

Visual Processing of Chinese Characters by Foreign
Language Learners

by

Ting (Alice) Xia

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Cognitive Science

in

Cognitive Science

Carleton University
Ottawa, Ontario

© 2018, Ting (Alice) Xia

Abstract

The constituent parts of compound Chinese characters provide crucial semantic and phonetic information for deciphering the character as a whole. Exploring sub-character level processing may help improve our understanding of the mechanisms of Chinese vocabulary storage and retrieval, especially in foreign language learners. Yet, empirical research on adult Chinese-as-a-foreign-language (CFL) learners in this area remains limited. Do CFL learners use decomposition strategies in visual character recognition? If so, how does the combination of familiarity with an individual character and character frequency contribute to this process? How do CFL learners and Chinese native speakers differ? To explore these questions, a masked priming character decision task was used. Results show that CFL learners do not decompose familiar high-frequency characters nor unfamiliar low-frequency characters. However, there was evidence of decomposition for unfamiliar high-frequency characters and familiar low-frequency characters. In comparison, native speakers were found to process compound characters holistically.

Acknowledgements

I am deeply grateful to my advisor, Dr. Kumiko Murasugi. This thesis was the product of countless hours of brainstorming and discussion with her. Dr. Murasugi's faith in my academic abilities not only contributed to the success of this project, but also to my personal growth. Despite nearly two years of exploring language and cognition, words still seem inadequate for expressing how thankful I am to have had her continuous guidance and support.

I would also like to thank the members of my examination committee, Dr. John Logan, Dr. David Wood, and particularly, Dr. Olessia Jouravlev, who believed in my potential, perhaps even more than I did at times. Dr. Jouravlev invested much time and energy into helping me become a better researcher than I was the day before.

Lastly, I would like to thank Laura Luo and her students, many of whom took time out of their busy schedules to participate in my study.

Table of Contents

Abstract.....	ii
Table of Contents	iv
1 Chapter: Introduction	1
1.1 Research background	1
1.2 Research questions	2
2 Chapter: Chinese Character Processing.....	5
2.1 Multilevel interactive-activation model	5
2.2 Lexical constituency model.....	7
2.3 Parallel distributed processing model	9
2.4 Character processing in Chinese native speakers.....	11
2.5 Character processing in CFL learners	13
2.6 Theories of word recognition in alphabetic languages.....	16
2.7 Summary	18
2.8 Predictions for character processing method	19
3 Chapter: Method.....	22
3.1 Participants	22
3.1.1 CFL learners	22
3.1.2 Chinese native speakers	23
3.2 Materials.....	24
3.2.1 Stimuli selection and design	25
3.2.1.1 <i>Experimental items</i>	25
3.2.1.2 <i>Filler items</i>	27

3.3	Experimental design.....	29
3.3.1	Character decision task.....	29
3.3.2	Character familiarity test and language background questionnaire.....	31
3.4	Procedure.....	32
3.5	Statistical analyses.....	33
4	Chapter: Results.....	34
4.1	Data cleansing.....	34
4.1	Character familiarity.....	34
4.2	Overview of reaction time results.....	36
4.3	Overview of accuracy rate results.....	39
4.4	Mixed effects models.....	43
4.4.1	Reaction time data analysis.....	44
4.4.1.1	<i>Fixed effects</i>	44
4.4.1.2	<i>Interaction terms</i>	45
4.4.1.3	<i>Random effects</i>	45
4.4.1.4	<i>Planned comparisons</i>	46
4.4.1.4.1	CFL learners group.....	47
4.4.1.4.2	Native speakers group.....	48
4.4.2	Accuracy data analysis.....	50
4.4.2.1	<i>Fixed effects</i>	51
4.4.2.2	<i>Interaction terms</i>	51
4.4.2.3	<i>Planned comparisons</i>	52
4.4.2.3.1	CFL learners group.....	52
4.4.2.3.2	Native speakers group.....	53

5 Chapter: Discussion	55
5.1 Review of the research questions and predictions	55
5.2 Interpretation of the findings	57
5.2.1 Character processing by CFL learners	57
5.2.2 Character processing by Chinese native speakers	60
5.2.3 Misidentification of real characters	61
5.2.4 Summary	63
5.3 Limitations of the study	64
5.4 Conclusion.....	65
Appendices.....	67
Appendix A	67
Appendix B	71
Appendix C	72
Appendix D	74
Appendix E.....	76
References	83

1 Chapter: Introduction

1.1 Research background

Chinese is considered a logographic language in which characters are the basic units of orthography (DeFrancis, 1989). Compound characters, which typically contain a semantic radical and phonetic component, represent 80% of the most commonly used Chinese characters (Hoosain, 1991). The semantic radical usually indicates the conceptual category that a character belongs to. For example, in the character that means “cat” (猫), the semantic radical on the left side (犹) suggests a relation to four-legged animals. The same radical is also found in the characters for “monkey” (猴), “dog” (狗), and “wolf” (狼), to name a few. In this sense, a semantic radical is comparable to a morpheme in an alphabetic language. Semantic radicals that are also characters in their own right are comparable to free morphemes in English. For example, in the character for “mother” (妈), the semantic radical on the left signals a relation to female-related concepts, but it also means “woman” when it stands alone (女). The phonetic component, on the other hand, only exactly matches the pronunciation and tone of a character one-third of the time (Fan, Gao, & Ao, 1984). Due to the combination of these sub-character components, Chinese has been labeled a mixed morpho-syllabic system, in which some elements point to semantic categories and others point to sound (Dehaene, 2009). This raises interesting questions regarding levels of mental representation for characters and sub-character components, and specifically, how they contribute to cognitive processes underlying Chinese character processing.

This thesis explores the effect of semantic radicals on the visual processing of

semantic-phonetic compound characters, in order to better understand how the “part-whole” distinction in these types of characters contribute to reading in Simplified Chinese. Research on character processing in Chinese native speakers has shown that high-frequency characters are processed holistically, while low-frequency characters are broken down into more informative parts (Tao & Healy, 2002; Ding, Peng, & Taft, 2004). The main interest of this thesis is to examine whether patterns of compound character processing can also be found in Chinese-as-a-foreign-language (CFL) learners who are native speakers of an alphabetic language.

1.2 Research questions

While psycholinguistic research on CFL learners has increased in recent years (see Williams, 2010; Yang, Wang, Shu, & Zevin, 2011; Shen & Jiang, 2013; Zhang & Li, 2016), it remains unclear whether CFL learners decompose semantic-phonetic compounds or access them holistically. It is also uncertain how factors such as vocabulary knowledge and frequency may affect character processing. Furthermore, studies that have examined CFL learners seldom explicitly measure the effect of item-level vocabulary knowledge (exception being Shen & Ke, 2007), which has been considered potentially more important than word – or in this case, character – frequency for second-language learners (Gass & Selinker, 2008).

This thesis therefore aims to address the following questions:

1. How do subjective character familiarity and whole-character frequency affect whether CFL learners decompose or holistically process Chinese semantic-phonetic compound characters?

2. Do CFL learners differ from Chinese native speakers during the visual processing of semantic-phonetic compound characters, and if so, how?

In order to better understand how compound characters may be encoded in the mental lexicon and retrieved under certain task demands, a character decision task with masked priming of a semantic radical was used. The task was given to CFL learners of various levels of proficiency, as well as adult Chinese native speakers who served as controls. Previous priming studies with native speakers did not require participants to rate their familiarity with characters items (e.g., Zhou & Marslen-Wilson, 1999). This may be problematic because results from earlier visual and auditory word recognition experiments in English showed that subjective word familiarity ratings were strongly correlated with slower reaction times for low-frequency words in the visual paradigm (Gernsbacher, 1984; Connine, Mullennix, Shernoff, & Yelen, 1990). Since character familiarity was not accounted for in previous research, and research has shown that this may be especially problematic for low-frequency lexical items, Chinese native speakers were included in this study as controls. The relationship between lexical decision task results and subjective item familiarity has not been sufficiently explored in this population.

In the character decision task used in this study, participants were required to judge as quickly and accurately as possible whether a target they saw was a real character or a fake character. The masked prime was a semantic radical that either matched or did not match the radical in the target. The distinct feature of masked priming, compared to the standard long-term priming paradigm, is the very short interval between onset of the prime and onset of the target. Typically, a prime is only presented for 50 to 60

milliseconds (ms), a period of time so brief that participants usually cannot identify the nature of the prime, yet it can be registered by the visual system (Kinoshita & Lupker, 2003). The objective of the character decision task is to measure whether or not participants involuntarily decompose a character into its constituent parts under certain conditions. The masked prime therefore serves as an effective mechanism for preventing participants from engaging with each character in a pre-determined manner, such as in the component search task used by Tao & Healy (2002).

To judge the target character either holistically or non-holistically, stimulus must be processed at the holistic level; if this process and its consequent judgment are affected by a prime below that of the whole-character level, then it can be assumed that a lower sub-character level was passed through to reach the whole character. Similarly, it is assumed that if a prime below the whole-character level does not contribute to any differences in judgment, then holistic processing has occurred. Since decomposition is defined in relation to responses elicited from the character decision task, little or no difference in reaction time or accuracy rate in matched versus unmatched conditions would suggest holistic processing. A significant difference in responses between the two conditions would be taken as evidence of decomposition.

2 Chapter: Chinese Character Processing

This chapter will begin with an overview of a few established models of Chinese character processing that are relevant to this thesis: the multilevel interactive-activation model (Taft, 1994); the lexical constituency model (Perfetti, Liu, & Tan, 2005); and the parallel-distributed-processing model (Yang, Shu, McCandliss, & Zevin, 2013). Previous research on Chinese processing at both the whole character and sub-character levels in native speakers and CFL learners will then be presented and discussed. Models of word recognition and processing in alphabetic languages will also be briefly discussed towards the end, for comparison. Finally, based on relevant findings from the existing literature, predictions of performance in the character decision task will be presented.

2.1 Multilevel interactive-activation model

The principle notions of the multilevel interactive-activation (IA) model proposed by Taft (1994) are borrowed from McClelland and Rumelhart's (1981) connectionist model of letter perception. This earlier model outlined layers of units within the visual system that are divided by feature units at the bottom, letter units in the middle, and word units at the top. It is claimed that word processing (in English) involves the discrete feed-forward activation of each of these layers, beginning with the feature units that comprise individual letters, and ending with the whole word.

The IA model of Chinese character processing suggests a similar cognitive mechanism containing discrete hierarchical levels that are activated by specific elements of the Chinese writing system (Fig. 1). Taft and Zhu (1997) define individual strokes as the feature-level units in this model. The second layer includes radicals. ("Radicals" in

this chapter is used as an umbrella term for both semantic radicals and phonological components). The third layer contains whole characters that are connected to semantic concepts.

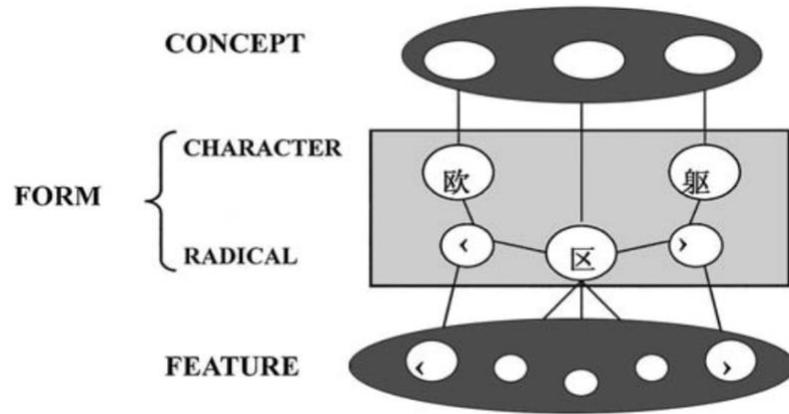


Figure 1. Schematic form of the multilevel interactive-activation model of Chinese character processing. Taken from Ding, G., Peng, D., & Taft, M. (2004). The Nature of the Mental Representation of Radicals in Chinese: A Priming Study. *Journal of Experimental Psychology: Learning Memory and Cognition*, 30(2), 530–539.

Based on the figure above, the IA model assumes that activation begins at the feature-level of individual strokes before entering the orthographic (form) sub-system. Both radical- and character-level units are directly linked to meaning. An additional assumption of the model not shown in the figure relates to the positional regularity of radicals. A series of priming experiments showed that facilitation effects were found when primes and targets shared the same radical in the same position (e.g., 躯 and 枢), and not in different positions (e.g., 欧 and 枢), leading to the conclusion that positional information about sub-character components is also stored at the radical level of representation (Ding et al., 2004).

There are several implications of the IA model. First, compound characters (e.g., 时, meaning “time”) are represented at a level above simple characters (e.g., 日, meaning “sun”), since simple characters may also act as radicals, and vice versa. Second, there is no distinction between a radical that can only be a radical (cf. bound morpheme in English), and a radical that can also be a character (cf. free morpheme in English). That is, the same level is activated for the character 日 when it stands alone, and when it appears as part of another character such as 时. Third, there is a distinction between radicals that exist in different positions, even if the radical itself remains the same.

2.2 Lexical constituency model

In the lexical constituency (LC) model proposed by Perfetti, Liu, and Tan (2005), inputs begin at the radical level, not the stroke level (see Fig. 2). Radical inputs (e.g., 亻, indicating a relation to people) activate whole characters (e.g., 他, meaning “he” or “him”) at the orthography level. Whole characters then activate the semantic layer, which is similar to the conceptual level in the IA model. The whole character units at the orthographic level, and semantic units that contain a unique meaning for each character, simultaneously activate phonological units, which are single syllables for the vast majority of Chinese characters.

The LC model assumes that radicals that can be characters in their own right (such as the example of 日 provided earlier) are not distinct from whole characters, and thus share the orthography level with whole characters. The model also assumes that there is no direct link between the radical units and phonological units, with the exception of these types of radicals that can stand alone as characters.

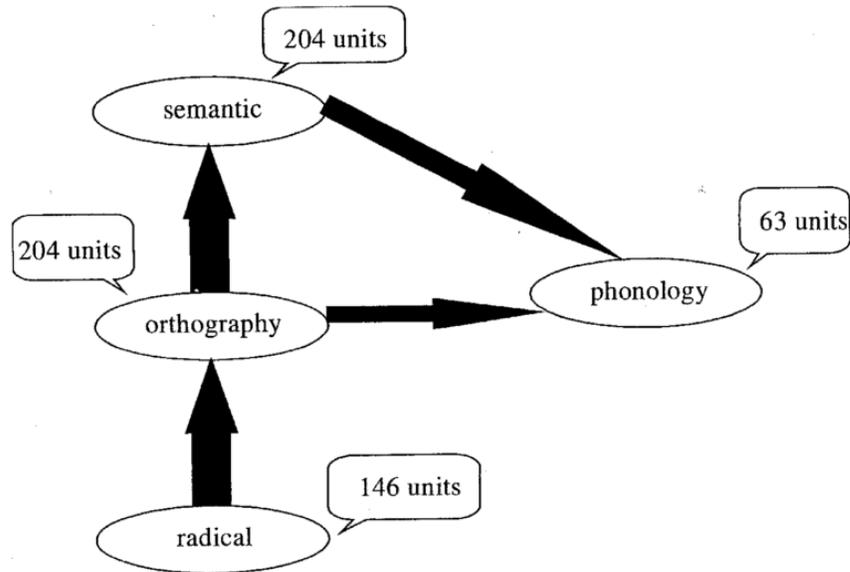


Figure 2. Schematic form of the lexical constituency model of Chinese character processing. Adapted from Perfetti, C. & Liu, Y. (2006). Reading Chinese characters: Orthography, phonology, meaning and the lexical constituency model. *Handbook of East Asian Psycholinguistics*.

There are two main assumptions about Chinese character processing that come out of the LC model: 1) phonological activation is obligatory; and 2) there is no direct phonological activation of sub-character components. Given that on average, there are five homophones per character in Chinese, a distinct phonology layer may not be warranted. In response to how this distinct feature of Chinese may affect the reliability of the LC model, Perfetti, Liu, and Tan (2005) argue that the strength of the relationship between written and spoken word forms may result in obligatory phonological processing, even if it does not ultimately feed into a semantic network. Additionally, the authors claim that the automatic activation of character pronunciations may be important in serving post-lexical memory and comprehension processes.

2.3 Parallel distributed processing model

The parallel distributed processing (PDP) model of Chinese reading proposed by Yang, Shu, McCandliss, and Zevin (2013) was established through computational simulations of three-way mappings among orthography, meaning and sound. The PDP model assumes that Chinese reading involves three levels of representation: orthographic, phonological, and semantic (Fig. 3). All three are linked via the semantic and the non-semantic (orthographic) pathways. The non-semantic pathway allows for phonological representations at both the character and sub-character levels, which is not the case in the LC model.

The PDP model suggests that in Chinese, orthography-to-meaning mappings are learned much faster than orthography-to-sound mappings. When compared to the “triangle” model of reading that accounts for spelling-to-sound occurrences in English (Seidenberg & McClelland, 1981), the PDP model is better able to capture variation in reading ability as a result of the statistical regularities in a particular writing system. It aims to describe how such regularities might affect the “division of labor” between semantic and phonological processing.

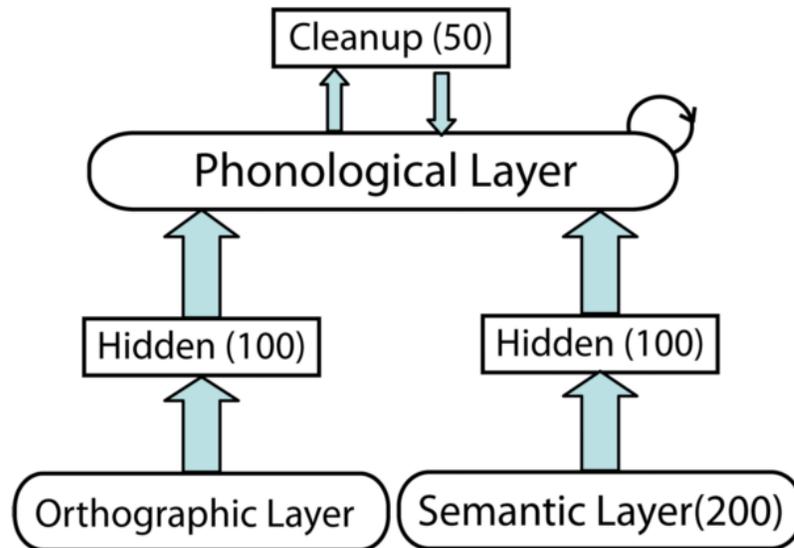


Figure 3. Schematic illustration of the PDP architecture used to simulate Chinese reading. Taken from Yang, J., Shu, H., McCandliss, B. D., & Zevin, J. D. (2013). Orthographic influences on division of labor in learning to read Chinese and English: Insights from computational modeling. *Bilingualism: Language and Cognition*, 16(2), 354-366.

The different theoretical frameworks introduced in this section reveal a lack of consensus on the exact levels of representations in Chinese character processing, along with unresolved debates about the importance of phonological access. However, at minimum, the idea that there is a direct link between radical and character seems to have been established. Regarding the more specific question of how radicals may inhibit or excite character processing, previous behavioural studies have attempted to provide some answers.

2.4 Character processing in Chinese native speakers

Research on character processing in Chinese native speakers has been driven largely by the debate over whether there is direct access to meaning, or whether access to meaning is mediated by phonology. Findings from behavioural research have differed in whether there is a stronger phonological or semantic advantage in reading Chinese words. In two experiments examining form-form (orthographic) versus form-meaning relationships in a naming task, Perfetti and Tan (1998) found that activation of graphemic information occurred within 43 ms after prime onset, without phonological or semantic activation; homophonic primes facilitated target character identification only after 57 ms, but semantic priming was still absent until 115 ms after the prime was presented. The results contributed to the construction of the LC model, which considers phonological activation as an automatic process during character identification. However, in a replication of these experiments, Chen and Shu (2001) not only failed to find a preference for early phonological activation, but in fact found that the semantic radical of a character was consistently activated before the phonetic component. This was unexpected, especially since a naming task was used, which would seem to bias participants towards the use of phonological rather than semantic information. Such results support the traditional “direct semantics” view of reading (e.g., Seidenberg, 1985; Taft & Zhu, 1997; Zhou & Marslen-Wilson, 1999).

Using a character decision task, Taft and Zhu (1997) provided evidence of left-to-right serial decomposition of compound characters through character decision tasks that were designed to examine variables such as radical frequency, position, and function. Semantic radical frequency was found to have an effect on character recognition speed,

but only when the position of the radical was taken into account. These results contributed to the theoretical framework of the IA model (Taft & Zhu, 1997).

Feldman and Siok (1997) called into question left-to-right serial processing by pointing out a neglected aspect of the experiments by Taft and Zhu (1997): semantic radicals tend to be positioned on the left side of a character, and phonetic components tend to be positioned on the right side. Without controlling for the function of these sub-character components, a full conclusion about compound character processing cannot be drawn. By manipulating both combinability (the number of combinations a sub-character component enters into to form a whole character) and position, Feldman and Siok (1999) found that reaction times in a character decision task were delayed for semantic radicals on the left, but not for phonetic components on the right. No significant interaction was found between the phonetic components and their position within a character. These results showed that: 1) semantic radical combinability facilitates character processing; and 2) the importance of radical position is more ambiguous when semantic versus phonetic function is taken into account. Importantly, these studies support a view of character decomposition (or analytic processing), in which reading involves making distinctions between the parts of a character and the character as a whole.

Other researchers have argued for holistic character processing, particularly with high-frequency items (Tao & Healy, 2002), which echoes the word superiority effect (Reicher, 1969). Using a component detection task, Tao and Healy (2002) found that native speakers made more errors on high-frequency Chinese characters than on low-frequency characters. These findings suggest that sub-character components were not fully processed because attention was directed to the whole character level of analysis.

High-frequency characters may reach their recognition threshold rapidly, leaving insufficient time to observe any effects facilitated by its component parts (Ding et al., 2004). The IA model potentially supports this interpretation, since characters (as well as radicals that can act as characters) have direct pathways to semantic knowledge. In other words, certain radicals do not have to be processed first. This “unitization effect” suggested by the Tao and Healy (2002) implies that the cognitive mechanism of Chinese character processing is a hybrid of analytic and holistic processing. Other researchers have also found that native speakers tend to process high-frequency words as whole units and low-frequency words as component morphemes (Taft & Zhu, 1995). This fits with previous research suggesting that radical priming effects in character decision tasks are observed only in low-frequency single characters, but not in high-frequency single characters (Ding et al., 2004).

2.5 Character processing in CFL learners

The acquisition of Chinese orthography has been considered especially challenging for CFL learners for a number of reasons. First, Chinese orthography is visually complex, with up to 44 variations of stroke patterns that combine to form over 400 sub-character components (Chang, Perfetti, Zhang, & Chen, 2014). Second, grapheme-to-phoneme translation is more opaque in Chinese than in most alphabetic languages, since there is no connection between the orthographic form of a character and its pronunciation.

In comparison to Chinese native speakers, CFL learners are more likely to use all possible available cues that are relevant for Chinese character acquisition, thus explaining

how they gradually become perceptive of the position and function of radicals in compound characters, without explicit instruction (Tong & McBride-Chang, 2010). While native speakers learn the orthographic form of characters after they have been exposed to the phonological patterns of Chinese, CFL learners usually acquire knowledge of the Chinese writing system in tandem with learning the pronunciation of characters (Liu, Wang, & Perfetti, 2007). Awareness of orthographic regularities in Chinese therefore develops in different ways between the two groups, which makes it interesting to examine whether there is also a difference in the visual character processing method preferred by native speakers and CFL learners.

Ke (1998) has proposed a three-stage model of character acquisition by CFL learners. First, there is a pre-component processing stage, where all the characters are learned as whole. In the second, component processing stage, learners begin to realize that each character can be broken down into parts, and are thus able to identify radicals and their functions in compound characters. The last stage is the automatic component processing stage, where the learners possess a level of orthographic awareness akin to Chinese native speakers. Lu (2002) argued that it typically takes two years for learners to enter the second, component processing stage. However, Jackson, Everson and Ke (2003) claimed that by the end of the first year of study, CFL learners would be able to decompose characters into sub-character components, but that this ability is limited mostly to radicals that can stand alone as single characters.

CFL learners seem to show a stronger inclination towards semantic radicals when asked to choose between pseudo-characters with either a real semantic radical or a real phonetic component in correct positions (Tong & Yip, 2014). Additionally, Taft and Zhu

(1999) showed that non-native Chinese participants with no background in the language became more sensitive to semantic radicals after receiving explicit instructions about the importance of the sub-character components in providing information about the character. These results speak to the robustness of the semantic radical and its contribution to character processing, even in CFL learners.

The studies mentioned above have focused mostly on beginner CFL learners. Shen and Ke (2007) compared learners at four different levels of proficiency to examine the correlation between semantic radical knowledge and radical application skills as a function of overall increased Chinese language skills. Participants ranged from first-year students who had only received one month of instruction, to fourth-year students, who had received three years of instruction. They were first tested for accuracy in radical recognition and function, and then asked to correctly re-create a character that was missing its semantic radical. In both experiments, the English equivalent of the character was provided. Results showed that while radical knowledge increased rapidly and linearly with more instruction, radical application skills only increased by approximately 3% between the first-year and second-year student participants. While these results are informative about the relationship between proficiency and radical awareness, the experimental procedures required participants to access English vocabulary for meaning, which may weaken interpretations that relate strictly to semantic radical knowledge.

The existing literature on CFL learners has focused mostly on learners' abilities to engage directly with sub-character components. The question of whether acquired vocabulary items are accessed holistically or in a decomposed manner remains unresolved.

2.6 Theories of word recognition in alphabetic languages

Current models of Chinese character processing have been derived largely from the conceptual frameworks established by word recognition data from speakers of alphabetic languages. Since there are some rough equivalencies between character processing in Chinese and word recognition in, for example, English, it is important to review existing research on the recognition and processing of words in alphabetic languages.

The dual-route cascaded (DRC) model of reading (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) has had a particularly enduring impact on the field of visual word recognition. Despite its name, the model suggests that three routes are activated during reading: the lexical semantic route, the lexical non-semantic route, and the “grapheme–phoneme conversion” (GPC) route. Similar to the models discussed above, each route in the DRC model is composed of a number of interacting layers with units that represent the smallest component of that layer, for example, individual letters at the letter input level. There are two ways in which the units of different layers interact: through inhibition, in which case the activation of one unit hinders the activation of other units; or through excitation, where activation of one unit contributes to the activation of other units. Horizontal inhibition may also occur between units at the same level. The “cascaded” assumption of the model simply claims that visual stimuli are continuously fed forward until the semantic level is reached. More specifically, it suggests that all words with shared orthographic features receive activation, and hence, the semantic properties of those words must also be activated. Competition at the semantic level would

persist until only one orthographic unit (the target word) remains. Central to the purpose of this thesis is that this horse-race style DRC model supports automatic morphological decomposition.

At a comparatively broader level, the dual-coding theory (Paivio, 1971) postulates that mental representation comprises two distinct systems: verbal and non-verbal. The former is responsible for processing linguistic information while the latter processes non-linguistic objects and events. This theory has recently been adopted to account for reading processes. According to Sadoski and Paivio (2013), reading involves three distinct dimensions of processing: representational, associative, and referential. During representational processing, visual input is initially activated (e.g., recognizing familiar words or parts of words), but semantics may or may not be involved. In contrast, associative and referential processing always activate meaningful components. Applied to Chinese, this theory assumes that characters with radicals are more likely to trigger both verbal and non-verbal activation, compared to characters without radicals. More importantly, it would be assumed that the presence of radicals always facilitates character processing and acquisition (Kuo, Sadoski, & Kim, 2014).

In terms of the visual processing of morphologically complex words, researchers have found evidence of early decomposition in French (Longtin & Meunier, 2005), English (McCormick, Rastle, and Davis, 2009), and Spanish (Beyersmann, Duñabeitia, Carreiras, Coltheart, & Castles, 2013), regardless of whether the target was a true compound word (e.g., farmer) or a pseudo-compound word (e.g., corner). This suggests that decomposition is sensitive to the surface structure of morphologically complex words, even when their constituents are not true morphemes. Such results support the

hypothesis that affixes have distinct representations in the mental lexicon that are separate from and below the lexical level (Taft & Forster, 1975).

More recently, the edge-aligned embedded word activation theory proposed by Grainger and Beyersmann (2017) has provided a new perspective for interpreting the early morphological processing of pseudo-compound words. First, the theory suggests that word recognition starts with the morphological processing of words embedded at the initial or final positions of a string of letters, i.e., the “edges”. Second, affixes play a secondary role in word decomposition, and aids decomposition only when the activation of an embedded real morpheme is hindered. Since it has been found that embedded words, such as “hat” in “chat”, are processed during visual word recognition in English, the authors argue that stems do not require a separate representation in the mental lexicon. That is, embedded stems can be activated without first removing any affixes. Affixes, on the other hand, are assumed to require a dedicated level of representation that does not provide the kind of facilitation effect that stems do.

2.7 Summary

Previous research on character processing in Chinese native speakers and CFL learners have focused mostly on semantic or phonological activation triggered by character sub-components. While studies on native speakers have examined the effect of frequency on processing, the effects of such measures on CFL learners are unknown. Although experiments have been done with Chinese-English bilinguals, using items in both Chinese and English in the same study (e.g., Wang, Perfetti, & Liu, 2005; Wang, Cheng, & Chen, 2006), no study has simultaneously tested and compared native speakers

and CFL learners. The obvious challenge in such a design is levelling the playing field between the two groups; it can be safely assumed that in every aspect of Chinese language ability, native speakers surpass CFL learners. However, it is unclear exactly how the two groups differ. This study therefore attempts to fill the gap in the literature by examining behavioural differences in just one ability – the visual processing of compound characters. To account for the expected imbalance in vocabulary knowledge between native speakers and learners, both frequency and individual character familiarity will be investigated as factors that may affect whether participants process characters holistically or in a decomposed manner.

Keeping in mind the relevant findings from previous behavioural studies, the section below presents predictions for the character processing method that CFL learners and Chinese native speakers will use under specific circumstances in the character decision task.

2.8 Predictions for character processing method

Since Gernsbacher (1984) and Connine et al. (1990) found that “experiential familiarity” was a better predictor of reaction time in English lexical decision tasks, the different range of character familiarity ratings between CFL learners and native speakers is expected to contribute to response differences between the two groups in the character decision task. In addition, a difference is expected between high- and low-frequency characters, based on Tao and Healy (2002). However, the interaction between individual character familiarity and character frequency has not been examined in past research. Familiarity is a factor with two levels: familiar and unfamiliar. Frequency is also factor with two levels: low and high. Since the interaction between these two factors is expected

to have an effect on task performance in both participant groups, predictions are given for each of the resulting four subsets of characters (Table 1).

Table 1

Predictions for character identification method employed by each participant group

	F + HF	F + LF	U + HF	U + LF
CFL learners	Holistic	Holistic or decomp.	Decomposition	Decomposition
Native speakers	Holistic	Holistic	N/A	Decomposition

Note: F = familiar; U = unfamiliar; HF = high-frequency; LF = low-frequency.

It is predicted that CFL learners will process familiar high-frequency characters holistically. Since high-frequency characters seem to be processed holistically by native speakers (Taft & Zhu, 1995), there is reason to believe that CFL learners will also holistically process characters that have been learned and reinforced through frequent exposure. The same prediction applies to native speakers.

Native speakers are also expected to process familiar low-frequency characters holistically. Familiarity may be a stronger predictor of fast automatic processing since it reflects exposure to lexical items through language production and not just their written form (Connine et al., 1994). This implies that holistic processing most likely occurs with lexical items that are known. It is unclear whether the same principle can be applied to CFL learners, since they constitute a heterogeneous group with lower language production abilities compared to native speakers. It is unclear how much familiarity can facilitate processing of a low-frequency character that is not seen very often in its print

form. Therefore, no specific prediction is made for CFL learners and familiar low-frequency characters.

It is predicted that CFL learners will decompose unfamiliar high-frequency characters. Despite the possible effects of frequency mentioned above, if certain characters have not been taught and/or learned, the likelihood of holistic processing is low. CFL learners may therefore default to using a decomposition strategy when making a judgment on an unfamiliar character. It is unlikely that native speakers will be unfamiliar with high-frequency characters used in this study. Predictions are therefore not applicable to this subset of characters for native speakers.

The decomposition method is predicted for both participant groups when processing unfamiliar low-frequency characters. Given the relative “novelty” of the stimuli in that category, participants are expected to search for clues within the characters that would help them decide whether the character is real or fake.

3 Chapter: Method

3.1 Participants

This study recruited two groups of participants: CFL learners and Chinese native speakers, who served as the control group. The sections below describe the recruitment methods that were employed, as well as a profile of second-language proficiency and exposure for each group.

3.1.1 CFL learners

A total of 28 CFL learners ranging in age from 18 to 40 years ($M = 22.1$, $SD = 4.92$) were recruited for this study. Recruitment took place through the following methods: 1) in person, from four different levels of Mandarin Chinese courses offered at Carleton University; 2) using the snowballing technique to find students who had previously taken Chinese language courses at Carleton with the same instructor but were no longer enrolled; 3) through the researcher's personal network. Compensation was given in the form of a \$10 Tim Hortons gift card.

Table 1 shows the number of participants from the highest course-level that each participant took at Carleton by the time of recruitment, as well as other details related to those courses. It excludes two CFL participants who had received formal language instruction only in China. The courses were taught by the same instructor using *Integrated Chinese Level 1, Part 1* (Liu & Yao, 2008) for first- and second-year students; *Integrated Chinese Level 1, Part 2* (Liu & Yao, 2009) for third-year students; and the first half of *Integrated Chinese Level 2, Part 1* for fourth-year students.

Table 2

Number of participants from each Chinese course-level and other related information

	Intensive 1 st year	2 nd year	3 rd year	4 th year
No. of participants	9	7	2	7
Total semesters completed	0	1	2	3
Hours of instruction	8 hrs/week	4 hrs/week	4 hrs/week	4 hrs/week

Cumulative formal instruction time among the CFL learners ranged from three to 48 months ($M = 18.4$ months). The average age of Chinese language acquisition was 18.4 years. In terms of language diversity, 18 were English monolinguals; six were bilingual in English and another alphabetic language; four were Chinese heritage speakers who reported mean proficiency scores of 2.35 in reading and 2.03 in writing Chinese.

Based on a language background questionnaire (Appendix C) adapted from Li, Zhang, Tsai, & Puls, (2014), CFL learners reported a mean score of 4.77 (1 = very poor, 7 = excellent) for their own general Chinese learning abilities, relative to their classmates. Using the same scale, learners reported mean proficiencies scores of 3.57 in listening; 3.51 in reading; 3.44 in speaking; and 3.03 in writing. In terms of overall language exposure, learners reported spending an average of 2.5 hours per day engaged in various activities in Chinese outside of class, including course-related reading and writing.

3.1.2 Chinese native speakers

25 Chinese native speakers were recruited as controls for this study. All were expected to know how to read and write Simplified Chinese. Participants ranged in age

from 18 to 54 years ($M = 24$, $SD = 7.76$) and were recruited through: 1) the Carleton University Institute of Cognitive Science SONA research participation credit system; 2) study advertisements posted around the Carleton University campus; 3) the snowballing method using social media; 3) the researcher's personal network. Participants who signed up for the study through SONA received a 1% course credit for their participation. All others received a \$10 Tim Hortons gift card as compensation.

All participants in this group indicated Chinese as their only native language. However, since most participants were recruited from a Canadian university and are assumed to be in an immersive English-language environment most of the time, it was important to capture elements of their English-language proficiency and use. Using a Chinese version of the same language background questionnaire that was administered to the CFL learners, it was found that the average amount of formal English-language instruction among participants in this group was 161.5 months (or 13.5 years), and mean age of English acquisition was 8.98 years. Despite this, mean self-rated proficiency level in English was 4.75 (out of 7), which was very similar to the mean score reported by CFL learners for Chinese. Compared to CFL learners with Chinese, the native speakers group reported slightly higher scores in their English-language abilities: 5.23 in listening, 4.73 in speaking, 4.41 in reading, and 4.64 in writing (out of 7). Participants in the native speakers group spend on average 2.63 hours a day speaking English, but 7.01 hours reading or writing in English.

3.2 Materials

This study consisted of three components: an on-line character decision task, an

off-line computerized character familiarity rating test, and an off-line written language background questionnaire. The sections below provide details about the experimental stimuli and design of each of these components.

3.2.1 Stimuli selection and design

3.2.1.1 Experimental items

Due to the high number of combinable orthographic features that are possible in Chinese (discussed earlier in the Introduction), the following criteria were used to constrain the selection of all experimental items: 1) left-right spatial configuration; 2) left-side semantic radical; 2) high semantic transparency, i.e., strong conceptual relatedness between the semantic radical and the whole character; and 3) no right-side component with the exact same pronunciation as the whole character, in order to avoid phonetic component effects. Spatial configuration and position of the semantic radical were considered more important factors to control for. Three of the 96 characters were lower in semantic transparency compared to all others but were still included in the final set of stimuli (see “child”, “wine”, and “action” in Appendix A).

Ninety-six semantic-phonetic compound Simplified Chinese characters were selected. In order to reflect the range of learners from four different course-levels, two-thirds of the characters were selected from the *Integrated Chinese* textbooks (Yao et al., 2009) to increase the possibility that CFL learners would be exposed to a balanced set of familiar and unfamiliar characters. Twenty-five characters were selected from *Integrated Chinese Level 1, Part 1*. Thirty-five characters were selected from *Integrated Chinese Level 1, Part 1*. Four characters were selected from *Integrated Chinese Level 2, Part 1*.

The uneven distribution was a result of the intentional focus on semantic-phonetic compound characters only. Additionally, using the selection criteria listed above, an equal number of characters from each textbook level could not be achieved without severely reducing the number of experimental items.

Aside from characters taken from the CFL learners' textbooks, the remaining 32 items were selected from the portion of low-frequency characters in the SUBTLEX-CH database (Cai & Brysbaert, 2010). This was done to ensure, to the best of the researcher's ability, that the native speakers group would also be exposed to characters that were likely to be unfamiliar, which could not be achieved using the textbook items alone.

The final 96 experimental character items had the following features: 1) a semantic radical on the left side that was shared by two characters in the set, resulting in only 48 unique radicals; 2) no repeated right-side component; 3) right-side components that generally did not have strong phonetic cues. A caveat to this last feature regards a general principle of low-frequency compound characters. Low-frequency characters tend to have more reliable phonetic components, whereas the opposites hold true for high-frequency characters (Shu, Chen, Anderson, Wu, & Xuan, 2003). Therefore, it was inevitable that some of the right-side components within the lowest frequency characters used in this study would have stronger phonetic cues. This was not fully controlled for, given the other constraints in the stimuli selection process that were considered more relevant to the research questions.

Following Ding et al. (2004), who also used radicals and masked priming in their experiments, characters lower than 80 per million words were defined as low-frequency; characters higher than 80 per million words were defined as high-frequency. This study

uses \log_{10} frequency as a standardized measure for all real characters to remove problems associated with frequency per million counts (e.g., corpus size and characters with frequency counts below one). Based on the SUBTLEX-CH database (Cai & Brysbaert, 2010), 80 per million words is equivalent to a \log_{10} frequency of 3.57. In this study, high frequency characters were thus defined as being above the \log_{10} frequency of 3.57, and low frequency characters were defined as those below. Applying this classification to the characters used in this study, 46 characters were high-frequency and 50 characters were low-frequency.

All characters were first listed in Microsoft Word in 72-point KaiTi font, which is the same font used in the learners' textbooks. Using the screenshot function, a JPEG image of each character was produced and trimmed down to 200 x 200 pixels, with the character centered. These images were used as the target stimuli in the character decision task, which is discussed in greater detail below.

In order to prevent pattern matching during the task, all characters were again listed in Microsoft Word, but in 60-point bold HeiTi font. The same method described above was used; the 48 unique radicals were then cropped from half of the characters and centered on a 200 x 200 pixels white canvas. All editing was done using the computer application Pixen (version 3.5.9b2; 2016), an open-source image editor.

3.2.1.2 *Filler items*

Ninety-six fake radicals and characters were also included in this study as filler primes and targets. Using the same image editor mentioned above, filler items, i.e., fake characters, were created using replicated images of the real characters. The semantic

radical in the image of a real character was rotated 180 degrees. If it still looked the same as the real radical, the positions of some strokes were shifted to ensure dissimilarity between the fake and real radicals. As for the right-side component, an average of three strokes was removed. This process was repeated using all 96 images of the real characters. Fake radicals were also created by rotating images of the real radicals 180 degrees. Table 3 shows examples of both the experimental and filler stimuli used in the character decision task. See Appendix A for a complete list of the stimuli.

Table 3

Examples of character decision task stimuli

	Matched condition			Unmatched condition		
	Prime	Target	N	Prime	Target	N
Target character	纟	练	48	纟	校	48
Fake character	丯	纟	48	丯	服	48

Fake characters were used as fillers. Filler items in character decision tasks from other studies were typically pseudo-characters created from recombined real sub-character components (e.g., in Zhou & Marslen-Wilson, 2000; Myers, 2016). In contrast, the filler items used in this study did not contain any real stroke combinations. Since CFL learners may not be as sensitive to illegal stroke combinations as native speakers are, it was considered undesirable for the filler items to contain meaningful sub-character components that learners could potentially use to incorrectly judge filler items as real characters. This would be especially problematic since decomposition would become

obligatory in order for CFL learners to make a judgment on whether or not the character was real. Though the meaningful components of real characters were manipulated to create the filler items, the square shape of the Chinese character was generally retained.

3.3 Experimental design

A 2 (CFL learners vs. native speakers) x 2 (real vs. fake character) x 2 (matched vs. unmatched condition) mixed design was used for the study. In the character decision task, stimuli were counterbalanced between two lists. In each list, a semantic radical prime was presented twice, and a target character was shown only once. If the character was in the matched condition in List 1, it would be in the unmatched condition in List 2. To control for pattern matching, the prime differed from the target in terms of font type (HeiTi vs. KaiTi), font size (60-point vs. 72-point) and typographic emphasis (bold vs. normal) in order to avoid visual matching. To summarize, participants were presented with the following stimuli: a) 48 semantic radical primes, once in a matched condition and once in an unmatched condition, resulting in 96 real target characters; b) 48 fake radical primes, once in a matched condition and once in an unmatched condition, resulting in 96 filler characters. Each trial in the character decision task contained one prime-target pair, resulting in a total of 192 trials.

3.3.1 Character decision task

A character decision task (CDT) was used to examine the main research questions in this study since it requires fast processing and judgment. Another common way to measure Chinese character processing is the naming task, which involves recording the

speed and accuracy of participants' pronunciation of characters (e.g., Perfetti & Liu, 2005; Shen & Ke, 2007). However, it is often used to examine the function of phonetic components and was therefore not chosen for this study.

Masked priming was applied to further prevent participants from developing an awareness of patterns in the experimental design. The objective of the CDT was to measure the extent to which CFL learners naturally decompose a character into constituent parts in order to help them process the whole character. The masked priming paradigm was therefore considered suitable for the purpose of the experiment, since primes presented at longer intervals would change how CFL learners interact with target characters.

The CDT was designed in PsychoPy2 (version 1.85.4) (Peirce, 2009) using built-in components. Presentation times for all stimuli in the task were based on the number of frames, given a screen refresh rate of 60 Hz, that would be equivalent to the desired millisecond display time. Each frame was equivalent to 16.667 milliseconds.

Each session in the CDT began with ten practice trials, followed by 192 experimental trials. In each trial, a black fixation cross was presented against a white background on the computer screen for 30 frames (500 ms), followed by a forward mask consisting of two “at” symbols (@@), also for 30 frames (500 ms). The semantic radical prime was presented for three frames (50 ms), in keeping with similar previous experiments using masked priming of morphemes in Chinese as well as alphabetic languages (e.g., Perfetti & Zhang, 1991; Heathcote, Nation, Castles, & Beyersmann, 2017). A backward pattern mask (@@) then replaced the prime and was displayed for two frames (33 ms). Afterwards, the target appeared and remained on the screen for 180

frames (3000 ms). A blank screen was presented as an inter-trial interval preceding the next trial and was displayed for 30 frames (500 ms) Appendix B provides an illustration of the task presentation timeline.

3.3.2 Character familiarity test and language background questionnaire

The purpose of the character familiarity test (Appendix C) was to measure the degree to which participants knew each of the 96 real characters that appeared in the character decision task. The test was administered using Microsoft Excel. The following scoring scale was given to the participants: “1” for characters that participants did not know at all; “2” for characters that participants had seen multiple times but do not remember the meaning of; and “3” for characters participants know and can use in a sentence or phrase.

A language background questionnaire was adapted from the Language History Questionnaire 2.0 from Ping Li’s Brain, Language, and Computation (BLC) Lab at Pennsylvania State University. The purpose was to create a profile of demographic and language exposure. A Simplified Chinese version of the questionnaire was also made available by the authors. It was also adapted and provided to the Chinese native speakers in this study.

The original questionnaire comprised of 22 questions in four modules pertaining to users’ linguistic background, proficiency in first or other languages; context and habits of language use; and language dominance. The authors of the questionnaire invite researchers to select questions and modules that are most prevalent to the research questions at hand. Questions selected for the adapted version of the questionnaire used in

this study related mostly to linguistic background; self-reported proficiency in either Chinese for CFL learners or English for Chinese native speakers; and daily language use and exposure. Considering the focus of the study, questions related to language dominance were not included. The adapted English-version questionnaire is provided in Appendix C.

3.4 Procedure

The CDT task was completed in PsychoPy2 using a 13-inch MacBook Air laptop. Participants were tested in an office or a quiet room. Participants were first provided with a consent form describing the low-risk nature of the experiment and their right to withdraw at any time. After providing informed consent, participants read a set of instructions on the screen for how to perform the CDT. Participants were asked to stay focused on the + sign in the center of the screen throughout the entire experiment. They were asked to press P if the target character they saw was real, whether they knew the meaning of the character or not. They were asked to press Q if they believed that the target character was not a real Chinese character. Participants were asked to make their judgments as fast and as accurately as possible. After a practice round, participants were given a chance to ask the researcher for clarifications regarding the task before the actual testing began. The task took on average 20 minutes to complete.

Following the CDT, participants were given a list of all 96 real characters on the same laptop and were asked to rate their familiarity using the scoring scale described above. Afterwards, participants were asked to fill out a printed version of the language background questionnaire, adapted from Li, Zhang, Tsai, and Puls (2014).

After completing all three components of the study, the research explained the full purpose of the experiment to each participant. When asked whether or not a radical prime was visible during the CDT, participants reported only seeing flashing “at” signs (the forward and backward masks).

3.5 Statistical analyses

All statistical analyses were completed in R (R Core Team, 2012) with RStudio (version 1.1.442; 2015). Using the R package *lme4* (Bates, Maechler, & Bolker, 2012), a series of mixed effects models were fitted with the following fixed effects: group, condition, character frequency, and character familiarity. The random effects structure included by-participant, by-condition, and by-item variation. Chi-square likelihood ratio tests were to determine goodness-of-fit for selecting the most suitable models. For examining the dependent variable of reaction time, linear mixed effects models (LMMs) were used. Analyses were run only on correct responses to real characters. For examining accuracy rates, generalized linear mixed effects models (GLMMs) were used with a binomial variance assumption. Analyses were run only on responses to real characters. Appendix E contains the R script used to run all mixed effects models.

4 Chapter: Results

4.1 Data cleansing

Data points from one participant in the CFL learners group and three participants in the native speakers group were removed due to data recording errors caused by software malfunction during the character decision task. In addition, data was missing from 28 trials, which were also subsequently removed. Finally, reaction times (RTs) were logarithmically transformed to correct skewness of distribution, and outliers 2.5 standard deviations away from the mean were trimmed (2.92% of data).

The scale used in the character familiarity test was reconsidered for the data analysis. Given the instructions for the test, participants should have scored characters that they were not familiar with as either 1 (not familiar) or 2 (somewhat familiar). A score of 2 was used to define characters that participants had seen before but did not know how to use. However, it was considered more important to isolate semantic knowledge in order to make judgments about the effect of condition, or more precisely, effect of the semantic radical prime. Scores of 1 and 2 were therefore combined.

4.1 Character familiarity

Character familiarity was an important part of the study design since similar measures used in previous related studies did not explicitly test participants' knowledge of individual items, but rather focused on overall language or vocabulary proficiency (e.g., Shen, 2007; Andrews & Lo, 2013).

Table 4 below shows that, as expected, CFL learners and native speakers had noticeably different mean character familiarity scores. The scores were computed from

the individual scores assigned by each participant to each character, and then averaged across each participant group.

Table 4

Mean character familiarity score by participant group (max = 1.00)

	Mean familiarity	SD	N	SE
CFL	0.37	0.48	2278	0.01
NS	0.83	0.38	1962	0.01

Since no subjective familiarity scores were collected and used for the selection of character items, a disparity in the ratio of familiar to unfamiliar characters between the two participant groups was expected, given the assumed level of proficiency in CFL learners and assumed level of vocabulary knowledge in native speakers. Figure 4 below compares character familiarity between learners and native speakers. The knowledge imbalance found between the two groups justifies the use of character frequency to account for unfamiliar high-frequency characters in the CFL learners group, and familiar low-frequency characters in the native speakers group.

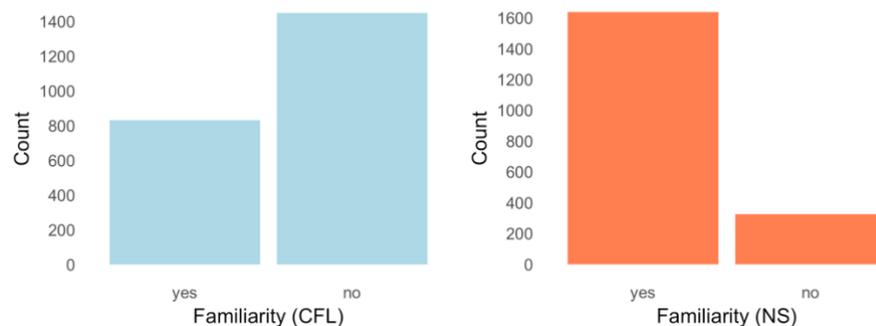


Figure 4. Number of familiar vs. unfamiliar characters in CFL learners (left) compared to native speakers (right)

4.2 Overview of reaction time results

Overall, average RTs from the character decision task show that CFL learners had a wider distribution of responses, which is likely due to the heterogeneity of this group compared to the native speakers. Additionally, median RTs differed between conditions for CFL learners, but were almost exactly the same for native speakers (Fig. 5).

Averaging across all variables, CFL learners were slower than native speakers. Mean RT for CFL learners was 727 milliseconds (ms), while mean RT for NS was 623 ms. For real characters in both groups, mean RTs were fastest in the matched condition and slowest in the unmatched conditions.

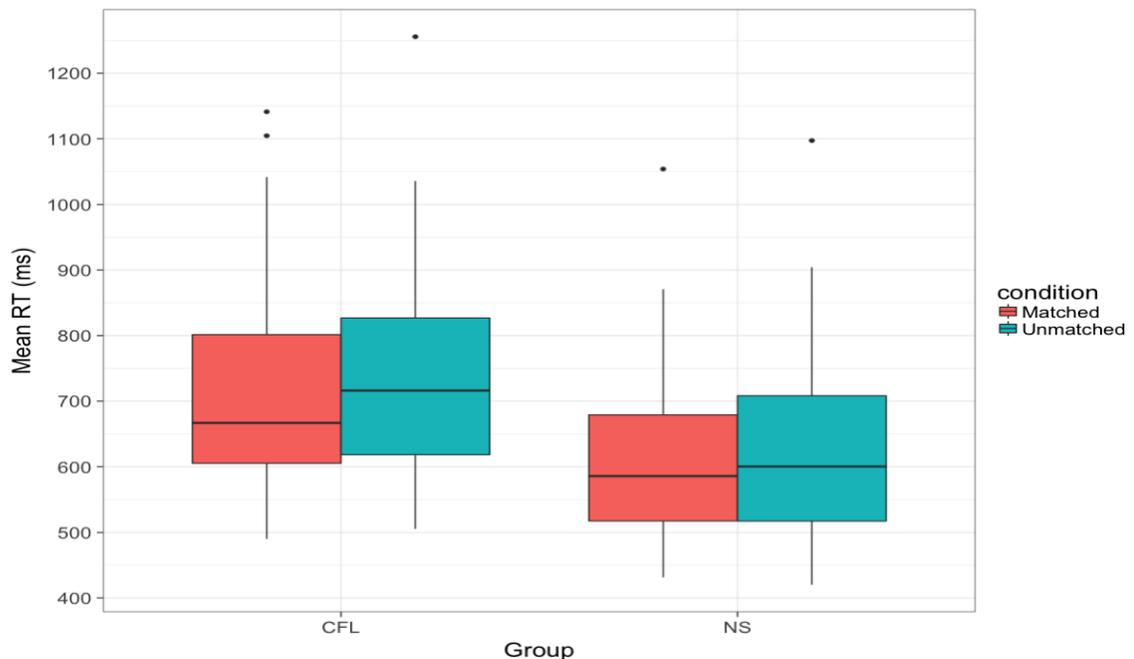


Figure 5. RT distribution of by group and condition

Figure 6 further shows differences between CFL learners and native speakers. Condition appeared to have a greater effect on unfamiliar characters for native speakers than for CFL learners. Similarly, condition produced a slightly bigger difference in mean RT for familiar characters in CFL learners than for native speakers.

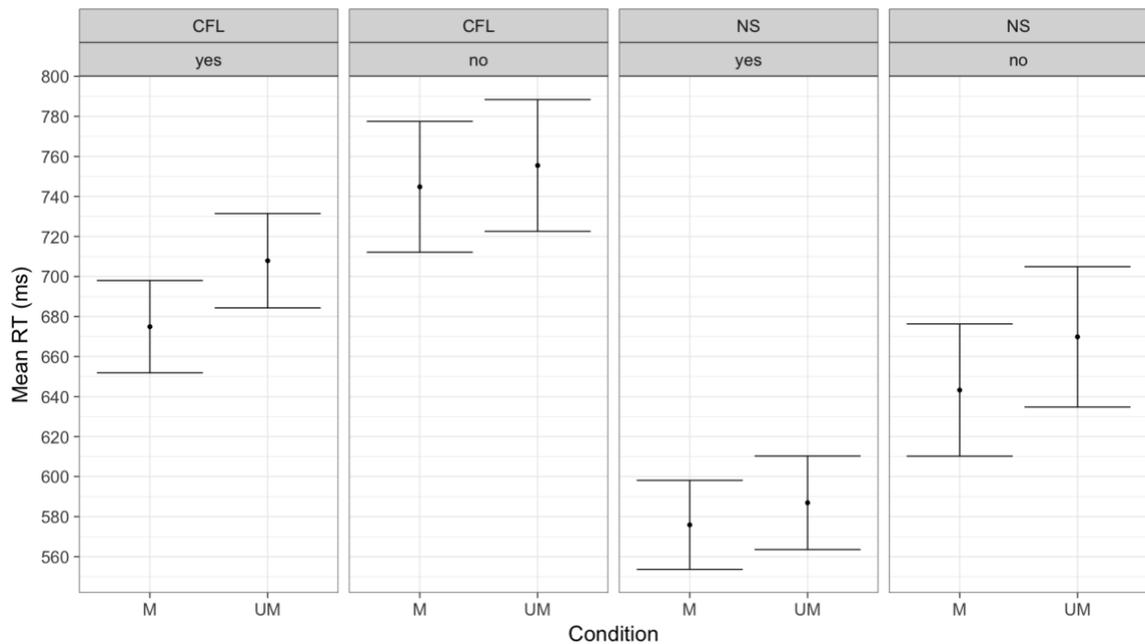


Figure 6. Mean RT by condition, group, and familiarity, with standard errors. (Note: M = matched; UM = unmatched; yes = familiar; no = not familiar.)

Differences between the two groups are further revealed by observing responses for familiar versus unfamiliar characters. In CFL learners, mean RT was 693 ms for familiar characters, and 761 ms for unfamiliar characters (-68 ms). In the native speakers group, mean RT was 583 ms for familiar characters and 663 ms for unfamiliar characters (-80 ms). While mean RTs for familiar versus unfamiliar characters followed a similar pattern in both groups, regardless of condition, the effect of familiarity was slightly larger for the native speakers group.

In CFL learners, mean RT was 699 ms for high-frequency characters and 748 ms for low-frequency characters. In the native speakers group, mean RT was 612 ms for high-frequency characters, and 600 for low-frequency characters. Although there is no evidence that vocabulary instruction in the Chinese language courses from which participants were sampled follows any kind of character frequency distribution, it seems that frequency has an effect on CFL learners' RT responses, especially compared to native speakers.

Table 5

Mean RT (ms) by group, target type, familiarity, frequency and condition

	Target	Fam.	Freq.	Condition	Mean	SD	Med	Min	Max	SE			
CFL	Real	Yes	High	M	674	120.	638	502	935	23.10			
				UM	689	109.27	688	488	948	21.03			
			Low	M	702	161.86	679	477	1140	35.32			
				UM	769	179.72	758	506	1070	37.47			
		No	High	M	706	180.64	659	477	1154	35.43			
				UM	752	183.13	710	500	1116	36.63			
			Low	M	780	194.92	741	495	1180	37.51			
				UM	763	185.43	732	515	1331	35.69			
				Filler	NA	NA	M	706	150.81	658	502	1093	29.02
							UM	705	144.38	681	516	1030	27.79
NS	Real	Yes	High	M	563	99.30	560	432	825	21.17			
				UM	573	106.50	576	418	861	22.70			

			M	594	114.71	593	430	918	24.46
		Low	UM	610	119.77	612	454	957	25.53
	No	Low	M	648	153.60	625	452	1054	32.75
			UM	676	168.06	673	412	1097	35.83
Filler	NA	NA	M	587	82.58	576	472	755	17.61
			UM	593	83.17	587	476	747	17.73

Note: M = matched; UM = unmatched

Based on Table 5, RT differences between the two conditions for filler items were marginal in both groups, suggesting that filler items were not treated the same way as real characters. However, the marginal difference (4 ms) in mean RT between the two conditions for unfamiliar characters in the CFL learners was also notable, suggesting that the priming effect might be the same for unfamiliar characters and fake characters for this group.

4.3 Overview of accuracy rate results

In terms of the percentage of correct responses elicited from the character decision task, the native speakers group performed most consistently in the matched condition and had the lowest percentage of errors. The distribution of correct responses shows that CFL learners had the highest number of errors overall. Additionally, the difference in the distribution of correct responses was greater between the matched conditions in both groups than in the unmatched conditions (Fig. 7).

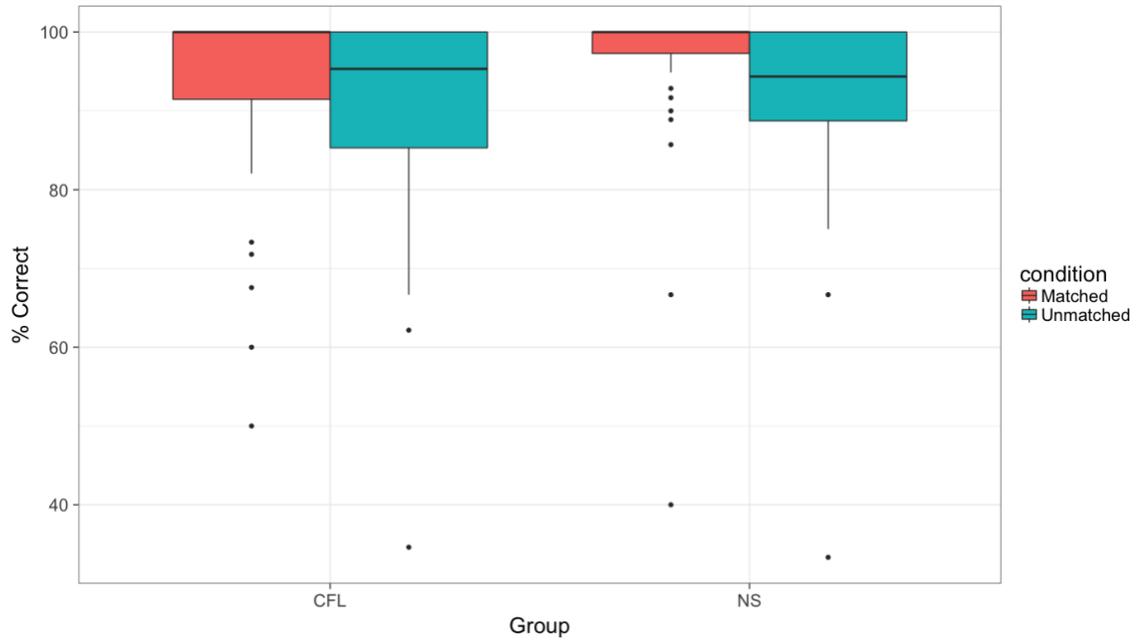


Figure 7. Accuracy rate distribution by group and condition

Familiarity seemed to have an effect on the accuracy of responses particularly in the native speakers group (Fig. 8). Unfamiliar characters resulted in greater individual variance, regardless of condition. Differences in mean accuracy rates between the two conditions for familiar characters was also slightly greater in native speakers compared to CFL learners.

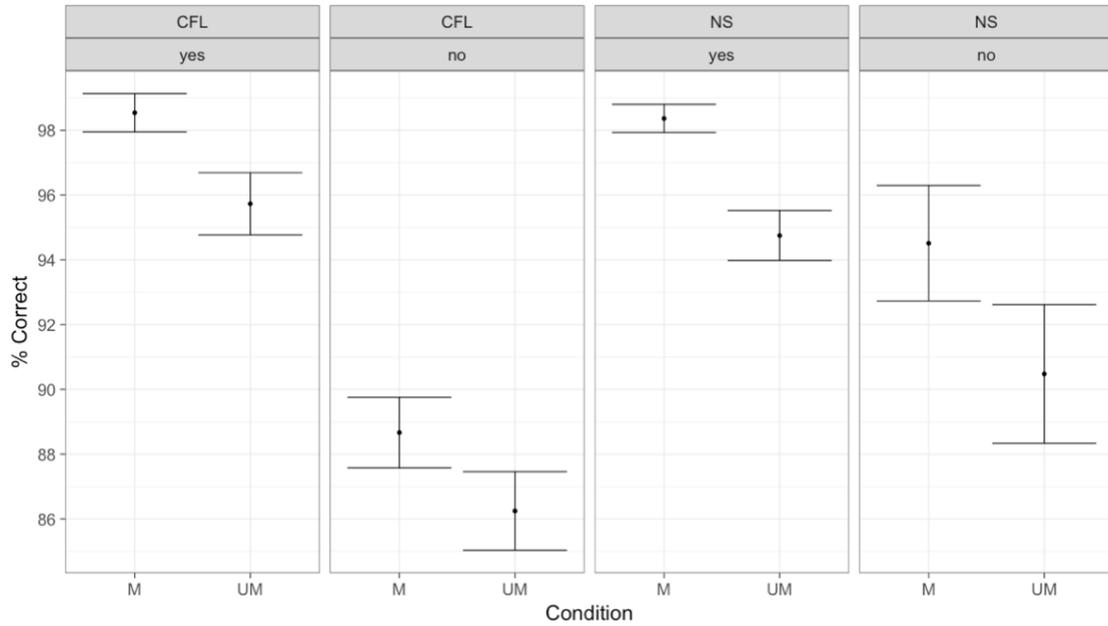


Figure 8. Mean accuracy rate by condition, group, and familiarity, with standard errors.

(M = matched; UM = unmatched; yes = familiar; no = not familiar).

Performance on high-frequency characters paralleled the patterns that were found in the RT responses. In CFL learners, accuracy rate for high-frequency characters was 94%, but only 89.7% for low-frequency characters. On the other hand, for the native speakers group, accuracy rate was 96.9% for high-frequency characters and 94.3% for low-frequency characters. As with RT responses, accuracy rates were highest for familiar high-frequency characters in both groups (Table 6).

While accuracy rates were similar between CFL learners and native speakers for real characters, accuracy rates of filler items were lower than that of unfamiliar characters in the CFL learners group. This suggests that perhaps the distinction between unfamiliar characters and fake characters may not be as clear for CFL learners as it is for native speakers, particularly when performing a task that requires fast judgment.

Table 6

Mean accuracy (%) by group, target type, familiarity, and condition

Group	Target	Fam.	Freq.	Condition	Mean	SD	SE		
CFL	Real	Yes	High	M	99.1	9.5	.5		
				UM	97.9	14.5	.8		
			Low	M	96.3	18.9	2.1		
				UM	89.8	30.4	2.8		
	Real	No	High	M	92.4	26.5	1.6		
				UM	86.6	34.2	2.1		
			Low	M	86.7	34.0	1.4		
				UM	86.1	34.6	1.5		
	Filler	NA	NA	M	80.6	39.6	1.1		
				UM	81.7	38.7	1.1		
NS	Real	Yes	High	M	98.8	10.9	.5		
				UM	95.0	21.9	1.0		
			Low	M	97.8	14.8	.8		
				UM	94.4	23.0	1.2		
	Real	No	Low	M	94.5	22.9	1.8		
				UM	90.5	29.4	2.1		
			Filler	NA	NA	M	83.0	37.6	1.2
						UM	84.1	36.6	1.2

Note: M = matched; UM = unmatched

4.4 Mixed effects models

LMMs were used to examine reaction time data, and GLMMs were used for accuracy rate analyses. The main advantage of using mixed effects models is that it revolves the issue of non-independence in a data set, especially when multiple data points are collected from each participant. The random effects structure allows model estimates to include variation due to idiosyncrasies in participants and experimental items. Since the estimates reflect how participants individually reacted to certain manipulations, the results allow for better generalization to the population.

P-values were obtained by chi-square likelihood ratio tests comparing a null model without the fixed effects of interest to a model with the effects of interest. Both fixed and random effects were added to the model using a backward stepwise model selection procedure with a regular maximum likelihood parameter estimation. If a significant interaction was found, p-values are reported for the interaction effect since significance of the main effects are uninterpretable due to the lack of independence among them (see e.g., Zar, 1999). For both the LMMs and GLMMs, significant interactions were followed by planned comparisons to investigate factor-level differences.

The random effects structure for all models used in the data analysis included both random intercepts and random slopes. Random intercepts inform the model of multiple responses per participant. Random slopes inform the model that each participant should be assigned its own adjusted slope for the matched and unmatched conditions, in order to account for the fact that each participant might respond differently between the two conditions.

The independent variables of group, condition, character frequency, and character familiarity were centered in order to reduce collinearity, which can be particularly problematic for interpreting interaction results. Centering these variables also helps adjust the model to reflect the grand mean of the independent variables, rather than the value of one baseline independent variable. The intercept of the model can then be understood as the expected value of the dependent variable (reaction time or accuracy rate) when the independent variables are set to their means as opposed to zero.

The sections below first present reaction time analyses, before moving on to accuracy rate analyses. For each set of data, fixed main effects are reported first, followed by results from interaction models. Lastly, results from planned comparisons between the CFL learners and native speakers will be presented.

4.4.1 Reaction time data analysis

4.4.1.1 Fixed effects

The random structure for the first model included by-participant intercepts and by-item (character) intercepts. Likelihood ratio tests showed that random slopes for condition significantly improved the random intercepts-only model ($\chi^2(2) = 8.45, p = .015$). Therefore, by-condition random slopes were added to the by-participant random intercepts.

A model with the main effect of group and random effects was compared to a null model with only random effects. There was a significant difference in reaction time

between the CFL learners and native speakers ($\chi^2(1) = 20.27, p < .001$). The native speakers were faster by about 135 ms \pm 38.52 (standard errors).

A model with the fixed effect of condition was also significant when compared to the null model ($\chi^2(1) = 9.98, p = .002$). Responses in the unmatched condition were estimated to be about 13 ms \pm 5.91 (standard errors) slower than in the matched condition.

Familiarity was found to be a significant main effect in a model that was compared again to the null model ($\chi^2(1) = 50.06, p < .001$). Unfamiliar characters elicited responses that were slower by about 47 ms \pm 6.56 (standard errors), compared to familiar characters.

Log₁₀ character frequency was the final factor that was tested. A significant effect of frequency was found in the model that was compared to the null model with random effects only ($\chi^2(1) = 44.13, p < .001$). Reaction times for low-frequency characters were about 52 ms \pm 7.61 (standard errors) slower compared to high-frequency characters. In sum, significance was found for each of the fixed effects of group, condition, familiarity and frequency.

4.4.1.2 Interaction terms

A four-way interaction model was used to investigate the relationship between group, condition, familiarity and frequency. The four-way interaction model significantly improved models with only three-way or two-way interactions, and more importantly, also the model without interaction terms ($\chi^2(1) = 86.37, p < .001$).

4.4.1.3 Random effects

Before presenting results from the planned comparisons for each participant group, the random structure of the interaction model will be briefly examined in this section. Figure 10 below is a plot of the random intercepts for participants, and random slopes in the matched condition (left) and unmatched condition (right), which are ordered by increasing value. The figure suggests that for both CFL learners and native speakers, the unmatched condition results in greater error around each estimate, while the matched condition results in a wider distribution of variance overall. This may be caused by the heterogeneity of the CFL learners group; specifically, there may be wider differences in baseline reaction times among the learners than among native speakers.

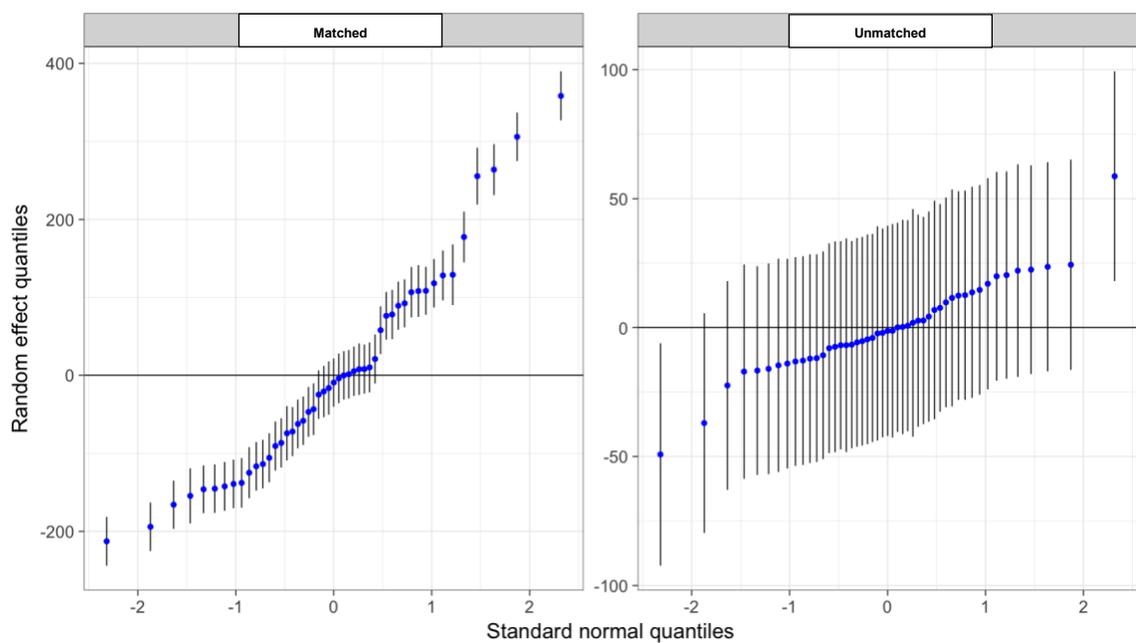


Figure 10. Participant-level variance in matched and unmatched conditions.

4.4.1.4 *Planned comparisons*

One of the main questions addressed in this thesis is whether or not CFL learners behave differently than native speakers in the visual recognition of compound Chinese

characters. Given this interest in potential differences between the two groups, planned comparisons were used to further investigate the effect of the interaction between condition, familiarity and frequency on reaction time in each participant group separately.

4.4.1.4.1 CFL learners group

A three-way interaction model with condition, familiarity and frequency was first fitted to the subset of data on CFL learners only. This model was significantly better than a model with the three variables as fixed effects only ($\chi^2(4) = 45.25, p < .001$). Model criticism was performed by first plotting the residuals against the fitted values. Given the skewness of the distribution, data points greater than or less than 2.5 standard deviations from the mean were removed. The model was then refitted with 60 data points removed (2.63% of total).

Conditional comparisons were then calculated using the *emmeans* package (Lenth, 2018) in R (Table 8). Figure 11 shows the direction of interactions between each level of familiarity and frequency, and the estimated reaction time values.

Table 8

Simple contrasts for condition in CFL reaction time model

	Estimate	SE	<i>df</i>	<i>t</i>	<i>p</i>
Contrast	Familiar, high-frequency				
M - UM	-7.37	12.37	133.28	-.60	.55
	Familiar, low-frequency				

M - UM	-61.33	22.51	731.41	-2.73	.007
Unfamiliar, high-frequency					
M - UM	-56.66	13.98	197.20	-4.05	<.001
Unfamiliar, low-frequency					
M - UM	5.02	10.66	89.03	.47	.64

Note: M = matched; UM = unmatched

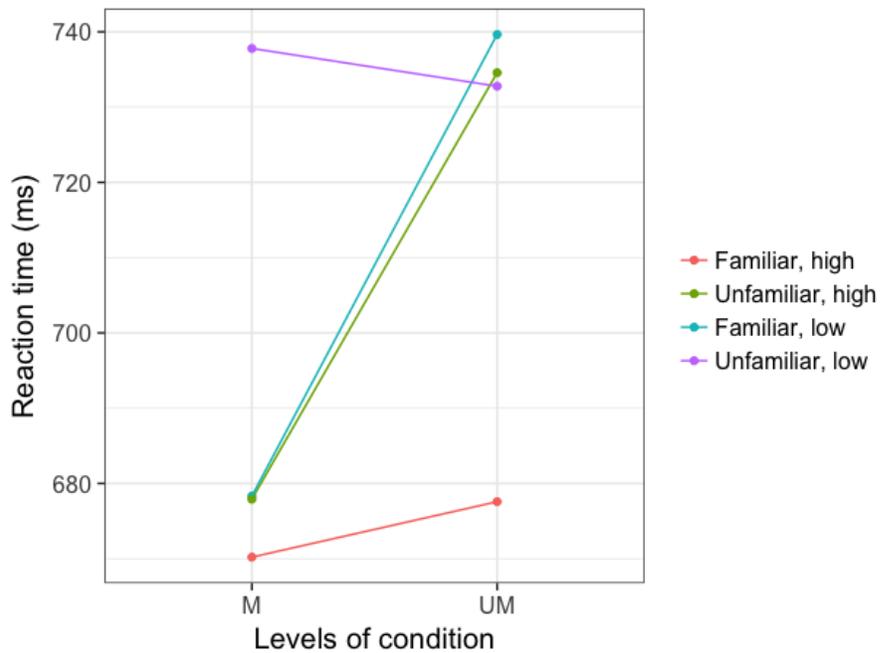


Figure 11. Interaction between familiarity and frequency on estimated RT in both conditions for CFL learners

4.4.1.4.2 Native speakers group

A model with the same interaction terms of condition, familiarity and frequency

was fitted for the native speakers group data and compared to a model with only fixed effects. The interaction model was significantly better ($\chi^2(4) = 28.09, p < .001$).

The same procedures performed for model criticism with the CFL learners group was applied to the interaction model for the native speakers group. After assessing the distribution of the residuals, the model was refitted with 55 data points (2.8% of total) removed.

Table 9 presents results of simple contrasts between matched and unmatched conditions for each combination of the two levels in familiarity and frequency. No significant contrasts were found. Figure 12 confirms these findings; estimated reaction time values in this group were not affected by interactions between condition, familiarity and frequency. No estimates are shown for the unfamiliar high-frequency combination since all native speakers in this study reported to know all high-frequency characters.

Table 9

Simple contrasts for condition in NS reaction time model

	Estimate	SE	<i>df</i>	<i>t</i>	<i>p</i>
Contrast	Familiar, high-frequency				
M - UM	-8.87	7.16	59.82	-1.24	0.22
	Familiar, low-frequency				
M - UM	-13.83	8.34	107.30	-1.66	0.10
	Unfamiliar, low-frequency				
M - UM	-14.45	11.54	309.56	-1.25	0.21

Note: M = matched; UM = unmatched

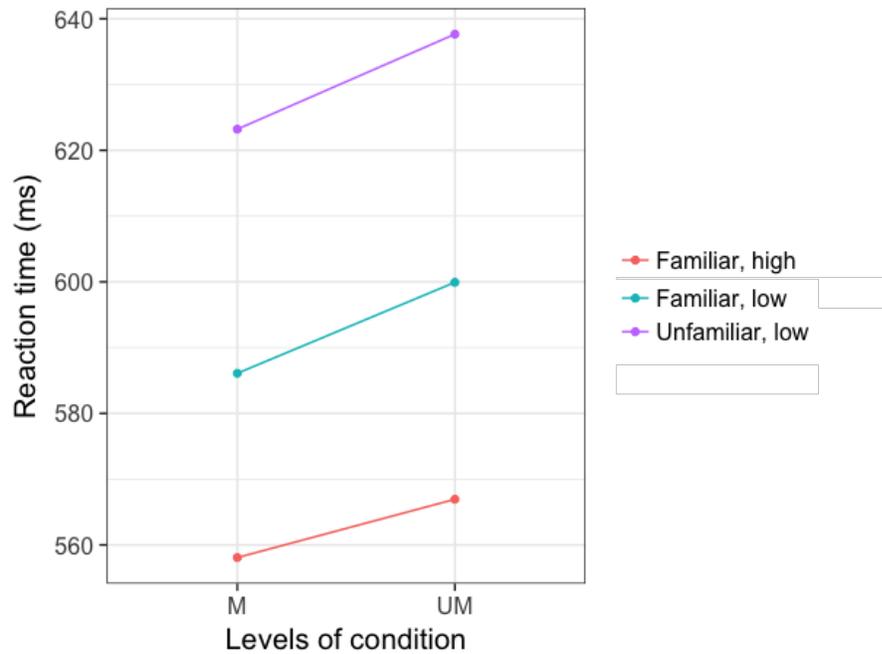


Figure 12. Effect of familiarity and frequency on estimated reaction time in both conditions for native speakers

4.4.2 Accuracy data analysis

As previously mentioned, accuracy rates were analysed using GLMMs, which are considered optimal for categorical data (Jaeger, 2008). Inclusion of by-participant and by-item random intercepts and random slopes were tested for goodness-of-fit using chi-square likelihood ratio tests. While the model for reaction time data included by-condition random slopes for by-participant intercepts, the most suitable random structure for the accuracy data model included by-participant random intercepts without random slopes, and by-condition random slopes for by-item random intercepts. This was a significant improvement to the model with random intercepts only ($\chi^2(4) = 9.15, p = .010$).

4.4.2.1 Fixed effects

A significant main effect of group was found when compared to the null model without any fixed effects ($\chi^2(1) = 4.96, p = .026$). Native speakers were more accurate in their responses during the character decision task ($z = 2.28, p = .023$), compared to CFL learners.

A significant fixed effect of condition was also found when compared to the null model ($\chi^2(1) = 10.15, p = .001$). Accuracy rates were lower in the unmatched condition ($z = -3.12, p = .002$) for both participant groups.

A model with familiarity as the fixed effect was significantly better than the null model ($\chi^2(1) = 55.11, p < .001$). Unfamiliar characters resulted in lower accuracy rates ($z = -8.16, p < .001$).

The effect of frequency was significant in the model, as compared to the null model ($\chi^2(1) = 18.90, p < .001$). Low-frequency characters produced more errors in participants' response compared to high-frequency characters ($z = -4.84, p < .001$). In sum, similar to the RT analysis, significant fixed effects of group, condition, familiarity and frequency were found for the accuracy data.

4.4.2.2 Interaction terms

A model with a four-way interaction between group, condition, familiarity and frequency was a significantly better fit of the data compared to the model with only individual fixed effects ($\chi^2(1) = 33.66, p < .001$). More specifically, a significant interaction was found between the native speakers group and frequency ($z = 2.64, p = .008$), as well as between frequency and familiarity ($z = 2.38, p = .017$).

4.4.2.3 *Planned comparisons*

4.4.2.3.1 CFL learners group

A model with a three-way interaction between condition, familiarity and frequency was compared to a two-way interaction model with frequency and familiarity, which was then compared to a model with fixed effects only. The two-way interaction model was considered the best fit for the data ($\chi^2(1) = 7.88, p = .005$). Specifically, the interaction between frequency and familiarity was significant ($z = 2.88, p = .004$). However, no significant effect of condition was found. Figure 13 shows that frequency had a greater effect on familiar characters compared to unfamiliar characters. Possible explanations for this pattern are provided in the discussion section.

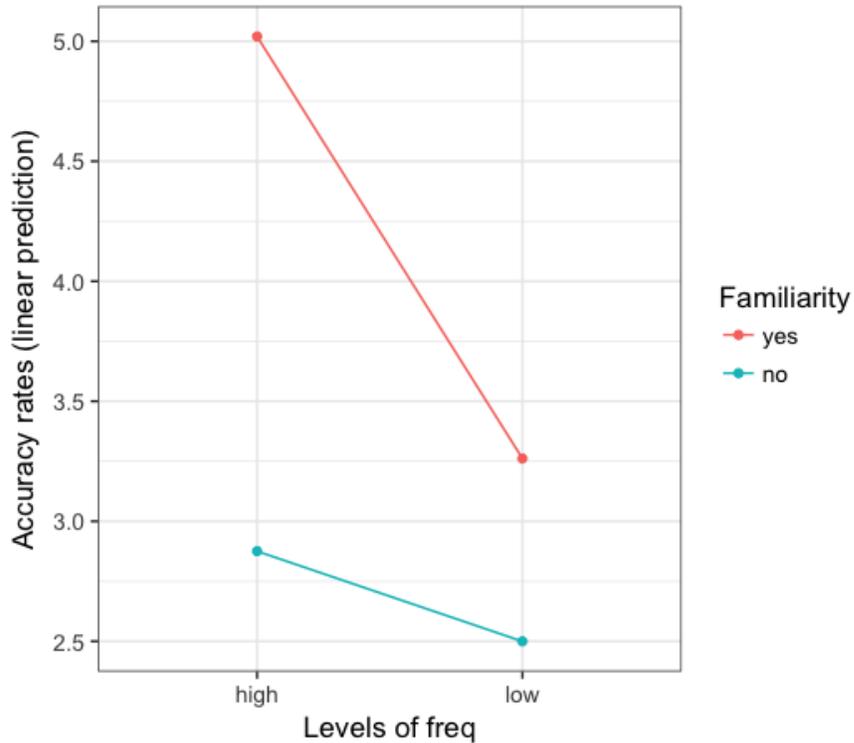


Figure 13. Interaction between frequency and familiarity on predicted accuracy rates in CFL learners

4.4.2.3.2 Native speakers group

For the native speakers, the model with fixed effects only was significantly better than the three-way interaction model with condition, familiarity, and frequency ($\chi^2(1) = 2.91, p = .09$). In terms of fixed effects, condition significantly affect accuracy rates ($z = -3.05, p = .002$), as did familiarity ($z = -2.54, p = .01$). There was no significant effect of frequency. Figure 14 provides an odds ratio plot (95% confidence interval) of the relationship between the fixed effects of condition, familiarity, and frequency on accuracy rates in native speakers. The predicted values are based on the intercept estimates of each fixed effect. The vertical grey line indicates no relationship between the

fixed effects and accuracy responses. The odds ratio for each fixed effect represents the chances that that effect will affect accuracy rate.

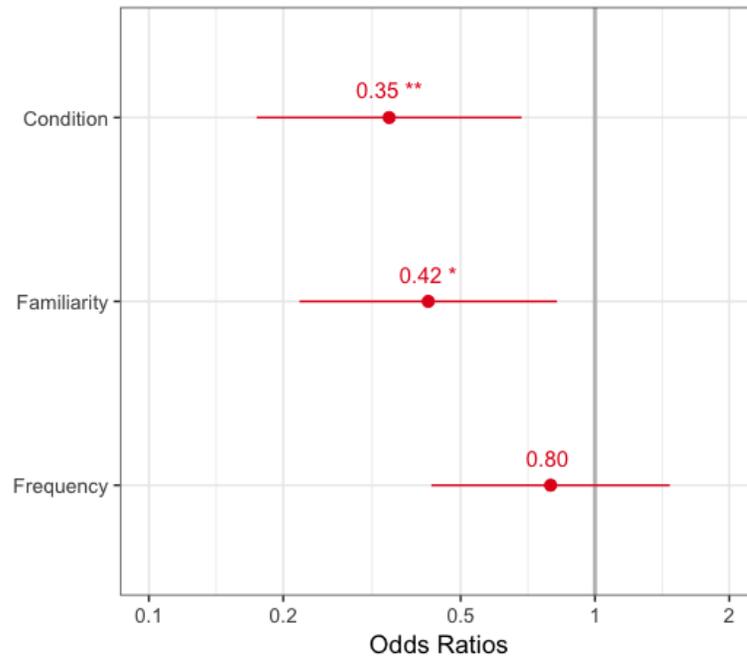


Figure 14. Individual factor effects on accuracy rates in native speakers

5 Chapter: Discussion

5.1 Review of the research questions and predictions

The purpose of this thesis was to examine how semantic radicals affect the visual processing of left-right configured semantic-phonetic compounds Chinese characters in CFL learners and native speakers. A character decision task with masked priming of a semantic radical that either matched or did not match the semantic radical in the target character was used to determine whether participants processed characters holistically or used a decomposition method. If a significant difference was found between the matched and unmatched prime conditions, then it was indicative that two different types of processing methods were used. Specifically, slower reaction times following an unmatched prime would suggest that there was interference in the character identification processing, likely resulting in decomposition. If no difference in responses existed between a matched or unmatched semantic radical prime, then holistic processing was assumed. The performance of CFL learners was compared to that of native speakers. The goal of this comparison was to understand how two distinct groups of Chinese speakers process the same type of compound character, and whether any processing differences could be observed using a character decision task.

Aside from examining the “part-whole” distinction in character processing, it was also of interest to examine whether the method of visual processing would change depending on participants’ familiarity with a character. It was predicted that for the most part, unfamiliar characters would be decomposed into sub-component parts as a way of seeking informative clues about the meaning of the character, even though the character

decision task did not explicitly measure character comprehension. At the same time, it was predicted that familiar characters would be processed holistically due to explicit instruction and intentional (as opposed to incidental) learning, which would presumably establish a stronger conceptual link between the visual identification of the character and access to its meaning. However, while familiarity can tell us whether a lexical item can be mapped onto its semantic representation in the mental lexicon, it cannot be used to judge the probability of exposure to certain characters. This may be important for knowing the degree to which the meaning of a character has been solidified. Character frequency was therefore included as a variable of equal importance and interest. For the purpose of this thesis, it functions as an available measure of the estimated amount of exposure to a lexical item. A character frequency list based on film and television series subtitles was considered more appropriate than using traditional books-based frequency counts, since the latter would be less applicable to CFL learners.

Interaction between character familiarity and whole-character frequency was used to explore whether these factors could further explain any significant differences between processing methods. In particular, CFL learners and native speakers were expected to holistically process familiar high-frequency characters, and to decompose unfamiliar low-frequency characters. It was unclear whether CFL learners would holistically process or decompose familiar low-frequency characters. This ambiguity pertains to the exact nature of the relationship between familiarity and frequency. In particular, it is difficult to know with certainty whether being familiar with a character is sufficient for a stable semantic encoding of that character, and to what extent greater exposure to its written form can help solidify this knowledge, if at all.

5.2 Interpretation of the findings

Reaction time analyses were performed only on correct responses to real characters, and accuracy rate analyses were performed on all responses to real characters. The sections below provide some possible explanations for the trends that were observed in each participant group. The main question of decomposition or holistic processing of certain subsets of characters in each group will also be addressed.

5.2.1 Character processing by CFL learners

Analyses of CFL learners' reaction times in the character decision task revealed several distinct patterns. In terms of familiar high-frequency characters, the difference between matched and unmatched prime conditions was very marginal, thus confirming the prediction that this subset of characters would be processed holistically.

Compared to familiar high-frequency characters, overall reaction times were slower for familiar low-frequency characters. Model estimates suggest that reaction time difference was the most extreme for these characters, which implies that CFL learners used a decomposition method when making judgments about characters that have been learned but were subsequently not seen very often. Based on these results, it seems that for CFL learners, simply knowing the meaning of a character is insufficient for holistic processing; how frequently learners come across a certain character in their textbook or other printed Chinese materials may matter too. Neuroimaging studies have found that even native speakers mentally re-trace the spatial sequence of stroke patterns in a character to aid vocabulary learning and subsequent memorization (Siok, Perfetti, Jin &

Tan, 2004), thus engaging in decomposition, to a degree. It would seem plausible for CFL learners to also employ decomposition as the default method of interacting with familiar, less commonly seen characters that would require greater efforts to learn and memorize.

Model estimates for differences between conditions in unfamiliar high-frequency characters were less drastic compared to familiar low-frequency characters. However, the effect of condition was greater for this subset of unfamiliar characters, although this may be due to the imbalance in vocabulary knowledge in the particular group of CFL learners that was tested in this study. The ratio of unfamiliar to familiar characters was 2:1, which potentially skewed the significance of the results towards unfamiliar characters. Despite this, the findings still suggest that decomposition was being used to process these characters, as predicted.

Looking only at the matched condition, results for unfamiliar high-frequency and familiar low-frequency characters were extremely similar. Although the main goal of the data analysis was to examine differences between experimental conditions within subsets of characters—and not on differences between subsets of characters on one condition – it should be mentioned that matched semantic radical primes facilitated responses to two arguably distinct groups of lexical items (familiar vs. unfamiliar) in almost the exact same way. Model estimates showed that reaction times were about 56 ms faster in the matched condition for familiar low-frequency character, while reaction times were about 61 ms faster in the matched condition for unfamiliar high-frequency characters. For CFL learners, the cognitive mechanism responsible for the visual processing of familiar or

unfamiliar characters seems to be mediated by frequency, even though there has been no evidence vocabulary instruction follows a frequency distribution.

Reaction times differed the least between conditions for unfamiliar low-frequency characters. It is assumed that CFL learners have not learned or come across these characters, and therefore rely only on the orthographic regularities of the characters to help make their judgments in the task. It was predicted that CFL learners would decompose unfamiliar low-frequency characters, but holistic processing was observed instead. It has been argued that after six months of study, CFL learners possess a basic knowledge of orthographic constraints of the Chinese writing system, such as whole-character configurations (e.g., left-right, top-down, enclosed, etc.) and positional regularities of sub-character components (Tong & Yip, 2014). Since the CFL learners tested in the study had received a cumulative mean of approximately 18 months of formal language instruction, it can be assumed that most participants had awareness of the elements that make up a real Chinese character. While reaction times between conditions did not differ significantly, responses to unfamiliar low-frequency characters were still slower compared to the filler items that fully violate the rules of Chinese orthography. Therefore, it could be argued that unfamiliar low-frequency characters are processed holistically by automatic activation of regularities such as the overall shape and structure of a character. This explanation is in line with research on CFL learners claiming that incidental learning through reading passages benefits awareness of the form of characters, more than learners' ability to use the characters (Ho, Ng, & Ng, 2003).

Although not statistically significant, CFL learners processed unfamiliar low-frequency characters slightly slower in the matched condition. In all other cases, slower

responses were found in the unmatched condition. Rather than viewing this anomalous trend as the consequence of a semantic radical prime that somehow facilitated decision-making when it did not match the radical in the target character, it seems more tenable to explore the idea that a matched prime inhibited processing of the target character. In a previous character decision task with native speakers, participants were primed with radical primes that were either semantically opaque or transparent when they were part of a target character, and inhibition was found for the semantically opaque target when it was followed by a semantically transparent radical prime (Feldman & Siok, 1999). A similar type of inhibition may be occurring in CFL learners when a prime and/or a sub-character component of an unfamiliar character are actually familiar to the participant.

5.2.2 Character processing by Chinese native speakers

Native speakers tended to use holistic processing for familiar characters, regardless of frequency. Pairwise contrasts between the matched and unmatched conditions for familiar characters confirmed the predictions. Based on the averaged reaction time data, low-frequency characters were processed slightly slower than high-frequency characters, but only by about 20 ms. On the other hand, reaction times were slowest for unfamiliar low-frequency characters. The difference between matched and unmatched conditions for these characters was the largest (28 ms), though again, not significant.

The model estimates further confirmed that native speakers consistently processed characters holistically. For native speakers, none of the matched-unmatched estimated contrasts presented earlier in the results section were significant. While a minor

processing cost was incurred for unfamiliar low-frequency characters, the overall processing method did not change.

Based on the results in this study, there was no evidence suggesting that native speakers used different character processing methods. However, unfamiliar characters represented only about one-fifth of all the characters that were seen by native speaker participants in the character decision task. It would be interesting to see if the same results would remain robust when tested with a larger set of unfamiliar low-frequency characters.

5.2.3 Misidentification of real characters

Only accuracy data on real characters were analyzed. Therefore, incorrect responses reflect trials in the character decision task where participants misidentified a real character as a fake one. Similar patterns of accuracy rate were found in CFL learners and native speakers. Both groups produced more errors in the unmatched condition, which follows previous behavioural results showing that a matched, semantically transparent radical prime facilitates correct character identification while unmatched primes, even if semantically transparent, produced greater error rates in native speakers (Feldman & Siok, 1999).

Condition did not affect accuracy responses in CFL learners, but a significant interaction between familiarity and frequency was found. Specifically, familiar low-frequency characters produced greater errors in responses compared to unfamiliar low-frequency characters. This suggests that, irrespective of the semantic radical prime, learners were affected by frequency, which supports results from the reaction time

analysis. Among the characters that learners claimed to know, it is possible that low-frequency characters elicited greater errors because they have not been solidified in memory the way that high-frequency characters have been. In contrast, responses by native speakers were affected more by familiarity with the characters, rather than frequency.

There were no priming effects on CFL learners' responses to unfamiliar low-frequency characters. Collapsing across both matched and unmatched conditions, mean accuracy rate for these characters was 86.4%, which was still higher than CFL learners' mean accuracy rate of 81.2% for filler items. It may be the case that CFL learners were treating these characters in a way that was similar to the filler items, suggesting insufficient awareness of certain stroke combinations.

While no predictions were made about processing method (holistic vs. decomposition) in relation to accuracy data results, it is still notable that despite using more time to make a judgment on unfamiliar low-frequency characters (since RTs were slowest), more errors were made in both groups. One way to explain this finding relates to the high combinability of Chinese orthographic components. Researchers have argued that responses to pseudo-characters with recombined real sub-character components are more difficult to interpret than responses to non-words in alphabetic languages, due to the vast possibilities of orthographic to phonological mapping in Chinese (Yang et al., 2011). If this is assumed to be true of native speakers, then perhaps the same view can be extended to CFL learners and unfamiliar low-frequency characters. These characters may appear to them in the same way that pseudo-characters appear to native speakers.

5.2.4 Summary

The main findings of this study suggest that familiar high-frequency characters are holistically processed by both CFL learners and native speakers. While on the surface, familiar high-frequency characters may be treated equally by both groups, frequency may play a more important role than familiarity in CFL learners, especially since previous research has claimed that the incidental learning of new characters occurs mainly through reading (as opposed to oral practice, for example) (Shu et al., 1995). More exposure to a character offers more chances to become familiar with the orthographic patterns of a character, its meaning, and diverse ways of using it. Holistic processing of these characters is therefore unsurprising.

In contrast, the findings showed a general processing cost associated with unfamiliar characters, though native speakers were less affected. For CFL learners, frequency remained an important factor. Within the set of characters that learners do not know, decomposition was used only for high-frequency characters; within the set of characters that the learners do know, decomposition was used only for low-frequency characters. Based on these results and the discussion above, it seems that decomposition is useful for helping us understand how orthographic and semantic elements of a mixed morpho-syllabic language system are organized in a foreign language learner, particularly one whose native language is not logographic. Decomposition also serves as a tool for examining whether such elements form a hierarchical structure in the mental lexicon, and if so, whether this structure revolves more around semantic consolidation or around incidental, statistical learning.

5.3 Limitations of the study

There were several limitations in the study. First, since only left-right horizontally-configured compound characters were tested, conclusions cannot be drawn about how CFL learners process compound characters with other configurations, e.g., vertical (杏) or enclosed (国). Second, earlier studies on gestalt aspects of Chinese characters have found that the constituent components of left-right horizontally-configured compound characters were named faster compared to compound characters with other configurations, suggesting that the types of compound characters used in this study may have been more inherently vulnerable to decomposition (Huang & Wang, 1992). Though this result has not been widely re-tested, it serves as a reminder that the perceptual dynamics within a sub-group of compound characters may vary significantly.

Third, the phonetic component in compound characters, which is an important aspect of character recognition and processing, were not taken into account. Examining the combination of semantic and phonological information while comparing across familiarity levels would produce a more comprehensive picture of vocabulary acquisition in CFL learners.

Sub-character level factors that may be considered relevant to the masked primes used in this study were not considered. Specifically, the following attributes of semantic radical primes were not controlled for: semantic radical token frequency (how often a radical appears either as a constituent of a character or a standalone character); semantic radical type frequency (how many characters the same semantic radical appears in); and semantic radical combinability (ability of a radical to combine with

phonetic components to form a compound character).

Participants' familiarity with the semantic radical primes was also not examined. As mentioned at the beginning of the thesis, Chinese characters can contain morphemes but can also be morphemes in themselves. 33 of the 48 semantic radical primes are considered free morphemes that can stand alone as single-character words. The true nature of the priming effect cannot be elucidated without knowing how the primes may be being registered by learners with varying levels of Chinese morphological awareness.

5.4 Conclusion

Measuring lexical decomposition not only reveals how visual stimuli take on meaning in a particular writing system, but it is also a way to understand the cognitive processes that allow the encoding and storage of the smallest units of meaning in a language, for flexible manipulation and use (Dehaene, 2009). Since it is possible to break down the majority of the most commonly used characters in Chinese into meaningful constituent parts (Hoosain, 1991), it is important to examine the conditions under which character decomposition would be obligatory, and whether this differs between learners of Chinese compared to native speakers of Chinese.

Research on this topic may help contribute to our understanding of how novice foreign language learners develop processing mechanisms to help them become more skilled readers in that language. For example, a study that tracks the progress of the same group of beginner CFL learners in separate phases would be useful for more directly testing preferred character processing methods as a function of increased vocabulary size and language proficiency. Future research should also use

neurophysiological techniques such as eye-tracking to determine, with greater precision, how learners visually interact with various sub-lexical elements in Chinese.

Lastly, in terms of language pedagogy, it hoped that this study may contribute to existing evidence in favour of structuring courses and instructional material around informative sub-character components when teaching Chinese vocabulary to foreign language students.

Appendices

Appendix A

List of all stimuli used in character decision task

No.	Radical primes (n = 48)	Target characters (n = 96)	Filler primes (n = 48)	Filler targets (n = 96)
1	女 woman	姐 older sister 妹 younger sister	𡇗	祖 姊
2	口 mouth	叫 to call 喝 to drink	𠂔	副 謁
3	亻 person	你 you 位 measure word for person	亻	你 位
4	忄 heart	快 quickly 怕 scared	忄	快 怕
5	目 sight	睡 sleep 眼 eye	目	睡 眼
6	讠 speech	语 language 说 speak	讠	语 说
7	土 earth	地 land 场 open space	土	地 场
8	日 sun	暖 warm 明 bright	日	昨 明
9	纟 silk	绸 silk 络 net	纟	绸 络
10	衤 clothes	衬 shirt 裤 pants	衤	衬 裤
11	钅 gold	钱 money 铁 steel	钅	钱 铁

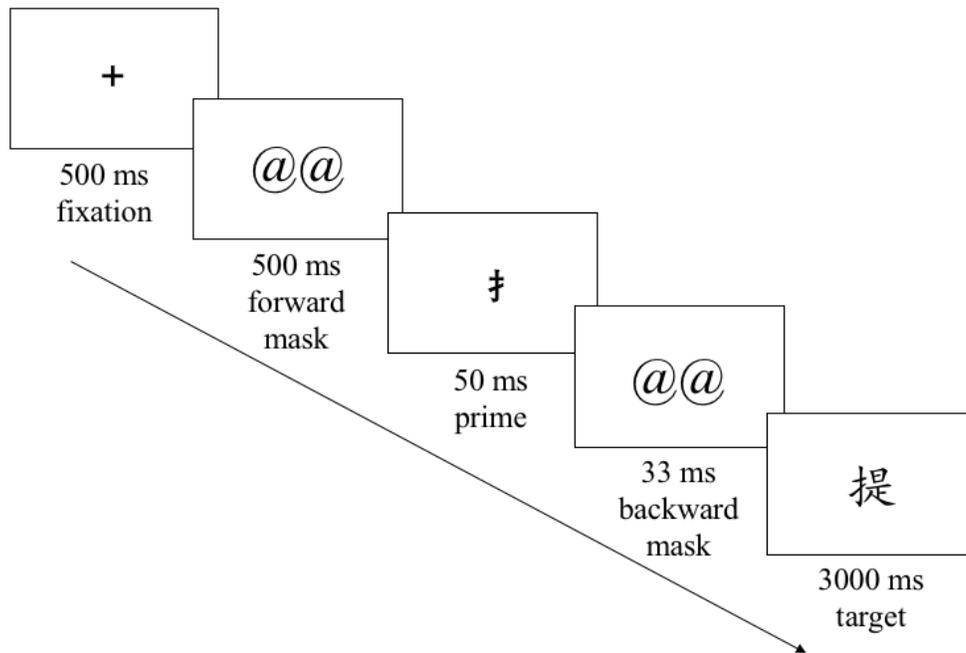
12	氵	water	汁 juice	洗 wash	氵	汁	洗
13	木	tree	枝 branch	植 to plant	木	枝	植
14	足	foot	跳 jump	踢 to kick	足	跳	踢
15	扌	hand	授 receive	换 exchange	扌	授	换
16	玉	jade	玩 to play (historically: “antiques”)	珠 pearl	玉	玩	珠
17	酉	wine	酷 cruel	酸 sour	酉	酷	酸
18	又	action	鸡 chicken	戏 game	又	鸡	戏
19	车	cart	辆 measure word for vehicles	转 to transfer	车	转	轻
20	舟	boat	航 navigate by sea	船 ship	舟	航	船
21	石	rock	确 firm, resolute	础 foundation	石	确	础
22	米	rice	料 feed	精 fine quality rice	米	料	精
23	月	flesh	脑 brain	脸 face	月	脑	脸
24	阝	mound	院 courtyard	际 border	阝	院	际
25	片	flat thin piece	版 board	牌 plate	片	版	牌

26	禾	agriculture	秋 autumn	租 rent	𥝌	秋	租
27	火	fire	灯 light	烘 bake	𠂇	灯	烘
28	冫	ice	冰 ice	凉 chilly	冫	冰	凉
29	舌	tongue	乱 chaotic	甜 sweet	呈	乱	甜
30	彳	pacing	行 alright	往 go towards	彳	往	行
31	犴	four-legged animal	猫 cat	狗 dog	犴	猫	狗
32	饣	food	馆 restaurant	饮 beverage	饣	馆	饮
33	犛	cow	犛	犛 give a bonus to	犛	犛	犛
34	虫	insect	蛭 mussel	蛭 parasitic worm	虫	蛭	蛭
35	弓	bow	弭 to repress	弭 red lacquer bow	弓	弭	弭
36	孑	child	孺 child	孜 hard-working	孑	孺	孜
37	身	body	軀 growth (of a body)	裸 nude	身	軀	裸
38	豸	ancient beast	貌	豸 fox cub	豸	貌	豸
39	马	horse	骋 to hasten	驭 chariot driver	马	骋	驭

40	耳	ear	聃 ears without rim	聃 ear disease	耳	聃	聃
41	歹	death	殇 to die in childhood	殄 to exterminate	歹	殄	殄
42	山	mountain	岵 mountain with vegetation	崄 tall and erect	山	岵	崄
43	赤	red	赭 deep red	赧 blushing with shame	赤	赧	赧
44	白	white	皑 white (of snow)	皤 white flower	白	皑	皤
45	鱼	fish	鲉 anchovy	鲐 minced fish	鱼	鲉	鲐
46	巾	cloth	帔 woman's short embroidered cape	幄 tent	巾	帔	幄
47	缶	earthenware	斝 small earthenware plate	匱 ancient coin box	缶	斝	匱
48	贝	wealth	贻 to bequeath	赊 to buy or sell on credit	贝	贻	赊

Appendix B

Stimuli presentation timeline in PsychoPy



Appendix C

Character Familiarity Rating Test

Please indicate (请指示) :

1 = You do not know or recognize the character at all

1 = 完全不认识这个字

2 = You have seen the character, but do not know or remember its meaning

2 = 见过这个字, 但是不知道或者不记得它是什么意思

3 = You know the character, and would be able to use it in a sentence or phrase

3 = 认识这个字, 并且可以用它造句或组成短语

- | | | |
|-------------|-------------|-------------|
| 1. 你 _____ | 15. 眼 _____ | 29. 行 _____ |
| 2. 怕 _____ | 16. 说 _____ | 30. 猫 _____ |
| 3. 语 _____ | 17. 场 _____ | 31. 馆 _____ |
| 4. 络 _____ | 18. 明 _____ | 32. 牌 _____ |
| 5. 衬 _____ | 19. 绸 _____ | 33. 租 _____ |
| 6. 钱 _____ | 20. 裤 _____ | 34. 烘 _____ |
| 7. 汁 _____ | 21. 铁 _____ | 35. 凉 _____ |
| 8. 踢 _____ | 22. 洗 _____ | 36. 甜 _____ |
| 9. 授 _____ | 23. 枝 _____ | 37. 往 _____ |
| 10. 玩 _____ | 24. 跳 _____ | 38. 狗 _____ |
| 11. 妹 _____ | 25. 换 _____ | 39. 饮 _____ |
| 12. 喝 _____ | 26. 珠 _____ | 40. 酸 _____ |
| 13. 位 _____ | 27. 院 _____ | 41. 戏 _____ |
| 14. 忙 _____ | 28. 冰 _____ | 42. 辆 _____ |

43. 船 _____

44. 确 _____

45. 精 _____

46. 脸 _____

47. 际 _____

48. 蝮 _____

49. 殄 _____

50. 嶮 _____

51. 赧 _____

52. 肥 _____

53. 鲑 _____

54. 幄 _____

55. 鲈 _____

56. 矧 _____

57. 犒 _____

58. 蛲 _____

59. 孳 _____

60. 孜 _____

61. 躐 _____

62. 豕 _____

63. 驺 _____

64. 瞠 _____

65. 姐 _____

66. 叫 _____

67. 睡 _____

68. 地 _____

69. 暖 _____

70. 植 _____

71. 酷 _____

72. 鸡 _____

73. 转 _____

74. 航 _____

75. 础 _____

76. 料 _____

77. 脑 _____

78. 版 _____

79. 秋 _____

80. 灯 _____

81. 乱 _____

82. 牝 _____

83. 蛭 _____

84. 弭 _____

85. 孺 _____

86. 艤 _____

87. 骋 _____

88. 聃 _____

89. 殄 _____

90. 岵 _____

91. 赧 _____

92. 皕 _____

93. 鯉 _____

94. 悻 _____

95. 鉢 _____

96. 贻 _____

Appendix D

Language Background Questionnaire (adapted from Li, Zhang, Tsai, & Puls, 2014)

1. Age: _____

2. Sex (circle one): Male / Female

3. Indicate your native language(s):

4. Indicate all **FOREIGN** language(s) you know or have studied:

Language	Age of acquisition	Cumulative amount of instruction (in months)

5. Rate your current ability in terms of listening, speaking, reading, and writing in **Chinese**:

Very poor	Poor	Limited	Functional	Good	Very good	Native-like
1	2	3	4	5	6	7

Language	Listening	Speaking	Reading	Writing
	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7	1 2 3 4 5 6 7

6. Estimate how many hours per day you spend engaged in the following activities in Chinese:

Watching television: _____(hrs)

Listening to radio: _____(hrs)

Reading for fun: _____(hrs)

Reading for school/work: _____(hrs)

Writing emails to friends: _____(hrs)

Writing for school/work: _____(hrs)

7. Estimate how many hours per day you spend speaking Chinese with the following groups of people.

Family members: _____(hrs)

Friends^a: _____(hrs)

Classmates: _____(hrs)

Coworkers^b: _____(hrs)

8. Rate your Chinese language learning skill. How good do you feel you are at learning Chinese, relative to your classmates and/or others? (circle one)

Very poor	Poor	Limited	Functional	Good	Very good	Native-like
1	2	3	4	5	6	7

Appendix E

R Script for Mixed Effects Models

```
library(lme4)
library(lmerTest)

## RT models ##

# random effects structure
lme.null0 = lmer(RT_ms ~ 1 + (1|participant) + (1|item), data = rt_cont)
lme.null = lmer(RT_ms ~ 1 + (1 + cond_cont|participant) + (1|item), data = rt_cont)
lme.null1 = lmer(RT_ms ~ 1 + (1 + cond_cont|participant) + (1 + cond_cont|item), data =
rt_cont)

anova(lme.null0, lme.null, refit = F)
anova(lme.null, lme.null1, refit = F)

# main effects

# effect of group_cont
lme.group_cont = lmer(RT_ms ~ group_cont + (1 + cond_cont|participant) + (1|item),
data = rt_cont)
anova(lme.null, lme.group_cont, refit = F)
summary(lme.group_cont)

# adding effect of cond_cont
lme.cond = lmer(RT_ms ~ cond_cont + (1 + cond_cont|participant) + (1|item), data =
rt_cont)
anova(lme.null, lme.cond, refit = F)
summary(lme.cond)

# adding effect of fam_cont
lme.fam = lmer(RT_ms ~ fam_cont + (1 + cond_cont|participant) + (1|item), data =
rt_cont)
anova(lme.null, lme.fam, refit = F)
summary(lme.fam)

# adding effect of freq_cont
lme.freq = lmer(RT_ms ~ freq_cont + (1 + cond_cont|participant) + (1|item), data =
rt_cont)
anova(lme.null, lme.freq, refit = F)
summary(lme.freq)

# four-way interaction model
```

```

RT1 = lmer(RT_ms ~ group_cont * cond_cont * freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT2 = lmer(RT_ms ~ group_cont * cond_cont * freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT3 = lmer(RT_ms ~ group_cont * freq_cont * fam_cont + cond_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT4 = lmer(RT_ms ~ group_cont + freq_cont * fam_cont * cond_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT5 = lmer(RT_ms ~ group_cont * cond_cont + freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT6 = lmer(RT_ms ~ group_cont * freq_cont + cond_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT7 = lmer(RT_ms ~ group_cont * fam_cont + freq_cont + cond_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT8 = lmer(RT_ms ~ group_cont + cond_cont * freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT9 = lmer(RT_ms ~ group_cont + cond_cont * fam_cont + freq_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT10 = lmer(RT_ms ~ group_cont + cond_cont + freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)
RT11 = lmer(RT_ms ~ group_cont + cond_cont + freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = rt_cont)

```

```

anova(RT1, RT2, refit = F)
anova(RT1, RT3, refit = F)
anova(RT1, RT4, refit = F)
anova(RT1, RT5, refit = F)
anova(RT1, RT6, refit = F)
anova(RT1, RT7, refit = F)
anova(RT1, RT8, refit = F)
anova(RT1, RT9, refit = F)
anova(RT1, RT10, refit = F)
anova(RT1, RT11, refit = F)

```

```
summary(RT1)
```

```
# by participant group
```

```
# three-way interaction model for CFL learners
```

```

CFL.RT.1 = lmer(RT_ms ~ cond_cont * freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = CFL_rt)
CFL.RT.2 = lmer(RT_ms ~ cond_cont * freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = CFL_rt)
CFL.RT.3 = lmer(RT_ms ~ cond_cont + freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = CFL_rt)

```

```

CFL.RT.4 = lmer(RT_ms ~ cond_cont * fam_cont + freq_cont + (1 +
cond_cont|participant) + (1|item), data = CFL_rt)
CFL.RT.5 = lmer(RT_ms ~ cond_cont + fam_cont + freq_cont + (1 +
cond_cont|participant) + (1|item), data = CFL_rt)

anova(CFL.RT.1, CFL.RT.2, refit = FALSE)
anova(CFL.RT.1, CFL.RT.3, refit = FALSE)
anova(CFL.RT.1, CFL.RT.4, refit = FALSE)
anova(CFL.RT.1, CFL.RT.5, refit = FALSE)

summary(CFL.RT.1)

anova(CFL.RT.1)
# This indicates that we have 3 significant simple/main effects
# pairwise comparison of main effects and interactions (compared to nothing)

# plot residuals
plot(density(scale(resid(CFL.RT.1))))

# residuals are skewed; plot fitted values against residuals
plot(fitted(CFL.RT.3), scale(resid(CFL.RT.1)))
abline(h=c(-2.5, 2.5), col="red")

# refit the model, excluding the outliers
CFL.RT.1_trim = lmer(RT_ms ~ cond_cont * freq_cont * fam_cont +
(1 + cond_cont|participant) + (1|item), data=CFL_rt,
subset=abs(scale(resid(CFL.RT.1)))<2.5)

# double-check summary
summary(CFL.RT.1_trim)

# three-way interaction model for native speakers

NS.RT.1 = lmer(RT_ms ~ cond_cont * freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = NS_rt)
NS.RT.2 = lmer(RT_ms ~ cond_cont * freq_cont + fam_cont + (1 +
cond_cont|participant) + (1|item), data = NS_rt)
NS.RT.3 = lmer(RT_ms ~ cond_cont + freq_cont * fam_cont + (1 +
cond_cont|participant) + (1|item), data = NS_rt)
NS.RT.4 = lmer(RT_ms ~ cond_cont * fam_cont + freq_cont + (1 +
cond_cont|participant) + (1|item), data = NS_rt)
NS.RT.5 = lmer(RT_ms ~ cond_cont + fam_cont + freq_cont + (1 +
cond_cont|participant) + (1|item), data = NS_rt)

anova(NS.RT.1, NS.RT.2, refit = FALSE)
anova(NS.RT.1, NS.RT.3, refit = FALSE)

```

```

anova(NS.RT.1, NS.RT.4, refit = FALSE)
anova(NS.RT.1, NS.RT.5, refit = FALSE)

summary(NS.RT.1)

plot(density(scale(resid(NS.RT.1))))
plot(fitted(NS.RT.1), scale(resid(NS.RT.1)))
abline(h=c(-2.5, 2.5), col="red")
nrow(NS_rt)-nrow(NS_rt[abs(scale(resid(NS.RT.1)))<2.5,])
(1-(nrow(NS_rt[abs(scale(resid(NS.RT.1)))<2.5,])/nrow(NS_rt)))*100

NS.RT.1_trim = lmer(RT_ms ~ cond_cont * freq_cont * fam_cont +
  (1 + cond_cont|participant) + (1|item),
  data = NS_rt, REML = F,
  subset=abs(scale(resid(NS.RT.1)))<2.5)

summary(NS.RT.1_trim)

## Accuracy rate models ##

# random effects structure

glme.acc.null0 = glmer(acc ~ 1 + (1|participant) + (1|item), data = acc_cont, family =
"binomial")
glme.acc.null = glmer(acc ~ 1 + (1|participant) + (1 + cond_cont|item), data = acc_cont,
family = "binomial")

anova(glme.acc.null0, glme.acc.null, refit = F)

# effect of group_cont
glme.acc.group = glmer(acc ~ group_cont + (1|participant) + (1 + cond_cont|item), data =
acc_cont, family = "binomial")
anova(glme.acc.null, glme.acc.group, refit = F)
summary(glme.acc.group)

# adding effect of cond_cont
glme.acc.cond = glmer(acc ~ cond_cont + (1|participant) + (1+ cond_cont|item), data =
acc_cont, family = "binomial")
anova(glme.acc.null, glme.acc.cond, refit = F)
summary(glme.acc.cond)

# adding effect of fam_cont
glme.acc.fam = glmer(acc ~ fam_cont + (1|participant) + (1+ cond_cont|item), data =
acc_cont, family = "binomial",
  glmerControl(optimizer="bobyqa", optCtrl = list(maxfun = 100000)))
anova(glme.acc.null, glme.acc.fam, refit = F)

```

```

summary(glme.acc.fam)

# adding effect of freq_cont
glme.acc.freq = glmer(acc ~ freq_cont + (1|participant) + (1 + cond_cont|item), data =
acc_cont, family = "binomial")
anova(glme.acc.null, glme.acc.freq, refit = F)
summary(glme.acc.freq)

# four-way interaction model

acc1 = glmer(acc ~ group_cont * cond_cont * freq_cont * fam_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc2 = glmer(acc ~ group_cont * cond_cont * freq_cont + fam_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc3 = glmer(acc ~ group_cont * freq_cont * fam_cont + cond_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc4 = glmer(acc ~ group_cont + freq_cont * fam_cont * cond_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc5 = glmer(acc ~ group_cont * cond_cont + freq_cont + fam_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc6 = glmer(acc ~ group_cont * freq_cont + cond_cont + fam_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc7 = glmer(acc ~ group_cont * fam_cont + freq_cont + cond_cont + (1|participant) +
(1+ cond_cont|item),
          data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

```

```

acc8 = glmer(acc ~ group_cont + cond_cont * freq_cont + fam_cont + (1|participant) +
(1+ cond_cont|item),
      data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc9 = glmer(acc ~ group_cont + cond_cont * fam_cont + freq_cont + (1|participant) +
(1+ cond_cont|item),
      data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa", optCtrl
= list(maxfun = 100000)))

acc10 = glmer(acc ~ group_cont + cond_cont + freq_cont * fam_cont + (1|participant) +
(1+ cond_cont|item),
      data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa",
optCtrl = list(maxfun = 100000)))

acc11 = glmer(acc ~ group_cont + cond_cont + freq_cont + fam_cont + (1|participant) +
(1+ cond_cont|item),
      data = acc_cont, family = "binomial", glmerControl(optimizer="bobyqa",
optCtrl = list(maxfun = 100000)))

anova(acc1, acc2, refit = F)
anova(acc1, acc3, refit = F)
anova(acc1, acc4, refit = F)
anova(acc1, acc5, refit = F)
anova(acc1, acc6, refit = F)
anova(acc1, acc7, refit = F)
anova(acc1, acc8, refit = F)
anova(acc1, acc9, refit = F)
anova(acc1, acc10, refit = F)
anova(acc1, acc11, refit = F)

summary(acc1)

# by participant group

# three-way interaction model for CFL learners

CFL.acc.1 = glmer(acc ~ cond_cont * freq_cont * fam_cont + (1|participant) + (1+
cond_cont|item), data = CFL_acc,
      family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))
CFL.acc.2 = glmer(acc ~ cond_cont * freq_cont + fam_cont + (1|participant) + (1+
cond_cont|item), data = CFL_acc,
      family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))

```

```

CFL.acc.3 = glmer(acc ~ cond_cont + freq_cont * fam_cont + (1|participant) + (1+
cond_cont|item), data = CFL_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))
CFL.acc.4 = glmer(acc ~ cond_cont + freq_cont + fam_cont + (1|participant) + (1+
cond_cont|item), data = CFL_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))

anova(CFL.acc.1, CFL.acc.2, refit = FALSE)
anova(CFL.acc.1, CFL.acc.3, refit = FALSE)
anova(CFL.acc.3, CFL.acc.4, refit = FALSE)

summary(CFL.acc.3)

# three-way interaction model for native speakers

NS.acc.1 = glmer(acc ~ cond_cont * freq_cont * fam_cont + (1|participant) + (1+
cond_cont|item), data = NS_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))
NS.acc.2 = glmer(acc ~ cond_cont * freq_cont + fam_cont + (1|participant) + (1+
cond_cont|item), data = NS_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))
NS.acc.3 = glmer(acc ~ cond_cont + freq_cont * fam_cont + (1|participant) + (1+
cond_cont|item), data = NS_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))
NS.acc.4 = glmer(acc ~ cond_cont + freq_cont + fam_cont + (1|participant) + (1+
cond_cont|item), data = NS_acc,
family = "binomial", glmerControl(optimizer="bobyqa", optCtrl = list(maxfun
= 100000)))

anova(NS.acc.1, NS.acc.2, refit = FALSE)
anova(NS.acc.2, NS.acc.3, refit = FALSE)
anova(NS.acc.2, NS.acc.4, refit = FALSE)

summary(NS.acc.4)

```

References

- Andrews, S., & Lo, S. (2013). Is morphological priming stronger for transparent than opaque words? It depends on individual differences in spelling and vocabulary. *Journal of Memory and Language*, *68*(3), 279–296.
<https://doi.org/10.1016/j.jml.2012.12.001>
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4. R package version 1.1–7. 2014.
- Beyersmann, E., Duñabeitia, J. A., Carreiras, M., Coltheart, M., & Castles, A. (2013). Early morphological decomposition of suffixed words: Masked priming evidence with transposed-letter nonword primes. *Applied Psycholinguistics*, *34*(5), 869–892.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*(6).
<https://doi.org/10.1371/journal.pone.0010729>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.
- Connine, C. M., Mullennix, J., Shernoff, E., & Yelen, J. (1990). Word familiarity and frequency in visual and auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*(6), 1084–1096.
- DeFrancis, J. (1989). *Visible speech: The diverse oneness of writing systems*. University of Hawaii Press.
- Ding, G., Peng, D., & Taft, M. (2004). The nature of the mental representation of radicals

- in Chinese: a priming study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(2), 530–539. <https://doi.org/10.1037/0278-7393.30.2.530>
- Fan, K. Y., Gao, J. Y., & Ao, X. P. (1984). Pronunciation principles of the Chinese character and alphabetic writing scripts. *Chinese Character Reform*, 3, 23-27.
- Feldman, L. B., & Siok, W. W. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology-Learning Memory and Cognition*, 23(3), 776-781.
- Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language*, 40, 559–576. <https://doi.org/10.1006/jmla.1998.2629>
- Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course*. Routledge.
- Gernsbacher, M. A. (1984). Resolving 20 years of inconsistent interactions between lexical familiarity and orthography, concreteness, and polysemy. *Journal of Experimental Psychology: General*, 113(2), 256–281.
- Grainger, J., & Beyersmann, E. (2017). Edge-aligned embedded word activation initiates morpho-orthographic segmentation. *Progress in Brain Research*, Vol. 232. <https://doi.org/10.1016/bs.plm.2017.03.009>
- Heathcote, L., Nation, K., Castles, A., & Beyersmann, E. (2017). Do “blacheap” and “subcheap” both prime “cheap”? An investigation of morphemic status and position in early visual word processing. *The Quarterly Journal of Experimental Psychology*, (October), 1–34. <https://doi.org/10.1080/17470218.2017.1362704>

- Ho, CSH; Ng, TT; Ng, W. (2003). A “radical” approach to reading development in Chinese: The role of semantic radicals and phonetic radicals. *Journal of Literacy Research, 53*(3), 849–878. <https://doi.org/10.1017/CBO9781107415324.004>
- Huang, J. T., & Wang, M. Y. (1992). From unit to gestalt: perceptual dynamics in recognizing chinese characters. *Advances in Psychology, 90*(C), 3–35. [https://doi.org/10.1016/S0166-4115\(08\)61885-3](https://doi.org/10.1016/S0166-4115(08)61885-3)
- Jackson, N. E., Everson, M. E., & Ke, C. (2003). Beginning readers’ awareness of the orthographic structure of semantic-phonetic compounds: Lessons from a study of learners of Chinese as a foreign language. *Reading development in Chinese children, 141-153*.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language, 59*(4), 434-446.
- Ke, C. (1998). Effects of language background on the learning of chinese characters among foreign language students. *Foreign Language Annals, 31*(1), 91–102. <https://doi.org/10.1111/j.1944-9720.1998.tb01335.x>
- Kinoshita, S., & Lupker, S. J., eds. (2003). *Masked priming: The state of the art*. NY: Taylor and Francis.
- Kuo, L. J., Li, Y., Sadoski, M., & Kim, T. J. (2014). Acquisition of Chinese characters: The effects of character properties and individual differences among learners. *Contemporary Educational Psychology, 39*(4), 287–300. <https://doi.org/10.1016/j.cedpsych.2014.07.001>
- Lenth, R. (2018). emmeans: estimated marginal means, aka least-squares means. R

package version 1.1.2. <https://CRAN.R-project.org/package=emmeans>

- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(3), 673-680.
- Liu, P.D., Chung, K. K. H., McBride-Chang, C., & Tong, X. (2010). Holistic versus analytic processing: Evidence for a different approach to processing of Chinese at the word and character levels in Chinese children. *Journal of Experimental Child Psychology*, 107(4), 466–478. <https://doi.org/10.1016/j.jecp.2010.06.006>
- Liu, Y., Wang, M., & Perfetti, C. A. (2007). Threshold-style processing of Chinese characters for adult second-language learners. *Memory & Cognition*, 35(3), 471–480. <https://doi.org/10.3758/BF03193287>
- Longtin, C. M., & Meunier, F. (2005). Morphological decomposition in early visual word processing. *Journal of Memory and Language*, 53(1), 26-41.
- Myers, J. (2016). Knowing Chinese character grammar. *Cognition*, 147, 127–132. <https://doi.org/10.1016/j.cognition.2015.11.012>
- McBride, C. A. (2016). Is Chinese special? four aspects of chinese literacy acquisition that might distinguish learning chinese from learning alphabetic orthographies. *Educational Psychology Review*, 28(3), 523–549. <https://doi.org/10.1007/s10648-015-9318-2>
- McBride-Chang, C., Wat, C. P., Shu, H., Zhou, A., & Wagner, R. K. (2003). Morphological awareness uniquely predicts young children’s Chinese character recognition. *Journal of Educational Psychology*, 95(4), 743–751. <https://doi.org/10.1037/0022-0663.95.4.743>

- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88(5), 375.
- McCormick, S. F., Rastle, K., & Davis, M. H. (2009). Adore-able not adorable? Orthographic underspecification studied with masked repetition priming. *European Journal of Cognitive Psychology*, 21(6), 813-836.
- Perfetti, C. & Liu, Y. (2006). Reading Chinese characters: Orthography, phonology, meaning and the lexical constituency model. *Handbook of East Asian Psycholinguistics*.
- Perfetti, C., Liu, Y., & Tan, L. H. (2005). The lexical constituency model: some implications of research on Chinese for general theories of reading. *Psychological Review*, 112(1), 43–59. <https://doi.org/10.1037/0033-295X.112.1.43>
- Peirce, J.W. (2009) Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.* 2:10. doi:10.3389/neuro.11.010.2008
- Pixen [Computer software]. (2016). Retrieved from <https://pixenapp.com/>.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, 81, 275–280.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- RStudio Team (2015). RStudio: Integrated development for R. RStudio, Inc., Boston, MA. URL <http://www.rstudio.com/>.
- Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and*

writing. Routledge.

- Seidenberg, M. S. (1985). The time course of phonological code activation in two writing systems. *Cognition*, *19*(1), 1–30. [https://doi.org/10.1016/0010-0277\(85\)90029-0](https://doi.org/10.1016/0010-0277(85)90029-0)
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523.
- Shen, H. H., & Jiang, X. (2013). Character reading fluency, word segmentation accuracy, and reading comprehension in L2 Chinese. *Reading in a Foreign Language*, *25*(1), 1.
- Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *Modern Language Journal*, *91*(1), 97–111.
<https://doi.org/10.1111/j.1540-4781.2007.00511.x>
- Siok, W. T., Perfetti, C. A., Jin, Z., & Tan, L. H. (2004). Biological abnormality of impaired reading is constrained by culture. *nature*, *431*(7004), 71.
- Shu, H., Anderson, R. C., & Zhang, H. (1995). Incidental learning of word meanings while reading: A Chinese and American cross-cultural study. *Reading Research Quarterly*, *76*-95.
- Shu, H., Chen, X., Anderson, R. C., Wu, N., & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, *74*(1), 27–47.
<https://doi.org/10.1111/1467-8624.00519>
- Taft, M. (1994). Interactive-activation as a framework for understanding morphological processing. *Language and Cognitive Processes*, *9*(3), 271-294.
- Taft, M., & Chung, K. (1999). Using radicals in teaching Chinese characters to second language learners. *Psychologia*, *42*(4), 243-251.

- Taft, M., & Zhu, X. (1995). The representation of bound morphemes in the lexicon: A Chinese study. *Morphological Aspects of Language Processing*, 293-316.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23(3), 761–775.
<https://doi.org/10.1037/0278-7393.23.3.761>
- Tao, L., & Healy, A. F. (2002). The unitization effect in reading chinese and english text. *Scientific Studies of Reading*, 6(2), 167–197.
https://doi.org/10.1207/s1532799xssr0602_03
- Tong, X., & McBride-Chang, C. (2010). Developmental models of learning to read Chinese words. *Developmental Psychology*, 46(6), 1662–1676.
<https://doi.org/10.1037/a0020611>
- Tong, X., & Yip, J. H. Y. (2014). Cracking the Chinese character: Radical sensitivity in learners of Chinese as a foreign language and its relationship to Chinese word reading. *Reading and Writing*, (August 2015), 159–181.
<https://doi.org/10.1007/s11145-014-9519-y>
- Wang, M., Perfetti, C. A., & Liu, Y. (2005). Chinese-English biliteracy acquisition: Cross-language and writing system transfer. *Cognition*, 97(1), 67–88.
<https://doi.org/10.1016/j.cognition.2004.10.001>
- Williams, C. H. (2010). Semantic vs. phonetic decoding strategies in non-native readers of Chinese. *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 71(5–A), 1623.
- Yao, T.C., Liu, Y., Ge, L., Chen, Y.F., Bi, N.P., Wang, X., & Shi, Y. (2009). *Integrated Chinese, Level 1, Part 1* (3rd Ed.). Boston: Cheng & Tsui Company.

- Yao, T.C., Liu, Y., Ge, L., Chen, Y.F., Bi, N.P., Wang, X., & Shi, Y. (2009). *Integrated Chinese, Level 1, Part 2* (3rd Ed.). Boston: Cheng & Tsui Company.
- Yao, T.C., Liu, Y., Ge, L., Chen, Y.F., Bi, N.P., Wang, X., & Shi, Y. (2009). *Integrated Chinese, Level 2, Part 1* (3rd Ed.). Boston: Cheng & Tsui Company.
- Yang, J., Shu, H., McCandliss, B. D., & Zevin, J. D. (2013). Orthographic influences on division of labor in learning to read Chinese and English: Insights from computational modeling. *Bilingualism: Language and Cognition*, 16(2), 354-366.
- Yang, J., Wang, X., Shu, H., & Zevin, J. D. (2011). Brain networks associated with sublexical properties of Chinese characters. *Brain and Language*, 119(2), 68–79.
<https://doi.org/10.1016/j.bandl.2011.03.004>
- Zar, J. H. (1999). *Biostatistical analysis* (4th ed.). New Jersey: Prentice Hall.
- Zhou, X. L., & Marslen-Wilson, W. (1999). The nature of sublexical processing in reading Chinese characters. *Journal of Experimental Psychology-Learning Memory and Cognition*, 25(4), 819–837. <https://doi.org/10.1037/0278-7393.25.4.819>
- Zhou, X., & Marslen-Wilson, W. (2000). The relative time course of semantic and phonological activation in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(5), 1245.