to what has been done in monosyllable homophone, there are some research probing meaning access of ambiguous spoken words for two-syllable homophone word (Shu & Zhang, 2000; Zhou et al., 2003).

Shu et al. (2000) employed a cross-model semantic priming-lexical decision task. In their Exp. 1, the homophone was presented alone as a prime which was followed by a visual probe, participants were asked to make lexical decision tasks to the probe. They manipulated two variables, (a) the frequency of the homophone (dominant and subordinate condition, note all the two-syllable words they used contain only two homophone mates; (b) the ISI between prime and probe (0 and minus 150 ms).Sample stimuli they used are shown below.

	Prime	Probe
Dominant	/dian4yuan2 / (电源)	开关
Subordinate	/dian4yuan2 / (店员)	伙计
Unrelated	/ lun2tai1 / (轮胎)	开关
Unrelated	/chan3pin3 / (产品)	伙计

In their Exp. 2, the same presentation method was adopted, except that the prime was embedded into a sentence. And the context bias factor was added. The results were as follows, (1) in isolated word presentation experiment, compared to non-homophonic control condition, the facilitatory priming effects were observed in both the dominant and subordinate homophone conditions, regardless of the ISI level; (2) In the sentence presentation experiment, priming effects were found only when the sentence context was consistent with the meaning of the homophone; both dominant and subordinate homophone priming effects were found at 0 ms ISI; but only dominant homophone priming effect was found at the -150ms ISI condition. The overall results indicate that both meanings are activated for isolated two-disyllable homophones. When in the sentence context, the frequency effect occurs very early during spoken word processing. That is, context plays a role in helping select the proper meaning of the homophone and this is consistent with findings for monosyllable words.

A different view was proposed by Zhou (2003) for meaning access of ambiguous spoken two-disyllable homophone. Employing the same task as in Shu et al., Zhou used an auditory word and a sentence as context in the Exp. 1 and 2 respectively. The ISI variable level of 0 ms was changed to 50 ms. All other settings were the same as in Shu et al. (2000).

As the example shown below, for Exp. 1, the prime consisted of two words with the first word serving as the context for the second prime word.

prime pair	probe			
/台灯 — 电源/	开关	杂志	服务	危险
/柜台 — 店员/	开关	杂志	服务	危险

For the 50ms ISI condition, their results replicated the 0ms ISI condition in Shu et al. (2000). But for the -150ms ISI condition, they found no priming effect for dominant and subordinate homophone in neither the sentence context nor the word context condition. When the context was consistent with the homophone's proper meaning or unrelated with either of the homophone meanings, the subordinate meaning was inhibited.

In sum, although both studies show that the two meanings of a disyllabic homophone word are activated, the homophone frequency effect is not consistent. In the Shu's study, the subordinate meaning in the consistent context was not activated, and in the Zhou et al. study, the subordinate meaning in the unrelated context was inhibited.

Chapter 2

The present studies

The literature review indicates a complex situation regarding the effects of homophone frequency and homophone density on spoken word processing in Chinese. For disyllabic homophone words, the basic pattern is that there is a frequency effect for the dominant meaning of a homophone but the nature of the effect seems to be different in different studies.

For monosyllabic homophone words, there is even less consistency. The dominant frequency effect was significant in isolate word experiments, and in ambiguous contexts. In either case, there was no additional information in word meaning selection. The effect was not present when the context does bias a particular meaning.

The effect of homophone density effect is less obvious. As can be seen from the review, most studies failed to observe any density effect in sentence context (Li P. & Yip, 1996; Yip, 2000; 2007), except one (Yip, 2002). But the effect exists in all isolate word studies (Chen & Ning, 2005; Zhou & Shu, 2008).

There are several factors to be considered in understanding this pattern of results.

For all studies failing to show any density effect, the responding targets were visual characters semantically related to the homophone target word. Participants based their responses on a visual word instead of the critical auditory homophones. This kind of paradigm may not directly reveal processes about spoken word recognition. That is, one may be in fact studying visual but not spoken word recognition. It is well-known that the frequency effect is sensitive for visual word processing and it is also reported that the homophone density effect is not sensitive to visual word processing (Taft, 1997; Chen, 2005). In contrast, for all studies showing density effects, the tasks required participants responding to the critical homophone words by using dictation or phonologic matching or primed lexical task. Another factor unclear is that the density effect sometimes occurred for low frequency words (Chen, 2005, exp.2) and sometimes for high frequency words (Zhou & Shu, 2008).

Briefly, the inconsistencies concerning the effect of homophone density on monosyllabic spoken word processing may result from different paradigms, including stimulus presentation modality (e.g., visual, auditory, or cross-modal), different stimuli selection criteria. These factors need to be carefully considered in future research.

In a pilot study of mine, I tried to integrate the good features in the former paradigms to develop a better controlled task to minimize confounds. The task was an auditory-visual phonological matching task manipulating both homophone density (2 levels: sparse/dense) and frequency of homophone mate within each monosyllable (3 levels: high/medium/low). Given that the auditory-visual matching task unavoidably involves visual words recognition process, a separate visual word lexical decision task was used to covariate out any potential frequency effects from visual word processing. ANCOVA showed a main effect for both morpheme frequency and homophone density but no interaction between the two. The frequency effect was facilitative in that the high frequency words were responded to faster than low frequency words, and this effect was present for all pair-wise frequency comparisons (high vs. low, high vs. medium, medium vs. low). The density effect was inhibitory in that larger family size homophones were responded to slower than smaller family size homophones and the effect was present only for the high and middle frequency words. However, results from the visual lexical decision task alone showed only a main effect of frequency but no effect of homophone density or its interaction with frequency. Clearly, processing modality must be seriously considered in studying the density effect.

A missing factor is syllable frequency, also called phonological frequency, which refers to the cumulative frequency of all homophone mates of a word. Studies found significant syllable frequency effect in Chinese spoken disyllabic word recognition (Zhou & Marslen, 1994, 1995) and in written Chinese (Ziegler et al, 2000). However, no study has examined this factor for monosyllabic words. There is further the possibility that homophone density factor may interact with syllable frequency and research should include both factors in the same study.

Granted we can establish the homophone density effect clearly, we are still left with the task of explaining how it arises. The existing research on spoken word recognition generally assumes that more homophone mates lead to more intensive competition among homophones. However, there has been no study that addresses the specific mechanism of this competition, such as at which level does such competition occur and what is the nature of this inhibitory effect? Specifically, whether this competition involves the multiple orthographic codes or the multiple semantic codes associated with a given syllable?

Given the primacy of spoken language over visual language, much language research has been concerned with the impacts of spoken language processing on visual language processing, such as the extensively investigated role of phonology in visual word recognition. However, research over the past decade has documented an interesting feedback consistency effect showing impacts of orthography on phonology. One critical piece of evidence is the finding that words that are associated with several possible spellings (e.g., /-ip/ as -eap or –eep) are judged more slowly in auditory decision tasks than those that are consistently associated with one spelling (e.g., /-ʌ note here?? k/ always as -uck). Similar effects have also been demonstrated with other paradigms such as naming, rime detection, phonological priming, and with ERP measures (Taft et al., 2008; Ziegler, Petrova & Ferrand, 2008; Seidenberg & Tanenhaus, 1979; Ziegler et al., 2003; Ventura et al., 2004; Chereau et al., 2007).

One popular explanation of these results is that orthographic representations are activated online in spoken word recognition, which then affects phonological processing via a feedback loop from phonology to spelling (in reference to the feed-forward loop in reading from spelling to phonology). Inconsistent mappings along this loop lead to competitions among multiple orthographic representations and slow down the word recognition process in reaching system equilibrium. This line of research highlights the interactive nature of linguistic processing and presents particular challenges to existing models in spoken word recognition which generally does not assume any role of orthography.

In a recent ERPs study, Perre & Ziegler (2008) also found a stronger on-line orthographic activation in spoken word lexical decision task. In this study, they manipulated the orthography inconsistent positions at the early or late of monosyllabic

French words (early: rhume pronounced /rym/ vs. late: noce pronounced /nos/; where /ry/ and /os/ can be spelled more than one form, which is like /ost/ in English may be spelled as "toast-ghost"), with mean acoustic duration of stimuli being 644 ms.

The result showed that compared to the control consistent word, a N320 (190 ms after the onset of the inconsistency) as well as a N600 component was evoked by early inconsistent words, and late inconsistent only elicited a N600 (160 ms after the onset of the inconsistency) component. N320 indexes bidirectional mapping between orthography and phonology and N600 indexes the inconsistent words have been identified and discriminated from the normal word completely.

The results provide strong evidence that how a word is spelled matters in the auditory processing task suggesting their online activation as indexed by the various ERP components. More recently, Pattamadilok, & Ziegler et al (2009) also addressed the same question by employing a semantic categorization task with the ERP technique. The logic for this study is that the lexical decision task may require participants to 'consult' the orthography to make decisions, but a semantic task would not need to explicit analyze the orthography information. If the orthographic consistency effect is found in this task, it would provide evidence for the non-strategic nature of orthographic activation in spoken word recognition. They manipulated the whole word frequency and the orthography inconsistent positions at the first or second syllable of disyllabic French words (first syllable: champagne vs. second syllable: "printemps"; where the syllable "cham" and "temps" can be spelled more than one form) in the experiment. They found a first syllable inconsistent effect in 300-350 ms time window and a second syllable inconsistent effect in 400-425 ms time window as well as word frequency effect in 550-750 ms window. The results further support the view that orthography can be activated online non-strategically, or not because of task demand.

Parallel to such findings in alphabetical languages, a reasonable hypothesis is that homophone density effect in Chinese spoken word recognition may also be caused by competitions of multiple orthographical forms associated with a homophonic syllable if we show that such orthographic codes do get activated online during spoken word recognition.

However, there may be some differences deserving consideration in Chinese. The Chinese writing system is pictographic (logographic) where sound and spelling mapping is not highly consistent. By introspection, in daily life conversation, people seem to seldom be aware of any activation of orthography. I described earlier that even the homophone frequency effect is absent in biasing context. When there is context, typically the homophone density effect does not show up. So it is possible that whether the orthography would be on-line activated in spoken Chinese may be conditional. That is, it may occur for small units including monosyllables and disyllables, but may not for larger units, such as idioms and sentences.

Moreover, generally, people listen to speech mainly for the extraction of meaning and it is usually unnecessary to activate the orthography. Therefore, the possible reason for why orthography gets activated in processing monosyllable and disyllable spoken words may be due to lack of enough information, e.g., homophone ambiguity makes meaning access more difficult, and the activation of orthography information can play a role in helping classify the meaning of an input. During language acquisition, people may combine information from both written and spoken language to gradually establish strong links between the two and will be automatically utilized. Therefore, for proficient readers, even if orthography information is not necessary, orthography still gets activated due to these pre-established strong connections. This lends some reason to examine semantic related components N400 (N4) other than orthography related ERP component.

There are two major goals of the present study. The first is to see whether one can establish the presence of the homophone density effect in spoken Chinese word recognition. The second goal is to use event-related potentials, as in Perre & Ziegler (2008), to show that such family size effect, if present, does not arise from task demand and strategies, but rather reflect the consequence of online orthographic code activation. Competition between the multiple orthographic codes associated with a given syllable may under the behaviorally observed homophone density effect. Furthermore, in addition to activation of multi-orthography for homophone, there may also be multiple meanings activation for the homophone. The situation is different from that in alphabetic languages. For example, when listening to /gəust/, though both ‘ghost’ and ‘ghoast’ can be activated, but ‘ghoast’ is not a real word in English, thus only orthography but not semantic inconsistent effect could occur in alphabetic language. At the end, I will also explore whether the grain size of language input units affects the online orthographic activation.

Towards these ends, I designed the following six experiments. The auditory lexical decision task and a detection task would be used to study monosyllabic homophone word in Exp. 1 and 4, and disyllabic homophone words in Exp. 2 and 5. Exp. 3 and 6 would test the homophone density effect when the homophones are part of a disyllabic word, using both behavioral and ERP measures. Based on existing ERP studies, I expect for isolated word input unit, homophones with more homophonic mates may elicit a larger N320 and N600 component than that with less homophonic mates. Apart from the homophone density factor, as mentioned earlier, for some of multisyllabic spoken word units, the whole word frequency may be a relevant factor. Chen (2003) reported that orthography is more likely to activate for words of low whole word frequency. Pattamadilok et al. (2009) also found a significant frequency effect for spoken French word with ERP at the 550-750 ms time window. So I will also manipulate this factor together with the homophone density factor.

Special attention needs to be paid concerning the time window of N4. The N4 might not be measured only around 400ms after the onset of the stimuli. This is because different words units have different physical durations leading to different onset times of N4.

This set of study will be the first systematic research into homophone processing in spoken Chinese word recognition. It will reveal the impact of written system on speech comprehension. Current models of spoken word recognition ignoring the impact of written language need to be reconsidered if there are indeed impacts of orthography on speech comprehension.

Chapter 3

Experiments

3.1 *Exp. 1 Behavioral demonstration of the homophone density effect for monosyllabic homophones*

This experiment evaluates the hypothesis that the homophone density effect is present for Chinese monosyllabic syllable during spoken word processing.

3.1.1 Participants

Thirty-right-handed healthy college students who were fluent Mandarin speakers in South China Normal University were recruited (mean age 23 years; range 21-26; 14 males). All had normal hearing and normal or corrected-to-normal vision. Written informed consent was obtained in accordance with guidelines from IRB in the Chinese University of Hong Kong.

3.1.2 Stimuli and design

The auditory stimuli were single syllables selected from a Chinese syllable library made in the Neurophonetics Laboratory of South China Normal University. All syllables were read by a Native Chinese female with Standard Mandarin and recorded in a sound-proof room using the AKG C410 microphone system and a Sony portable digital recorder TCD-D8 (sampling frequency 16 bit/44.1 kHz). The Sound Forge software would be used in offline editing the sound form and the sound would be saved as wav format. And the duration of stimulus was shorter than 500 ms (mean duration was about 430 ms).

The design was a one-way repeated-measures design with two levels of homophone density, large vs. small. Here, large homophone density size syllables were defined as the number of homophone mate being larger than 9, and the number of small density size syllable ranged from 2 to 8. For better control, the following factors were matched across the two levels, syllable frequency (SF), frequency (HF) and stoke number of the highest homophone within the homophone family of a syllable, salience (ratio between frequency of the highest homophone and that of the second-to-highest homophone with the syllable). Thirty monosyllable words were selected for each density group (see Figure 1 for details). Another 60 non-words were included as fillers. All non-words, each including an onset, rime and tone parts, can be pronounced but do not exist in Chinese.

Table.1 the matched information of stimuli cross the conditions in exp.1

stimulus condition	density size	Syllable frequency	HF of syllable	Salience	Stoke number	Example stimuli	Mean duration
B	13. 3	9306. 2	6338. 3	5. 8	8. 1	Jiao3(15)	431
S	4. 9	9263. 6	6340. 3	5. 2	7. 5	Dui4(4)	426

Note: HF=frequency of the highest homophone within the homophone family of a syllable

The number in round bracket refers to the number of homophone mates in each syllable

3.1.3 Task and procedure

Subjects were asked to perform an auditory lexical decision task. The auditory stimuli were delivered with a pair of earphones binaurally by using the E-prime software. Participants sat comfortably before a computer screen inside a quiet testing room. The stimuli were presented in random order. After completing every 40 trials, participants were asked to take a 1-min break.

Each trial started with a cross fixation point (+) presented for 500 ms in the center of the screen, then the spoken stimulus was presented while the fixation cross remained on the screen for 3000 ms. Participants should press the response keys as quickly as possible after they heard the auditory stimulus. Once the subject pressed the response key, the fixation disappeared immediately. And inter-trial interval was 1000 ms (blank screen). No feedback was provided during the experiment. There were 20 practice trials with feedbacks before the experimental session.

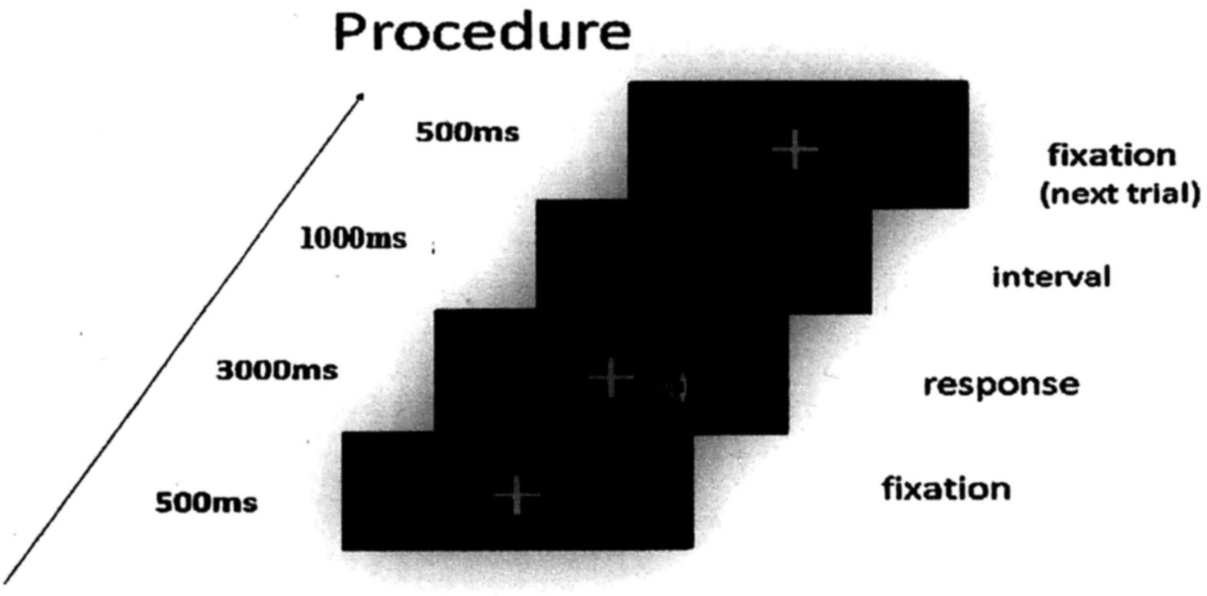

Fig. 1 The main procedure of one trial in Experiment 1

3.1.4 Data analysis

RTs were measured from target onset to response. RTs greater than three SDs beyond the overall mean of a participant were discarded (less than 1%). Reaction times to stimuli and accuracy of performance were analyzed by paired t-tests.

3.1.5 Results

Mean RT and ACC for large and small homophonic density size syllable were 836 (SD=72) vs. 794 (SD=71) ms, and 92.4% (SD=0.06) vs. 94.9% (SD=0.05), respectively. The results of the paired t-test showed a significant main effects of Homophone density in RT (t (29) = 4.15, $p<0.05$), and in ACC (t (29) = 2.03, $p=0.05$), indicating a processing advantage for small homophone density words.

Table 2 Mean RTs (ms) and accuracy rates for 2 conditions in Experiment 1.

	RT		ACC	
condition	B	S	B	S
	836.0	794.0	92.4	94.9
	(71.9)	(71.2)	(5.7)	(5.1)

Note: Standard deviations are shown in parentheses.

3.1.6 Discussion

The results established a valid paradigm to demonstrate a clear inhibitory homophone density effect for processing isolated monosyllable Chinese spoken words, which was consistent with the reported homophone inhibitory effect in auditory sentences (Yip, 2002) and in isolated auditory word priming task (Chen & Ning, 2005). And it suggests that when perceiving an isolated monosyllable spoken word, a word with a larger family size would active more homophones than a word with a small family size; and the more homophone mates there were, the stronger competition would be induced for processing the syllable. This result was different from the homophone density effect in English (Stephanie, 2006), where using the same auditory lexical decision task, Stephanie found that homophone words were responded to faster than non-homophone word.

3.2 *Exp. 2 Behavioral demonstration of the homophone density effect in disyllabic homophones*

Exp. 1 showed that monosyllable words with larger homophonic size have longer response time than that with smaller homophonic size in the auditory lexical decision task. Compared with the mono-syllabic homophones, the homophonic members decrease sharply for disyllabic words (the homophone members range from 2 to 6, and the mode is 2). Exp. 2 was intended to explore the effect of homophone density in Chinese disyllabic homophones.

3.2.1 Participants

Thirty college students who were fluent Mandarin speakers in South China Normal University were recruited (mean age 22 years; range 19-25; 13 males). All had normal hearing and vision or corrected-to-normal vision. Written informed consent was obtained in accordance with guidelines from IRB in the Chinese University of Hong Kong.

3.2.2 Stimuli and design

The homophone density size of the whole word was the only factor in this experiment. There were two levels to this factor, one involving disyllabic words with homophone mates and the other involving disyllabic words that are non-homophonic. The first condition was called the homophone condition (called the HW condition) and the second the non-homophone condition. The second condition was further broken down into two sub-conditions. One is the High frequency word control condition (called the HF condition) involving non-homophonic disyllabic words whose whole word frequency matches that of the homophone words with the highest frequency. The other was the cumulative whole word control condition (called the CF condition) involving non-homophonic disyllabic words whose whole word frequency matches that of the cumulative word frequency of the homophone words. In total, 174 disyllabic words were chosen in Exp. 2 with 58 words for each condition.

For better control, the following factors were matched across the three conditions: whole word frequency, the homophone density size of the first and second syllable of the word, morpheme frequency as well as stroke number of the first/second morpheme, cumulative syllable frequency of the first and second syllable. The detail information of stimuli was listed in the Table 3.

Table 3 Stimulus characteristics for all conditions in Exp. 2

Condition	homo density	Highest freq	Accumulate freq	1st MF	2nd MF	1st family	2nd family	1st SF	2nd SF	1st Highest HMF	2nd Highest HMF
HW	2.6	557.7	675.0	3999.7	4237.5	13.2	16.5	17062.2	33106.1	11121.2	23701.7
HF	1	557.1		3973.3	4287.1	13.0	16.0	17284.7	32057.8	12162.1	23077.8
CF	1		671.4	3950.6	4256.3	12.6	15.1	17139.3	32137.9	11523.3	23735.4

The same number of non-words were constructed and used as fillers. A non-word was constructed by two real monosyllables, but the conjunction was not a real Chinese word (e.g. /kun1luan3/坤卵).

The auditory stimulus recording procedure was the same as that in Exp. 1. And the mean duration of stimulus was 560 ms (the mean duration for HW, CF and YF stimuli were 547,566 and 568 ms, respectively).

3.2.3 Task and procedure

Both the task and procedure were identical as in Exp. 1.

3.2.4 Data analysis

RTs were measured from target onset to response. RTs greater than three SDs beyond the overall mean of a participant were discarded. These cutoffs led to the rejection of less than 1% of the observations. RTs to stimuli and ACC of performance were both analyzed by one-way repeated-measures analyses of variance (ANOVAs).

3.2.5 Results

The ANOVA revealed no significant main effect of condition in RT ($F (2, 58) = 2.06, p = 0.13$), but a significant effect in ACC ($F (2, 58) = 8.83, p < 0.001$). Accuracy in the HW condition was significantly higher than that in the CF condition ($t(0.5, 29) = 2.12, p < 0.5$), in the HW condition than in the HF condition ($t (0.5, 29) = 4.00, p < 0.001$), and in the CF condition than in the HF condition ($t (0.5, 29) = 2.23, p < 0.05$). The first two comparisons combined indicate an advantage for homophonic disyllabic word with higher accuracy. And the last comparison indicates a facilitatory word frequency effect for non-homophone disyllabic word.

Table 4 Mean RTs (ms) and accuracy rates (%) for 3 conditions in Experiment 2.

	RT			ACC		
condition	HW	CF	HF	HW	CF	HF
	810.79	811.77	820.77	91.88	89.77	87.95
	(70.47)	(71.25)	(66.56)	(0.05)	(0.06)	(0.06)

Note: Standard deviations are shown in parentheses.

3.2.6 Discussion

By comparing the HW condition with the other two control conditions (CF & HF), the results demonstrate a clear homophone density effect for processing of isolated disyllable Chinese spoken words. However, the homophone density showed a facilitatory effect in the current experiment, indicating that a disyllable spoken word with homophone mates could be more easily recognized than a non-homophone word, which was the same as that found in English (Stephanie, 2006).

Different from ACC, there was no condition effect with the RT index. The most likely reason is that the mean durations for words in different conditions were significantly different (F (2, 171) = 7.42, $p < 0.01$). The duration for the HW stimuli was significant shorter than that in both the CF and HF conditions ($ps < 0.01$; 547 vs. 566 vs. 568 ms). Such stimulus duration differences may contaminated the response times.

3.3 *Exp. 3 Behavioral demonstration of the effect homophone density and word frequency in disyllabic spoken words*

For most of the disyllabic words, their component syllables are usually homophonic. Also for such words, a critical factor to consider is the whole word frequency. In Exp. 3, I manipulated the density size of the constitutive syllable of the disyllabic word as well as the whole word frequency. The goal was to examine whether the homophone family size effect was present and how it might be affected by whole word frequency.

3.3.1 Participants

Thirty-two college students who were fluent Mandarin speakers in South China Normal University were recruited (mean age 22 years; range 19-26; 11 males). All have normal hearing and vision or corrected-to-normal vision. Written informed consent was obtained in accordance with guidelines from IRB in the Chinese University of Hong Kong.

3.3.2 Stimuli and design

A 2 (family size of constitutive syllable: large/small) x 2(whole word frequency: high/low) factorial within-subject design was used. For half of the participants, the size factor was manipulated for the first constituent syllable (Exp. 3a), for the other half, it was manipulated for the second constituent syllable (Exp. 3b).

For Exp. 3a, the stimuli included 34 sets (34*4 = 136 words in total) of disyllabic Chinese compound words, with both characters being free morphemes. Each set consisting of 4 words the first syllable of which was manipulated while holding their second syllable (character) constant, e.g., /jing4ai4/敬爱, /xi3ai4/喜爱, /bo2ai4/博爱, /pian1ai4/偏爱. The four words in each set fell into a 2 x 2 design crossing over the homophone density of the first syllable (large vs. small) and the frequency of the whole word (high vs. low). By using the same second character (syllable), irrelevant stimulus characteristics can be nicely controlled across the four conditions.

For Exp. 3b, the similar 2x2 design and stimuli control were adopted, except that the manipulated syllable was the second constituent syllable of the disyllabic word. The stimuli included 33 sets (33*4 = 132 words in total) of disyllabic Chinese compound words, with both characters being free morphemes. Each set consisting of 4 words the second syllable of which was manipulated while holding their first syllable (character) constant, e.g., /huo3yan4/火焰, /huo3zai1/火灾, /huo3lu2/火炉, /huo3tan4/、火炭. /

For better control, the following factors were matched across the four conditions: whole word frequency, the homophone density of the first (or second) syllable of the word, morpheme frequency as well as stroke number of the first/second morpheme, cumulative syllable frequency of the first (or second) syllable. The detail information of stimuli was listed in the Table 5 & 6. The auditory stimuli were recorded in the same way as in Exp. 2. The mean durations of stimuli were 641ms and 645ms for Exp. 3a and 3b respectively.

Table 5 the matched information of stimuli cross the conditions in exp.3a

condition	example	word frequency	1st family	2nd family	1st MF	2nd MF	1st SF	2nd SF	1st stoke	2nd stoke	duration
HB	敬爱	407.8	11.4	6.1	1102.2	3800.4	4710.0	6474.2	9.4	8.0	647
HS	喜爱	416.1	4.1	6.1	1158.2	3800.4	4740.7	6474.2	9.9	8.0	636
LB	博爱	30.1	12.9	6.1	1093.5	3800.4	4791.9	6474.2	8.9	8.0	655
LS	偏爱	30.4	3.6	6.1	1120.6	3800.4	4777.1	6474.2	9.1	8.0	641

Table 6 the matched information of stimuli cross the conditions in exp.3b

condition	example	word frequency	1st family	2nd family	1st MF	2nd MF	1st SF	2nd SF	1st stoke	2nd stoke	duration
HB	火焰	562.6	5.5	13.2	5554.5	1290.8	8554.7	5770.8	7.1	8.9	648
HS	火灾	560.8	5.5	4.8	5554.5	1351.6	8554.7	5790.4	7.1	8.8	633
LB	火炉	30.3	5.5	11.3	5554.5	1439.5	8554.7	5910.3	7.1	8.8	647
LS	火炭	29.2	5.5	4.1	5554.5	1339.0	8554.7	5963.8	7.1	8.8	636

3.3.3 Task and procedure

The task was spoken word lexical decision task. The procedure was the same as in Experiment 2.

3.3.4 Data analysis

RTs were measured from target onset to response. RTs greater than three SDs beyond the overall mean of a participant were discarded. These cutoffs led to the rejection of less than 1% of the observations. RTs to stimuli and ACC of performance were both analyzed in a 2x2 two-way repeated-measures analyses of variance (ANOVAs).

3.3.5 Results

Exp. 3a results

The ANOVA revealed a significant main effect of word frequency in RT (F (1, 31) = 59.52, $p < 0.001$), and ACC (F (1, 31) = 39.47, $p < 0.001$), indicating an advantage for high frequency words. In addition, a significant main effect for Homophone density was obtained in RT (F (1, 31) = 42.78, $p < 0.001$), and ACC (F (1, 31) = 11.09, $p < 0.01$), indicating an advantage for small homophone density words.

A significant interaction between word frequency and homophone family size was also found in RT (F (1, 31) = 22.09, $p < 0.001$) and ACC (F (1, 31) = 15.75, $p < 0.001$). Further simple effect analyses suggest that the homophone density effect was restricted to low frequency words in RT (F (1, 31) = 41.89, $p < 0.001$) and ACC (F (1, 31) = 19.59, $p < 0.001$), it was not obtained for high frequency word ($p> 0.2$). And the word frequency effect was only obtained for large homophone density words, both in RT (F

(1, 31) = 97.45, $p < 0.001$) and ACC (F (1, 31) = 48.34, $p < 0.001$).

Table 7 Mean RTs (ms) and accuracy rates for 4 conditions in Experiment 3a & 3b.

	exp.3a				exp.3b			
condition	HB	HS	LB	LS	HB	HS	LB	LS
RT	889.1 (107.8)	881.8 (100.9)	973.3 (107.4)	898.7 (109.7)	879.0 (96.4)	882.2 (96.8)	941.5 (102.9)	899.1 (110.9)
ACC	91.3 (0.05)	94.7 (0.05)	85.3 (0.07)	87.7 (0.08)	92.9 (0.04)	91.3 (0.06)	86.60 (0.06)	86.9 (0.06)

Note: Standard deviations are shown in parentheses.

Exp. 3b results

For the RT data, there was a significant main effect of word frequency ((F (1, 31) = 38.93, $p < 0.001$), indicating an advantage for high frequency words. In addition, a significant main effect of Homophone density was obtained (F (1, 31) = 8.23, $p < 0.01$), indicating an advantage for small homophone density words. A significant interaction between word frequency and homophone family size was also found (F (1, 31) = 23.86, $p < 0.001$). Further simple effect analyses suggest that the word frequency effect was obtained for both the large (F (1, 31) = 48.12, $p < 0.001$) and small (F (1, 31) = 5.06, $p < 0.05$) homophone density words. But the homophone ~~density~~ effect was restricted to low frequency words in RT (F (1, 31) = 37.15, $p < 0.001$), and not obtained for high frequency words ($p > 0.4$).

For the ACC data, there was only a main effect of word frequency ACC (F (1, 31) = 79.19, $p < 0.001$), with the high frequency words showing higher accuracy than the low frequency words. There was no main effect for family density or two-way interaction in ACC (p values > 0.5).

3.3.6 Discussion

Both Exp. 3a and 3b showed the same pattern of interaction effect between whole word frequency and family density size of constituent syllable. The family density size effect was found only in the low frequency words. The word frequency effect was found for the big family size word in both Exp. 3a and 3b. A significant word frequency effect was also found for the small family size words in Exp. 3b. The finding was consistent with what has been found for low frequency homophones in Chen's study (2005). The possible reasons were (1) the low word frequency led to more difficult access to the whole word; (2) listening to words with a larger density size homophonic syllable could activate multi-homophones, and the stronger competition between different homophonic orthography would make it harder to access the word meaning. Thus, the homophone density effect was more sensitive for the low frequency words.

Exp. 4 ERP evidence of homophone family size effect for monosyllabic homophones

Exp. 1 showed that monosyllable words with larger homophonic size showed longer response time than those with smaller homophonic size in auditory lexical decision. Exp. 2 was intended to use event-related potentials (ERPs) to reveal the neural basis of such homophone density effect. Specifically, I tried to understand the cognitive processes through which homophone size influences spoken word recognition, in particular, whether it involves competition among multi-orthographic or

multi-semantic activations associated with a particular homophone syllable.
Based on the literature, there are several ERPs components or time windows of interest.

Firstly, the previous auditory studies showed that N1 and P2 components were typically related to physical analysis of the stimuli; and normally, the enhancement of N1 often accompanied the decline of the P2 (Crowley & Colrain., 2004).

Secondly, Perre and Ziegler (2008) found an online activation of orthography information for the early orthography inconsistent during the 300-350 ms (190 ms after the onset of inconsistency) and the late inconsistent during the 575-625 ms (160 ms after the onset of inconsistency) which localized over the centro-posterior electrode sites (CP1, CP2, Pz, Oz, O1, O2), when perceiving monosyllable French spoken words. Later on, in a semantic task, Pattamadilok and Ziegler et al. (2009) found the same early orthography inconsistent effect in the 300-350 ms windows and late orthography inconsistent effect in 400-425ms and 525-575ms time windows when perceiving disyllabic spoken French words. One more question to be addressed was that the activation of orthography was far before the end of the spoken words offset (mean duration of spoken words are 640 ms and 650 in the last mentioned experiments).

Thirdly, the N400 component was often considered as a semantic related component (Danie B et al., 2001; Pritchard et al., 1991).Generally speaking, for spoken word processing, the duration of the N400 component were longer than N400 in visual word recognition; Normally, the auditory N400 starts from 300ms and lasts to 600ms or even more longer, such as to 800ms (Danie B et al., 2001; Petten et al., 1999).

Furthermore, the elicited N400 amplitude was more significant over the frontal than over the posterior sites (Liu et al., 2006; Chen Bai et al., 2006; Hagoort & Brown, 2000; McCallum et al., 1984; Holcomb & Neville, 1990). Besides, the frequency of words, which presented as a classical lexical access marker, also affected the amplitude of N400 (Petten & Kutas, 1990). The low frequency words induced a larger N400 than that in the high frequency words. For spoken word recognition, a word frequency effect was found from 450-750 ms time windows for disyllabic French words (Pattamadilok et al., 2009).

But it was unclear whether the ERPs component for orthography and the frequency effect in Chinese spoken word recognition is the same as that in French. So in Exp. 4, I included a non-homophone high vs. low frequency monosyllable word conditions as a control pair-group. The logic was that both high and low non-homophone words have only one word form for each words, the word frequency was the only difference for these two kinds of words. If there was any ERPs component difference between these two kinds of word, it could only be elicited by the words semantic information.

3.4.1 Participants

Twenty college students who were fluent Mandarin speakers in the Chinese University of Hong Kong were recruited (mean age 21 years; range 19-24; 8 males). Written informed consent was obtained in accordance with guidelines from IRB in the Chinese University of Hong Kong.

3.4.2 Stimuli and design

For homophone words, the stimuli and design were the same as in Exp. 1. For the non-homophone words, a word frequency factor with high and low frequency level was manipulated. Moreover, for high frequency word condition, its mean frequency was matched to the mean accumulate frequency of the homophone word conditions; and for low frequency word condition, its mean frequency was matched to the mean highest homophonic frequency in the homophone word conditions. The detail information of the non-homophone words is listed in the table 8.

Table 8 Stimulus characteristics for the high and low frequency non-homophone word in Exp. 4

stimulus condition	Syllable frequency	density size	Stoke number	Example stimuli	Duration (ms)
High frequency	9269.7	1	7.4	Che1(车)	424
Low frequency	6335.2	1	8.0	Mai2(埋)	430

3.4.3 Task and procedure

The same auditory lexical decision task was used.

The procedure was the same as in Exp. 1, except that the inter-trial interval was changed to 1500ms. Participants were asked not to blink or move while there was a fixation on the screen.

ERP recording

Scalp voltages were collected from 58 Ag/AgCl electrodes mounted in an elastic cap (ElectroCap International, Eaton, USA, 10-10 system). The left mastoid was used as recording reference. Eye movements and blinks were monitored with two electrodes providing bipolar recordings of the horizontal and vertical electro-oculogram (EOG). Inter-electrode impedances were kept below 5 K. EEG was filtered with an analogue bandpass filter of 0.01-50 HZ. The signals were sampled continuously throughout the experiment with a sampling rate of 1000 Hz, and digitally re-referenced to the linked mastoids. Continuous EEG data were divided off-line into epochs beginning 100 ms prior to and 800ms, time-locked to the onset of the auditory word.. The first 100 ms was used for baseline correction.

Average ERPs were based on trials with a correct response only. The rejection rate was bellow 5%, including rejections due to ocular and muscular artifacts. ERP analysis was performed on time windows of interest using 3-way repeated-measures ANOVAs. For homophone word conditions, the factors were hemisphere (left, right hemisphere), brain regions (anterior, medial, posterior regions) selected based on prior research, homophone family size (big, small). For control non-homophone conditions, the word frequency (high, low) and the same hemisphere, brain regions factors were analyze. The six regions of interest were left-anterior (F5, F3, F1, FC5, FC3, FC1), left-medial (C5, C3, C1, CP5, CP3, CP1), left-posterior (P5, P3, P1, PO5, PO3,PO1), right-anterior (F2, F4, F6, FC2, FC4, FC6), right-medial (C2, C4, C6, CP2, CP4, CP6) right-posterior (P2, P4, P6, PO2, PO4,PO6), midline- anterior (FCZ, FZ), midline-central (CZ, CPZ), and midline-posterior (PZ, POZ).

That is, I would examine the difference of ERP waveforms for different conditions in each 100 ms window extending from 300 to 800 ms.

3.4.4 Data Analysis

3.4.4.1 Behavior data analysis

Paired-sample t-tests would be applied to analyze the RT and ACC data.

Results

Results for the homophone conditions

Analysis of the RT data showed a significant effect of homophone density. Mean RT for large and small homophonic density size syllable were 868 ms and 820 ms, respectively (t (0.5, 19) = 2.76, $p<0.05$). Mean error rates for the two conditions were 3.5% and 4.6%, and there were no significant difference between the two (t (0.5, 19) = -0.98, $p = 0.34$). The results established a valid paradigm to demonstrate a clear inhibitory homophone density effect for processing of isolated monosyllable Chinese spoken words.

Results for the non-homophone control conditions

Analysis of the RT data showed a significant effect of frequency effect. Mean RT for high and low frequency syllable were 818 ms and 850 ms respectively (t (0.5, 19) = 2.19, p<0.05). Mean error rates for the two conditions were 7.8% and 5.5%, and there were no significant effects between the two (t (0.5, 19) = 1.95, p > 0.1). The results showed a clear facilitative frequency effect for processing of isolated non-homophone monosyllable Chinese spoken words.

3.4.4.2 ERP analysis

Figure 2 and 3 show the grand-averaged ERP waveforms of the two conditions. The averaged amplitudes and latencies of N1 between 80 and 180ms and P2 between 180 and 280ms), the mean amplitudes of 300-400 ms, 400-500ms, 500-600 ms, 600-700 ms and 700-800 ms time windows were analyzed separately with 3-way repeated-measures ANOVAs.

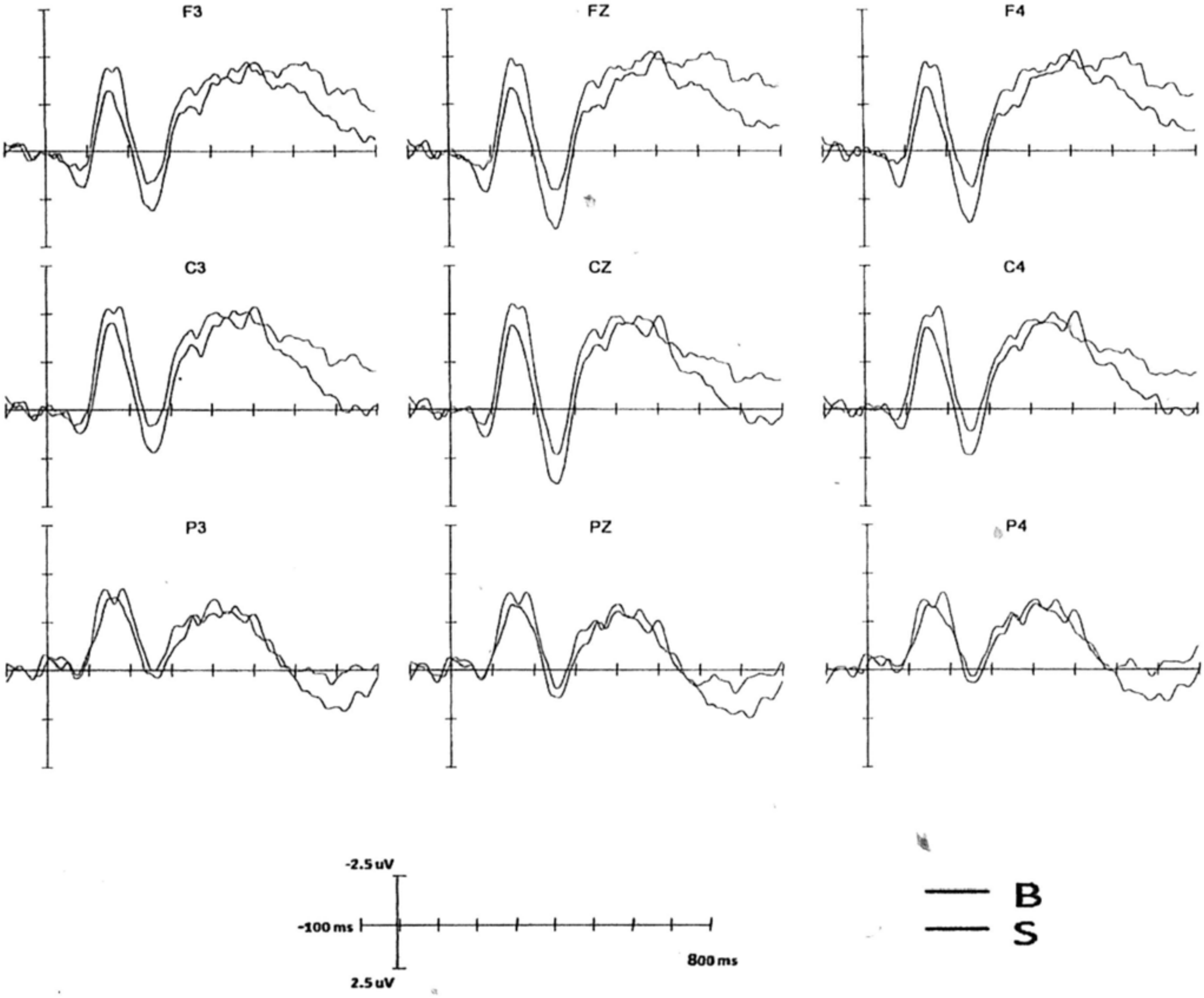

Figure 2 Exp 4: Grand average ERP waves of monosyllable homophone; the big family size homophone (B: green line) and the small family size homophone (S: red line)

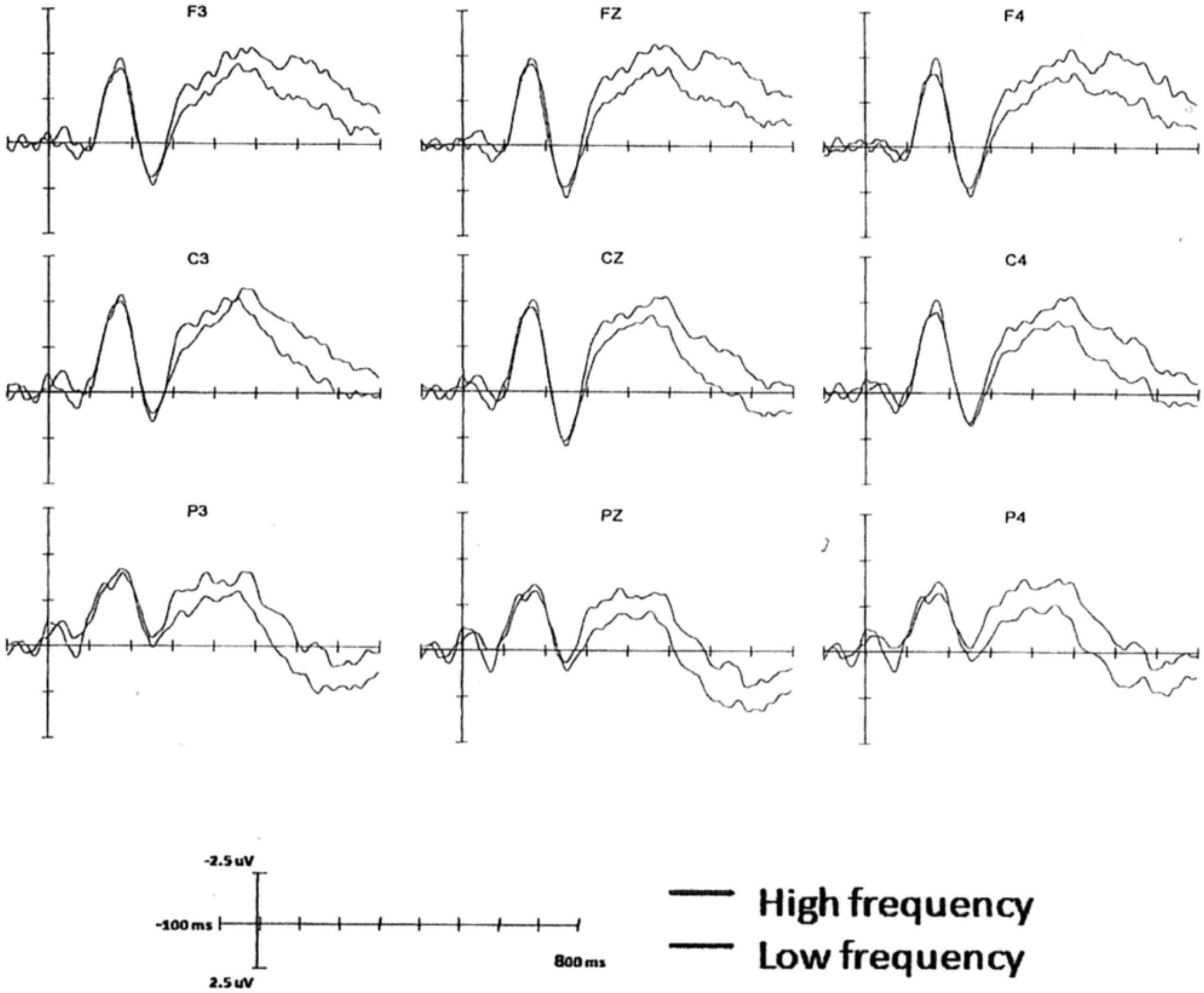

Figure 3 Exp 4: Grand average ERP waves of monosyllable non-homophone; high frequency word (green line) and low frequency word (red line)

N1 (80-180ms) peak to peak latency and peak amplitude

For homophone conditions

N1 (80-180ms) peak to peak latency and peak amplitude

N1 latency

ANOVA results did not show any significant main or interaction effects (mean latency for big and small family sizes were 159ms and 159ms respectively).

N1 amplitude

ANOVA results showed that a main effect of family size was significant (*F* (1, 19) = 8.37, $p < 0.01$). The big family size words elicited significantly larger negative ERP responses than the small family size words (-6.05 vs. -4.88µV).

P2 (160-270ms) peak to peak latency and peak amplitude

P2 latency

ANOVA results did not show any significant main or interaction effects (mean latency for the big and small family size words were 250ms and 248ms respectively).

P2 amplitude

ANOVA results showed that the main effect of family size and regions were significant and no other effect was found. For family, the small family size words elicited significantly larger positive ERP amplitudes than the big family size words (*F* (1, 19) = 10.41, $p < 0.01$; 3.44 vs. 1.90µV). For the region factor (*F* (2, 38) = 19.32, $p < 0.001$; 3.62 vs. 2.42 vs. 0.94µV), further comparison indicates that the mean amplitude for three brain regions were significantly different among each others, with mean amplitude being the largest in anterior region and the smallest in posterior regions.

N300-400ms, 400-500ms and 500-600ms time windows

No effect was found in each of these time windows (p values > 0.2).

N600-700ms time window

ANOVA results showed that the main effect of family size and region were significant and no other effect was found. For family size factor, the big family size words elicited significantly larger negative ERPs than the small size family words (*F* (1, 19) = 20.10, $p < 0.001$; -2.00 vs. -0.12µV). For region factor (*F* (2, 38) = 19.63, p <

0.001; -3.03 vs. -1.34 vs. 0.47μV), the comparison indicates that the mean amplitude for three brain regions were significant different among each others. The largest positive component was elicited in anterior regions, and the smallest one was elicited in posterior regions.

N700-800ms time window

ANOVA results showed that the main effect of family size and region were significant and no other effect was found. For family size factor, the big family size words elicited significantly larger negative ERP amplitudes than the small family size words (*F* (1, 19) = 8.35, p < 0.01; -1.56 vs. 0.32μV). For region factor (*F* (2, 38) = 10.94, p < 0.001; -2.00 vs. -.49 vs. 0.62μV), the comparison indicates that the mean amplitude for three brain regions were significantly different across each others. The largest positive component was elicited in anterior regions, and the smallest one was elicited in posterior regions.

For control non-homophone conditions

N1 and P2 peak to peak latency

ANOVA results did not show any significant main or interaction effects for both N1 and P2 peak latency (for N1, mean latency for high and low frequency were 160ms and 161ms respectively; for P2, mean latency for high and low frequency were 248ms and 248 ms respectively).

N1 and P2 peak to peak amplitude

N1 amplitude

ANOVA results showed a 3-way interaction of frequency by region by hemisphere (F (2, 38) = 4.68, p < 0.05). Further comparisons showed no frequency effects in any

ROI, p values > 0.3. And the main effect of hemisphere was significant ($F(1, 19) = 6.53$, $p < 0.05$; 5.52 vs. 5.10μV), indicating that the ERPs were of a larger amplitude in left hemisphere than that in right hemisphere.

P2 amplitude

ANOVA results did not show any significant main or interaction effects for P2 peak amplitude, p values > 0.3 (mean amplitude for high and low frequency were 2.32 and 2.29μV respectively).

N300-400ms time window

ANOVA results only showed a main frequency effect reaching significance ($F(1, 19) = 6.12$, $p < 0.05$). More comparison showed that low frequency words elicited significantly more negative ERPs than di9d the high frequency words (-2.91 vs. -1.83μV).No other effect was found (p values > 0.3).

N400-500ms time window

ANOVA results showed that the main effect of frequency and region were significant and no other effect was found. For frequency factor, the low frequency words elicited significantly larger negative amplitudes than the high frequency words ($F(1, 19) = 4.40$, $p = 0.05$; -3.16 vs. -4.15μV). For region factor ($F(2, 38) = 5.34$, $p < 0.01$; -4.31 vs. -3.93 vs. -2.72μV), more comparisons indicate that the mean amplitude for both anterior and medial regions were larger than that elicited in posterior region.

N500-600ms time window

ANOVA results showed that the main effect of frequency and region were significant and no other effect was found. For frequency factor, the low frequency words elicited significantly larger negative ERPs than the high frequency words (*F* (1, 19) = 8.66, $p < 0.01$; -1.52 vs. -3.15μV). For region factor (*F* (2, 38) = 20.23, $p < 0.001$; -3.82 vs. -2.49 vs. -0.68μV), more comparisons indicate that the mean amplitude for three brain regions were significantly different across each others, with the largest negative components elicited in anterior region, and the smallest one in posterior region.

N600-700ms time window

ANOVA results showed that a 2-way interaction of frequency and region was significant (*F* (2, 38) = 4.88, $p < 0.05$); further comparison indicated that the word frequency effect was only found in anterior region (*F* (1, 19) = 4.33, p=0.05; -3.04 vs. -4.02μV). Besides, a main effect of region was significant (*F* (2, 38) = 27.37, $p < 0.001$; -3.57 vs. -1.76 vs. 0.22μV). More comparisons indicate that the mean amplitude for three brain regions were significantly different across each other, with the largest negative components elicited in anterior region, and the smallest one elicited in posterior region.

N700-800ms time window

ANOVA results showed that the main effect of frequency, region and hemisphere were significant and no other effect was found. For frequency factor, the low frequency words elicited significantly larger negative amplitudes than the high frequency words (*F* (1, 19) = 9.04, $p < 0.01$; -0.78 vs. 0.69μV). For region factor (*F* (2, 38) = 18.56, $p < 0.001$; -1.71 vs. 0.25 vs. 1.34μV), more comparisons indicate that the mean amplitude for the three brain regions were significantly different across each other. The largest negative component was elicited in anterior region, and the smallest one was elicited in posterior region. For hemisphere factor (*F* (2, 38) = 3.71, $p < 0.05$; -0.17 vs. 0.30 vs. -0.26μV); More comparisons indicate that the mean amplitude for both left and right hemispheres were more negative than that was elicited in middle-line regions.

3.4.5 Discussion

When comparing the high and low frequency non-homophone words, I did not find any significant differences for the N1 and P2 components. This indicates that there was not any acoustic physical difference between these two types of words. Secondly, the word frequency effect where the low frequency words elicited enhanced negative component related to the high frequency words, was significant from 300ms to 800ms time windows, and the effect was more negative over anterior region across 400 to 800 ms time windows.

When comparing the big and small family size words, I found significant difference between the two conditions in N1 and P2 components, displaying a N1 enhancement and a P2 decline for big family size words;

Second, the results showed that the family size effect was elicited from 600 to 800 ms time windows as a larger negative response for the big family size words compared with the small family size ones. This effect was more significant in anterior region, especially in 600-700 ms time windows. However, the results did not show any effect before the offset time (430 ms) of the spoken words, not to mention time window related to orthography as observed for French spoken words recognition.

Further, for both the non-homophone and homophone words, the scalp distribution of the ERP responses was similar so that the negative components were more significant over the anterior regions than in posterior regions.

Exp. 5 ERP evidence for homophone density effect in disyllabic words

The goal of the present study was to exam the neural basis of the homophone density effect in disyllabic spoken Chinese words.

As the meaning of a two syllabic whole word cannot be determined until its second syllable was available, except for the N1 and P2 components, I would only analyze the time windows from 300 to 800 ms, based on the ERP time courses found in Exp. 4.

3.5.1 Participants

Twenty-four participants (8 males, mean age = 22, SD = 2.29) were recruited following the same criteria as in Exp. 4.

3.5.2 Stimuli and design

After an item analysis for Exp. 2, some stimuli, such as /xi1yi1/ 西医 and / chen2jin4/ 沉浸, were found to incur extremely high errors (85% and 65%), suggesting that most participants considered these words as non-words. Such words were deleted and several new words were prepared balancing various stimulus characteristics, forming a total of 168 disyllabic words, 56 words for each of the three conditions. The detailed information of stimulus characteristics is provided in Table 9.

Table 9 Details of stimulus characteristics for all conditions in Exp. 5.

condition	homo density	Highest freq	Accumulate freq	1st MF	2nd MF	1st family	2nd family	1st SF	2nd SF	1st Highest HMF	2nd Highest HMF
HW	2.8	545.0	660.8	4022.6	4147.9	13.7	16.5	17587.2	32676.0	11410.6	23367.1
HF	1	543.6		3968.6	4173.6	12.9	16.0	17556.1	32398.3	12041.1	23473.5
CF	1		662.7	3950.3	4150.0	12.6	15.1	17550.5	32284.7	11759.3	23417.1

3.5.3 Task and procedure

The task was auditory lexical decision and the procedure was similar to that in Exp. 4.

3.5.4 ERP recording

The setting and parameters of ERP recording was the same as in Exp. 4.

Data analysis

3.5.4.1 Behavior analysis

The statistics analysis method was the same as in Exp. 2.

The analysis method was the same as in Exp. 4.Behavior results

The mean RT and ACC averaged over all participants for the different conditions (HW, CF, HF) are shown in Table 10.

Table 10 Mean RTs (ms) and accuracy rates (%) for 3 conditions in Exp. 5

	RT			ACC		
condition	HW	CF	HF	HW	CF	HF
	876.3	878.0	869.5	94.4	92.2	92.1
	(65.3)	(74.8)	(71.2)	(3.6)	(3.4)	(4.7)

Note: Standard deviations are shown in parentheses.

ANOVA for.the RT data did not show any significant effect of condition ($p > 0.6$; 876, 877, 869 ms for HW, CF, and HF conditions respectively).

ANOVA for the ACC data revealed a significant effect of condition ($F(2, 46) = 3.66$, $p < 0.05$; HF vs. CH vs. YH = 5.6% vs. 7.8% vs. 8.0%). Pair-wise comparison showed that accuracy for the HW condition was significantly slower than for both the CF ($t(23) = -2.67$, $p<0.05$) and the HF conditions ($t(23) = -2.03$, $p = 0.05$). There was no difference between the CH and HF conditions ($p > 0.5$).

3.5.4.2 ERP results

Figure 4 shows the grand-averaged ERP waveforms of the three conditions. The averaged amplitudes and latencies of N1 (between 70 and 160ms) and P2 (between 150 and 260ms), the mean amplitudes of five later time windows (260-360ms, 360-460ms, 460-560ms, 560-660ms, and 660-800ms) were analyzed with 3-way repeated-measures ANOVAs.

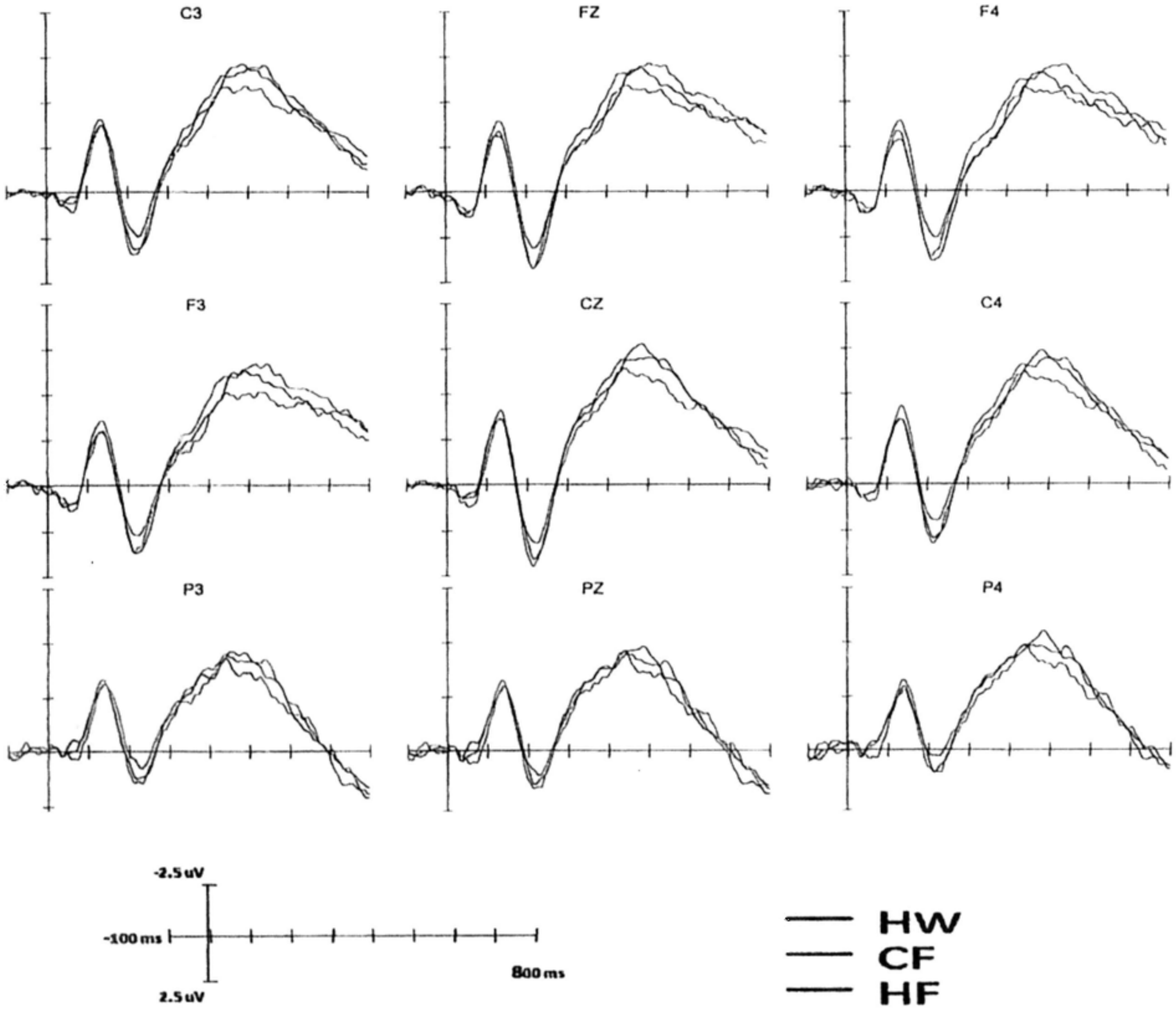

Figure 4 Exp 5: Grand average ERP waves of disyllable homophone (HW: green line), high frequency control word (HF: blue line) and accumulate frequency control word (CF: red line)

N1 (80-160ms) peak latency and amplitude

N1 latency

The ANOVA results did not show any significant main or interaction effects (mean latency for HF, CH and YH conditions were 130ms, 131ms and 131ms, respectively).

N1 amplitude

The ANOVA results showed that a main effect of condition was significant (*F* (2, 46) = 6.91, $p < 0.01$). The HW words elicited significantly more negative ERPs than did both the CF and HF words (ps <0.05, -3.87 > -3.69 = -3.63μV). There was no difference between the CF and HF words ($p > 0.3$).

P2 (160-280ms) peak latency and amplitude

P2 latency

The ANOVA results did not show any main or interaction effect. The mean P2 latency for the HW, CF and HF conditions were 217ms, 220ms and 219ms respectively.

P2 amplitude

The ANOVA results showed a main effect of condition (*F* (2, 46) = 3.35, $p < 0.05$). Both the HW and CF words elicited significantly larger positive ERP amplitudes than did the HF words (p values < 0.05, 3.42μV, 3.28μV and 2.67μV, respectively). There was no difference between the HF and CH words ($p > 0.5$).

N300-400ms time window

The ANOVA results did not show any significant main or interaction effects. The mean amplitude for the HW, CF and HF conditions were -1.56μV, -1.88μV, and -2.01μV, respectively.

N400-500ms time window

The ANOVA results did not show any significant main or interaction effects. The mean amplitude for the HW, CF and HF conditions were -4.16μV, -4.37μV, and -4.68μV, respectively.

N500-600ms time window

The ANOVA results showed a 2-way interaction of condition by region (*F* (4, 92) = 2.60, p < 0.05). Further analysis showed that the condition effect was significant in anterior and medial regions but not posterior regions. The HW words elicited a smaller negative amplitude than did the CF and HF words in anterior region (-5.08 vs.-5.85 vs. -6.48μV, p values <0.001), in medial region (-4.44 vs.-5.32 vs. -5.53μV, p values < 0.001), but not in posterior region (-2.72 vs. -3.06 vs. -3.26μV, p values > 0.1). The ANOVA results showed significant main effects of condition (*F* (2, 46) = 4.32, p < 0.05) and region (*F* (2, 46) = 12.97, p < 0.001). For the condition factor, further comparison show that both the HF and CH words elicited larger negative ERP response than did the HW words (p values < 0.05; -4.10 vs. -4.89 vs. -5.10μV). There was no difference between the HF and CF words (p = 0.5). For the region factor, the negative components elicited in both anterior and medial regions were significantly larger than that elicited in posterior regions (p values < 0.01; -5.90 vs. -5.09 vs. -3.01μV). There was no significant difference between anterior and medial regions (p = 0.07).

N600-700ms time window

The ANOVA results showed no significant effect for condition or its interaction with other factors (p values > 0.2).

N700-800ms time window

The ANOVA results did not show any significant effect for condition or its interaction with other factors (p values > 0.4).

3.5.5 Discussion

This experiment examined the ERP response time course to understand the nature of the homophone density effect in spoken disyllabic Chinese words recognition. The results showed a series of ERP components and time windows, including N1, P2, and 500-600 ms time window associated with the homophonic disyllable spoken word processing.

The relationship between N1 and P2 showed a consistent pattern cross the conditions where larger N1 was accompanied with a decreased P2, suggesting matched stimulus acoustic features across conditions.

Second, in the 500-600 ms time windows, the result showed homophone words evoked smaller negative responses than did the non-homophone words over the anterior and medial regions. However, in the same time window, there was no word frequency effect. This indicates the presence of the family size effect in ERP response, independent of word frequency. With a cross-modal priming task, Shu et al. (2000) found that both meanings of a disyllabic Chinese homophone were activated at least from -150ms offset time of the spoken word. In the present experiment, the 500ms was near the offset time of the stimulus mean duration (560ms) and within the time slot of -150ms offset time.

Taken these results together, it is suggested that the family density effect probably arises from activations of multiple homophonic meanings associated with the disyllabic homophones.

This result is very different from the facilitative effect found for auditory English words with polysemous meanings showing faster response time than non-ambiguous words (Rodd et al., 2002; Klepousniotou & Baum, 2007) or homophones (Stephanie, 2006).

Exp. 6 Interaction between homophone density and word frequency

For most disyllabic words, although these words themselves were not homophonic as a whole word, their component syllables are usually homophonic syllables. This experiment would follow Exp. 3 but record ERPs when participants were presented with such words to examine how homophonic syllables contribute to the processing of the whole word. In addition to the homophone density size of the constituent syllable, I also manipulated the whole word frequency.

3.6.1 Participants

Twenty-four college students were recruited following the same criteria as in Exp. 4.

3.6.1 Stimuli and design

The auditory stimuli were the same as in Exp. 3.

A 2 (density size of constituent syllable: large/small) x 2 (whole word frequency: high/low) factorial within-subject design was used. For half of the participants, the density size factor was manipulated for the first constituent syllable of the whole word (word-initial syllable, Exp. 6a), and for the other half, it was manipulated for the second constituent syllable (word-final syllable Exp. 6b).

3.6.2 Task and procedure

The task was a spoken word lexical decision task. The procedure was the same as in Exp. 5.

ERP recording

The recording procedure was the same as in Exp. 5.

3.6.3 Data analysis

The behavior analysis followed that in Exp. 3 and the ERP analysis followed that in Exp. 5.

Both density size and whole word frequency are within factors, and the statistic method is the same as in Exp. 5.

Behavior results

3.6.4.1 Results for Exp. 6a

Reaction times (RT) and accuracies (ACC) were analyzed in 2x2 repeated-measures analyses of variance (ANOVAs).

The ANOVA results showed a significant main effects of word frequency for RT (F (1, 23) = 31.22, $p < 0.001$) and for ACC (F (1, 23) = 23.10, $p < 0.001$), indicating an advantage for high frequency word. The main effect of homophone density was also significant for both RT (F (1, 23) = 66.89, $p < 0.001$) and ACC (F (1, 23) = 9.27, $p < 0.01$), indicating an advantage for words with its first syllable having a small homophone size.

A significant interaction between word frequency and family size was also found in both the RT data (F (1, 23) = 61.40, $p < 0.001$) and the ACC data (F (1, 23) = 17.03, p <

0.001). Further simple effect analyses suggest that the word frequency effect was only obtained for large homophone density words both in RT (F (1, 23) = 88.24, $p < 0.001$) and ACC (F (1, 23) = 39.09, $p < 0.001$). And the homophone density effect was restricted to low frequency words in RT (F (1, 23) = 104.15, $p < 0.001$) and ACC (F (1, 31) = 19.40, $p < 0.001$), but not obtained for high frequency word (p values > 0.3).

Table 11 Mean RT (ms) and accuracy (%) for all conditions in Exp. 6a & 6b

	exp.6a				exp.6b			
condition	HB	HS	LB	LS	HB	HS	LB	LS
RT	893.6 (98.40)	892.9 (94.30)	991.0 (96.90)	889.3 (107.50)	886.2 (95.9)	895.0 (93.8)	942.1 (101.8)	900.1 (109.0)
ACC	93.5 (0.05)	92.5 (0.05)	85.8 (0.07)	81.5 (0.04)	94.3 (0.05)	94.7 (0.04)	87.7 (0.06)	87.8 (0.07)

Note: Standard deviations are shown in parentheses.

3.6.4.2 Results for Exp. 6b

For the RT data, the ANOVA showed a significant main effects of word frequency (F (1, 23) = 19.56, $p < 0.001$), indicating an advantage for high frequency words. There was also a significant main effect of homophone density (F (1, 23) = 4.97, $p < 0.05$), indicating an advantage for words with its second syllable having small homophone density. A significant interaction between word frequency and homophone family size was also found (F (1, 23) = 18.10, $p < 0.001$). Further simple effect analyses suggest that the word frequency effect was only obtained in the large homophone density words (F (1, 23) = 26.20, $p < 0.001$). And the homophone density effect was restricted to low frequency words (F (1, 31) = 18.10, $p < 0.001$), but not obtained for high frequency word (p> 0.3).

For the ACC data, there was only a main effect of word frequency (*F* (1, 23) = 52.27, p < 0.001), showing more accurate performance for the high frequency words than the low frequency words.

ERP analyses

Figure 5 and 6 showed the grand-averaged ERP waveforms of two conditions in experiment5a and 5b. The averaged amplitudes and latencies of N1 between 80 and 160ms and P2 between 160 and 280ms), the mean amplitudes of 300-400ms, 400-500ms, 500-600ms, 600-700ms and 700-800ms time windows were analyzed separately with 4-way repeated-measures ANOVAs.

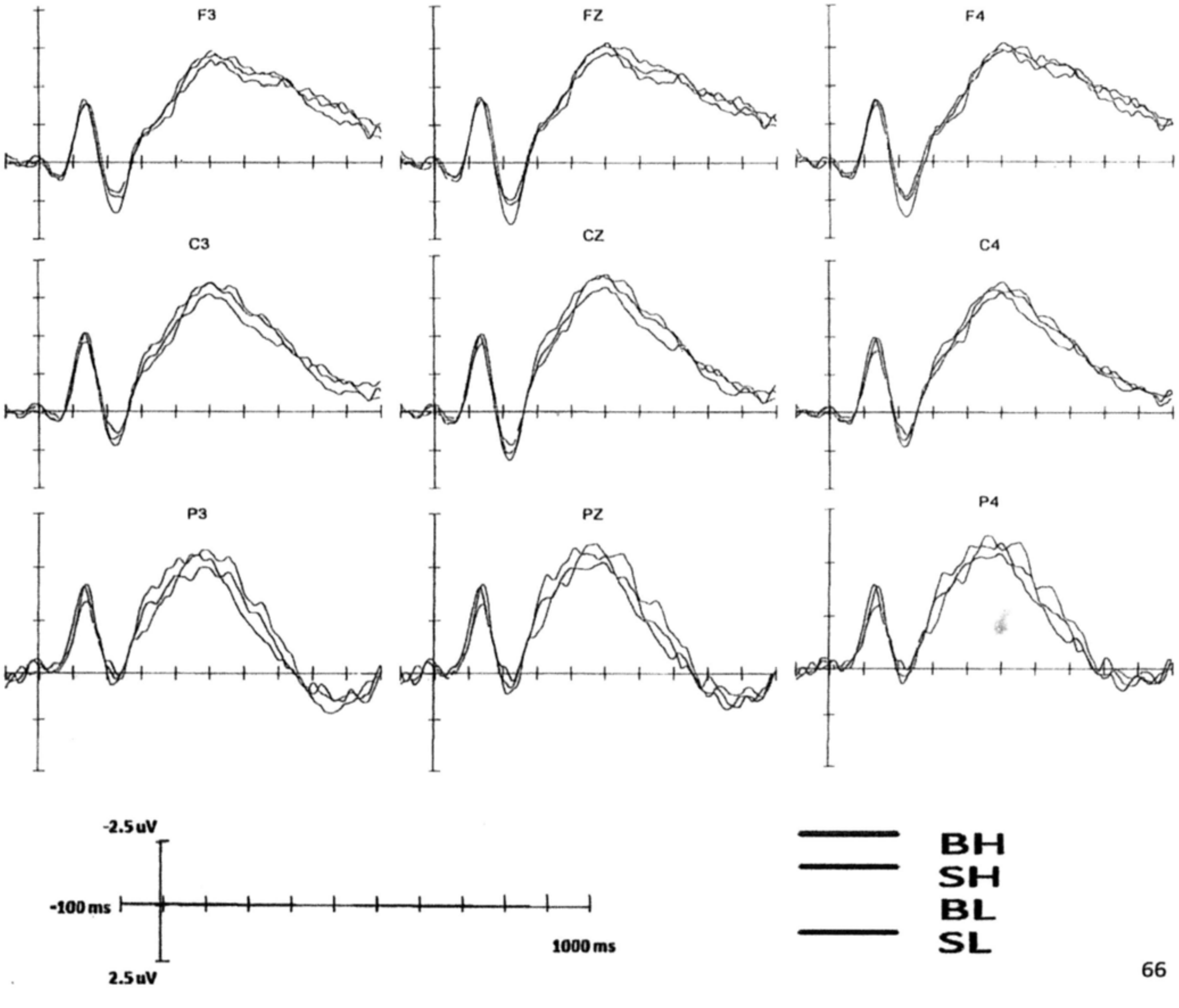

Figure 5 Exp 6a: Grand average ERP waves of homophone as word-initial constituent of disyllable word (e.g.敬爱,喜爱,博爱,偏爱); big homophonic family size-high frequency word (BH: green line), small homophonic family size-high frequency word (SH: blue line), big homophonic family size-low frequency word (BL: yellow line), small homophonic family size-low frequency word (SL: red line)

3.6.4.3 **For Exp. 6a**

N1 (80-160ms) peak latency and amplitude

N1 latency

The ANOVA results showed only a 2-way interaction of word frequency by family size (*F* (1, 23) = 6.51, $p < 0.05$). Further comparison did not show any simple effect (p values > 0.07). The mean latency for each condition was around 130ms. N1 amplitude

The ANOVA results showed no significant effect (p values > 0.1).

P2 (160-280ms) peak latency and amplitude

P2 latency

The ANOVA results showed no significant effect (p values > 0.1). The mean latency for each condition was around 220ms.

P2 amplitude

The ANOVA results showed a 2-way interaction of word frequency by region (*F* (2, 46) = 7.92, $p < 0.01$). Further comparison showed a significant word frequency effect in anterior regions, with the low frequency word elicited a more positive response than did the high frequency words (*F* (1, 23) = 11.06, $p < 0.01$; 3.54 vs. 4.28μV). Moreover, the ANOVA result also revealed significant main effect of for hemisphere and region. For hemisphere, the ERP amplitudes was more positive in the right hemisphere than in the left hemisphere (F (1, 23) = 7.09, $p < 0.05$; 2.85 vs. 2.42μV). For region, the anterior region showed the most positive response and the posterior region showed the smallest (F (2, 46) = 27.43, $p < 0.001$; 3.91 vs. 2.56 vs. 1.43μV).

N300-400ms time window

The ANOVA results only showed a main effect of family size (F (1, 23) = 8.13, $p < 0.01$; -2.88 vs. -3.57μV), indicating that the small family size words elicited a more negative response than did the large family size words. There was also a marginal significant main effect of region (F (2, 46) = 3.09, $p = 0.055$; -2.79 vs. -3.71 vs. -3.17μV) where medial regions showed the largest negative response among the three regions (p values < 0.05). There was no difference between the anterior and posterior regions ($p = 0.45$).

N400-500ms time window

The ANOVA results only showed a 2-way interaction of family size by region (F (2, 46) = 5.26, $p < 0.05$). Further comparison showed a family size effect in posterior regions where the large family size words elicited a more negative response than did the small family size words (F (1, 23) = 10.78, $p < 0.01$, -4.33 vs. -5.10μV).

N500-600ms time window

The ANOVA results showed a 2-way interaction of family size by region (F (2, 46) = 3.27, $p < 0.05$). Further comparison showed a significant family size effect in all three regions where the large family size words elicited a more negative response than did the small size words: anterior region, F (1, 23) = 9.67, $p < 0.01$; -7.63 vs. -6.68μV;

medial region, F (1, 23) = 20.48, $p < 0.001$; -7.46 vs. -6.16μV; and posterior region, F (1, 23) = 27.64, $p < 0.001$; -4.93 vs. -3.53μV.

Moreover, the ANOVA results also showed main effects for word frequency, family size, and region. For word frequency, the low frequency word elicited a more negative response than did the high frequency words (F (1, 23) = 8.05, $p < 0.01$; -6.51 vs. -5.62μV). For family size, the large family size words elicited a more negative response than did the small family size words (F (1, 23) = 20.11, $p < 0.001$; -6.68 vs. -5.46μV). For region, anterior and medial regions showed the most negative responses than posterior regions (F (2, 46) = 23.98, $p < 0.001$; -7.16 vs. -6.81 vs. -4.23μV).

N600-700ms time window

The ANOVA results showed a 2-way interaction of word frequency and family size was significant (F (1, 23) = 5.88, $p < 0.05$). Further comparison showed that it found a significant word frequency effect was restrict to big family size condition, with low frequency words elicited a larger negative amplitudes than that in high frequency words (F (1, 23) = 22.71, $p < 0.001$; -7.06 vs. -4.58μV); and it also found a significant family size effect in low frequency word condition, with big family size words elicited a larger negative amplitudes than that in small family size words (F (1, 23) = 60.85, $p < 0.001$; -7.06 vs. -4.24μV), and a marginal significant family size effect in high frequency word condition, with big family size words elicited a larger negative amplitudes than that in small family size words (F (1, 23) = 4.08, $p = 0.055$; -4.58 vs. -3.47μV).

Moreover, the ANOVA results also showed main effects for word frequency, family size, hemisphere, and region. For word frequency, the low frequency word elicited a more negative response than did the high frequency words (*F* (1, 23) = 27.47, $p < 0.001$; -4.03 vs. -5.65μV). For family size, the large family size word elicited a more negative response than did the small family size words (*F* (1, 23) = 40.32, $p < 0.001$; -5.82 vs. -3.86μV). For hemisphere, the ERP negative amplitudes in right size hemisphere was larger than that in left size hemisphere (*F* (1, 23) = 4.82, $p < 0.05$; -4.59 vs. -5.09μV). For region factor, the anterior regions showed the largest negative response and the posterior region the smallest (*F* (2, 46) = 33.13, $p < 0.001$; -6.47 vs. -5.36 vs. -2.69μV).

N700-800ms time window

The ANOVA results showed a 2-way interaction of word frequency and family size (*F* (1, 23) = 7.15, $p < 0.05$). Further comparison showed a significant word frequency effect for the large family size words where the low frequency words elicited a greater negative response than did the high frequency words (*F* (1, 23) = 17.17, $p < 0.001$; -5.04 vs. -2.58μV). There was also a significant family size effect for the low frequency words where the large family size words elicited a more negative response than did the small family size words (*F* (1, 23) = 29.43, $p < 0.001$; -5.04 vs. -2.44μV).

Moreover, the ANOVA results also showed main effects for family size, word frequency, hemisphere, and region. For word frequency, the low frequency words elicited a more negative ERP response than did the high frequency words (*F* (1, 23) = 14.76, $p < 0.001$; -3.75 vs. -2.45μV). For family size, the large family size words elicited a more negative response than did the small family size words (*F*(1, 23) = 14.95, $p < 0.001$; -3.81 vs. -2.38μV). For hemisphere, the right hemisphere showed a more negative response than did the left hemisphere (F (1, 23) = 5.84, $p < 0.05$; -3.39 vs. -2.80μV). For region, the anterior regions showed the most negative response and the posterior regions the least negative response (F (2, 46) = 45.14, $p < 0.001$; -5.21 vs. -3.34 vs. -0.73μV).

3.6.4.4 For Exp. 6b

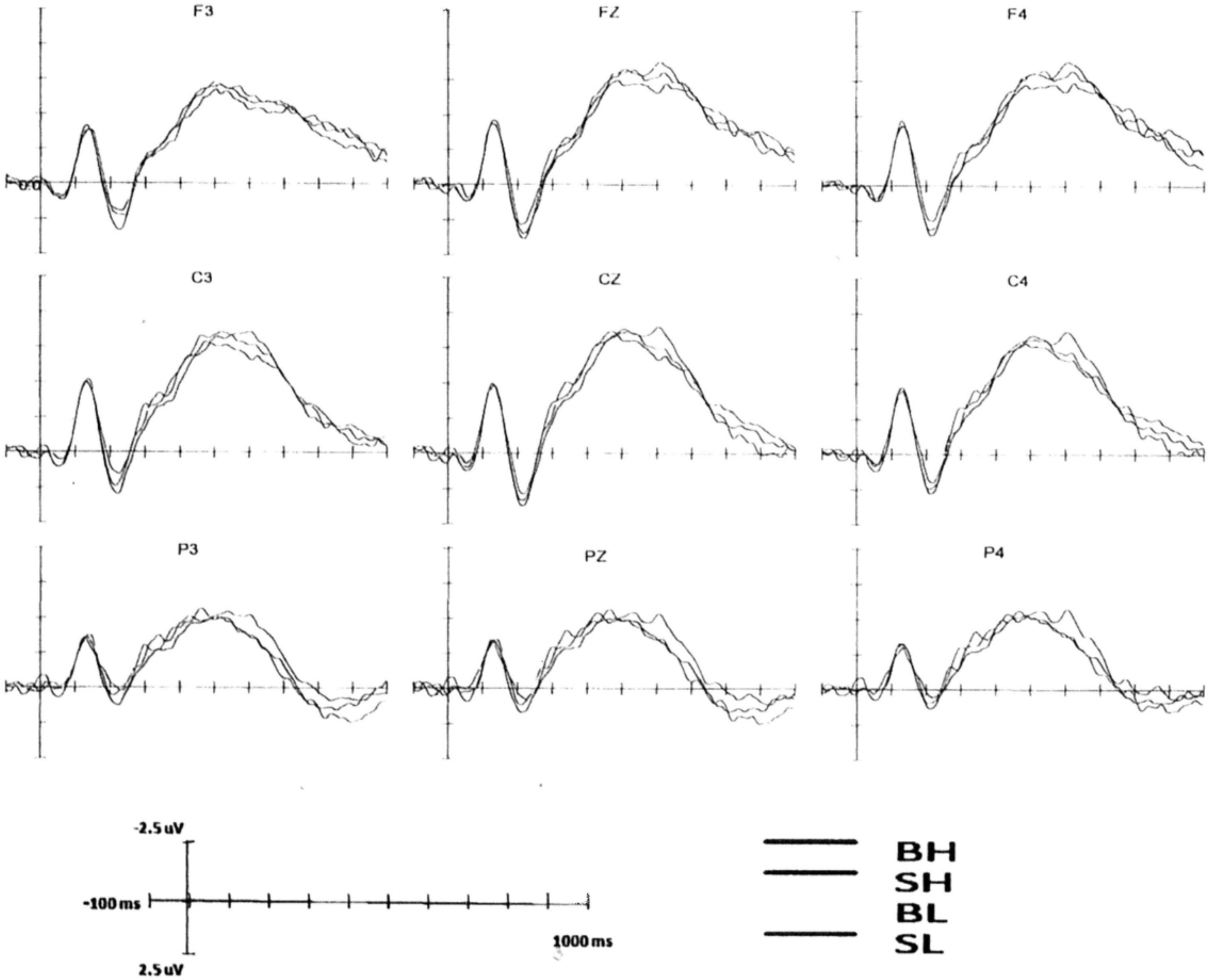

Figure 6 Exp 6b: Grand average ERP waves of homophone as word-initial constituent of disyllable word (e.g.火焰,火灾,火炉,火炭); big homophonic family size-high frequency word (BH: green line), small homophonic family size-high frequency word (SH: blue line), big homophonic family size-low frequency word (BL: yellow line), small homophonic family size-low frequency word (SL: red line)

N1 (80-160ms) peak latency and amplitude

N1 latency

The ANOVA results showed no significant effect (p values > 0.1).

N1 amplitude

The ANOVA results showed no significant effect (p values > 0.1).

P2 (160-280ms) peak latency and amplitude

P2 latency

The ANOVA results showed only a 3-way interaction of hemisphere by region by word frequency ($F(4, 92) = 4.00$, $p < 0.05$).

P2 amplitude

The ANOVA results showed no significant effect (p values > 0.1).

N300-400ms time window

The ANOVA results showed no significant effect (p values > 0.1).

N400-500ms time window

The ANOVA results showed no significant effect (p values > 0.1).

N500-600ms time window

The ANOVA results showed only a 3-way interaction of hemisphere by region by family size ($F(4, 92) = 2.85$, $p < 0.05$). Further comparison showed that large family size words elicited a more negative response than small family size words in both left-anterior area ($F(1, 23) = 6.11$, $p < 0.05$; -7.90 vs. -7.10μV) and left-medial area ($F(1, 23) = 4.24$, $p = 0.05$; -7.33 vs. -6.68μV). Moreover, the ANOVA results also showed a main effect of word frequency ($F(1, 23) = 15.26$, $p < 0.001$; -6.78 vs. -5.80μV), indicating that the low frequency words elicited a larger negative response than did the high frequency words. There was a main effect of region (*F* (2, 46) = 34.43, $p < 0.001$) where both anterior and medial regions showed a larger negative response than did the posterior region (*p values <0.001*; -7.68 vs. -7.02 vs. -4.17μV).

N600-700ms time window

The ANOVA results showed a significant 3-way interaction of hemisphere by region by family size (*F* (4, 92) = 3.67, $p < 0.01$). The large family size words elicited a greater negative response than did the small family size words only in left-medial areas (*F* (1, 23) = 4.77, $p < 0.05$; -7.53 vs. -6.89μV). There was a 2-way interaction of word frequency by region (*F* (2, 46) = 7.15, $p < 0.01$). The low frequency words elicited a more negative response than did the high frequency words in all three regions: posterior region, *F* (1, 23) = 8.55, $p < 0.01$; -7.75 vs. -6.83μV; medial region, *F*(1, 23) = 24.99, $p < 0.001$; -6.62 vs. -5.12μV; and posterior region, *F*(1, 23) = 27.90, $p < 0.001$; -3.62 vs. -1.94μV.

Moreover, the ANOVA results also showed a main effect of word frequency (*F* (1, 23) = 23.19, $p < 0.001$; -6.78 vs. -5.80μV), indicating that the low frequency words elicited a larger negative response than did the high frequency words. There was a main effect of region (*F* (2, 46) = 50.14, $p < 0.001$) where both the anterior and medial regions showed a more negative response than did posterior regions (*p values <0.05*; -7.68 vs. -7.02 vs. -4.17μV).

N700-800ms time window

The ANOVA results showed a 2-way interaction of word frequency by region (*F*(2, 46) = 12.12, $p < 0.001$). Further comparison showed that the low frequency words elicited a more negative response than did the high frequency words in all three regions: posterior region, *F* (1, 23) = 6.52, $p < 0.05$; -5.75 vs. -4.93μV, medial region, *F* (1, 23) = 20.20, $p < 0.001$; -3.94 vs. -2.49μV, posterior region, *F* (1, 23) = 27.04, $p < 0.001$; -1.08 vs. 0.6μV.

Moreover, the ANOVA results also showed a main effect of word frequency (*F*(1, 23) = 18.57, $p < 0.001$; -3.59 vs. -2.28μV), indicating that the low frequency words elicited a more negative response than did the high frequency words. There was a main effect of family size (*F* (2, 46) = 7.98, $p < 0.05$; -3.40 vs. -2.47μV), indicating that the words with larger family size elicited a more negative response than did the words with a small family size.

3.6.5 Discussion

This experiment examined the ERP correlates of the homophone density effect and how it interacts with word frequency during recognition of non-homophonic disyllabic Chinese words consisting of homophonic syllables. The results revealed several ERP components including N1, P2, and responses in the 300-800 ms time window associated with disyllabic spoken word recognition.

The N1 component was not affected by word frequency and homophonic family density size of the constituent syllable in both Exp. 6a and 6b. But for P2, the low frequency word elicited a more positive response in anterior regions than did the high frequency words when the first syllable was homophonic (Exp. 6a). It is possible that the lower level acoustic properties of the low frequency words did not match completely with that of the high frequency words, producing this early ERP response difference.

For the word-initial syllable condition (Exp. 6a), the homophonic family density size effect took effects in the 300 to 800 ms time windows, although it did not show consistent patterns in all time windows. In the 300-400 ms time window, the small family size words evoked a larger negative component than did in the large family size words. The family size effect changed into the opposite pattern from 400 to 800 ms time windows as the large family size words evoked a more negative response than did the small size words. Especially, in the 400-500 ms time window, the family size effect was only found in the posterior region more likely to be associated with orthographic activation (Perre et al., 2008; Pattamadilok et al., 2009).

The word frequency effect took effect from 500 to 800 ms time windows where the low frequency words evoked a more negative response compared with the high frequency words. And the time window for the frequency effect in the present study was similar to that found for French disyllabic spoken words (Pattamadilok et al., 2009). Interaction between word frequency and family density size was found from 600-800 ms time windows, and the results showed that the word frequency effect was restricted to the large family size words and the family size effect was restricted to the low frequency words.

For the word-final syllable condition (Exp. 6b), the homophonic family density size effect took effects in the 500 to 800 ms time window where the large family size words evoked a more negative response than did the small family size words. However, this effect was only found in left-anterior and left-middle areas in the 500-600ms time window, in left-middle areas in the 600-700ms time window, and over the whole brain regions in the 700-800ms time window.

The word frequency effect also took effect in the 500 to 800 ms time windows where the low frequency words evoked a more negative response than did the high frequency words. There was no interaction between word frequency and family density size in any of the time windows.

To summarize, the ERP results indicated orthographic activation in an early state of disyllabic word recognition when the first constituent syllable of the words were manipulated. This effect, however, was not present when the second syllable was manipulated. Besides, the family size effect and the word frequency effect were found in later stages (500-800 ms) of the processing of the spoken words in both experiments. In particular, the word frequency effect interacted with the homophonic family size effect as the family size effect was restricted to the low frequency words in the homophonic syllable was the first constituent of the disyllabic words, and these two effects were most significant over the frontal and (or) medial sites of the brain. The later effects in the 500-800 ms time windows were therefore likely to reflect semantic processing so that the more the homophone mates the more negative ERP response for the family size effect, and the lower the word frequency, the larger negative ERP response.

Chapter 4

General Discussion

In the present study, using a set of behavioral and ERP experiments, I investigated (1) whether the homophonic family size effect is present at different types of spoken Chinese words, including monosyllabic homophones, disyllabic homophones, and disyllabic words with theirs constituent syllable being homophonic; (2) whether the homophonic family size effect would interact with word frequency for disyllabic words; (3) to reveal the time course of the family size effect and find out whether it was due to the activation of orthographic information or semantic information during spoken word recognition.

The results indicate that the homophonic family density factor does play a role in spoken word recognition, but the effect appears to be different for different types of words. From the behavioral results, the homophonic family size effect showed up as an inhibitory effect for both monosyllabic and non-homophonic disyllabic words. The effect was restricted to the low frequency non-homophonic disyllabic word both when the first and the second syllable were homophonic. For disyllabic homophone words, the effect showed up as a facilitative effect. Beyong the behavioral findings, the ERPs studies also showed that the family size was evoked in different time windows for different types of spoken words. In the following sections, I will discuss the time course and the nature of the homophonic density size effect cross different types of spoken words.

4.1 Is the family size effect orthographic-based or meaning-based?

Monosyllable homophonic words

One central feature of the present study is that the homophonic family size was manipulated at a large and a small level, after a tight control of many confounds, such as syllable frequency, highest frequency within the homophonic family. With such control, theoretically, the critical difference between the large and small family size homophones was the number of visual forms and meanings associated a spoken word. The inhibitory family size effect could therefore be attributed to the competition among the multiple activations of orthographies or meanings or both of them within the homophone family.

Previous ERP studies on alphabetic language indicate that the orthography information can be detected online over the centro-posterior sites during the spoken words input (Perre et al., 2008; Pattamadilok et al., 2009). In comparison, the semantic information, normally indexed by the N400 component, usually shows a wide-spread scalp distribution with a frontal, medial, centro-posterior focus cross different kinds of tasks (Kutas & Federmeier,2009), being more prominent over frontal sites than posterior sites in auditory tasks (McCallum et al., 1984; Holcomb & Neville, 1990). These characteristics are normally are more salient for the late N400 component (Hagoort & Brown, 2000; Cummings et al. 2006; Desroches et al., 2009).

From the results in Exp. 4, the homophone density effect was only found in the 600-800 ms time window where the large family size monosyllable words evoked more negative responses than the small family size words. Further, the effect seems to be more negative in anterior site than in medial and posterior site. However, no effect was localized over centro-posterior sites in any time windows. Previous research on the processing of monosyllable homophone in spoken Cantonese suggests that both the dominant and subordinate meanings of a homophone were activated in sentence comprehension (Yip, 1998, 2007; Li and Yip, 1996). Specifically, in Yip (2002), the recognition time was longer for large family size words than small size words.

When the present results are combined with the literature findings, the inhibitory homophone density size effect is mostly likely attributed to increase competition between multiple meaning activations within the homophone family. However, the meanings information did not seem to be activated online, i.e., not during but after the physical presentation of the spoken word, as the family size effect occurred rather late in the ERP time course data after the offset time of the monosyllabic spoken words.

Disyllable homophones

When compared with the non-homophone disyllabic word, the disyllable homophone evoked a less negative response than did the non-homophone words over the anterior and medial site. Further, the waveform was more negative in anterior site than in medial site. It has been showed that both the dominant and subordinate meanings of a homophone were activated in meaning access of ambiguous spoken words (Shu et al., 2000; Zhou et. al., 2003). Hence, the homophone density effect is likely to be a semantics-based effect in disyllabic homophonic spoken words.

For majority of the disyllable words that are not homophonic themselves but with their first or second constituent syllable being homophonic, the results were as follows.

When homophone density size factor was manipulated in the word-initial syllable, the homophone family size effect showed up as different patterns in different time windows. In the 300-400 ms time windows, the small family size words evoked a more negative response than did the large family size words. Overall, the evoked waveform was more negative over the medial site. In the 400-500 ms time window, the family size effect was only localized over the posterior site where the large family size words evoked a more negative response than did the small size words. In the 500-800 ms time window, the elicited waveform for the family size effect was distributed all over the scalp where the large family size words elicited a more negative response than did the small family size words. Moreover, this family size effect was modified by the word frequency factor and was significant for low frequency words but not for high frequency words in the 600-800 ms time window.

When the homophone density size factor was manipulated for the word-final syllable, the homophone family size effect showed up so that the large family size words elicited a more negative response than did the small family size words, localized over the left-anterior and left-medial sites in the 500-700 ms time window. In the 700-800 ms time window, the scalp distribution was similar.

Therefore, as for the monosyllables, the family size effect found in the 500-800 ms time window in the present study is likely to be meaning-based. That is, the multiple meanings associated with a constituent homophone syllable were activated during the auditory input period.

Beyond that, the most interesting finding in the present study is that the medial-posterior site effect in the 300-500 ms time window for word-initial syllable condition was often considered as a hallmark of orthographic activation in spoken word recognition. What is more striking is that the pattern of ERP effects for the family size effect in the 300-400 ms time window was the opposite to that the pattern in other time windows. A possible explanation may be constructed based on a consideration of studies of the cloze probability in sentence comprehension where low cloze probability words elicited a larger ERP negativity compared with high cloze words and the scalp distribution was mostly over medial and posterior sites in 300-500 ms time window (Delong et al., 2011). The cloze probability refers the probability that a word can be filled in a sentence. For example, the word "apology" is high cloze probability word in the sentence "Dale was very sorry and knew he owed Mary a(n) ____." and the word "check" is low cloze probability word in the above sentence (Kutas & Hillyard, 1984).

In analogous to this situation, in the present study, very early in the recognition process, any word forms associated with a homophonic syllable was possibly a correct response, a large family size word, e.g., with 18 homophone mates, would therefore be easier to process as the probability was high that participants activated any of the 18 members, compared with a small family size word, e.g., with 3 homophone mates. That is, the orthographic activation would lead to reduced negative ERP responses, just as the high cloze words similar associated with a high access probability.

To summarize the family size effect for the different types of spoken Chinese words, it should be noted that during spoken words input, the ultimate aim for listeners is to get meaning from the input auditory strings for understanding. This explains why semantic-based family size effect was robust and present for all types of words investigated in the present study. Beyond this, the presentation of the first syllable may cue participants to activate the orthographic representations. That is, input of the first syllable did not offer enough information for a listener to gain the whole word meaning, an economical way from the point of view of the cohort model, was to activate the visual word forms of all the possible homophones as a preparation to facilitate the comprehension of the whole word once the second syllable was presented. This strategy would not work for the word-final syllable condition as once the second syllable was inputted, the whole word meaning could be accessed without having to using the orthographic information associated with the second syllable.

4.2 The ERP signature for the word frequency effect

Apart from the homophone family size effect, the present study also offers some ideas about the neural basis of the word frequency effect. As a controlled condition, non-homophone monosyllabic words were divided into the high and low frequency levels. It was found that the frequency effect was significant from 300-800 ms time windows. Furthermore, the effect was consistent with general expectation in that the low frequency word evoked more negative amplitudes. Overall, this effect was localized more over the frontal regions than the medial and posterior sites cross different time windows.

When studying the disyllabic words, words with relative high and low word frequency were chosen as the non-homophone disyllabic controls Exp. 5. However, there was no sign of any word frequency in any time window for these words. It is suspected that frequency difference for across the two stimuli groups may not be large enough to reveal the frequency effect (For Exp. 5, the mean frequency were 544 and 663 for the two groups).

The word frequency effect was clearly observed in Exp. 6a and 6b. For both the word-initial and the word-final syllable conditions, the word frequency effect showed up so that low frequency words evoked more negative responses during the 500-800 ms time windows than high frequency words for disyllabic spoken words. The time course of this effect was very similar to that found in French disyllable words (Pattamadilok et al., 2009).

Taken together the different time courses for the monosyllable and disyllable spoken words, the results suggest that the word frequency effect occurs very early, clearly before the end of acoustic length of a spoken word. This is an issue that has been addressed in several studies using the gating task (Magne et al., 2007; Li & Yip, 1998; Yip, 2002; Wu & Shu., 2003; Grosjean F., 1980, see a review). The results show that the time point for monosyllable word started around 300ms and for disyllable word it started around 500ms. This may have been caused by stimulus duration differences between the two types of words. Albeit this potential influence in the early time period, the later effects were all around 800 ms, similar across the two types of words.

4.3 Implications for spoken word recognition models

The interactive model for alphabetic language word recognition addresses the relationship among orthography, phonology and semantic information (meaning) during word recognition appliable to both visual and spoken word recognition (Stone & Van Orden, 1994; Stone et al., 1997; Van Orden, 2002; Harm & Seidenberg, 2004; Ventura et al., 2004). In this model, the connections among orthography, phonology and semantics are all bidirectional (see the figure 7). This means that the input of a spoken word would activate its corresponding orthography and semantic information. Accordingly, when a word rimes with multiple spellings, the auditory input of this word would activate multiple orthographic forms, such as */i:p/* to EEP and EAP or */meid/* to made and maid, as well as the corresponding meanings.

However, the above model does not entirely fit the Chinese data as well. According to the present study, the activation of the sound-to-orthography connection is conditional but the activation of the sound-to-meaning connection is compulsory. Along with other Chinese visual word recognition studies (Liu et al., 2003; Chen et al., 2006; Zhou et al, 1999), the pattern of connection three representations in Chinese should be modified as in Figure 8, being different from that for alphabetic language.

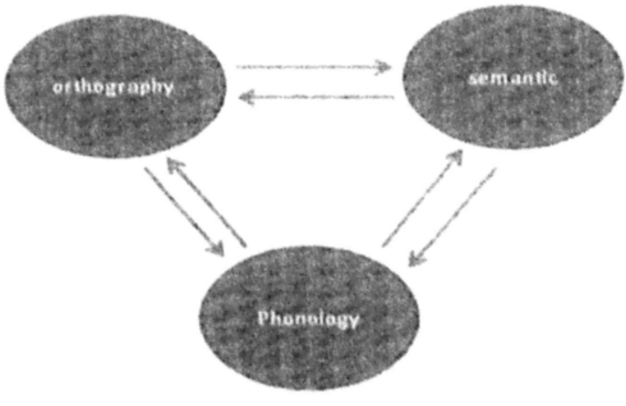

Fig. 7 the relationship of orthography, semantic and phonology in interactive model

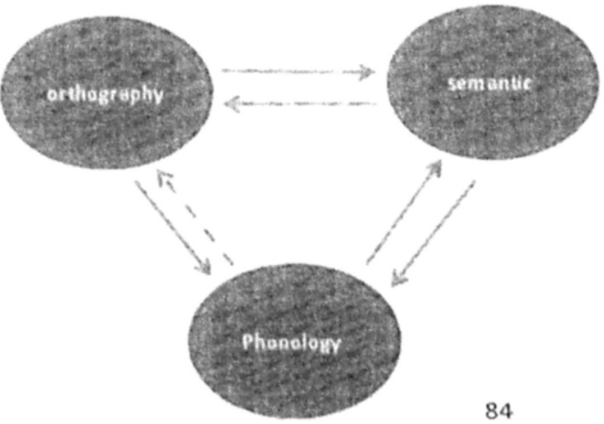

Fig. 8 the relationship of orthography, semantic and phonology for Chinese word

Currently, there is only one spoken word recognition model for Chinese, the Multi-level Cluster Representation model (Zhou & Marslen-wilson, 1994; 1995). The model includes representations at the word level, the morpheme level, and the syllable level for disyllabic compound words. Clearly the homophone family size factor needs to be explicitly incooperative in this model. The present data is insufficient to provide very specific directions as how such model revision may proceed.

4.4 Limitations

One limitation of the present study was that different types of words were presented for different amount of time, which may cause some problems in cross-experiment comparison. Even for the same type of words, there are different lengths of durations across different words that have different combinations of consonant, rhyme, and tone. Beyond within experiment variations, the mean duration for disyllable words were not equated between disyllabic homophones in Exp. 2 and 5, and disyllabic words in Exp. 3 and 6. So there is a potential that this stimulus characteristics may influence to some extent the overall results comparison across different experiments.

Secondly, in Exp. 2 and 4, when selecting non-homophone disyllable words as controls, their frequency were matched to that of the homophonic words' accumulate frequency or highest frequency. This matching process restricted the range of the mean frequency difference across the two types of non-homophonic disyllable words, and may be partially responsible for the failure in finding any word frequency effect in this experiment.

Another possible limitation is particularly experimental paradigm, the lexical decision task used in this study may influence the conclusions of spoken word recognition. The lexical decision task requires participants to make decision on whether a word is a real or pseudo word. This has never been done in natural settings of speech perception.

To conform with the artificial demand of this laboratory task, participants may adopt some special strategies, such as intentionally make use of orthographic information, and such strategies normally would not be used in daily communication. The use of these strategies hampers the effort to understand speech perception as an automatic process that occurs automatically and implicitly. There has been many studies showing that explicit and implicit task requirements lead to very different regions of activations in the brain, even when the stimuli were held constant (Thuy et. al., 2004; Chen et al., 2009; 2003).

4.5 Future directions

As mentioned in the limitations section, the experiment task may influence participants' response such as adopting special task strategies that would normally not be used in daily life communication. Therefore, it would more desirable in future studies to adopt more natural tasks, such as putting spoken words in a sentence or dialogue context.

The other direction would be to compare normal participants as studied here with illiterates who do not have orthographic knowledge attached to the homophonic words. This would be able to help single out more clearly the impact of orthography on spoken word recognition. For disyllable spoken words, only two factors were considered here, the whole word frequency and the homophone family size of the whole word or its constituent syllable. Another important factor missing is morpheme frequency. Although there was evidence that morpheme frequency did not affect response time and accuracy in disyllabic Chinese spoken word recognition (Zhou & Marslen-Wilson, 1994), a pilot experiment I conducted recently has shown that when morpheme frequency and homophone family size were manipulated together for constituent syllable, there was interaction between the two factors so that the morpheme frequency showed an inhibitory frequency effect for large family size homophones. It would be interesting to use ERP to trace this interactive effect as it exposes a separate aspect of homophone family size effect.

Chapter 5

Conclusion

In summary, the present set of experiments investigated the homophone family size effect in different types of spoken Chinese words using auditory lexical decision tasks. When homophone family size was manipulated in monosyllable words and word-initial or word-final syllable of a disyllabic word, large family size slowed down the recognition process. But this effect was reversed when the family size factor was manipulated for disyllabic homophone words, and large family size words led to better performance compared with small family size words.

These findings firmly establish that the family size effect plays an important role in spoken Chinese word recognition. They further indicate that whether the direction of the effect is facilitative or inhibitory depends on the type of spoken words. The ERP results showed that size effect is conditional orthography-based, i.e., resulting from competition of multiple orthographic code activation in a relatively early processing stage, for homophonic syllables as the first constituent of a disyllabic word. There is also a component of the family size effect that is semantics-based, i.e., resulting from competition of multiple semantic code activation in a relative later processing stage for all types of spoken words. Finally, the family size effect is not independent of the whole word frequency when the first syllable of the disyllabic word is homophonic. These findings present valuable insights into spoken word recognition and highlight the importance of semantic information for Chinese speech perception extensively involving homophones.

References

Chen, B., Wang, L. & Peng, D. (2003). The time course of graphic,phonological and semantic information processing in Chinese character recognition (Ⅱ). Acta Psychologica Sinica, ,35 (5) :576 ~ 581

Chen, B. & Ning, H. (2005). Homophone Effects in the Recognition of Chinese Character : the Evidence of Phonology Influencing the Graphic Processing of Chinese Character. Psychological exploration,2005,4:35-39 [in Chinese]

Chereau, C., Gaskell, M.G., Dumay, N. (2007). Reading spoken words: orthographic effects in auditory priming. Cognition 102, 341–360.

Li, P., & Yip, M. C. W. (1996). Lexical ambiguity and context effects in spoken word recognition: Evidence from Chinese. In G. Cottrell (ed.), Proceedings of the 18th Annual Meeting of the Cognitive Science Society, 228-232.

Li, P., & Yip, M. C. W. (1998). Context Effects and the Processing of Spoken Homophones. Reading and Writing, 10, 223-243.

Liu yuan. (1990). Common Frequency Dictionary of Modern Chinese. Aerospace Press.

Pattamadilok, C., Perre, L., Dufau, S., & Ziegler, J. C. (2009). On-line Orthographic Influences on Spoken Language in a Semantic Task. Journal of Cognitive Neuroscience, 21(1), 169-17

Perre, L., & Ziegler, J. C., (2008). On-line activation of orthography in spoken word recognition. Brain Research, 1188, 132-138.

Pexman, P. M., Cristi, C., & Lupker, S. J., (1999). Facilitation and interference from formally similar word primes in a naming task. Journal of Memory and Language, 40, 195–229.

Pexman, P.M., Lupker, S.J., , Jared, D. (2001). Homophone effects in lexical decision. Journal of Experimental Psychology:Learning, Memory, and Cognition, 27, 139–156.

Shu H , Tang Y H , Zhang Y X. A study on the resolution of lexical ambiguity of two - syllable homophone in Chinese (In Chinese) Acta psycholgica Sinica. 2000 , 32 (3) : 247～252 .

Taft, M., & Zhu, X. (1997). A model of lexical processing: Orthography, Phonology, and Morphology. In D.Peng, H. Shu, & H.-C. Chen (Eds.). Cognitive research on Chinese language. Jinan, PRC: Shandong Education Press [in Chinese].

Taft, M., Castles, A., Davis, C., Lazendic, G., & Nguyen-Hoan, M.,(2008). Automatic activation of orthography in spoken word recognition: Pseudohomograph priming. Journal of Memory and Language, 58, 366–379.

Ventura, P., Morais, J., Pattamadilok, C., Kolinsky, R., (2004). The locus of the orthographic consistency effect in auditory word recognition. Lang. Cogn. Processes 19, 57–95.

Yi, H., Shu H. & Zhang D., (2001).The Characteristics and Factors Affecting Children Character Output Under the Dictation Task. Development and Education of psychology, 1,01–05 [In Chinese]

Yip, M. C. W., (2000).Spoken word recognition of Chinese homophones: The role of context and tone neighbors. Psychologia, 43, 135-143.

Yip, M. C. W., (2002).Access to homophonic meanings during spoken language comprehension: Effects of context and neighborhood density. In the Proceedings of ICSLP'2002, 1665-1668.

Yip, M. C. W., (2007).Spoken word recognition of Chinese homophones: A further investigation. Proceedings of Interspeech'2007, 362-365.

Zhang D., Zhang H., Zhou X.& Shu H.,(1999). Production of Chinese characters under dictation.

Applied Linguistics, 29 (1), 67-70 [In Chinese]

Zhou Z J ,Chen Y.M ,Yang L.X ,Chen Hsuan–Chih, (2003). The inhibition mechanism for resolving lecical ambiguity of two-syllable homophones in Chinese. Acta psycholgica Sinica,35(3), 1-8 [in Chinese]

Zhou, H. & Shu, H. (2008) The Homophone Family Size Effect and Semantic Transparency Effect in Chinese P-O Mapping. Psychological Science, 31, 852 – 855

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

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

Zhou, X., & Marslen-Wilson, W.D. (1995). Morphological structure in the Chinese mental lexicon. Language and Cognitive Processes, 10, 545-601.

Ziegler, J. C., & L. H. Tan., (2000). Phonology matters: The phonological frequency effect in written Chinese. Psychological Science, 3: 234-238

Ziegler, J. C., Muneaux, M., & Grainger, J., (2003). Neighborhood effects in auditory word recognition: Phonological competition and orthographic facilitation. Journal of Memory and Language, 48, 779–793

Ziegler, J. C., Petrova, A., & Ferrand, L. (2008). Feedback consistency effects in visual and auditory word recognition: where do we stand after more than a decade? Journal of Experimental Psychology: Learning Memory & Cognition, 34, 643-661.

McCallum WC, Farmer SF, & Pocock PV, (1984). The effects of physical, semantic incongruities on auditory event-related potentials. Electroencephalography, Clinical Neurophysiology, 59:477-488.

Holcomb PJ, Neville HJ., (1990). Auditory visual semantic priming in lexical decision: a comparison using event-related brain potentials. Language Cognitive Processes, 5, 281-312.

Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. Memory and Cognition 18(4), 380-393.

Van Petten, C., Coulson, S., Rubin, S., Plante, E., & Parks, M. (1999). Time course of word identification and semantic integration in spoken language. Journal of Experimental Psychology, Learning, Memory, and Cognition, 25, 394–417.

Pritchard, W. S., Shappell, S. A., & Brandt, M. E. (1991). Psychophysiology of N200/N400: A review and classification scheme. In J. R. Jennings, P. K. Ackles, & M. G. H. Coles (Eds.), Advances in psychophysiology, vol. 4 (pp. 43–106).Greenwich, CT: JAI Press.

Danie B, Colin M. B, & Peter H, (2001). Electrophysiological Evidence for Early Contextual Influences during Spoken-Word Recognition: N200 Versus N400 EffectsJournal of Cognitive Neuroscience, 7, 967–985

P. Hagoort, C.M. Brown, (2000). ERP effects of listening to speech: semantic ERP effects Neuropsychologia, 38, 1518-1530

Desroches, A. S., Newman, R. L., & Joanisse, M. F., (2009). Investigating the Time Course of Spoken Word Recognition: Electrophysiological Evidence for the Influences of Phonological Similarity. Journal of Cognitive Neuroscience, 21(10), 1893–1906

Marta Kutas and Kara D. Federmeier, (2009). N400. Scholarpedia, 4(10):7790 doi:10.4249/scholarpedia.7790

A. Cummings, R. Čeponienė, A. Koyama, A.P. Saygin, J. Townsend, & F. Dick, (2006). Auditory semantic networks for words and natural sounds. BRAIN RESEARCH, 1115, 92–107

K. A. DELON, G.P. URBACH, D. M. GROPPE, & M. KUTAS., (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. Psychophysiology, 1 – 5.

Kutas, M., & Hillyard, S. A., (1984). Brain potentials during reading reflect word expectancy and semantic association. Nature, 307, 161 – 163.

Wu N. & Shu H., (2003). The gating paradigm and spoken word recognition of Chinese. Acta Psychologica Sinica, 1(2), 105-109

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. Perception and Psychophysics, 28, 267-283.

Chen,H.-C., Vaid, J., & Wu, J.-T. (2009). Homophone Density and Phonological Frequency in Chinese Word Recognition. Language and Cognitive Process, 24, 967-982.

Chen, H.-C., Vaid, J., & Wu, J.T., (2004). Phonology is not universal: Phonological frequency and phonological neighborhood size effects in Chinese single character recognition and naming. Poster presented at Fourth International Mental Lexicon Conference, University of Windsor, Ontario

Thuy DH, Matsuo K, Nakamura K, Toma K, Oga T, Nakai T, Shibasaki H, Fukuyama H. (2004) Implicit and explicit processing of kanji and kana words and non-words studied with fMRI. Neuroimage. 23(3):878-89

Chen, B.G., Liu, Y., Peng, D.L. & Perfetti, C. (2005). Chinese character frequency affects the timing of graphic, phonological and semantic information: an Event-Related Potential study. Paper presented at the 11th international Conference on processing Chinese and other East Asian languages.

Liu, Y., Perfetti, C. A.& Hart, L.(2003) ERP evidence for the time course of graphic, phonological, and semantic information in Chinese meaning and pronunciation decisions. Journal of Experimental Psychology: Learning, Memory, and Cognition, 29, 1231–1247.

Harm, M. W. & Seidenberg, M. S., (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. Psychological Review, 111 (3), 662-720.

Stone, G. O., Vanhoy, M., & Van Orden, G. C. (1997). Perception is a two-way street: Feedforward and feedback phonology in visual word recognition. Journal of Memory and Language, 36, 337-359.

Stone, G. O. & Van Orden, G. C. (1994). Building a resonance framework for word recognition using design and system principles. Journal of Experimental Psychology: Human Perception and Performance, 20 (6), 1248-1268.

Ventura, P., Morais, J., Pattamadilok, C., & Kolinsky, R. (2004). The locus of the orthographic consistency effect in auditory word recognition. Language and Cognitive Processes, 19 (1), 57-95.

張學新 (已接受). 漢字拼義理論：用認知心理學揭示漢字的本質. 《華南師範大學學報》(社科版).

Appendix A stimuli for Experiment 1& 4

ID	syllable	homophone size	highest hom. word	ID	syllable	homophone size	highest hom.word	ID	nonword	ID	nonword
1	bei4	12	被	1	ban4	7	半	1	hang1	31	pian3
2	bo2	15	薄	2	bian1	6	边	2	na1	32	pen3
3	cheng2	12	成	3	bo1	7	波	3	rui1	33	pang3
4	dian4	14	电	4	chan3	4	产	4	re1	34	nen3
5	e4	14	恶	5	chu1	2	出	5	ru1	35	pei3
6	ge2	12	格	6	chuan2	4	船	6	nan1	36	zei3
7	gu3	12	古	7	jia4	6	架	7	mian1	37	pa3
8	hui4	18	会	8	ding4	5	定	8	nuo1	38	le3
9	jiao3	15	角	9	du2	6	读	9	lian1	39	tou3
10	jing1	13	经	10	dui4	4	对	10	kuo1	40	tai3
11	jue2	14	决	11	fang2	5	防	11	nv1	41	kuang3
12	li2	12	离	12	gen1	2	跟	12	cei1	42	cei3
13	li3	11	里	13	gou1	4	沟	13	nei1	43	xia3
14	lu4	16	路	14	jia3	6	甲	14	mei1	44	nie3
15	mu4	12	木	15	jun1	5	军	15	nuan1	45	que3
16	qi1	13	七	16	lian4	5	练	16	te2	46	xun3
17	qian1	11	千	17	mai4	4	卖	17	bang2	47	shei3
18	quan2	13	全	18	mao4	6	冒	18	che2	48	qiong4
19	shan1	12	山	19	mei3	3	每	19	dou2	49	shei4
20	shan4	11	善	20	mu3	5	亩	20	jun2	50	jiong4
21	shu1	16	书	21	qing1	6	清	21	shuan2	51	keng4
22	si4	13	四	22	sheng1	7	生	22	hei2	52	zei4
23	wei3	13	委	23	sheng4	4	胜	23	ri2	53	xiong4
24	xiao1	15	消	24	shou3	3	手	24	sai2	54	teng4
25	xie4	15	谢	25	ti2	5	提	25	zhua2	55	qun4
26	xia2	11	霞	26	tian2	3	田	26	re2	56	tian4
27	yi2	18	移	27	ze2	6	则	27	dui2	57	tun4
28	yin1	11	因	28	zhu2	6	逐	28	le2	58	shuang4
29	ying2	12	营	29	zu3	4	组	29	run2	59	nang4
30	zhen1	13	真	30	zuo4	8	作	30	kuai2	60	ting4
Mean density size		13.3				4.9					

Appendix B stimuli for Experiment 2& 5

Disyllable Homophone (HW)				High frequency control word (HF)				Cumulate frequency control word (CF)			
ID	Pinyin	homophone size	example words	ID	Pinyin	homophone size	word	ID	Pinyin	homophone size	word
1	bao4fu4	2	报复、抱负	1	chun1lian2	1	春联	1	ba1xi1	1	巴西
2	bi3shi4	2	鄙视、笔试	2	dai4ci2	1	代词	2	bai3she4	1	摆设
3	dan1jia4	2	单价、担架	3	yan4fan2	1	厌烦	3	ban1jia1	1	搬家
4	dao3dan4	2	导弹、捣蛋	4	jiao3bian4	1	狡辩	4	bi4ye4	1	毕业
5	fu2li4	2	福利、浮力	5	dao4li3	1	道理	5	chao1zai4	1	超载
6	gang1cai2	2	刚才、钢材	6	fa3yi1	1	法医	6	chen2jiu4	1	陈旧
7	gong1ke4	2	攻克、功课	7	wen1he2	1	温和	7	chu4jiao3	1	触角
8	heng2liang2	2	衡量、横梁	8	gan1yan2	1	肝炎	8	dian4liu2	1	电流
9	jian1gu4	2	坚固、兼顾	9	hong2qi2	1	红旗	9	gao1zhong1	1	高中
10	jiao1ji2	2	焦急、交集	10	kai1xin1	1	开心	10	guan1yuan2	1	官员
11	jie2jing4	2	洁净、捷径	11	wu1ya1	1	乌鸦	11	gui1mo2	1	规模
12	jing1ying2	2	经营、晶莹	12	qi2quan2	1	齐全	12	hu2xu1	1	胡须
13	mi4feng1	2	密封、蜜蜂	13	ji4xiang4	1	迹象	13	hui1huo4	1	挥霍
14	mu4di4	2	目的、墓地	14	jie4jian4	1	借鉴	14	jian3yan4	1	检验
15	po4li4	2	魄力、破例	15	ge2jue2	1	隔绝	15	jiao4shi1	1	教师
16	qian2cheng2	2	前程、虔诚	16	xu1wei3	1	虚伪	16	jin1bi4	1	金币
17	qiang2bao4	2	强暴、墙报	17	lian2xu4	1	连续	17	jin4bu4	1	进步
18	qing1dan4	2	氢弹、清淡	18	lan2qiu2	1	篮球	18	jiu1zheng4	1	纠正
19	qing1dao3	2	青岛、倾倒	19	ling2zhi1	1	灵芝	19	ju4li2	1	距离
20	sheng1yu4	2	生育、声誉	20	bian4hu4	1	辩护	20	kui2wu2	1	魁梧
21	shou3zhang3	2	首长、手掌	21	ji1chu3	1	基础	21	li2ming2	1	黎明
22	shu4mu4	2	数目、树木	22	fan4wei2	1	范围	22	lu4xian4	1	路线
23	xiao1shou4	2	销售、消瘦	23	jie1duan4	1	阶段	23	sui4yue4	1	岁月
24	xing2cheng2	2	形成、行程	24	she4ji1	1	射击	24	tong1dao4	1	通道
25	yan3shi4	2	掩饰、演示	25	tai4yang2	1	太阳	25	tuo2bei4	1	驼背
26	yi2qi4	2	仪器、遗弃	26	xiang1wei4	1	香味	26	wu2tong2	1	梧桐
27	yi4li4	2	毅力、屹立	27	xie2e4	1	邪恶	27	xi1shu1	1	稀疏
28	ying1wu3	2	英武、鹦鹉	28	peng2zhang4	1	膨胀	28	xia2gu3	1	峡谷
29	you2ji4	2	游记、邮寄	29	fen4dou4	1	奋斗	29	xian4dai4	1	现代
30	zhan4you3	2	占有、战友	30	xing4ge2	1	性格	30	xin1ling2	1	心灵
31	zhi2wu4	2	植物、职务	31	yan2shen1	1	延伸	31	yuan2ze2	1	原则
32	zhu3fu4	2	嘱咐、主妇	32	you4zhi4	1	幼稚	32	zao4yin1	1	噪音
33	zhu4yuan4	2	祝愿、住院	33	zhen1cha2	1	侦查	33	zhi1ji3	1	知己
34	shan1dong4	2	煽动、山洞	34	zheng4shi2	1	证实	34	zhi3zhen1	1	指针
35	you2tong3	2	油桶、邮筒	35	zhi4hui4	1	智慧	35	zong1jiao4	1	宗教
36	gong1li4	6	功利、公历	36	huo4bi4	1	货币	36	jian1yu4	1	监狱
37	shi4li4	6	势力、视力	37	jin4bu4	1	进步	37	you4zhi4	1	幼稚
38	gong1yi4	5	工艺、公益	38	huang2di4	1	皇帝	38	shu4e2	1	数额

39	fu4shu4	4	复数、复述	39	li3fu2	1	礼服	39	wei4ju4	1	畏惧
40	bao3jian4	3	保健、宝剑	40	ji2he2	1	集合	40	qi2yi4	1	奇异
41	shou3shi4	3	手势、首饰	41	yan4hui4	1	宴	41	xian4dai4	1	现代
42	xi1li4	4	西历、犀利	42	ji1ji2	1	积极	42	bu4wei4	1	部位
43	yuan2xing2	4	圆形、原形	43	bi3ji4	1	笔记	43	you2zhi1	1	油脂
44	cheng2xian4	3	呈现、呈献	44	zuo4jia1	1	作家	44	mao2yi1	1	毛衣
45	ci2xing4	3	磁性、词性	45	pi1jian1	1	披肩	45	xiang1si4	1	相似
46	gong1ji1	3	攻击、公鸡	46	yi1jiu4	1	依旧	46	ji4cheng2	1	继承
47	jian3pu3	3	简谱、俭朴	47	hong2qi2	1	红旗	47	qi4xi1	1	气息
48	jing1li4	3	经历、精力	48	yi4shi2	1	意识	48	wu3qi4	1	武器
49	qi2shi4	3	歧视、骑士	49	jia1shu3	1	家属	49	zhi2gong1	1	职工
50	quan2li4	3	权利、权力	50	li4xi1	1	利息	50	bian4cheng2	1	变成
51	shi1shou3	3	失守、尸首	51	jiu4ye4	1	就业	51	tong2shi4	1	同事
52	yin1su4	3	因素、音速	52	lian2yi1	1	涟漪	52	hui4yuan2	1	会员
53	yu4shi4	3	预示、浴室	53	qian1yi2	1	迁移	53	di4li3	1	地理
54	yuan2zhu4	3	援助、圆柱	54	jian4yi4	1	建议	54	jue2shi2	1	绝食
55	zhong1zhi3	3	中指、终止	55	jin1yu2	1	金鱼	55	jiao3zheng4	1	矫正
56	zhu4shou3	3	助手、驻守	56	gao1yuan2	1	高原	56	ji4xu4	1	继续
Mean density size		2.8				1				1	

Appendix C stimuli for Experiment 3a & 6a

ID	BH		SH		BL		SL	
group	**Pinyin**	**word**	**Pinyin**	**word**	**Pinyin**	**word**	**Pinyin**	**word**
1	jing4ai4	敬爱	xi3ai4	喜爱	bo2ai4	博爱	pian1ai4	偏爱
2	yun4cang2	蕴藏	yin3cang2	隐藏	yan3cang2	掩藏	duo3cang2	躲藏
3	zi1chan3	资产	cai2chan3	财产	wu4chan3	物产	nan2chan3	难产
4	ci2chang3	磁场	zhan4chang3	战场	qiu2chang3	球场	kai1chang3	开场
5	wei2chi2	维持	bao3chi2	保持	xie2chi2	挟持	jiang1chi2	僵持
6	yi2dong4	移动	ruo4dian3	弱点	wu1dian3	污点	ran2dian3	燃点
7	rong2dian3	熔点	bai3dong4	摆动	chu4dong4	触动	huang4dong4	晃动
8	e4du2	恶毒	bing4du2	病毒	mei2du2	梅毒	hen2du2	狠毒
9	hu2du4	弧度	hou4du4	厚度	ling2du4	零度	nan2du4	难度
10	shi1ge1	诗歌	chang4ge1	唱歌	shan1ge1	山歌	er2ge1	儿歌
11	yang2guang1	阳光	pu4guang1	曝光	zheng1guang1	争光	zhan1guang1	沾光
12	fu2hao4	符号	bian1hao4	编号	jian3hao4	减号	an4hao4	暗号
13	yong1hu4	拥护	fang2hu4	防护	bi4hu4	庇护	shou3hu4	守护
14	xian1hua1	鲜花	huo3hua1	火花	lan2hua1	兰花	lang4hua1	浪花
15	hu4kou3	户口	shang1kou3	伤口	dong4kou3	洞口	xiong1kou3	胸口
16	ke4ku3	刻苦	tong4ku3	痛苦	su4ku3	诉苦	wa1ku3	挖苦
17	ping2lun4	评论	miu4lun4	谬论	xu4lun4	绪论	gai4lun4	概论
18	yan2mi4	严密	jin3mi4	紧密	xie4mi4	泄密	mao4mi4	茂密
19	lu4mian4	路面	hua4mian4	画面	feng1mian4	封面	lian3mian4	脸面
20	yin1mou2	阴谋	can1mou2	参谋	yu4mou2	预谋	mi4mou2	密谋
21	jie2mu4	节目	mang2mu4	盲目	zhang4mu4	账目	er3mu4	耳目
22	ya1pian4	鸦片	ka3pian4	卡片	zhi3pian4	纸片	wa3pian4	瓦片
23	yong3qi4	勇气	qing1qi4	氢气	sha1qi4	杀气	du2qi4	毒气
24	xun2qiu2	寻求	zhui1qiu2	追求	ke1qiu2	苛求	yang1qiu2	央求
25	huan4suan4	换算	tui1suan4	推算	gu1suan4	估算	hua2suan4	划算
26	xu1xian4	虚线	hang2xian4	航线	zhen1xian4	针线	zhuan1xian4	专线
27	lian2xiang3	联想	meng4xiang3	梦想	xia2xiang3	遐想	cai1xiang3	猜想
28	jue2xin1	决心	dan1xin1	担心	shen1xin1	身心	cao1xin1	操心
29	guan4xing4	惯性	gan3xing4	感性	wu4xing4	悟性	nai4xing4	耐性
30	yao2yan2	谣言	xuan1yan2	宣言	ge2yan2	格言	nuo4yan2	诺言
31	ying2yang3	营养	pei2yang3	培养	liao2yang3	疗养	shou1yang3	收养
32	zhou4ye4	昼夜	hei1ye4	黑夜	xiao1ye4	宵夜	ao2ye4	熬夜
33	qu1zhe2	曲折	cuo4zhe2	挫折	gu3zhe2	骨折	zou4zhe2	奏折
34	han4zu2	汉族	zang4zu2	藏族	li2zu2	黎族	miao2zu2	苗族

Appendix D stimuli for Experiment 3b & 6b

ID	BH		SH		BL		SL	
group	Pinyin	word	Pinyin	word	Pinyin	word	Pinyin	word
1	an1jing4	安静	an1ding4	安定	an1wei1	安危	an1xiang2	安详
2	bao1wei2	包围	bao1kuo4	包括	bao1xiang1	包厢	bao1za1	包扎
3	bao4zhi3	报纸	bao4kan1	报刊	bao4xiao1	报销	bao4chang2	报偿
4	biao3yang2	表扬	biao3xiang4	表象	biao3ge1	表哥	biao3bai2	表白
5	chu1lu4	出路	chu1xi2	出席	chu1mo4	出没	chu1na4	出纳
6	chun1feng1	春风	chun1qiu1	春秋	chun1lian2	春联	chun1guang1	春光
7	da3sui4	打碎	da3po4	打破	da3mo2	打磨	da3lie4	打猎
8	dong4yao2	动摇	dong4tai4	动态	dong4ci2	动词	dong4mai4	动脉
9	du2su4	毒素	du2hai4	毒害	du2qi4	毒气	du2she2	毒蛇
10	dui4dai4	对待	dui4chen4	对称	dui4zhe2	对折	dui4ce4	对策
11	fa1she4	发射	fa1cai2	发财	fa1mei2	发霉	fa1shao1	发烧
12	fang1zhen1	方针	fang1an4	方案	fang1zhang4	方丈	fang1cun4	方寸
13	feng1shan4	风扇	feng1jing3	风景	feng1sha1	风沙	feng1shuang1	风霜
14	gong1si1	公司	gong1gong4	公共	gong1jia1	公家	gong1zhang1	公章
15	hai3xia2	海峡	hai3wan1	海湾	hai3shen1	海参	hai3mian2	海绵
16	huo3yan4	火焰	huo3zai1	火灾	huo3lu2	火炉	huo3tan4	火炭
17	jia1shu3	家属	jia1ting2	家庭	jia1pu3	家谱	jia1fang3	家访
18	jiao4tang2	教堂	jiao4lian4	教练	jiao4huang2	教皇	jiao4tu2	教徒
19	ke3ai4	可爱	ke3pa4	可怕	ke3yi2	可疑	ke3guan1	可观
20	lao3han4	老汉	lao3ye2	老爷	lao3bo2	老伯	lao3tao4	老套
21	mi4shu1	秘书	mi4mi4	秘密	mi4jue2	秘诀	mi4fang1	秘方
22	mian4mu4	面目	mian4mao4	面貌	mian4rong2	面容	mian4pang2	面庞
23	ning2jie2	凝结	ning2gu4	凝固	ning2ju4	凝聚	ning2zhong4	凝重
24	nong2ye4	农业	nong2yao4	农药	nong2fu4	农妇	nong2huo2	农活
25	pi2xie2	皮鞋	pi2fu1	皮肤	pi2yan2	皮炎	pi2bao1	皮包
26	qi4xi1	气息	qi4fen1	气氛	qi4gong1	气功	qi4nao3	气恼
27	qing2xu4	情绪	qing2xing2	情形	qing2di2	情敌	qing2qu4	情趣
28	shen2xian1	神仙	shen2sheng4	神圣	shen2yun4	神韵	shen2cai3	神采
29	te4zheng1	特征	te4bie2	特别	te4qu1	特区	te4chan3	特产
30	ti3yu4	体育	ti3xi4	体系	ti3ge2	体格	ti3tie1	体贴
31	wu2gu1	无辜	wu2chi3	无耻	wu2liao2	无聊	wu2lai4	无赖
32	xin1ling2	心灵	xin1zang4	心脏	xin1gan1	心肝	xin1yuan4	心愿
33	xu1wei3	虚伪	xu1jia3	虚假	xu1wu2	虚无	xu1huan4	虚幻