1-1-2012

The effect of musical training on verbal and tonal working memory

Ching-i Lu
Wayne State University,

THE EFFECT OF MUSICAL TRAINING ON VERBAL AND TONAL WORKING MEMORY

by

CHING-I LU

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2012

MAJOR: COMMUNICATION
SCIENCES & DISORDERS

Approved By

____
Advisor Date

____

____

____

ACKNOWLEDGMENTS

I would like to thank all the people who have helped me in the process of completing this dissertation. First, I would like to express my gratitude to my advisor, Dr. Margaret Greenwald, for her support and guidance in academia and everyday life throughout my five years in the doctoral program. I would also like to thank my committee members: Dr. Anthony Cacace, an expert in audiology, who offered me an opportunity to learn TMS; Dr. Jean Andruski, an expert in linguistics, who has always nurtured me and offered me to attend conferences around the world; Dr. Mark Greenwald is the only person who was able to use a metaphor, such as the *Leitmotif* (leading motive) of Richard Wagner, to explain academic writing to me. My committee members have always offered insights that not only challenged my thinking, but also encouraged me to further improve my work.

I would like to thank Dr. Susan Bowyer for offering me an opportunity to learn MEG; Dr. Li Hsieh, as a friend, who has always supported me. Dr. Derek Daniels, as a collaborator, who allowed me access to his dataset for different analysis; I have also learned much about stuttering from him.

Special thanks to my friends who have been supporting me over these five years. Many thanks to Siny at UVA; when I felt frustrated, you have sent many cute pictures of Marbles to make me smile. Thank you so much. Many thanks to Jack at UW, Seattle, who came to the States the same year as I did. We encouraged and complained to each other by phone every day during our first year in our graduate program and traveled around to so many National Parks in US and Canada during these five years. Many thanks for Lai-Yee and Ting-Yi: we met each other in Detroit. I will always miss having dinner with you on Friday nights. We encouraged and supported each other; you are my friends forever. Many thanks for Fang-Tzu at BGSU, who is a long time friend from my college always supporting me. Thank you, Vanessa, for giving me a ride to grocery shopping every week when I did not have a car in my first year in Detroit. Thank you, Daniel at Yale U, your statistical and psychology knowledge has helped me clarify many concepts. Many thanks to Andrea, Carol, Denise, Nicole, and Sean. Once again, thank you, all my friends; I could not have achieved so much without your help.

I would also like to thank my younger brother and my sister-in-law, who took care of my parents when I was away from home and supported me financially. Without your help, I could not have completed everything in the US.

I would like to dedicate my dissertation to my parents. Thank you for your support and encouragement to follow my dreams; my dream would not have come true without you.

TABLE OF CONTENTS

Acknowledgments ii

List of Tables vi

List of Figures vii

Chapter 1 Introduction 1

Chapter 2 Review of the Literature 3

Cognitive-Linguistic Processes in Reading 3

A theoretical model of English reading 4

The Chinese writing system 8

A theoretical model of Chinese reading 12

A theoretical model of reading music 16

Components of Verbal Working Memory 17

The phonological loop 18

Characteristics of the phonological loop 19

Other factors affecting the phonological loop 26

The central executive 33

Working Memory for Tone Information 34

Working memory and reading Mandarin Chinese 35

Working memory for musical tone 37

Central executive function in musical working memory 41

Chapter 3 Method 47

Participants 47

Experimental Materials 48

Design of the Study 49

Procedure 55

Scoring and Data Screening 57

Chapter 4 Results 59

Comparisons of Performance Accuracy 59

Comparisons of Performance Speed 68

Comparing the Effects of Visually Similar Stimuli 75

The English Rhyming Task: Effects of Visual Similarity on Accuracy 77

The Mandarin Homophone Task: Effects of Visual Similarity on Accuracy 79

The Music Task: Effects of Visual Similarity on Accuracy 81

The English Rhyming Task: Effects of Visual Similarity on Reaction Time 85

The Mandarin Homophone Task: Effects of Visual Similarity on Reaction Time 87

The Music Task: Effects of Visual Similarity on Reaction Time 89

Chapter 5 Discussion 92

The Effect of Musical Training on Verbal and Tonal Working Memory 92

The Effect of Visual Similarity of Target Stimuli 98

Conclusion 100

Limitations and Future Studies 101

Appendix All Stimuli 104

References 112

Abstract 128

Autobiographical Statement 129

LIST OF TABLES

Table 3.1: Demographic Data of Musicians and Non-musicians 48

Table 3.2: Response Patterns of 1-back and 2-back Tasks 57

Table 4.1: Mean Accuracy Rates of 1-back Tasks by Group 60

Table 4.2: Mean Accuracy Rates of 2-back Tasks by Group 61

Table 4.3: Results of Three-Way Mixed Model ANOVA: Accuracy for Group by Task by Difficulty Level 62

Table 4.4: Mean Reaction Times of 1-back Tasks by Group 69

Table 4.5: Mean Reaction Times of 2-back Tasks by Group 69

Table 4.6: Results of Three-Way Mixed Model ANOVA: Reaction Times for Group by Task by Difficulty Level 70

Table 4.7: Accuracy for Four Categories of Target Stimuli by Group 77

Table 4.8: Reaction Time for Four Categories of Target Stimuli by Group 85

LIST OF FIGURES

Figure 2.1 An example of musical notation 16

Figure 2.2 Schematic of the phonological loop with visual and auditory input (Vallar, 2006). 30

Figure 3.1 Example stimuli for the Musical note task. 51

Figure 3.2 Number of items under each category in 1-back paradigm. 52

Figure 3.3 An Example of 1-back paradigm shown in a graphic format. 53

Figure 3.4 Number of items in each category in 2-back paradigm. 54

Figure 3.5 An Example of the 2-back paradigm. 55

Figure 4.1 Accuracy for all participants in the 1-back vs. 2-back tasks. 63

Figure 4.2 Accuracy rates for musicians and non-musicians in the 1-back tasks. 65

Figure 4.3 Accuracy rates for musicians and non-musicians in the 2-back tasks. 67

Figure 4.4 Reaction times for all participants in the 1-back tasks vs. 2-back tasks 71

Figure 4.5 Reaction times of musicians and non-musicians for the 1-back tasks. 73

Figure 4.6 Reaction times of musicians and non-musicians for the 2-back tasks. 74

Figure 4.7 Accuracy rates for each stimulus category for English rhyming task in both groups. 78

Figure 4.8 Accuracy rates for each stimulus category for Mandarin homophone task in both groups. 80

Figure 4.9 Accuracy rates for each stimulus category for the music task in both groups. 82

Figure 4.10 Reaction time for each stimulus category for the English rhyming task in both groups. 86

Figure 4.11 Reaction time for each stimulus category for the Mandarin homophone task in both groups. 88

Figure 4.12 Reaction time for each stimulus category for the music task in both groups. 90

CHAPTER 1

INTRODUCTION

Music and language require complex cognitive systems of the human mind. Both types of information processing depend on temporal memory storage, and there is some evidence that working memory systems for language and music may overlap. Despite an increasing interest in the comparison of working memory for language and music, there is a lack of behavioral paradigms for directly comparing these forms of working memory, especially as related to reading. No study has examined working memory for linguistic tone information in a tonal language, and few studies have examined working memory for musical tone (Berz, 1995; Ockelford, 2007), especially comparisons between musicians and non-musicians. There is little research on how working memory for phonological information compares to working memory for tone information, both of which are involved in working memory for tonal languages such as Mandarin Chinese (Jing & Lu, 2009; Lu & Zhang, 2007; So & Siegel, 1997; Xu & Li, 2009).

Studies addressing the relationship between verbal working memory and musical working memory have focused on the presumed effect of musical training on central executive function (Degé, Kubicek & Schwarzer, 2011; Hargreaves & Aksentijevic, 2011; Moreno, Bialystok, Barac, Schellenberg, Cepeda & Chau, 2011; Schellenberg, 2011). There is some evidence that musical training results in improved executive function, inhibitory control and verbal ability (Degé, Kubicek & Schwarzer, 2011; Moreno et al., 2011). Musical training appears to alter working memory for music, such that in recalling musical tones musicians are better able to discriminate pitches that are close together as compared to non-musicians (Williamson, Baddeley & Hitch, 2010). Musical skill training has been found to influence working memory capacity (Lee, Lu & Ko, 2007). There is some evidence that musical training not only changes musical working memory, but also verbal working memory (Franklin, 2008). Theoretical models of reading in English include verbal working memory as a necessary component of oral reading that operates prior to and during pronunciation. In English reading models, the same verbal working memory mechanism is thought to be involved in short-term storage of lexical (whole-word) or sublexical phonological information derived from print. There have been few studies of how verbal working memory functions in Mandarin Chinese, and no studies of working memory for linguistic tone. Musical working memory has received little research attention to date, particularly as related to how musical tone information is derived from visual musical notation.

The current study focused specifically on working memory in four visual recognition tasks, in which working memory load (*n*-back) was varied orthogonally while subjects judged English rhyming, Mandarin homophones, Mandarin tones, and musical notes. Musicians and non-musicians were included to examine the influence of musical training on task performance. The effects of increased task demands on working memory were also examined across tasks.

CHAPTER 2

REVIEW OF THE LITERATURE

This dissertation explored the translation from print to sound of the tonal language Mandarin and other reading stimuli in healthy volunteers. The performance of musicians and non-musicians was compared across a variety of reading tasks to examine whether musical training can facilitate reading of Mandarin tone. The effects of increasing working memory load on reading performance across tasks were also examined. This chapter will review current theoretical models of reading, verbal working memory, and working memory for tone information.

Cognitive-Linguistic Processes in Reading

Most research regarding cognitive-linguistic processes in reading has involved English reading (Acheson & MacDonald, 2009a; Binder & Borecki, 2008; Grainger, et al., 2005; Soto & Humphreys, 2007; Tydgat & Grainger, 2009). Theoretical models of reading have been based on studies of normal reading in healthy volunteers and evidence for how reading breaks down in brain-damaged individuals (Greenwald & Berndt, 1999; Greenwald, 2004; Rapp & Goldrick, 2000; Goldrick & Rapp, 2007; Chialant & Caramazza, 1998; Miozzo & Caramazza, 1998; Rosazza, Appollonio, Isella & Shallice, 2007). Patterns observed in acquired dyslexia provide a window into reading processes that are difficult to isolate in healthy volunteers for whom translation of print to meaning or sound occurs very rapidly. Normal reading and patterns of acquired dyslexia have also been simulated in computational models of normal reading and its disorders (Rapp & Goldrick, 2000; Jacobs & Grainger, 1992; Rey, Dufau, Massol & Grainger, 2009; Oberauer & Lange, 2009). To date, there have been relatively few studies of how specific reading subprocesses may be differentially impaired in acquired dyslexia in languages other than

English. Some current theoretical models of English reading are reviewed below, and then theories of reading in Mandarin Chinese are discussed.

A theoretical model of English reading. There are a variety of approaches to describing how print is translated to meaning and/or sound in English. In general, reading comprehension of words is thought to involve visual perception, visual word recognition, and semantic comprehension of word meaning. Oral reading is thought to require activation of an abstract phonological word form that is subsequently held in verbal working memory during pronunciation. However, there are different descriptions of how visual, semantic and phonological information are activated during reading, with various degrees of interactivity incorporated into theoretical models of reading.

Different versions of a dual-route model of reading English have been put forth, with the basic distinction between English reading by a whole-word (i.e., lexical-semantic) route and English reading via a grapheme to phoneme (i.e., nonlexical or sublexical) route that supports direct translation of print to sound without recognition or comprehension of the whole word (e.g., Marshall & Newcombe, 1973). The lexical-semantic route often has been described as supporting reading of all real words, both regularly spelled (e.g., PLANT) and irregular (e.g., YACHT) (Greenwald, 2004). The sublexical route (i.e., grapheme-to-phoneme conversion) has been described as supporting oral reading of regularly spelled words and nonwords (e.g., BERK) (Beeson & Hillis, 2001). There are differences in the way theoretical reading models represent the interaction of lexical and sublexical information, with variations of the dual-route models depicting the routes as distinct (e.g., Coltheart, Curtis, Atkins et al., 1993) and other models allowing for substantial influence of the lexical route on the sublexical route (e.g., Patterson & Hodges, 1992). The dual-route reading model has been challenged and in some models of reading there are not separate routes for lexical and sublexical processing (e.g., Plaut, McClelland & Seidenberg et al., 1996).

There is evidence for a third route in English reading: a lexical non-semantic route. For example, Schwartz, Saffran and Marin (1980) described an individual with dementia who was able to read aloud irregularly spelled words without comprehending their meaning. This pattern suggests that the patient was using a lexical route to pronounce the irregular words, but without semantic mediation. Although many individuals with semantic disruption may retain partial semantic knowledge of irregular words (Hillis & Caramazza, 1992), there is evidence for this lexical non-semantic reading route from other patients with severe damage to semantic memory (Greenwald & Berndt, 1998).

In addition to differences in the way that lexical and sublexical information are represented across English reading models, there are also differences in the degree to which activation of information of one type or "level" of cognitive processing is thought to influence activation of other levels of information. In a description of normal reading processes, Goldrick and Rapp (2007) reviewed five mechanisms of interactivity: 1) cascading activation: the flow of activation to a later level before a computation has been completed at the earlier level; 2) feedback: the flow of information from later to earlier levels; 3) connectivity distance: the directness of connections between levels; 4) domains of interactivity: the number and types of cognitive processes that are assumed to interact; and 5) seriality: the degree to which there are processing steps or decision points between input and output.

In the Interactive Activation Model (Rey et al., 2009), cognitive processing among reading subcomponents is conceived to be cascaded and interactive, including feedback of information to previous levels. There is some evidence that in cascaded-interactive processing early letter activation feeds forward and partially engages word representations, and feedback to the letter level influences subsequent processing (e.g., Behrmann, Plaut & Nelson, 1998). Based on behavioral and event-related potential (ERP) results from healthy volunteers, Rey and colleagues (2009) described letter perception with feedforward excitatory connections from the feature to the letter levels, as well as lateral inhibition at the letter level and excitatory feedback from the feature levels.

In contrast, restrictive models of lexical interactivity depict little or no interaction of one level to the next. For example, in the Discrete Feedforward Account (DFA), as reviewed by Goldrick and Rapp (2000), semantic, lexical, and phoneme information is activated in a strictly forward direction. That is, stimuli are confined to the current processing stage only, and only the item selected at the end of a given stage is processed at the following stage. On this account, there is a clear selection point at the end of each processing stage. For example, conceptual information is selected at the semantic level (e.g., furry, feline, pet, and warm), subsequently the corresponding word form is selected at the lexical level (e.g., CAT), and next the corresponding phonemes are selected (e.g. /k/ /æ/ /t/). Roelofs and colleagues (Levelt, Roelofs & Meyer, 1999; Roelofs, 2004) have argued for such restricted interaction among semantic and phonological processes in verbal production tasks. To date, researchers have yet to agree on the degree of interactivity among cognitive subprocesses involved in reading

Despite differences in the way that lexical and sublexical information are represented across these different theoretical models of English reading, and the extent to which the models include interactivity among reading subprocesses, it is generally accepted that reading comprehension requires activation of word meaning in semantic memory and that oral reading requires activation of abstract phonological information needed for pronunciation. Lexical (i.e., whole-word) phonological information also can be used to perform other related tasks such as deciding whether two printed words rhyme with one another. The English rhyming task has often been used to examine the integrity of an individual's "phonological output lexicon," that is, the memory store of abstract whole-word phonological forms (Beeson & Hillis, 2001). The rhyming task consists of deciding whether two words rhyme or not, and the stimuli can be two written words or two pictures. This task would be difficult for individuals who fail to activate a full lexical phonological representation from the visual stimulus (Raymer & Rothi, 2001). In the rhyming task, the examiner would want to control the visual similarity of stimuli to ensure that the individual is not performing the task based solely on the visual similarity of the printed words. For example, stimuli should include word pairs that rhyme but are visually dissimilar (e.g., loose vs. juice) as well as word pairs that do not rhyme but are visually similar (e.g. fever vs. never).

The rhyming task described above requires verbal working memory in that two phonological word forms must be held in memory temporarily while their phonological characteristics are compared. Working memory refers to the capacity to hold information online for some purpose, such as understanding conversation (Crosson, 2000). The concept of working memory assumes "… a limited capacity system, which temporarily maintains and stores information, supports human thought processes by providing an interface between perception, long-term memory and action" (Baddeley, 2003, p. 829). Verbal working memory also has been termed verbal short-term memory (Campoy & Baddeley, 2008; Lee, Lu & Ko, 2007; Tree, Longmore & Besner, 2011) or the phonemic buffer (Beeson & Hillis, 2001; Blumstein, 1998). Verbal working memory is a component of most theoretical models of English reading in that once the abstract phonological information is obtained (either lexically or sublexically) it must be held in working memory during pronunciation.

Verbal working memory is thought to be required for pronunciation of any reading stimulus, whether English or a different language. However, comparatively few studies have examined reading and verbal working memory in languages other than English. The unique characteristics of some writing systems, such as Chinese, may not be captured fully by theoretical models of reading in alphabetic languages. Details of the Chinese writing system are reviewed below, as well as theoretical models of Chinese reading.

The Chinese writing system. The modern Chinese writing system is considered to be logographic, in that the basic unit or symbol (i.e., the character) of Chinese is associated with a unit of meaning (i.e., morpheme) in the spoken language (Weeks, Chen & Gang, 1997). Chinese includes different dialects, such as Mandarin and Cantonese. Mandarin is the official language in Mainland China and Taiwan and is used most all over the world. Cantonese is mainly spoken in Guangdong, Guangxi, and Hong Kong, and is the most extensively spoken Chinese dialect after Mandarin. The current study will be focused on Mandarin Chinese. Mandarin is known as a tonal language in that every syllable is differentiated according to one of four tones or voice inflections (Lin, Wu, Ting & Wang, 1996; Shriberg & Kent, 2003; Taft & Chen, 1992; Wang, Jongman & Sereno, 2001).

Written Chinese is always presented using characters having a square shape as a basic writing unit. The principles of formation of Chinese characters are derived from the six categories of characters: pictographs, ideographs, compound ideographs, loan characters, analogous characters, and phonetic compounds (So & Siegel, 1997; Leong, 1986). Pictographs mean form imitation; that is, stylized drawings of the objects they represent. Ideographs, so-called indicatives, express an abstract idea through an iconic form, including iconic modification of pictographic characters. Compound ideographs, so-called compound indicatives, mean two or more pictographic or ideographic characters combined to form a third meaning. Loan characters, so-called borrowed characters, are “borrowed” to write another homophonous or near-homophonous morpheme (i.e., a morpheme pronounced the same or nearly the same as another morpheme with a different meaning). Analogous characters, so-called derived characters, have similar meanings and often the same etymological root, but diverge in pronunciation.

The sixth category of Chinese written character is the phonetic compound, so-called phono-semantic compound characters. The majority of all Chinese characters are phonetic compounds, accounting for over 80% of all Chinese characters (So & Siegel, 1997; Leong, 1986). Phonetic compounds consist of two parts, a phonetic component radical and a semantic radical. The phonetic radical is an existing character pronounced approximately the same as the new target word, and the semantic radical is an often graphically simplified element with the same general meaning as the new target word. The phonetic radical of the compound acts as a rhyming clue, whereas the semantic radical part suggests the semantic domain of the word. For example, a compound character (e.g. 櫻 /Ying1/, cherry) typically consists of a semantic radical (e.g.木, wood) that provides the clue about the meaning of the character, and a phonetic radical (e.g. 嬰 /Ying1/) that gives information about the character’s pronunciation (Shu et al., 2005).

The square shape of the written Chinese character conveys two parts of phonological information: phonetic segments (consonant and vowel) and a suprasegmental phonological feature (tone) (Hallé, Chang & Best, 2004; Lee, Tao & Bond, 2008; Leong, 2002; Siok & Fletcher, 2001; Spinks, Liu, Perfetti & Tan, 2000, Tong, Francis & Gandour, 2007; Taft & Chen, 1992; Wang, Jongman & Sereno, 2001; Wang, et al., 2008). Each character corresponds to one syllable and one tone, making up a Chinese monosyllable. The typical Chinese monosyllable consists of three elements: 1) the onset (an initial consonant preceding the vowel), 2) the rime (at least one vowel and any consonant sounds that come after the vowel), and 3) the tone (Siok & Fletcher, 2001). However, some syllables consist only of an initial sound (onset) or a final sound (rime) and a tone.

Word meanings in tonal languages vary as a function of the tone associated with each syllable (Shriberg & Kent, 2003). Tone (i.e., lexical tone) includes many phenomena that determine the patterns of pitch rises and falls in a language. Lexical tone refers to how pitch is used to convey semantic meaning in speech. Lexical tone is part of the phonemic structure used to distinguish words. It indicates the regulation of fundamental frequency to produce contrast, such as falling pitch, rising pitch, rising-falling pitch (a rising pitch segment followed by a falling pitch segment) or level pitch (no change in vocal pitch).

In Mandarin Chinese, four different tones are used as suprasegmental phonological features that change the pitch of the syllable and provide lexical contrast (Siok & Fletcher, 2001). Based on the pattern of pitch contour, the 1st tone is the high-level tone, the 2nd tone is mid-rising tone, the 3rd tone is mid-falling-rising tone, and the 4th tone is the high-falling tone. Mandarin also includes one neutral tone (referred to as Tone 5). This 5th tone does not have a specific pattern of pitch contour (Lin et al., 1996).

Taft and Chen (1992) provided evidence that words bearing the same rime with different tones may sound perceptually more similar to the Chinese listener than words with different rimes. This implies that Chinese listeners may find it more difficult to differentiate and identify words with different tones and the same rime (e.g. 曲 /qu 3 /, song vs.去 /qu 4/, go) than those with the same tone but different rime (e.g.氣 /qi 4 /, air vs.去 /qu 4/, go).

Because spoken Chinese has only 1200 syllables but over 5000 commonly used morphemes, characters often have the same pronunciation (i.e., they are homophones). Somewhat like

English homophones (e.g., “allowed” and “aloud”), these Chinese homophones are pronounced exactly the same way, with the same phonetic information and the same tone. To understand the meaning of a character, the Chinese reader has to distinguish among homophonic morphemes. For example, in Chinese characters, homophonic morphemes indicate 是 (/shi4/ yes), 試(/shi4/ try), 室(/shi4/ room), and 市 (/shi4/ city) (Shu et al., 2005). The meaning of the homophone is determined based on how the character looks and also based on sentence context.

Although it has been argued that orthography plays a more important role than phonology in Chinese reading (Lu & Zhang, 2007), the importance of phonological skills has been emphasized in several studies of Chinese reading (Li & Ho, 2011; Taft & Chen, 1992; So & Siegel, 1997; Shu et al., 2005). Skill level in oral reading of Chinese has been found to be highly correlated with performance in tone and rhyme discrimination tasks (So & Siegel, 1997).

Learning to read Chinese appears to progress from an early logographic stage in which visual-orthographic skill is paramount, to a later orthographic-phonological stage in which conversion of the visual stimulus to phonology plays an important role. Soik and Fletcher (2001) investigated the roles of phonological awareness and visual-orthographic skills in Chinese reading development. Phonological awareness is the ability to conceive of spoken words as smaller units of sound segments, including syllable onsets and rimes. Based on evidence from Chinese children in Grades 1 to 5, Soik and Fletcher (2001) suggested that phonological awareness and visual skill both are important factors in Chinese reading development and that the overall success of Chinese reading can be predicted by the level of phonological awareness (Soik & Fletcher, 2001; Tan, Hoosain & Siok, 1996).

Wang, Jongman and Sereno (2001) studied how native Mandarin Chinese-speaking children perceived and learned Mandarin tones. Their results indicated that the lexical tone system, along with other pitch-related abilities such as intonation, was acquired before the segmental system of consonants and vowels. Also, the children in this study first perceived and processed lexical Mandarin tone as part of the intonation system of the tonal language.

A theoretical model of Chinese reading. Due to the characteristics of the Chinese writing system, theoretical models of reading in alphabetic languages may not provide an adequate framework for understanding normal reading and dyslexia in Chinese. Tan and colleagues (1996) argued that because written Chinese is a meaning-based logography, the reader first must know the meaning associated with the Chinese written character to activate its correct pronunciation. On this account, Chinese characters map onto the morpheme (meaning) and cannot be pronounced by direct access to phonology. Because Chinese characters represent morphemes, awareness of morphemes as linguistic units is necessary for understanding the Chinese writing system (Shu, Meng, Chen, Luan & Cao, 2005).

However, Weekes, Chen and Gang (1997) proposed that oral reading of Chinese characters does not require initial access to meaning prior to activation of phonology but can be accomplished via an alternative non-semantic route. Based on evidence from a Chinese patient with impaired oral picture naming (58% accurate) but perfect performance in oral reading of Chinese, Weekes and colleagues argued that Chinese readers can map orthographic units directly onto phonological output without semantic mediation. In contrast, picture naming requires mapping onto orthographic units via semantic pathway before verbal production. The pattern of performance of the patient in this study was interpreted as reflecting impairment to the connection between the semantic system and phonology, which reduced oral picture naming performance but did not impair oral reading.

Weekes and colleagues (1997) have proposed a model of Chinese reading based on the dual-route model of English reading described above. In this model, normal reading of Chinese is conceptualized as involving three interconnected levels of representations: orthographic, semantic, and phonological. As in the English reading models, normal visual recognition precedes activation of the meaning of the visual stimulus with the semantic system, or activation of phonological codes directly via a non-semantic route. Chinese radicals and characters are represented in this model as independent orthographic units in the Chinese word recognition system. This is because many radicals are pronounced differently than the whole characters that contain them, and some dyslexic Chinese readers are able to retrieve the names of radical components but not the names of the whole characters that contain them. According to Weekes and colleagues (1999), it is possible that morphemes are also represented as independent units in the Chinese word recognition system. They noted that to read a Chinese character correctly, the reader must know the pronunciation that is associated with the character as a whole-word.

The relationship between semantic activation and phonological activation in Chinese reading is further described in a semantic priming study by Zhou and Marslen-Wilson (1999a). The semantic priming design allows experimenters to measure the extent to which semantic processing of a target stimulus is aided by the prior presentation of a related or unrelated stimulus (i.e., the prime). In a semantic priming task with healthy volunteers, Zhou and Marslen-Wilson (1999a) examined whether phonological and orthographic information in the primes affected semantic activation in reading of Chinese. They found that semantic activation of base words was influenced by the phonological characteristics of the prime, but that it was the interactions among phonology, orthography and morphology that determined semantic activation. They suggested that interactive processes in semantic activation also involve feedback from semantic activation to orthographic and phonological activation. From their point of view, this study does not support the hypothesis that the Chinese reader must understand the meaning of the written character prior to accessing phonology.

There is evidence from eye movements during Chinese reading that phonological codes influence the identification of Chinese characters. Pollatsek, Tan and Rayner (2000) studied eye movements during oral reading of target Chinese characters. Participants were presented with a "preview" stimulus (i.e., a single Chinese character) in the parafovea to the right of visual fixation. When the participant moved their eyes to the character location, a target character replaced the preview during the eye movement to the target. The relationship between preview and target characters was controlled to be visually similar or dissimilar, homophones or nonhomophones, synonyms or non-synonyms, sharing a phonetic radical or sharing a semantic radical, and considered to be a phonetically regular compound character (i.e., the phonetic radical in isolation was pronounced the same as the compound character) or a phonetically irregular character (i.e., the phonetic radical in isolation was pronounced differently from the compound character).

Participants obtained a similar degree of preview benefit from both the orthographically similar and the orthographically dissimilar homophones, which is evidence that the effect cannot be due to the orthographic similarity of preview and target. However, the regular Chinese characters were pronounced faster than the irregular Chinese characters, suggesting that phonological codes were involved in the identification of the Chinese target characters. Based on the results of this experiment, Pollatsek and colleagues (2000) hypothesized that for phonetically regular characters, phonological processing of phonetic radicals cooperates with phonological processing of whole characters, leading to a faster activation of phonological information than for phonologically irregular characters.

The Chinese reading performance of three children with developmental dyslexia described by Shu and colleagues (2005) provided support for the model of Weekes, Chen and Gang (1997). The results showed that the development of the semantic pathway and the non-semantic pathway can be selectively delayed in learning to read Chinese. The dissociation between the two types of developmental dyslexia suggested in these cases provided evidence that a semantic and a non-semantic pathway exist independently in Chinese reading. Shu and colleagues (2005) hypothesized that the reading model of Weekes and colleagues can be extended to account for deficits in morphological and phonological awareness that impeded normal developmental of the semantic and non-semantic reading pathways in these three children.

Li and Ho (2011) reported that tone discrimination and tone production were found to correlate significantly with Chinese word reading ability. They found that Chinese dyslexic children show weaknesses in tone awareness. They concluded that accurate tone processing appears to clarify or eliminate confusions between phonological and morphological representations. Moreover, some Chinese characters, such as homophonic words with different tones, may confuse non-skilled readers, such as dyslexic children.

As noted above in regard to English reading, it is generally believed that phonological information is held short-term in a verbal working memory store prior to and during oral reading. Few empirical studies have addressed the issue of verbal working memory with phonemic information (consonant and vowel) in a tonal language, as discussed below. No empirical study has been reported to address the issue of suprasegmental feature (tone) in verbal working memory. Translation of tone information from print to sound also can be accomplished in reading of musical notation. A theoretical model of translation of tone in music reading is discussed in the following section.

A theoretical model of reading music. Reading of musical notation (Western music) involves figure-ground integration or perception. As shown in Figure 2.1, the ground reveals a stave, whereas figures are other symbols, including clefs, time signatures, types of notes, types of rests, other symbols (e.g., sharp, fermata, repeat), loudness level of notes (e.g., *ff*), jargon affecting notes (e.g.. *legato*), and jargon affecting speed of notes (e.g., *rit.*). Most of these symbols are in the time-based domain and the pitch-based domain, whereas spatial representation indicates pitch height. The pitch-based domain is used to map vertically organized spatial location, whereas the time-based domain is used to map horizontally to a sequence of musical events from left to right on a stave (Steward, Walsh & Frith, 2004). Deriving pitch from reading musical notes requires translation of the visual stimulus into pitch.

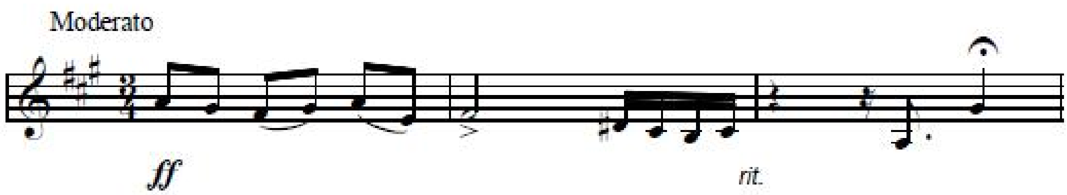


Figure 2.1 An example of musical notation.

Tone in music means pitch. As defined by Deutsch (1999): “Pitch is the most characteristic property of tones, both simple (sinusoidal) and complex. Pitch systems (like the diatonic-chromatic and the 12-tone systems) are among the most elaborate and intricate ever developed in Western and non-Western music. Pitch is related to the frequency of a simple tone and to the fundamental frequency of a complex tone (pp. 93-94).”

The Western tonal music system is highly structured and contains a set of twelve semitones in an octave. The twelve semitones are combined into subsets of seven tones that make up a scale; a scale contains seven pitches per octave. Pitches separated by an octave are heard as very similar and are typically given the same name, referred to as the pitch class (e.g., all the notes called “A” on a piano keyboard) (Krumhansl & Toiviainen, 2003; Patel, 2008).

Schön (2002) offered a preliminary model of music sight-reading that includes three phases: visual encoding, transcoding, and production levels. However, this model does not include a detailed description of abstract internal representations that would be involved in the transcoding of print to sound, and there are no other theoretical models of cognitive processes involved in reading musical notation. Based on the English reading model described above, it could be hypothesized that transcoding from musical notation to sound involves visual perception, visual recognition, comprehension, and musical working memory before production.

Although there are no theoretical models of reading musical notation in the literature that include a music working memory component, Gudnubdsdittur (2010) claimed that sight reading required general mental capacities such as working memory and mental speed. Lehman (2007) also suggested that the task of music reading is demanding of short-term (i.e., working) memory. Furthermore, for singing or instrumental production, tone information would need to be held in working memory as is needed for verbal production of English or Chinese in the reading models above. Verbal working memory is described below, and the nature of working memory for tones in music also will be addressed below.

Components of verbal working memory

The multi-component model of working memory developed by Baddeley and colleagues (Baddeley & Hitch, 1974; Baddeley, 2003) includes four main components: the phonological loop (which deals with sound, phonological and verbal information); the visual-spatial sketchpad (which temporally stores spatial and visual information such as color and shape information); the episodic buffer (which links different domains such as spatial, visual, sound and verbal information); and the central executor (a flexible system for controlling and regulating cognitive processes). Of these four components, the phonological loop has been linked to language function more often than the other components (Baddeley, 2003; Repovs & Baddeley, 2006; Williamson, Baddeley & Hitch, 2010).

In verbal production tasks, including oral reading of English and Chinese, abstract phonological information is generally believed to be held short-term in a verbal working memory store prior to and during verbal production. Verbal working memory is a key component for verbal production tasks in general, and plays an important role in language learning (Song & Cheng, 2006; Gottardo, Chiappe, Yan, Siegel & Gu, 2006; Wang, 2001; Yeung, 2007). Verbal working memory is often described as including the phonological loop and the central executor (i.e., the central executive; Crosson et al., 1999; Mueller et al., 2003; Nadeau et al., 2000)

The phonological loop. The phonological loop is proposed to be the verbal working memory subsystem that is used to hold verbal or tone information, whether the stimulus is presented in auditory form or from other input modalities (Baddeley, 2003; Sweet et al., 2008). Three main phenomena of working memory are proposed to be associated with the phonological loop (Acheson & MacDonald, 2009b; Buchsbaum & D’Esposito, 2008; Mueller, Seymour, Kieras & Meyer, 2003). The first of these (i.e., phonological coding of stored word sequences) is based on the assumption that during verbal working memory tasks word sequences have to be memorized, coded and stored as temporary phonological representations. The second phenomenon (i.e., information loss through time-based decay) is based on the assumption that it requires time to refresh the memory trace for storing items in verbal working memory. The third phenomenon (i.e., memory-trace retention by strategic articulatory rehearsal) is based on the assumption that articulatory rehearsal can refresh working memory and that articulatory durations measured for word sequences are reliable predictors of memory spans.

The phonological loop has been described as being composed of two parts: 1) a phonological store; and, 2) an articulatory rehearsal process (Repovs & Baddeley, 2006; Awh et al., 1996; Chen & Desmond, 2005; Baddeley & Larsen, 2007; Buchsbaum & D'Esposito, 2008; Baddeley, 2003). The phonological store, or buffer, has been described as that component of verbal working memory that interfaces with other perceptual and mnemonic systems. It is also known as an input storage system, which can hold speech related information for one to two seconds (Rudner & Rönnberg, 2008; Vallar, 2006). The phonological store is thought to enable the formation and maintenance of multidimensional representations (Rudner & Rönnberg, 2008). It is assumed to have a limited capacity of about four chunks or episodes, and to be accessible through conscious awareness (Baddeley, 2010). Auditory memory traces in the phonological store are thought to decay over a period of one to two seconds, unless refreshed by articulatory rehearsal (Rudner & Rönnberg, 2008). The articulatory rehearsal process (i.e., the output rehearsal process or articulatory mechanism) serves to refresh contents of the phonological store subvocally, thus allowing the system to maintain short sequences of verbal items in memory for an extended interval.

Characteristics of the phonological loop. Findings from a variety of working memory experiments have allowed researchers to elaborate descriptions of the phonological loop. These findings include the phonological similarity effect, the word length effect, the serial position effect, the irrelevant sound effect, and the concurrent articulation effect, described below. Many researchers argue that all five effects support the assumption of a separable phonological store and an articulatory mechanism (Acheson & MacDonald, 2009a; Acheson & MacDonald, 2009b; Mueller et al., 2003) though Buchsbaum and D'Esposito (2008) maintained that the serial position effect is not related to the phonological loop.

Experimental analysis of the phonological store usually depends on measures of memory capacity and interference tasks (Repovs & Baddeley, 2006). Jones, Hughes and Macken (2007) defined the phonological store concept in terms of six key characteristics: 1) encoding of verbal materials in phonological form, 2) direct access to the store by auditory stimuli, 3) indirect access to the store via a grapheme-to-phoneme conversion process of visual-verbal stimuli, 4) rapid loss of phonological representation through decay, 5) restoration of phonological representations, and 6) loss of information in the store through interference based on phoneme similarity.

Experimental analyses of the articulatory mechanism have incorporated a variety of tasks in which experimenters examine the effects of disruption to articulatory rehearsal with concurrent articulation, described below. For example, concurrent articulation has been observed to interfere with repetition (Coltheart, 1993), rhyme judgments and homophone judgments (Tree, Longmor & Besner, 2011), immediate serial recall (Gupta & MacWhinney, 1995), reading in the Stroop task (Chmiel, 1984), and reproduction of rhythm (Saito & Ishio, 1998).

The phonological similarity effect. There is evidence that it is much more difficult to recall a set of phonologically similar than dissimilar words (Acheson & MacDonald, 2009a; Baddeley, 2003; Jones, Hughes & Macken, 2007; Lobley, Baddeley & Gathercole, 2005; Mueller, Seymour, Kieras & Meyer, 2003). This effect is thought to result because items that are specified by each cue are encoded phonologically, with similar items having fewer distinguishing cues (Baddeley, 2003). The phonological similarity effect has been observed in a listening span task in healthy volunteers, and interpreted as reflecting reliance on a phonological code during information retention in the working memory store (Lobley, Baddeley & Gathercole, 2005).

When phonologically similar stimuli are presented, a common type of speech error is an onset exchange in which the initial consonants in a bi-syllable exchange their serial positions (e.g. *She sells seashells…*as *She shells sea sells*). Another variation of exchange error in recalling phonologically similar stimuli is to misorder initial phonemes (e.g., incorrectly saying /piy/ and /siy/ for the target utterance /siy/ and /piy/) (Acheson & MacDonald, 2009a). Acheson and MacDonald (2009a) argued that phonological similarity is a type of contextual similarity, reflecting interaction between phonological similarity and phoneme position. On this account, serial ordering errors in different syllable positions emerge when there is greater activation for an incorrect phoneme than for a correct one at a given position.

After reviewing data from several studies in which the phonological effect was observed (e.g., Caplan et al., 1992; Caplan and Waters, 1994; Lovatt et al., 2000), Mueller and colleagues (2003) concluded that different experimental task instructions induced different experimental results. For example, instructions that discouraged participants from using verbal rehearsal to perform the serial recall task but encouraged participants to adopt other nonverbal rehearsal strategies caused relatively little difference between serial recall accuracy for phonologically similar versus phonologically dissimilar words (Caplan, Rochon & Water, 1992; Mueller et al., 2003). In contrast, instructions that encouraged participants to use articulatory rehearsal resulted in significantly lower serial recall accuracy for phonologically similar than for phonologically dissimilar words (Caplan & Water, 1994). Other linguistic variables such as frequency, familiarity, number of phonemes, and semantic associations also modified the phonological similarity effect (Lovatt et al., 2000). With high error rates (e.g., > 50%) the phonological similarity effect tends to disappear, indicating that subjects are abandoning the loop for alternative strategies such as semantic or visual coding (Baddeley, 2003).

The word length effect. It is easier to recall a set of short words than a set of longer words (i.e., with length defined as number of phonemes) (Acheson & MacDonald, 2009a). The most widely accepted account of this word length effect is that it reflects time-based decay of information in the phonological store, in that longer words take longer to produce and thus allow more decay (Acheson & MacDonald, 2009a; Neath, Bireta & Surprenant, 2003). Caplan and colleagues (1992) investigated the articulatory determinants of the word length effect on memory span tasks. They found that when words are equated for number of phonemes, it is the phonological structure of a word, not features of its actual articulation, that determines the word length effect in span tasks. However, they noted that if overt or subvocal articulatory rehearsal is used as an optional memory strategy it also can result in a word length effect.

Neath, Bireta and Surprenant (2003) suggested that the word length effect might be an artifact of the particular set of stimuli used in the recall task. They used stimuli from previous studies and their own new stimuli to examine the word length effect by pronunciation time. They found that only one set of stimuli from Baddeley et al. (1975) showed the expected word length effect. The stimuli from Caplan and colleagues (1992) showed a reverse word length effect (i.e., long items recalled more than short items), and the stimuli from Lovatt and colleagues (2000) and from Neath and colleagues (2003) resulted in no significant differences in recall for short versus long words.

Mueller and colleagues (2003) reported that most evidence for the word length effect in verbal serial recall derives from studies in which articulatory duration (i.e., the time taken to pronounce a word) and phonological complexity (i.e., the number of phonemes and syllables in a word) were confounded. Their review indicated that stimuli that differed in both number of phonemes and articulatory duration produced a word length effect, whereas the stimuli that differed only in their articulatory duration and not in number of phonemes did not produce a word length effect.

Another factor that can impact the word length effect is list length. Baddeley (2010) noted that when list length is increased from four to eight words, and several learning trials are allowed, the expected pattern reverses such that the longer word list is recalled more accurately than the shorter word list (Baddeley & Hitch, 1974; Baddeley, Chincotta, Stafford & Turk, 2002). They attributed this finding to the influence of long-term memory in that repeated learning trials appear to force greater reliance on long-term memory, making word meaning the crucial factor in facilitating recall.

The serial position effect. The serial position effect is a phenomenon in which recall of auditory stimuli presented in a list is affected by list position. Stimuli in the first and second positions and stimuli in the final position are remembered better than stimuli in the middle list positions (Baddeley, 1986; Acheson & MacDonald, 2009a). Better recall of the initial stimulus in the list (i.e., the primacy effect) is thought to result because the first one or two items receive a greater amount of rehearsal than the latter stimuli and thus are stored in long-term memory. Superior recall of the last stimulus in the list as compared to middle stimuli (i.e., the recency effect) is thought to result because the last stimulus has less decay due to its more recent presentation (Baddeley, 1986; Medin, Ross & Markman, 2004).

In a recent review, Acheson and MacDonald (2009a) described three important characteristics of the serial position effect, and described it as being involved in recall of word list positions and also in recall of phoneme positions within a syllable. First, they noted positional constraints that appear to affect error production in list recall. That is, a common error in serial recall tasks is the production of a correct element but in an incorrect serial position only one or two positions earlier or later than its correct location. Second, they noted primacy and recency effects in syllable position and noted that onset and offset syllable positions are “edges” of the syllable, which means that both of these (first and last) positions in the syllable are particularly distinct because there are no items preceding or succeeding them. Lastly, these authors note that because the first and last positions have fewer positions over which transposition errors can occur, their recall is better than the middle items. These three characteristics of the serial position effect appear to reflect features of the phonological store, which has a limited capacity to hold information.

In a study of visual identification of letters, digits, and symbols presented in strings, Tydgat and Grainger (2009) found that viewing fixation, viewing position, and visual field interact with the serial position effect to influence letter and word recognition. Experimental manipulations of visual presentation revealed that letters and digits were recalled with much higher accuracy than symbols in the first position of the string, and there was a final-position advantage for recall of letters and digits compared with symbol stimuli. The authors concluded that the receptive field size of retinotopic letter and digit detectors has adapted to the need to optimize the processing of letter and digit strings. In other words, the size and shape of the visual receptive fields has changed as a result of experience in reading words and numbers; for example, the smaller the receptive field the less interference there would be from neighboring characters in the string.

The irrelevant sound effect. The irrelevant sound effect occurs when performance in immediate list recall is significantly impaired by the presence of to- be- ignored irrelevant sound (i.e., an irregularly changing acoustic stream) during presentation of list items visually (Johns & Macken, 1993; Schendel, 2006) or auditorily (Schendel, 2006). Jones and colleagues (1993) found that this effect can be induced with a variety of forms of irrelevant stimuli including variable tones, speech or music (Repovs & Baddeley, 2006; Acheson & MacDonald, 2009a).

The irrelevant sound effect is thought to reflect the results of competition between the irrelevant sound and the presented list stimuli, disrupting storage and representation of serial order within the phonological store (Baddeley, 2003; Repovs & Baddeley, 2006). It appears that acoustic information that varies irregularly (e.g., a foreign language) may be noisier and thus more disruptive to serial list recall than "regular" sound (e.g., white noise or one's native language) (Colle & Welsh, 1976).

Alley and Greene (2008) explored the effects of three different types of irrelevant sound on verbal working memory: vocal music (i.e., music with lyrics), equivalent instrumental music (i.e., the same melody of vocal music without lyrics), and irrelevant speech. In a digit span task, participants showed the best recall performance in the silent (control) condition, followed by the instrumental music condition and then the irrelevant speech condition, whereas the worst performance was demonstrated in the vocal music condition. Because the two most difficult conditions involved language (i.e., lyrics and speech), the authors concluded that the participants used their phonological store to memorize verbal information and therefore irrelevant speech or lyrics presented simultaneously interfered with recall.

The concurrent articulation effect. Concurrent articulation, also known as articulatory suppression, occurs when an individual says continuous irrelevant speech sounds (e.g., "the, the, the…") while simultaneously listening to stimuli in an immediate serial recall task (e.g., Gupta & MacWhinney, 1995). Concurrent articulation can disrupt performance in nonspeech tasks such as rhyme judgments, homophone judgments (Tree, Longmor & Besner, 2011) and reproduction of rhythm (Saito & Ishio, 1998). Concurrent articulation is thought to block rehearsal and thus to prevent the transformation and storage of information in the phonological store (Baddeley, 2010). This effect supports the idea that subvocal articulation in real time refreshes the decaying memory trace within the phonological store (Repovs & Baddeley, 2006). Also, overt concurrent articulation may provide acoustic masking of the stimuli that could disrupt perception of phonologically similar letters, such as *CDGPTV* (Baddeley & Larsen, 2007). As noted below, the effect of concurrent articulation can vary depending on stimulus modality and type.

Other factors affecting the phonological loop. The five characteristics of the phonological loop described above can diverge under different experimental conditions, such as presentation of stimuli through different sensory modalities. Long-term memory also interacts with operation of the phonological loop, as described below.

Visual-orthographic factors. The results of verbal working experiments can differ for word stimuli presented in written versus auditory format. Concurrent articulation erases the phonological similarity and word length effects in serial recall for visual stimuli but not auditory stimuli (Acheson & MacDonald, 2009a; Baddeley & Larsen, 2007; Rudner & Ronnberg, 2008). That is, for visual stimuli there is no significant difference in performance given phonologically similar versus dissimilar stimuli and no significant difference given short versus long words lists.

One explanation for this pattern is that concurrent articulation blocks visual information from being recoded into phonological form and entering the phonological store whereas auditory-verbal stimuli have direct access to the phonological store without requiring recoding from orthography (Acheson & MacDonald, 2009a; Baddeley & Larsen, 2007). This may be a learning effect in that adults compared with Children have more experience in mapping from auditory stimuli to meaning or articulation than they do in mapping from orthographic stimuli. There is some evidence that speech production processes may be involved in developing the phonological code for written input, and concurrent articulation would interfere with this process (Acheson & MacDonald, 2009a).

To examine the concurrent articulation effect in the visual modality, Tree, Longmore and Besner (2011) conducted experiments with healthy volunteers involving phonological judgments of print. They found that in a rhyme judgment task, concurrent articulation increased errors for both printed word and nonword stimuli, whereas it slowed reaction times for words but not for nonwords. The authors interpreted these results as reflecting the dual-route reading model in that whole words read via the lexical route must be phonologically segmented in short-term memory, while nonwords read via the sublexical grapheme-to-phoneme conversion route are already segmented. Concurrent articulation is thought to disrupt the segmentation of words as reflected in slower reaction times.

In the Tree et al. (2011) study, participants demonstrated faster reaction times in making auditory rhyme judgments of similarly spelled word pairs (e.g., COT-POT) as compared to word pairs that rhymed but had less orthographic overlap (e.g., YACHT-POT). When given visual stimuli in the rhyme judgment task, the participants were more likely to respond ‘Yes’ to non-rhymes that were orthographically similar as compared to orthographically dissimilar stimuli, and they were more likely to respond ‘No’ to rhymes that were orthographically dissimilar as compared to orthographically similar. Both of these patterns reflect reliance on visual-orthographic information rather than on phonological information in performing rhyme judgments.

Orthographic similarity also affected performance in the homophone judgment task in the same study. With concurrent articulation false positive rates (i.e., incorrectly judging word pairs to be homophones) increased for orthographically similar items as compared to orthographically dissimilar non-homophone pairs, and false negatives increased with orthographically dissimilar items (i.e., incorrectly rejecting visually dissimilar homophone pairs) as compared to orthographically similar homophone pairs.

Tree and colleagues (2011) demonstrated that there is a complex interaction between phonological and orthographic codes and the degree of phonological disruption from concurrent articulation. They suggested that researchers take more notice of how orthographic similarity can affect performance in phonological working memory experiments. Baddeley and Larsen (2007) also suggested when examining verbal working memory, researchers should consider the results based on visual versus verbal input.

Stimuli controlled for visual similarity at the word level (e.g., similar: *fly, cry*; dissimilar: *lie, sigh*) and at the letter level (e.g., similar: *Kk, Ww*; dissimilar: *Dd, Rr*) were incorporated into a verbal serial recall experiment with healthy volunteers (Logie, Della Sala, Wynn & Baddeley, 2000). At both the word and letter levels, participants recalled fewer visually similar items, with or without concurrent articulation. When working memory experiments involve visual stimuli (e.g., random letter sequences), normal healthy subjects may use more than one form of coding for retention. Use of multiple (visual, phonological, and semantic) codes may augment recall performance, relative to use of single codes. Logie and colleagues (2000) assumed that sequential orthographic stimuli can be held in some kind of visual temporary memory store.

Logie and colleagues noted that participants report various strategies in performing verbal serial recall, including visual imagery, semantic coding, or subvocal rehearsal. In part, this heterogeneity may reflect task instructions in that subjects may use a visual code if they are required to retain information about visual form, but in contrast rely heavily on a phonological code if they are asked to recall information about letter identity (Logie et al., 2000). Slower stimulus presentation rates also lead participants to use a wider range of strategies.

Baddeley and Larsen (2007) conducted three experiments in which healthy volunteers attempted to recall sequences of similar or dissimilar consonants (6 letters), presented visually or auditorily under concurrent articulation. They found that the phonological similarity effect occurred for auditory stimuli but was erased when the stimuli were visual, an outcome pattern discussed above. However, an unexpected result from this study was that performance was better with visual than with auditory presentation, whereas the typical trend is for the opposite. The authors suggested that the concurrent articulation and phonological similarity may have encouraged participants to use supplementary visual or semantic codes, which possibly are more readily accessible with visual presentation.

Vallar (2006) and Baddeley (2010) hypothesized that there is a visual short-term store before orthographic to phonological recoding or grapheme-to-phoneme conversion, in which visual stimuli are encoded in terms of shape. This visual short-term store (sometimes referred to as iconic memory) is depicted in Figure 2.2 (Vallar, 2006). In this theoretical model, verbal input is analyzed phonologically and then has direct and automatic access to the phonological store and to long-term memory. A rehearsal process is thought to involve recirculation of the memory trace between the phonological store and a phonological output mechanism (i.e., buffer) before speech production. In contrast to verbal input, visual input must be converted to phonology prior to accessing the phonological output buffer and interacting with the phonological store and long-term memory.

The schematic in Figure 2.2 can be used as a framework for interpreting performance in working memory tasks involving visual stimuli. For example, if participants are influenced by the visual-orthographic characteristics of the stimuli (i.e., orthographic similarity effect), one could hypothesize that they are attempting to perform the task using visual short-term memory and not phonological information. In contrast, if participants are not influenced by visual similarity and instead are influenced by phonological similarity then one could hypothesize that they are relying on the phonological output buffer in performing the working memory task.

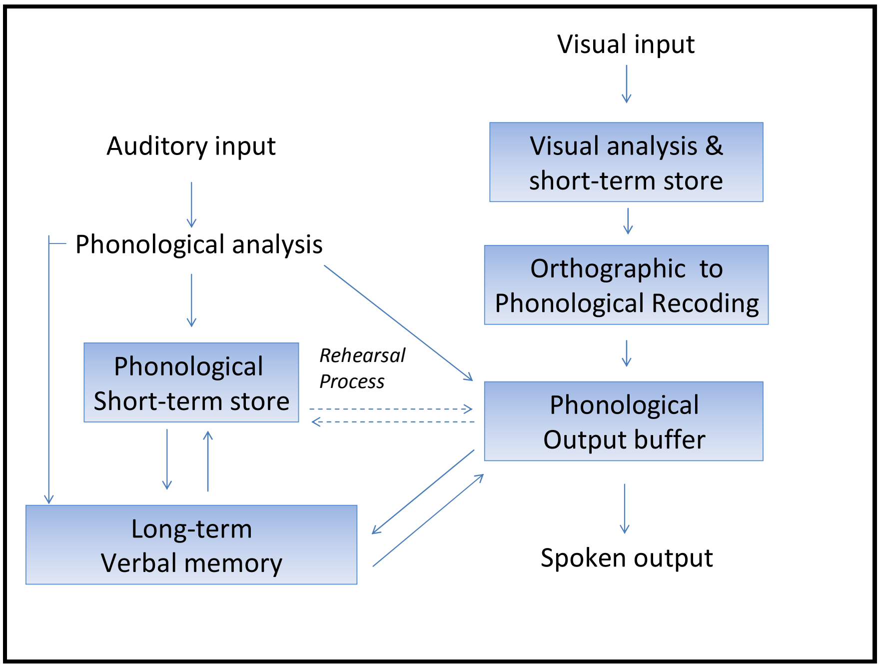


Figure 2.2 Schematic of the phonological loop with visual and auditory input (Vallar, 2006).

The influence of long-term memory. In addition to the characteristics of verbal working memory described above, there is evidence that verbal working memory can be influenced by long-term memory (LTM). Long-term memory refers to the retention of information for more than 30 seconds with larger capacity than working memory. Forgetting from long-term memory is usually due to retrieval failure caused by interference from other knowledge (Medin, Ross & Markman, 2004).

Language is a learning process, and is therefore impacted by top-down knowledge, which refers to high-level information (e.g., stored phonological, lexical and semantic representations) that guides the search for lower-level information (e.g., letters from that word) (Medin, Ross & Markman, 2004). It is widely accepted that long-term memory interacts with working memory, but the crucial question is how they interact (Baddeley, 2010). As depicted above in Figure 2.2, the relationship between working memory and LTM includes bidirectional influence between the phonological short-term store and LTM.

One example of how long-term knowledge can be used to aid immediate recall is that performance in immediate recall of nonwords is better when stimuli are similar in phonetic structure to the native language as compared to a non-native language (Baddeley, 2003). Similarly, it is easier for bilingual speakers to memorize a telephone number using their native language rather than using a second language or unfamiliar language, reflecting the importance of long-term phonological knowledge in short-term verbal memory (Baddeley, 2003; 2010).

Other evidence for the influence of long-term memory on verbal working memory is that participant strategies strongly impact the effects of phonological similarity and word length on working memory, as mentioned above. For example, Campoy and Baddeley (2008) instructed participants to use a phonological strategy, a semantic strategy or no strategy in a serial recall paradigm and found evidence that the semantically instructed group did attempt to use the strategy of associating items based on word meaning to remember the information. This semantic associative information consists of rich long-term linguistic knowledge that can influence verbal working memory.

Long-term linguistic knowledge is also reflected in a variety of lexical effects observed in verbal working memory experiments, such as improvements in recall due to lexicality (words > nonwords), lexical frequency (high frequency > low frequency words), and lexical-semantic representation (concrete > abstract words; and high image > low image words) (Acheson & MacDonald, 2009a). Sublexical factors also affect performance in verbal working memory experiments. Long-term knowledge about syllable structure influences the serial ordering of verbal information in verbal working memory, and errors reflect syllable position constraints.

Other types of long-term sublexical phonological knowledge affect performance in verbal working memory experiments, including phonotactic frequency and phonological neighborhood density (Acheson & MacDonald, 2009a). Phonotactic frequency refers to the frequency with which sounds are combined in a language (Davidson, 2011). In verbal recall, nonwords with high phonotactic frequency (i.e., composed of common verbal sounds and sound combinations) are easier to recall than are those with low phonotactic frequency. Phonological neighborhood density is defined as the number of words that differ from a target by only one phoneme in the same position (Acheson & MacDonald, 2009a). In verbal working memory, words that come from dense phonological neighborhoods (i.e., with more phonologically related words in the language) are recalled more accurately than are those from sparse neighborhoods. Both phonotactic frequency and phonological neighborhood effects reflect top-down knowledge (i.e., long-term memory).

Skilled performers such as musicians and chess players appear to have an expanded working memory capacity that can activate acquired knowledge and special memory skills (i.e., long-term memory) (Chase & Simon, 1973; Stigler, 1984). Ericsson and Walter (1995) proposed the notion of "long-term working memory" to describe the large demands on working memory during these forms of expert performance or in tasks such as text comprehension.

In Figure 2.2 above (Vallar, 2006), connections from verbal LTM to the phonological short-term store and the rehearsal process illustrate the support of LTM systems in aspects of immediate retention. In addition to LTM influencing performance in working memory tasks, there is evidence from different subject populations that the phonological short-term store contributes to long-term learning. For example, the capacity of phonological memory in children is a main predictor of vocabulary acquisition (Gathercole & Baddeley, 1993).

The central executive. Based on Baddeley's model of working memory (2003), the central executive component is a supervisory activating system that controls behavior by habit patterns or schemas. In other words, the central executive is a cognitive control function that regulates the encoding, retrieval, and integration or manipulation of information entering working memory from different sensory storage systems or from long-term memory (Miller & Kupfermann, 2009).

Hedden and Yoon (2006), based on results from Miyake et al. (2000), suggested that executive function can be decomposed into three distinctive processes: shifting among multiple tasks or mental sets, updating and monitoring of representations in working memory, and inhibition of responses. According to these authors, shifting among multiple task demands involves loading a new goal set into working memory and inhibiting prior, non-relevant goal sets and their corresponding task dimensions. Updating involves the manipulation of representations stored in working memory, such as reordering or recombining according to task demands. Inhibition involves the suppression of unwanted or irrelevant representations, goals, and responses.

Bao and colleagues (2006) discussed the concept of shifting and suggested that normal adults have two abilities of shifting: the goal of stability, which is to complete one response while resisting the tendency to jump to another one; and the goal of flexibility, which refers to switching from one goal or action plan to another when necessary. They noted that shifting is accompanied by inhibition of the previous response set, which is referred to as "backward inhibition." They provided evidence that backward inhibition occurs not only during response switching but also during working memory attention switching, which implies that backward inhibition could be a general mechanism that serves to reduce interference from all potentially competing cognitive stimuli (Bao et al., 2006).

A commonly used measure of central executive function is the *n*-back task (Baddeley. 2003). In the *n*-back task, participants are presented with a series of stimuli and are asked to indicate whether the current stimulus matches the stimulus presented *n* stimuli back in the series, where *n* equals a number between 0 and 3 (Simmons, 2000). The *n*-back task can be structured to include various types of stimuli, including letters (Lieberman & Rosenthal, 2001; Sweet et al., 2008), words (Crosson et al., 1999), and viewed objects (Christensen & Wright, 2010). The tendency of participants to accept or reject various types of distractor items (e.g., orthographically, semantically or phonologically similar or dissimilar) can be informative about varying linguistic processing demands in the task and strategies the participants may be using to complete the task (e.g., Crosson et al., 1999).

Evidence for the components of verbal working memory (i.e., the phonological store, the articulatory mechanism, and the central executive) primarily has come from research involving non-tonal languages such as English. Differences between the Mandarin Chinese and English writing systems may influence the results of verbal working memory experiments. However, there has been little work examining working memory for the tonal component in Mandarin reading. Another unresolved issue of working memory is the relationship of tone information in tonal language versus music. In the next section, working memory for Mandarin tone and for music will be discussed.

Working memory for tone information

There are different definitions of tone information in the field of linguistics as compared to music. Few studies have directly examined how linguistic tone and musical tone compare, especially in the working memory field. No research has been conducted on links between lexical tone and verbal working memory in Chinese or other tonal languages. Studies of how working memory relates to other aspects of reading in Mandarin Chinese are reviewed below, in addition to studies of working memory for musical tone.

Working memory and reading Mandarin Chinese. There is evidence that working memory plays an important role in reading Chinese. Deficits in both linguistic and working memory processes appear to contribute to the difficulties of poor readers of Chinese (So & Siegel, 1997). Leong, Tse, Loh and Hau (2008) examined the relationships among children's performance on several cognitive tasks, including verbal working memory tasks (memory span and tongue twister tasks), text comprehension, and Chinese pseudoword[1] reading. The results showed that children with better verbal working memory performed much better in nonword reading and text comprehension.

Chung, Ho, Chan, Tsang and Lee (2011) assessed cognition in Chinese Cantonese-speaking adolescents with versus without developmental dyslexia and found that the dyslexic group was weaker than controls in rapid naming, visual-orthographic knowledge, morphological awareness, and verbal working memory. These four cognitive measures were also associated with word reading, word spelling, and reading comprehension.

Xu and Li (2009) manipulated the content of Chinese word lists used in a serial recall task to examine the effects of working memory in Chinese reading. The expected word length effect (greater recall of short words than long words[2]) was observed when word lists were made up of only short words or only long words. However, when word lists consisted of mixed short and long words, the results were inconsistent. That is, when one long word was embedded in a list of

---

[1] A Chinese pseudoword reading consists of a meaningless two-character word; for example from Leong, et al. (2008): 炮喻 [pau3 jy4].

[2] Short words refer to two-characters, whereas long words refer to four-character words. Xu and Li (2009) used country names such as Switzerland /瑞士/ (two characters) vs. Australia /澳大利亞/ (four characters).

four short words, the long word was recalled more accurately than a short word embedded in a list of four long words. Through a variety of such manipulations of word list stimuli and presentation, Xu and Li (2009) also found that total pronunciation time influenced recall performance. They also examined the effect of visual encoding on the memory process by using a technique of output delay. That is, after the final item in the word list was presented on a computer screen, a white cross was presented in the same location for two seconds. This manipulation disrupted the normal word length effect, suggesting that aside from phonological encoding, visual encoding also contributed to the working memory process in Chinese reading. Based on these results, the authors concluded that working memory may involve multiple, parallel, and different encodings that are both competitive and complementary in memory processing.

Mandarin Chinese speakers have demonstrated greater memory spans on forward digit span and spatial span than English speakers (Chen et al., 2009). Chincotta and Underwood (1997) reported that Chinese speakers obtained a larger digit span without concurrent articulation than English, Finnish, Greek, Spanish and Swedish speakers, but that with concurrent articulation there were no differences across languages. With concurrent articulation, the normal translation of visual stimuli into phonological codes was prevented, thus reducing the contribution of the phonological loop. This indicated that superior digit span in Chinese speakers was determined by phonological loop function, as concurrent articulation eliminated the advantage of Chinese over other languages. The high performance by Chinese speakers in digit span may be a feature of the language itself. Digit names in Chinese are monosyllabic and shorter in terms of articulation duration than other languages (Hoosain, 1984). An alternative explanation is that the digit span of the Chinese speakers was mediated by the phonological loop functioning at a faster rate of subvocal rehearsal (Chincotta & Underwood, 1997).

Lu and Zhang (2007) investigated the role of the phonological loop in Chinese reading comprehension. Participants read text samples normally or under conditions of concurrent articulation or irrelevant sound, and then judged whether the meaning of the text was acceptable or not. Selected words in the text passages were controlled for word frequency (high or low) and for phonological or orthographic similarity. They found that accuracy in the normal reading condition was higher than the concurrent articulation condition or the irrelevant sound condition, suggesting that the phonological store and articulatory rehearsal play an important role in Chinese reading. Not surprisingly, accuracy was higher given high frequency (i.e., more common) words as compared to low frequency words. Accuracy for the text containing correct words was higher than text containing the phonologically similar distracters. Text comprehension was worse given the orthographically similar distracter words as compared to the correct target words or the phonologically similar distracters. Moreover, within the difficult low frequency words, accuracy and reaction time given the orthographically similar words were significantly worse as compare to the phonologically similar words. This study provided evidence for the importance of the visual-orthographic code in Chinese text comprehension.

As noted above, there have been no studies reported about working memory for lexical tones. Studies of working memory in Mandarin Chinese have been limited to other aspects of reading skill as described above. Unlike working memory for lexical tones, musical working memory has been the focus of a number of studies over the past two decades. In the next section, features of musical tone and musical working memory will be reviewed.

Working memory for musical tone. Although no theoretical models of reading musical notation have included a working memory component, Berz (1995) was the first investigator who presented a theoretical model of musical working memory. This theoretical model of music was centered around the central executive controller with “loops” of different types of information interacting with it, including multiple sensory loops (such as smell and taste components), a phonological loop (verbal component), a visual-spatial sketchpad (visual component), and a music memory loop (musical component). Berz made two major assumptions: 1) that the central executive is a key component of musical working memory; and 2) that there are two different loops to support language and music: a phonological loop and a musical loop. The musical memory components from Berz (1995) were very similar to the phonological loop of Baddeley’s (1990) verbal working memory model with a musical store (i.e., similar to the phonological store) and an articulatory mechanism based on musical inner speech. However, Berz did not provide empirical evidence to support this theoretical model.

There are some parallels in the characteristics of musical working memory and verbal working memory in terms of how performance patterns observed experimentally may relate to subcomponents of working memory. That is, the effects of phonological similarity, serial position, irrelevant sound and concurrent articulation in verbal working memory are similar to the effects of pitch similarity, serial position, irrelevant sound, and concurrent articulation in musical working memory. However, unlike verbal working memory, there is a lack of a word length effect (i.e., note length) in studies of musical working memory.

The pitch similarity effect. The effect of pitch similarity in serial recall of tones is to make recall of tones that are similar in pitch more difficult than recall of tones that are not close in pitch. This pitch similarity (i.e., pitch proximity) effect was observed in non-musicians who participated in a study of serial recall of verbal and musical materials (Williamson, Baddeley & Hitch, 2010). Musicians who also participated in this study did not exhibit the pitch proximity effect. Both musicians and non-musicians demonstrated the expected phonological similarity effect in recalling verbal materials in this study. Based on these findings, the authors proposed that verbal and musical information are stored separately but share the same articulatory mechanism. They also hypothesized that musicians may use a different strategy than non-musicians in memorizing tonal information, such as multidimensional auditory and visual codes.

The serial position effect. In several studies the possibility of a serial position effect was examined in tasks of recall with auditory stimuli, specifically musical tone information (Leshowitz & Hanzi, 1974; Surprenant, Pitt & Crowder, 1993; Silverman, 2007). Consistent with previous research with verbal information, a serial position effect was observed for musical tone stimuli such that tones in earlier (primacy) and final (recency) positions were recalled more accurately than tones in the middle positions of a sequence.

The irrelevant sound effect. As noted above in regard to verbal working memory, the irrelevant sound effect means that immediate recall is impaired by the concurrent or subsequent presentation of irrelevant verbal or tonal information. In a study of serial recall for digits and for tones, Schendel (2006) examined whether the irrelevant sound effect is a result of the 'similarity of content' (i.e., acoustic overlap) or 'similarity of process' (i.e., the changing state of the auditory sequence) between the to-be-remembered information and the to-be-ignored information. He reported evidence that both similarity of content and similarity of process reduced accuracy of serial recall in digit recall and in tone recall. In tone recall, greater pitch overlap resulted in greater interference of the irrelevant sound on recall of the to-be-remembered information. Based on the similar patterns of results in working memory experiments involving language or music stimuli, Schendel (2006) argued that working memory for language and music are controlled by a single acoustic loop. This acoustic loop or acoustic store would take the place of the phonological store working memory for language. Because singing or listening to music impairs performance in language and working memory tasks, Schendel also argued that there is a single articulatory rehearsal mechanism for language and music.

The concurrent articulation effect. As described above in relation to verbal working memory, concurrent articulation tasks require participants to repeat a certain phrase or word aloud while simultaneously doing a working memory task. This concurrent articulation requires that participants suppress the articulatory mechanism. Few studies have examined this effect in musical working memory; however, Koelsch and colleagues (2009) used concurrent articulation in a study of brain imaging involving working memory for tones. The brain regions that underlie working memory for verbal or tonal information have been examined in several brain imaging studies (Hickok, Buschsbaum, Humphrise & Muftuler, 2003; Koelsch, Schulze, Sammler, Fritz, Müller & Gruber; 2009; Schulze, Zysset, Mueller, Friederici & Koelsch, 2011). Koelsch and colleagues (2009) presented healthy volunteers with strings of sung syllables and asked them to remember either the pitch (tonal information) or the syllable (verbal information) from the German alphabet under six different experimental conditions: 1) verbal (syllable) rehearsal, 2) verbal suppression, 3) tonal (pitch) rehearsal, 4) tonal suppression, 5) no memorization, rehearsal or singing, and 6) no memorization but sing a song. In the rehearsal conditions, participants covertly rehearsed either syllables or pitches. During the suppression conditions (i.e., concurrent articulation), participants covertly sang a children's song while trying to maintain either verbal or tonal information. Based on patterns of brain activation across conditions, the authors suggested that both rehearsing verbal and musical tonal information, as well as storage of verbal and musical tonal information relied on overlapping neural networks.

In summary, based on the few studies of musical working memory to date, it appears that musical working memory is similar to verbal working memory in terms of having a storage component (i.e., possibly a tonal store similar to the phonological store) and an articulatory mechanism. Like verbal working memory, musical working memory appears to be influenced by stimulus similarity, serial position, irrelevant sound, and concurrent articulation. There is some evidence that music and language may share elements of the same working memory process.

Central executive function in musical working memory. Given that musical working memory appears similar to verbal working memory in having a storage component and an articulatory mechanism, one can hypothesize that musical working memory also includes a central executive. As noted above, the central executive is thought to be a cognitive control function that regulates the encoding, retrieval, and integration or manipulation of information entering working memory from different sensory storage systems or from long-term memory (Miller & Kupfermann, 2009).

Ockelford (2007) offered a new construct of musical working memory based on previous theoretical models of Baddeley (1986) and Berz (1995). Unlike previous studies, he described a “musical executive” component of musical working memory. He hypothesized that this musical executive could be related to the central executive (Baddeley, 1986). On this account, the musical executive processes perception and strategic encoding of notes in memory. However, this report by Ockelford (2007) is a case observation in which the participant demonstrated the ability to “listen and play” chromatic blues. Blues is a specific musical structure and includes a lot of improvisation. It is questionable how much we can infer about the Western musical system from this case observation.

Studies addressing the relationship between central executive function and musical working memory have focused on the musical training effect (Degé, Kubicek & Schwarzer, 2011; Hargreaves & Aksentijevic, 2011; Moreno, Bialystok, Barac, Schellenberg, Cepeda & Chau, 2011; Schellenberg, 2011). These studies examine whether the association between music lessons and intelligence (IQ) was mediated by executive function. In one of these studies, Degé, Kubicek and Schwarzer (2011) found that musical training could predict inhibitory control. There were three groups of children in this study: 1) no musical training, 2) one to four years of music instrumental training, and 3) more than four years of music instrumental training. All participants were administered tests of set shifting (i.e., an animal sorting task), selective attention (an auditory attention and response set task), planning and organization ability (a drawing condition; e.g., “draw a clock”), inhibition (inhibit automatic responses in favor of novel responses), and design fluency, which measured the ability to generate multiple unique designs by connecting dots presented in structured or random arrays. The fluid intelligence of participants was also measured as an index of IQ. The study results supported the hypothesized association between musical training and intelligence mediated by executive function. Selective attention and inhibition were the strongest contributors to the observed association. These results support the argument that daily practice of instrumental music enhances executive function and that musical training predicts inhibitory control.

In another recent study, short-term musical training was found to enhance verbal intelligence and executive function (Moreno, Bialystok, Barac, Schellenberg, Cepeda & Chau, 2011). Study participants were 48 preschool children between the ages of four and six years, and 24 children received musical-listening training while the other 24 received visual art training of visuospatial skills. The training programs consisted of two daily sessions of one hour each, five days a week for four weeks. After 20 days of training, only the children who had received the musical training exhibited enhanced performance on a measure of verbal intelligence (i.e., one vocabulary subset from the Wechsler Preschool and Primary Scale of Intelligence, WPPSI-III, Wechsler, 2002), with 90% of children in the music group showing this improvement. The investigators concluded that the music-listening training improved not only musical listening skills, but also transferred to improved verbal ability. Their explanation of their results is that music processing overlaps with cognitive mechanisms used in language.

Other studies of functional brain plasticity related to musical training include a study of the neuroarchitecture of verbal and tonal working memory in 17 non-musicians and 16 musicians using fMRI (Schulze, Zysset, Mueller, Friederici & Koelsch, 2011). The auditory stimuli were a spoken syllable and a sine wave tone presented simultaneously. Participants were asked to listen to sequences of five auditory (verbal + tonal) stimuli and to rehearse internally either syllables (during a verbal condition) or tones (during a tonal condition). At the end of each trial, a test stimulus was presented consisting of one syllable and one tone. Participants had to press a button to indicate whether the syllable (in the verbal condition) or the sine wave tone (in the tonal condition) had been presented in the initial sequence. The fMRI data showed overlapping brain regions hypothesized to be involved in both verbal and tonal working memory. Additionally musicians activated a specific region only during the verbal condition (right insular cortex) or only during the tonal working memory condition (right globus pallidus, right caudate nucleus, and left cerebellum). Thus, Schulze and colleagues (2011) suggested that two working memory systems may exist in musicians: a phonological loop supporting phonological information, and a tonal loop supporting musical tone information. Based on these results, Schulze and colleagues (2011) hypothesized that auditory (verbal- syllable and tonal- sine wave tone) working memory was not a unique system in that maintenance of pitch information appears to require a tonal loop just as maintenance of verbal information in working memory appears to require a phonological loop. They also provided results from participants' behavioral performance[3], which indicated somewhat more accurate performance of the musicians in the verbal and tonal conditions as compared to the non-musicians.

Based on the findings summarized here, it appears that musical training results in improved executive function, inhibitory control and verbal ability (Degé, Kubicek & Schwarzer, 2011; Moreno et al., 2011). Musical training presumably involves musical working memory. However, none of these previous studies have mentioned the relationship between central executive function and musical working memory specifically. As noted above, Berz (1995) proposed a theoretical model of music that was centered on a central executive controller and different forms of memory loops interacting with it, including multiple sensory loops, a phonological loop, a visual-spatial sketchpad, and a musical memory loop. In this framework, these memory loops interact with central executive function before production. However, it remains unclear whether central executive function is one component of musical working memory or if it is an independent system.

There is a general lack of research comparing musical working memory and verbal working memory. Lu, Greenwald and Bowyer (2010) investigated musical and verbal working memory using a task of musical transposition from print to sound in a brain activation study using magnetoencephalography (MEG). Transposition of print to sound was compared for musical notation versus written words in three healthy volunteers. Musical transposing from print

[3] Schulze and colleagues (2011) collected behavioral data during MRI scans. The auditory input stimuli consisted of a spoken syllable and a sine wave tone simultaneously. Participants subsequently listened to sequences of five auditory (verbal + tonal) stimuli and rehearsed internally (syllables in the verbal condition or tones in the tonal condition). At the end of each trial, a probe stimulus was presented, and participants had to indicate by pressing a button whether a stimulus has been presented (YES) or not (NO) in the initial sequence. Participants' responses were collected.

involves mental conversion of notes to a different key than what is written, by raising or lowering all the notes by a given interval. The influence of working memory was examined in this study by comparing subject performance on short versus long stimulus lists of musical or verbal information. Brain activation results of this study are preliminary, but the behavioral design of the study can be useful in further studies comparing musical and verbal working memory.

Behavioral paradigms for directly comparing musical working memory and verbal working memory are lacking, including paradigms for comparing memory for linguistic tone and musical tone. In the current dissertation, translation from print to sound was examined in healthy volunteers across four levels of phonological or tonal information: 1) phonological alone (English print); 2) phonological and tonal in a task emphasizing phonological information (Mandarin homophones); 3) phonological and tonal in a task emphasizing tonal information (Mandarin tones); and 4) tonal alone (musical notation). The performance of musicians and non-musicians was compared across reading tasks, in an attempt to examine whether musical training can facilitate reading or working memory of linguistic tone information. Also, the effects of increased working memory load across task were examined in musicians and non-musicians. Specifically, this study addressed the following research questions:

Research Question 1: How is accuracy of performance in musicians and non-musicians in four visual recognition tasks influenced by increased task demands on working memory? Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians on four visual recognition tasks with increased task demands on working memory.

Research Question 2: How is speed of performance in musicians and non-musicians in

four visual recognition tasks influenced by increased task demands on working memory?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians in four visual recognition tasks with increased task demands on working memory.

Research Question 3: How is accuracy of performance in musicians and non-musicians in visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians in visual recognition tasks with increased visual similarity of task stimuli.

Research Question 4: How is speed of performance in musicians and non-musicians in the visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians in visual recognition tasks with increased visual similarity of task stimuli.

CHAPTER 3

METHOD

The goal of the current study was to examine the translation from print to sound of the tonal language Mandarin and other reading stimuli in healthy volunteers. The performance of musicians and non-musicians was compared across a variety of reading tasks in an attempt to examine whether musical training can facilitate reading of Mandarin tone. The effects of increasing working memory load on reading performance across tasks were also examined. This chapter describes study participants, experimental materials, the study design and procedure, experimental tasks, and data screening.

Participants

Sixty participants from Taipei, Taiwan with the ability to speak and read English and Mandarin completed the study voluntarily. All participants had normal vision and hearing (with or without correction), and normal motor and cognitive abilities. Thirty participants (6 male, 24 female; 26 right-handed, 2 left-handed and 2 ambidextrous) self-reported no professional musical training. Their mean age was 22.97 (*SD* = 4.86); years of education ranged from 12-20 (*M* = 15.58, *SD* = 2.87). The other thirty participants (30 females; 28 right-handed and 2 ambidextrous) were musicians with at least 11 years of musical training who were able to read standard music notation. They reported a mean of 16.53 years of music training (range = 11-23 years). Their mean age was 22.27 (*SD* = 3.81); years of education ranged from 13-20 (*M* = 15.63, *SD* = 2.28). There were no significant differences between the two groups in gender (30 female musicians; 6 male and 24 female non-musicians), age, years of learning English, years of education, with the exception of years of learning music [$t\ (58) = 20.733$, $p < .000$**]. The demographic data for the two groups are listed in Table 3. 1.

Table 3.1

Demographic Data of Musicians and Non-musicians

	Musicians [a]		Non-musicians [b]	
	Mean	*SD*	*Mean*	*SD*
Age	22.27	3.81	22.97	4.86
Years of learning English	12.07	3.51	11.3	3.29
Years of Education	15.63	2.28	15.58	2.87
Years of learning Music*	**16.53**	**3.64**	**0.97**	**1.90**

[a] $n = 30$ [b] $n = 30$

Experimental Materials

Three types of visual stimuli were incorporated into the experimental tasks: English, Mandarin Chinese, and written musical notation. All English word stimuli (n = 280; 40 for the practice trial and 240 for the experimental trial) were chosen from previous studies (Reinarz, 1997; Pexman, Cristi & Lupker, 1999; Kielar & Joanisse, 2009; Rueckl et al., 1997, Binder & Borecki, 2008; Schwartz, Kroll & Diaz, 2006) and from the Psycholinguistic Assessments of Language Processing in Aphasia (PALPA; Kay, Lesser & Coltheart, 1997) to control for length (3-6 letters), orthographic similarity / dissimilarity, phonological similarity / dissimilarity, and word frequency (Francis & Kučera, 1982). To ascertain that the vocabulary level matched the English language skills of Taiwanese high school graduates, all selected stimuli were compared with the list of English Vocabularies compiled by the College Entrance Examination Center, Taiwan.

The first set of Mandarin Chinese stimuli was incorporated into a Mandarin homophone task (n=280; 40 for the practice trial and 240 for the experimental trial). These stimuli were chosen from previous studies (Bi, 2006; Tan, Hoosain & Siok, 1996; Tan & Perfetti, 1997; Zhou & Marslen-Wilson, 1999a; Zhou & Marslen-Wilson, 1999b; Xu, Pollatsek & Potter, 1999; Pollatsek, Tan & Rayner, 2000; Leong, Cheng & Tan, 2005) to control for orthographic and phonological similarity / dissimilarity. The frequencies of stimulus words were verified against the Mandarin Chinese Character Frequency List Based on National Phonetic Alphabets (CKIP, 1995).

The second set of Mandarin Chinese stimuli was incorporated into a Mandarin tone task, including the 1st tone (level), the 2nd tone (rising), the 3rd tone (dipping), and the 4th tone (falling) (Taft & Chen, 1992; Leong, 2002). All Mandarin tonal word stimuli (n = 280, 40 for the practice trial and 240 for the experimental trial) were chosen from previous studies (Tan, Hoosain & Siok, 1996; Pollatsek, Tan & Rayner, 2000; Hallé, Chang & Best, 2004; Lee, Tao & Bond, 2008; Mitterer, Chen & Zhou, 2011; Tong, Francis & Gandour, 2007; Wang et al., 2004; Malins & Joanisse, 2010) to control for phonological similarity / dissimilarity, homophone versus nonhomophone, and the same versus different tones. The frequencies of stimuli were checked against the Mandarin Chinese Character Frequency List Based on National Phonetic Alphabets (CKIP, 1995).

All musical notation stimuli (n = 280, 40 for the practice trial and 240 for the experimental trial) were chosen based on the Western musical system. This study was based on 12 major keys presented on the g-clef and the f-clef.

Design of the Study

This study used a three-factor mixed design: $2_{between}$ (group: musicians versus non-musicians) x 4_{within} (linguistic: English rhyming versus Mandarin homophone versus Mandarin tone versus Musical tone) x 2_{within} (task difficulty: easier versus difficult), with accuracy rate (AR) and reaction time (RT) as the dependent variables.

Stimuli in each of the English, Mandarin homophone, and music reading tasks were carefully controlled for two variables: phonological/pitch similarity and visual-orthographic similarity. Participants were instructed to judge phonological/pitch similarity. For these tasks, half of the correct responses and half of the incorrect target responses were similar to the target in terms of visual-orthographic features. This manipulation of experimental stimuli was included so that the effect of visual-orthographic similarity on task performance could be assessed.

Visual similarity for stimuli in the musical note task was defined by location on the musical staff, in that the same location was visually similar. Stimuli were divided into notes that were in the same location (LS) and having the same pitch names (PS), notes in a different location on different clefs but with the same pitch names with the same pitch names (LDPS), notes that were in the same location (LS) with different pitch names (PD), and notes that were in different locations with different pitch names (LDPD). (The same pitch only occurs in the same location. However, if both stimuli are in the same pitch class, which means they have the same pitch name but in a different octave, they have the same pitch name with different pitch height. In this task, pitch similarity was judged by whether it was the same pitch name, such as middle c and high C.) Examples of the musical notation stimuli are presented in Figure 3.1.

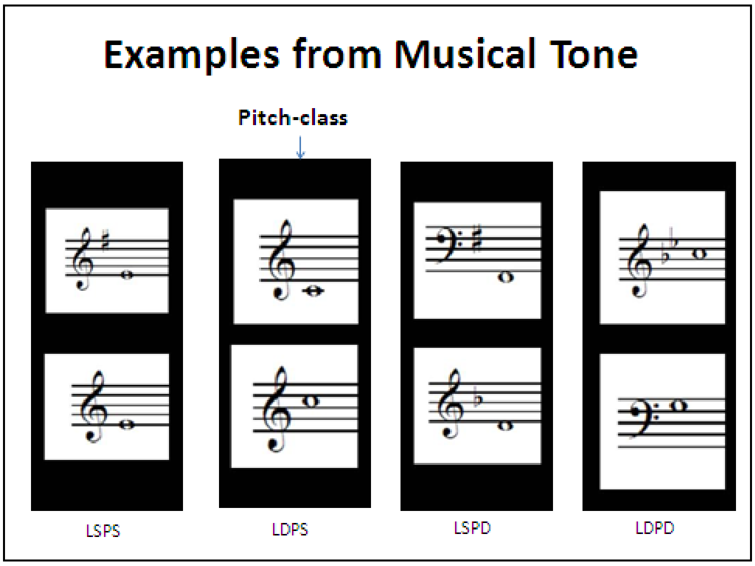


Figure 3.1 Example stimuli for the Musical note task. *Left to Right:* notes are in the same location with the same pitch name (LSPS), different locations (i.e., different clefs) with the same pitch name (LDPS), the same location with different pitch names (i.e., different clefs) (LSPD), and different locations with different pitch names (LDPD).

The stimuli in the Mandarin tone task were carefully controlled for two variables: homophone similarity and tone similarity. Participants were instructed to judge tone similarity. Stimuli were divided into homophones with the same tone (HTS), homophones with different tones (HTD), non-homophones with the same tone (nHTS), and non-homophones with different tones (nHTD). All stimuli are presented in the Appendix.

A modified version of the *n*-back task was adopted for use in this study and was similar to the one used by Kim and colleagues (2002). As described above (Baddeley, 2003), in the *n*-back paradigm participants are presented with a series of stimuli and are instructed to indicate whether the current stimulus matches the stimulus presented *n* stimuli back in the series, where *n* equals a number between 0 and 3 (Simmons, 2000). In the current study, each linguistic task was presented in both the 1-back and 2-back paradigms.

In the 1-back design (the design was the same across all linguistic tasks), each linguistic task consisted of 120 target items for which the participant had to respond. Of these, a 'yes' response was correct for 60 items and a 'no' response was correct for 60 items (see Figure 3.2). A correct 'yes' response indicated that the target item rhymes with or has the same homophonic or the same linguistic tone or the same pitch name (e.g., phonologically similar; PS) to the item immediately preceding it 1-back (i.e., the "probe"), and a correct 'no' response indicated that the target is different in sound (e.g., phonologically dissimilar, PD) compared to the probe item 1-back. Half of correct 'yes' responses looked similar (e.g.., orthographically similar, OS) to the probe item immediately preceding it 1-back, and half of correct 'yes' responses looked different (e.g., orthographically dissimilar, OD) to the probe item 1-back. Also, half of correct 'no' responses looked similar to the probe item, and half of correct 'no' responses looked different from the probe item.

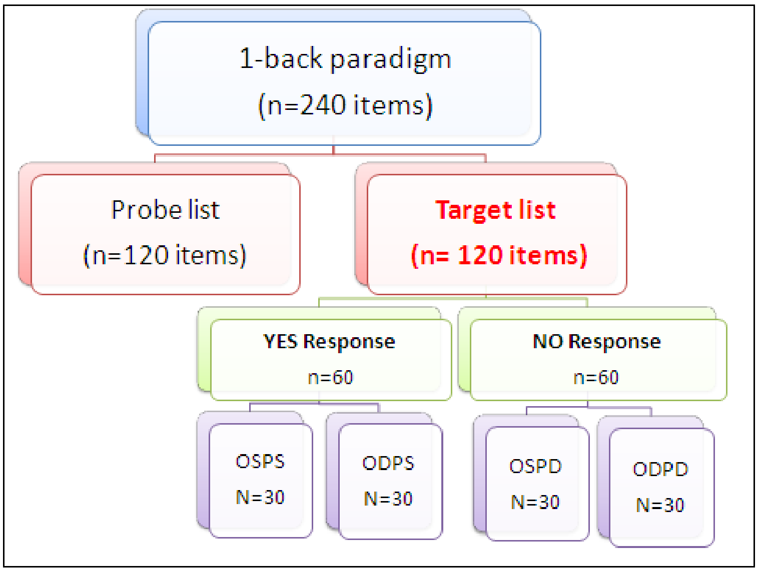


Figure 3.2 Number of items under each category in 1-back paradigm.

In the 1-back task, the total number of stimuli was 240, including the 120 probe items and 120 target items. Participants were asked to respond to each of the 240 items, which were presented in a pseudorandom sequence of pairs (probe → target). The correct response to the probe items was always 'no,' and the probe items were not included in calculations of performance accuracy or reaction time.

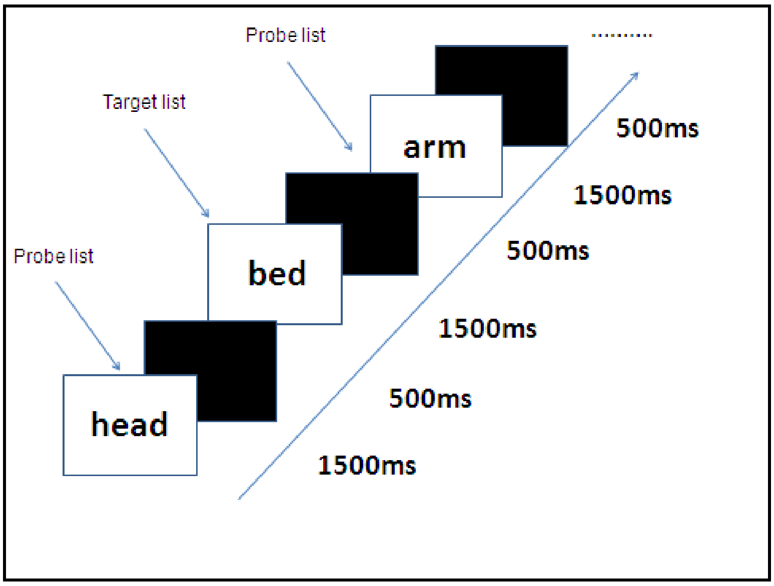


Figure 3.3 An example of 1-back paradigm shown in a graphic format.

Each item was presented in a fixed central location on a laptop monitor for 1500ms, followed by a blank screen for 500ms. The duration of each task was 8 minutes (2000ms * 240 items). To reduce fatigue, each task was divided into 2 blocks (4 minutes each) with a short (30 second) break in between. Thus, each linguistic task took a total of 8 minutes, 30 seconds (see Figure 3.3). The 2-back task was designed the same as the 1-back task, except that only half as many target words were included to keep the total number of stimuli and the total task duration the same as the 1-back task. In the 2-back task, more non-target words were included because probe words had to be 2-back from the target. The design was the same across all linguistic tasks. Because only 60 target words were included, consisting of 30 correct 'yes' responses and 30 correct 'no' responses, there were fewer subtypes of stimuli in the 2-back task as compared to the 1-back task. For example, for the English rhyming, Mandarin homophone, and musical note tasks in the 2-back paradigm, the correct 'yes' responses were all similar to the probe 2-back in terms of visual-orthographic features, and the correct 'no' responses were all dissimilar to the probe 2-back in terms of visual-orthographic features (see Figure 3.4). To reduce participants' expectations, additional distracting filler stimuli were inserted. The probe items and distracter and filler lists were not included in calculations of performance accuracy or reaction time.

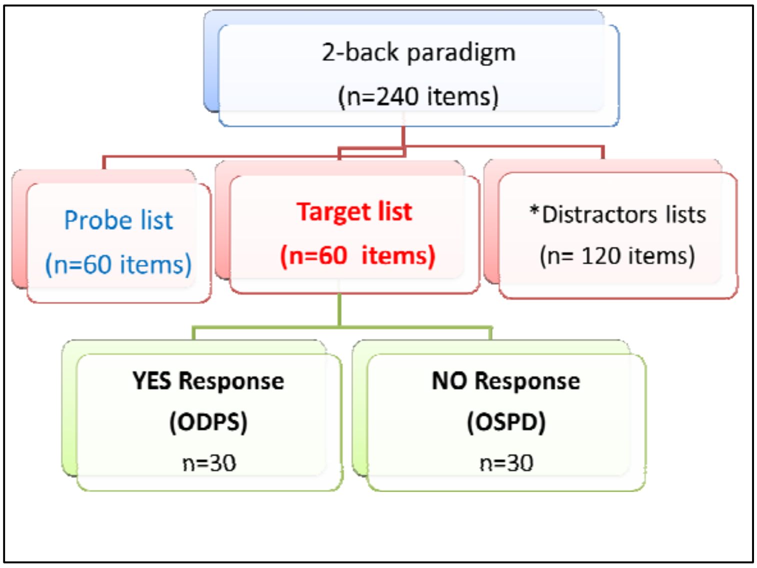


Figure 3.4 Number of items in each category in 2-back paradigm.

In the 2-back task, a probe stimulus was first presented, follow by a distracter; next a target stimulus was shown, followed by a distracter as well (see Figure 3.5). Other aspects of presentation and task duration were the same as the 1-back task.

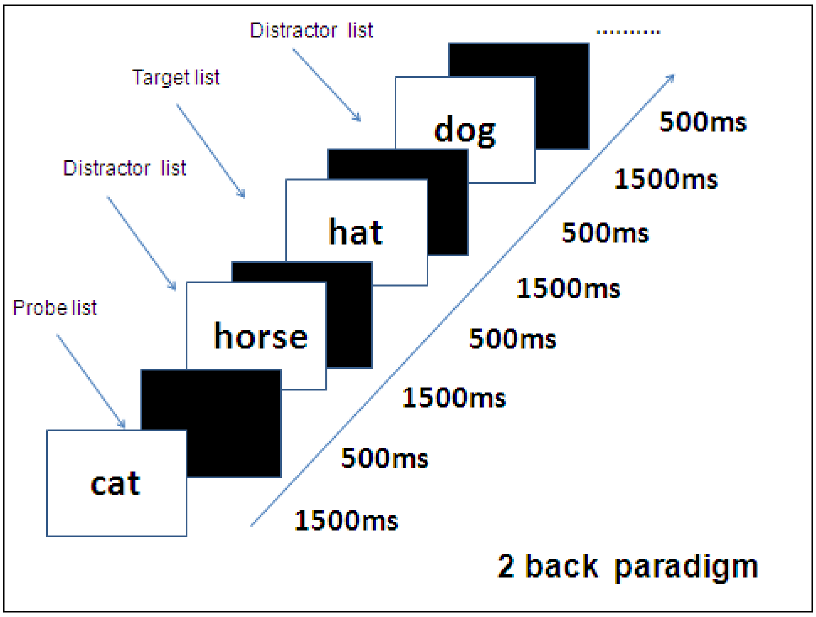


Figure 3.5 An example of the 2-back paradigm.

Procedure

Screening tasks. Screening measures included a questionnaire for self-report of demographic information, years learning English in Taiwan, and years of education. For musicians, the questionnaire also included self-report of years learning music and major instruments.

All participants were examined by the administrator with a vision screening, speech discrimination screening, the Edinburgh Handedness Inventory (Oldfield, 1971), Mini-Mental State Examination (MMSE; Folstein, Folstein & McHugh, 1975), which screened for cognitive impairment, and the digit span task from the Wechsler Adult Intelligence Scale-III (WAIS III; Wechsler, 1997), which was used to measure of working memory storage capacity. There were no significant differences between the two groups in Edinburgh Handedness Inventory [t (58) = 1.87, p = .067], MMSE [t (58) = -1.43, p = .160], and digit span [t (58) = .59, p = .558].

Participants also were asked to complete four discrimination tasks: rhyming discrimination

(English stimuli), Mandarin homophone discrimination (Chinese stimuli), Mandarin tone discrimination (Chinese stimuli), and note discrimination (musical notation). In each of the four tasks, two visual stimuli were presented simultaneously one above the other on the computer screen. Participants had to decide whether these two stimuli sounded the same in terms of rhyme / homophone / tone / or pitch name. The purpose of these discrimination tasks was to ensure that participants understood how to do all experimental tasks. The average accuracy rates for each discrimination task across all participants were as follows: English rhyming, 78%; Mandarin homophone, 97.5%; Mandarin tone, 91.5%; and Music, 82.5%.

Experimental Procedure. There was one experimental site: Taipei Jingmei Girls High School, Taipei, Taiwan. Each participant was tested individually in a quiet room. Informed consent was obtained from the participants upon their arrival. Next, participants were asked to complete the self-report questionnaire and the Edinburgh Handedness Inventory. Afterwards, the investigator administered the vision screening, speech discrimination, Mini-Mental State Examination and digit span tasks. The procedure including screening tasks and experimental tasks consisted of two visits of approximately one and a half hours each for a total of three hours during one month.

Experimental tasks. A total of eight experimental tasks were administered (i.e., each of the four experimental tasks presented in the 1-back and the 2-back paradigms). Stimuli were presented electronically using E-Prime Professional 2.0 software (Psychology Software Tools, Pittsburgh, PA) presented on an IBM ThinkPad R60e laptop with a 15" screen size (13.1” x 10.6”) monitor. An external number pad was connected with the laptop, providing “YES” and “NO” response keys. Participants sat in front of the laptop screen and they were instructed to use only the right index finger to press the buttons. (The few participants who were not right-handed were instructed to use either their right or left index finger to press the buttons. These participants did not make more mistakes than right-handed participants.) Participants were instructed to press any key when they were ready to start a trial, and to response within 2 seconds of each stimulus or their response for the particular trial would not be recorded. They also were instructed that if they were unable to respond within 2 seconds, they should skip the immediate stimulus pair and focus on the next one.

In both the 1-back and 2-back tasks for all four experimental tasks (rhyming / homophone / Mandarin tone / musical tone), participants were instructed to press the 'yes' button if the target stimulus sounded similar in rhyme / homophone / Mandarin tone / musical tone as the probe stimulus. Otherwise, participants were instructed to press the 'no' button. Descriptions of the response patterns for the 1-back and 2-back tasks are given in Table 3.2.

Table 3.2

Response Patterns of 1-back and 2-back Tasks

	1-back "YES" response	2-back "YES" response	"NO" response
Rhyming	If a word rhymes with the one that came before it	If a word rhymes with the one that came 2 before it	Others
Homophone	If a word is homophonic with the one that came before it	If a word is homophonic with the one that came 2 before it	Others
Mandarin tone	If a word has the same tone as the one that came before it	If a word has the same tone as the one that came 2 before it	Others
Music	If a note has the same pitch name as the one that came before it	If a note has the same pitch name as the one that came 2 before it.	Others

Scoring and Data Screening

As noted above, only responses to the target stimuli were included in calculations of accuracy and reaction time. The 1-back paradigm was computed for 120 target responses (60 correct 'yes' responses and 60 correct 'no' responses). The 2-back paradigm was computed for 60 target responses (30 'yes' responses and 30 'no' responses). To know about the performance of all subjects across all experimental tasks, the number of correct acceptance, correct rejection, incorrect acceptance, and incorrect rejection responses were calculated. Participants' accuracy rate in each task was computed by:

$$AR_{task} = \frac{\sum correct\ acceptance + \sum correct\ rejection}{Total\ number\ of\ target\ responses}$$

Speed (reaction time) was computed the same way as the accuracy rate. Participants' reaction time in each task was computed by:

$$RT_{task} = \frac{\sum RTs\ of\ correct\ acceptances + \sum RTs\ of\ correct\ rejections\ of\ targets}{Total\ correct\ responses\ to\ targets}$$

Attempts were made to ensure results were not biased due to missing data. Three non-musician participants exceeded 25% missing response rate in the 2-back music task (missing 18 to 21 items). Data from these three participants were removed from subsequent analysis of this task; thus, analyses of the 2-back music task were conducted on 30 musicians and 27 non-musicians. For all other tasks, analyses were conducted on 30 musicians and 30 non-musicians.

The remaining data were examined for normality, specifically kurtosis and skewness. The subsets of data for each task fitted the criteria for normality; kurtosis and skewness were within the acceptable range under a conservative alpha level (Z score between ± 3.29) (Tabachnick & Fidell, 2007) so that transforming the data was not necessary.

CHAPTER 4

RESULTS

Three general areas of interest were addressed in this study: 1) comparing print to sound translation of written English versus written Mandarin Chinese versus musical notation; 2) examining how performance on these reading tasks may relate to musical training (the only group variable in this study); and, 3) assessing how increased working memory load may affect performance across these reading tasks. The four measures of visual recognition used involved a range of visual stimuli that correspond to phonological information (i.e., written English), combined phonological and tone information in a task emphasizing phonological information (i.e., written Chinese in a Mandarin homophone task), combined phonological and tone information in a task emphasizing tone information (i.e., written Chinese in a Mandarin tone task), or musical tone information (i.e., Western musical notation system).

To address the research questions and hypotheses described in Chapter Two, the data obtained in this study were assessed with respect to the following comparisons: 1) differences in overall accuracy between musical training groups; 2) differences in overall reaction time between groups; 3) differences across tasks within each group, and 4) differences in error pattern between and within groups.

Comparisons of Performance Accuracy.

The first research question posed in Chapter Two addressed the accuracy of task performance of the musician group and the non-musician group. It was assumed that the 2-back tasks would place more demands on working memory than the 1-back tasks, and that this increased task difficulty would be reflected in lower accuracy scores.

Research Question 1: How is accuracy of performance in musicians and non-musicians in

four visual recognition tasks influenced by increased task demands on working memory?

Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians in four visual recognition tasks with increased task demands on working memory.

The mean number of correct responses in each 1-back task was computed for each group. See Table 4.1 for descriptive statistics. All accuracy scores reported in this chapter are the sum of correct acceptances and correct rejections of the target items.

Table 4.1

Mean Accuracy Rates of 1-back Tasks by Group

	Musicians [a]		Non-musicians [b]	
	Mean	*SD*	*Mean*	*SD*
English Rhyming	.72	.07	.70	.11
Mandarin Homophone	.96	.02	.94	.04
Mandarin Tone	.96	.02	.92	.04
Music	.94	.05	.70	.10

[a] $n = 30$ [b] $n = 30$.

The mean accuracy rates were computed for the musician group and the non-musician group for each 2-back task. See Table 4.2 for descriptive statistics.

Table 4.2

Mean Accuracy Rates of 2-back Tasks by Group

	Musicians [a]		Non-musicians [b]	
	Mean	*SD*	*Mean*	*SD*
English Rhyming	.50	.09	.49	.11
Mandarin Homophone	.91	.06	.88	.09
Mandarin Tone	.81	.09	.76	.10
Music	.79	.10	.50[c]	.09

[a] $n = 30$ [b] $n = 30$ [c] $n = 27$.

To examine the effects of group, visual task, and difficulty level on performance accuracy, a 2 x 4 x 2 factorial mixed model analysis of variance (ANOVA) was conducted. The between-subjects factor Group included musicians versus non-musicians; for the repeated measures factors, Task included English rhyming versus Mandarin homophone versus Mandarin tone versus Music tasks; and Difficulty Level included 1-back versus 2-back tasks. Table 4.3 shows the significant main effects for group, task, and difficulty level. The accuracy rates after increased task demands were significantly different in that the accuracy rate of 1-back tasks (M = .85, SD = .05) was higher than 2-back tasks (M = .71, SD = .07). There were two significant two-way interactions: between group and task, and between task and difficulty level. There was no significant three-way interaction.

Table 4.3

Results of Three-Way Mixed Model ANOVA: Accuracy for Group by Task by Difficulty Level

Source	*df*	*F*	η^2 [d]	*p*
Group [a]	1	57.87	.51	<.0005
Difficulty [b]	1	655.61	.92	<.0005
Group * Difficulty	1	1.66	.03	.203
Within-group-error	55			
Tasks [c]	3	301.21	.85	<.0005
Group* Tasks	3	56.39	.51	<.0005
Tasks * Difficulty	3	43.29	.44	<.0005
Group * Task * Difficulty	3	1.18	.02	.318
Within-group-error	165			

a. Group included musicians versus non-musicians;

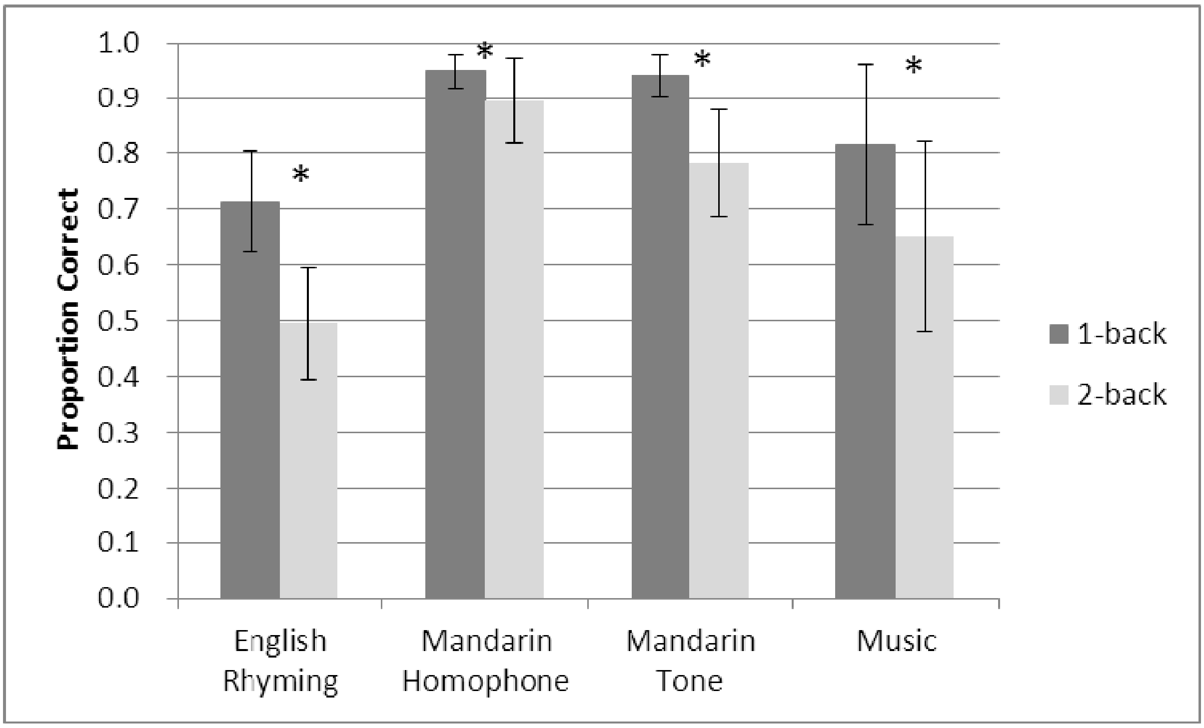
b. Difficulty included 1-back versus -2-back paradigm;

c. Task included English rhyming versus Mandarin homophone versus Mandarin tone versus Music tasks.

d. η^2 (eta-square) refers to effect size, reflecting the proportion of variance in a dependent variable associated with different level of an independent variable.

Four paired-samples t-tests were performed to compare the performance of all participants in the 1-back vs. the 2-back condition within each task, with accuracy rates as the dependent variable. Results indicated significant differences with increasing task difficulty across all participants in all tasks. Performance in the 1-back task was significantly more accurate than the

2-back task in the rhyming task [t (59) = 22.12, $p < .0005$], the Mandarin homophone task [t (59) = 5.69, $p < .0005$], the Mandarin tone task [t (59) = 13.32, $p < .0005$], and the music task [t (56) = 15.15, $p < .0005$] across all participants. Accuracy for all participants in the 1-back vs. 2-back tasks is depicted in Figure 4.1.



*$p < .05$

Figure 4.1 Accuracy for all participants in the1-back vs. 2-back tasks.

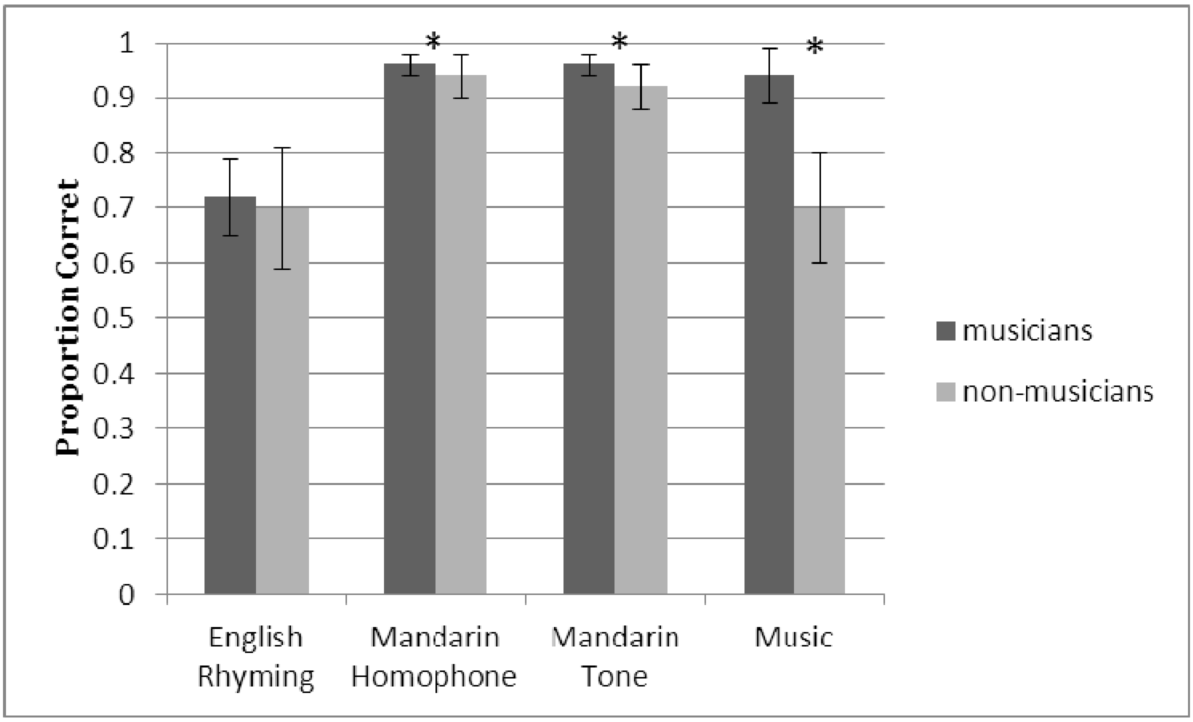
For the musician group, four paired-samples t-tests were performed to compare the 1-back to the 2-back tasks with accuracy rates as the dependent variable. Results indicated significant differences with increasing task difficulty for the musician group in all tasks. Performance in all 1-back tasks was significantly more accurate than in the 2-back tasks for the rhyming task [t (29) = 15.51, $p < .0005$], the Mandarin homophone task [t (29) = 4.59, $p < .0005$], the Mandarin tone task [t (29) = 10.72, $p < .0005$], and the music task [t (29) = 9.520, $p < .0005$] in the musician group.

Non-musicians showed a pattern similar to the musicians. For the non-musician group, four paired-samples t-tests were performed to compare the 1-back task to the 2-back task with accuracy rates as the dependent variable. Results indicated significant differences with increasing task difficulty in the non-musician group in all tasks. Their performance in the 1-back tasks was significantly more accurate than the 2-back condition in the rhyming task [t (29) = 15.68, $p < .0005$], the Mandarin homophone task [t (29) = 3.78, $p = .001$], the Mandarin tone task [t (29) = 8.58, $p < .0005$], and the music task [t (29) = 12.62, $p < .0005$] in the non-musician group.

To further compare the performance of musicians versus non-musicians within each task, a one-way Group ANOVA was conducted separately for each task. For each 1-back and 2-back task, a separate one-way ANOVA was conducted comparing the accuracy of musicians versus non-musicians on the task. The results of the 1-back task analyses are depicted in Figure 4.2, and the results of the 2-back analyses are depicted in Figure 4.3. As indicated in these figures, the musicians performed with significantly higher accuracy than the non-musicians in some tasks. Although the musicians were > 20% more accurate in the music task than the non-musicians, the overall pattern of accuracy across tasks was similar for the musician and the non-musician groups.

Accuracy data for the 1-back tasks were examined further using a 2 (Group) x 4 (Task) mixed model ANOVA. There were two significant main effects and one interaction effect. There was a significant main effect of group [F (1, 58) = 67.65, $p < .0005$] in that the accuracy rate of musicians (M = .89, SD = .03) was higher than for non-musicians (M = .81, SD = .04). There also was a significant main effect of task [F (3, 174) = 206.80, $p < .0005$] in that all participants showed the highest accuracy in the homophone task (M = .95, SD = .03), followed by the Mandarin tone task (M = .94, SD = .04), the music task (M = .82, SD = .14), and then the

English rhyming task (M = .71, SD = .09). There was a significant interaction of group and task [F (3, 174) = 48.16, p < .0005]. The sources of this interaction will be described below and in Figure 4.2.



*p < .05

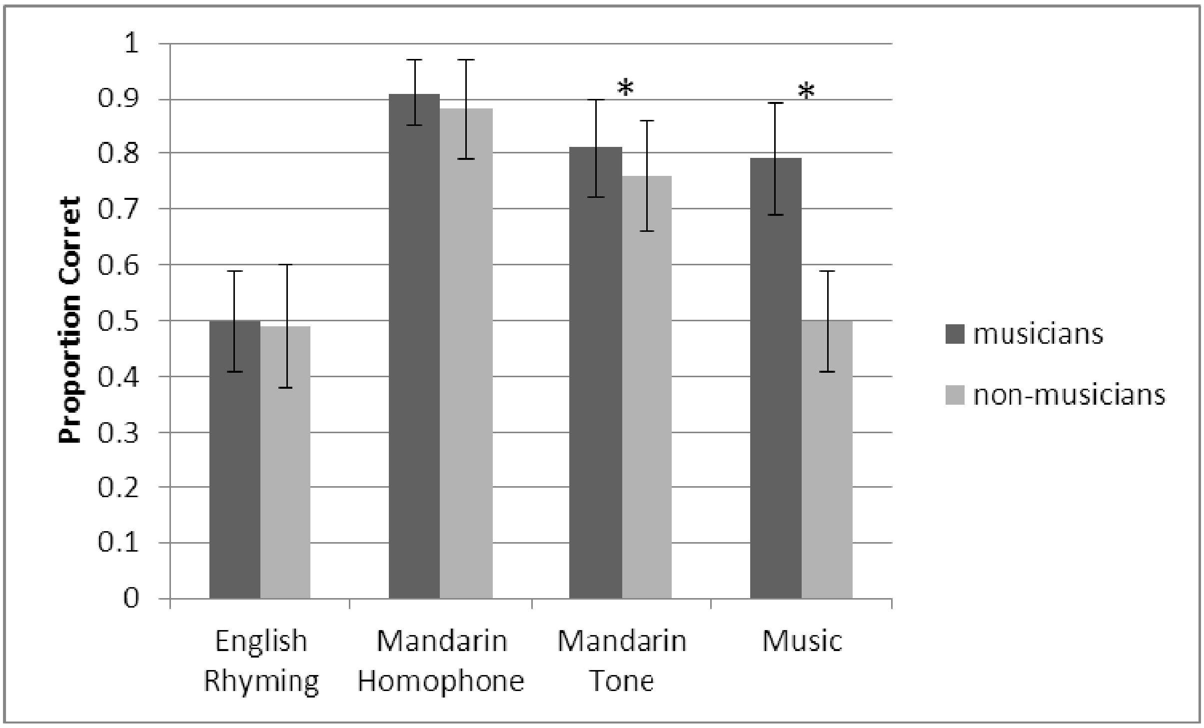
Figure 4.2 Accuracy rates for musicians and non-musicians in the 1-back tasks.

Among two groups, four one-way ANOVA were conducted. In the Mandarin homophone task, mean accuracy rates were significantly higher for the musicians than for the non-musicians in the 1-back task [F (1, 58) = 6.69, p = .012]. In the Mandarin tone task, the musicians' mean accuracy rate was significantly higher than the non-musicians' in the 1-back task [F (1, 58) = 17.69, p < .0005]. Finally, in the music task, mean accuracy rates were significantly higher for the musicians as compared to the non-musicians in the 1-back task [F (1, 58) = 138.40, p < .0005].

Within the musician group, accuracy for the 1-back tasks was examined using a one-way

ANOVA. There was a significant difference in the accuracy rates of musicians across 1-back tasks [F (3, 87) = 232.16, p < .0005]. Musicians showed the most accurate performance in the Mandarin homophone task (M = .96, SD = .02) and the Mandarin tone task (M = .96, SD = .02), followed by the music task (M = .94, SD = .05), and the English rhyming task (M = .72, SD = .07). Within the non-musician group, accuracy data across tasks was also examined using a one-way ANOVA. A significant difference in accuracy across all 1-back tasks was observed [F (3, 87) = 95.25, p < .0005] in that non-musicians were most accurate in the Mandarin homophone task (M = .94, SD = .04), followed by the Mandarin tone task (M = .92, SD = .04), the music task (M = .70, SD = .10), and the English rhyming task (M = .70, SD = .11). Mean accuracy scores for the 1-back tasks are shown above in Table 4.1

The accuracy results of the 2-back tasks were examined separately from the 1-back tasks. The 2-back task accuracy data were also subjected to a 2 (Group) x 4 (Task) mixed model ANOVA. There were two significant main effects and one interaction effect. Significant main effect of group [F (1, 55) = 39.00, p < .0005] in that the accuracy rate of musicians (M = .75, SD = .05) was higher than non-musicians (M = .67, SD = .06) across all 2-back tasks. There also was a significant main effect of task [F (3, 165) = 239.96, p < .0005], in that all participants showed the highest accuracy rates in the Mandarin homophone task (M = .90, SD = .08), followed by the Mandarin tone task (M = .78, SD = .10), the music task (M = .65, SD = .17), and the English rhyming task (M = .50, SD = .10). There was a significant interaction between group and task [F (3, 165) = 35.00, p < .0005]. The sources of this interaction will be described below and in Figure 4.3.



*$p < .05$

Figure 4.3 Accuracy rates for musicians and non-musicians in the 2-back tasks.

Among two groups, four one-way ANOVA were conducted. The musicians' mean accuracy rate was significantly higher than the non-musicians' in the 2-back Mandarin tone task [F (1, 58) = 4.70, p = .034]. Mean accuracy rates were also significantly higher for the musicians as compared to the non-musicians in the 2-back music task [F (1, 55) = 125.01, p < .0005].

Within the musician group, accuracy data in the 2-back tasks were examined using a one-way ANOVA. There was a significant difference in accuracy across tasks [F (3, 87) = 158.78, p < .0005], with musicians showing the most accurate performance in the Mandarin homophone task (M = .91, SD = .06), followed by the Mandarin tone task (M = .81, SD = .09), the music task (M = .79, SD = .10), and the English rhyming task (M = .50, SD = .09). Within the non-musician group, accuracy data in the 2-back tasks were also subjected to a one-way ANOVA. Significant differences in accuracy across all 2-back tasks were observed in the non-musicians [*F* (3, 78) = 119.63, *p* < .0005], with the most accurate performance in the Mandarin homophone task (*M* = .88, *SD* = .09), followed by the Mandarin tone task (*M* = .75, *SD* = .10), the music task (*M* = .50, *SD* = .09), and the English rhyming task (*M* = .49, *SD* = .11). Mean accuracy scores for the 2-back tasks are shown above in Table 4.2.

Comparisons of Performance Speed.

The second research question posed in Chapter Two addressed the speed of task performance of the musician group and the non-musician group. It was assumed that the 2-back tasks were more demanding of working memory than the 1-back tasks, and that this increased task difficulty would be reflected in longer (i.e., slower) reaction times.

Research Question 2: How is speed of performance in musicians and non-musicians in four visual recognition tasks influenced by increased task demands on working memory?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians in four visual recognition tasks with increased task demands on working memory.

The mean speed of correct responses in each 1-back task was computed for each group (see Table 4.4 for descriptive statistics). All reaction times reported in this chapter are based on correct responses only (i.e., correct acceptances and correct rejections of the target items). All reaction times reported are in milliseconds.

Table 4.4

Mean Reaction Times of 1-back Tasks by Group

	Musicians [a]		Non-musicians [b]	
	Mean	*SD*	*Mean*	*SD*
English Rhyming	939	120	919	130
Mandarin Homophone	735	83	779	94
Mandarin Tone	876	131	934	133
Music	901	134	943	159

[a] $n = 30$ [b] $n = 30$.

The mean accuracy rates were computed for the musician group and the non-musician group for each 2-back task. See Table 4.5 for descriptive statistics.

Table 4.5

Mean Reaction Times of 2-back Tasks by Group

	Musicians [a]		Non-musicians [b]	
	Mean	*SD*	*Mean*	*SD*
English Rhyming	1002	151	993	204
Mandarin Homophone	835	128	891	120
Mandarin Tone	972	193	1076	154
Music	1013	178	888[c]	188

[a] $n = 30$ [b] $n = 30$ [c] $n = 27$.

To examine the effects of group, visual task, and difficulty level on performance speed, a 2 x 4 x 2 mixed model analysis of variance (ANOVA) was conducted. The variable Group included musicians versus non-musicians; the variable Task included English rhyming versus Mandarin homophone versus Mandarin tone versus Music tasks; and the variable Difficulty Level included 1-back versus 2-back tasks. Table 4.6 shows the significant main effects for Task and Difficulty Level. The reaction times of 1-back tasks (*M* = 878, *SD* = 97) were faster than 2-back tasks (*M* = 960, *SD* = 133). There were two significant two-way interactions: between group and tasks, and between tasks and difficulty level. There was also a significant three-way interaction. The results of the three-way repeated-measures ANOVA are listed in Table 4.6.

Table 4.6

Results of Three-Way Mixed Model ANOVA: Reaction Times for Group by Task by Difficulty Level

Source	*df*	*F*	η^2	*p*
Group [a]	1	.49	.01	.486
Difficulty [b]	1	40.56	.42	<.0005
Group * Difficulty	1	1.07	.02	.306
Within-group-error	55			
Tasks [c]	3	43.24	.44	<.0005
Group* Tasks	3	6.28	.10	<.0005
Tasks * Difficulty	3	6.30	.10	<.0005
Group * Task * Difficulty	3	9.49	.15	<.0005
Within-group-error	165			

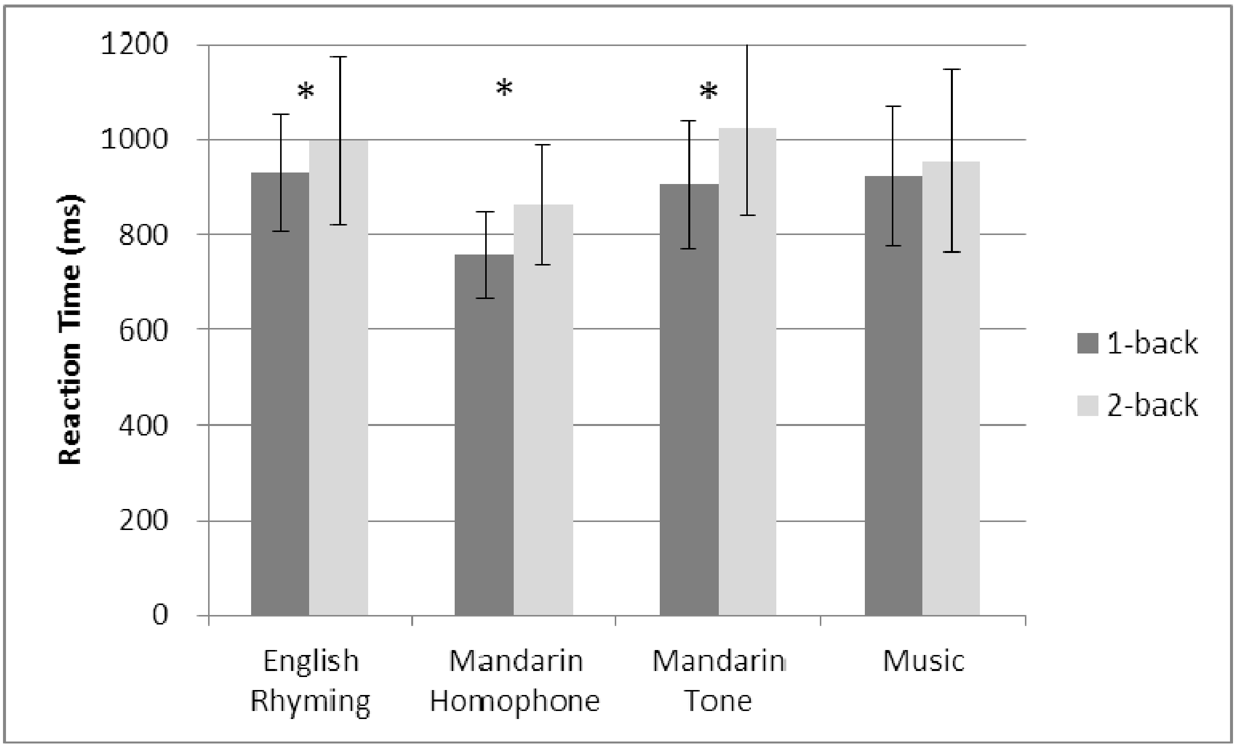
a. Group included musicians versus non-musicians;

b. Difficulty included 1-back versus -2-back paradigm;

c. Task included English rhyming versus Mandarin homophone versus Mandarin tone versus

Music

Four paired-samples t-tests were performed to compare the performance of all participants in the 1-back vs. the 2-back condition within each task, with reaction times as the dependent variable. Results indicated significant differences with increasing task difficulty across all participants in all tasks. Performance of all 1-back tasks was significantly faster than the 2-back condition in the rhyming task [t (59) = -3.93, p < .0005], the Mandarin homophone task [t (59) = -8.01, p < .0005], and the Mandarin tone task [t (59) = -6.05, p < .0005]. However, the music task showed no significant difference in speed between 1-back vs. 2-back [t (56) = -1.91, p = .061] across all participants. Reaction time for all participants in the 1-back vs. 2-back tasks is depicted in Figure 4.4.



*p < .05

Figure 4.4 Reaction times for all participants in the1-back tasks vs. 2-back tasks.

For the musician group, four paired-samples t-tests were performed to compare the 1-back to the 2-back tasks with speed as the dependent variable. Reaction time in the 1-back task was significantly faster than the 2-back condition in the rhyming task [t (29) = -2.89, p = .007**], the Mandarin homophone task [t (29) = -4.89, p < .0005], the Mandarin tone task [t (29) = - 3.95, p < .0005], and the music task [t (29) = - 5.10, p < .0005] in the musician group.

In the non-musician group, four paired-samples t-tests were performed to compare the 1-back to the 2-back tasks with speed as the dependent variable. Reaction times in the 1-back tasks were significantly faster than the 2-back condition in the rhyming task [t (29) = - 2.69, p = .012], the Mandarin homophone task [t (29) = -6.57, p < .0005] and the Mandarin tone task [t (29) = -4.61, p < .0005]. However, reaction times in the 1-back task were not significantly different from the 2-back condition in the music task [t (26) = 1.55, p = .133] in the non-musician group.

To further compare the performance of musicians versus non-musicians within each task, a one-way Group ANOVA was conducted for each task. For each 1-back and 2-back task, a separate one-way ANOVA was conducted comparing the reaction time of musicians versus non-musicians on the task.

Reaction time data for the 1-back tasks were examined using a 2 (Group) x 4 (Task) mixed model ANOVA. There was no significant main effect of group on speed [F (1, 58) = 1.54, p = .219] across all 1-back tasks. The overall reaction times of musicians (M = 863, SD = 103) did not significantly differ from non-musicians (M = 894, SD = 90). A significant difference in speed was observed across all 1-back tasks [F (3, 174) = 47.64, p < .0005] in that all participants showed the fastest speed in the Mandarin homophone task (M = 757, SD = 91), followed by the Mandarin tone task (M = 905, SD = 134), the music task (M = 922, SD = 147), and the English rhyming task (M = 930, SD = 124). There was not a significant interaction effect of 1-back tasks and group on speed [F (3, 174) = 2.17, p = .094]. The results of these 1-back task analyses are depicted in Figure 4.5.

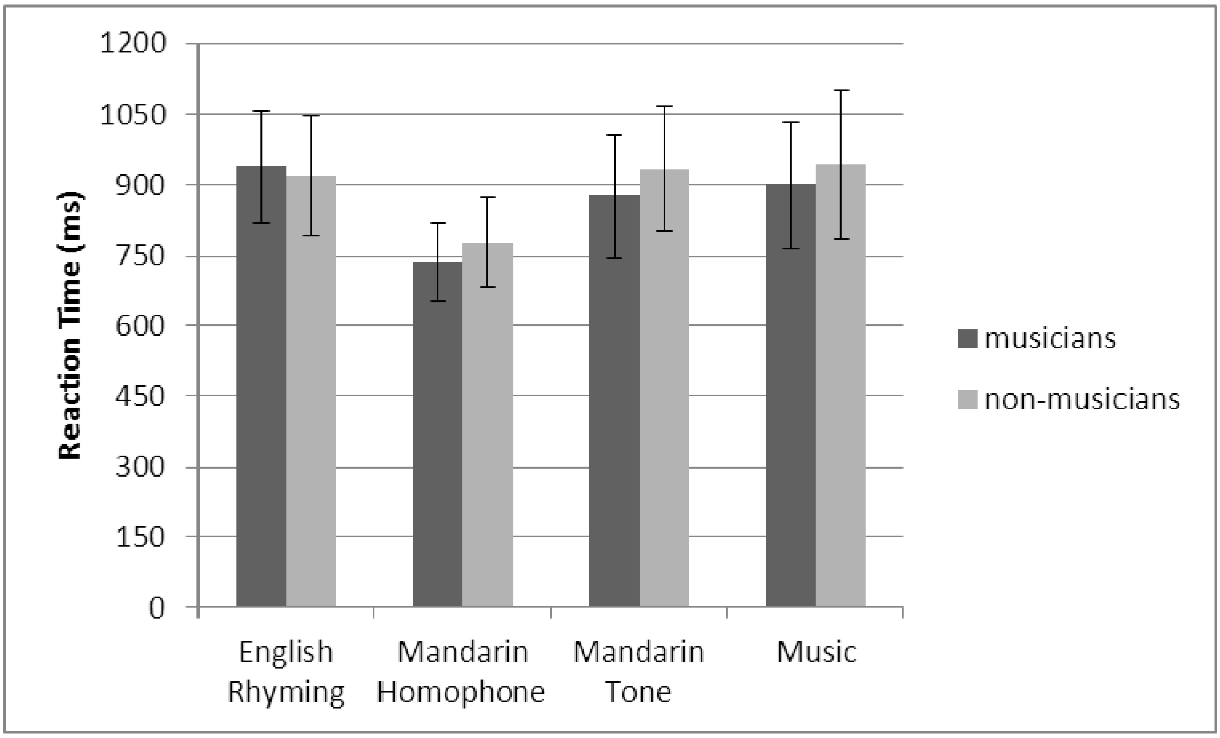
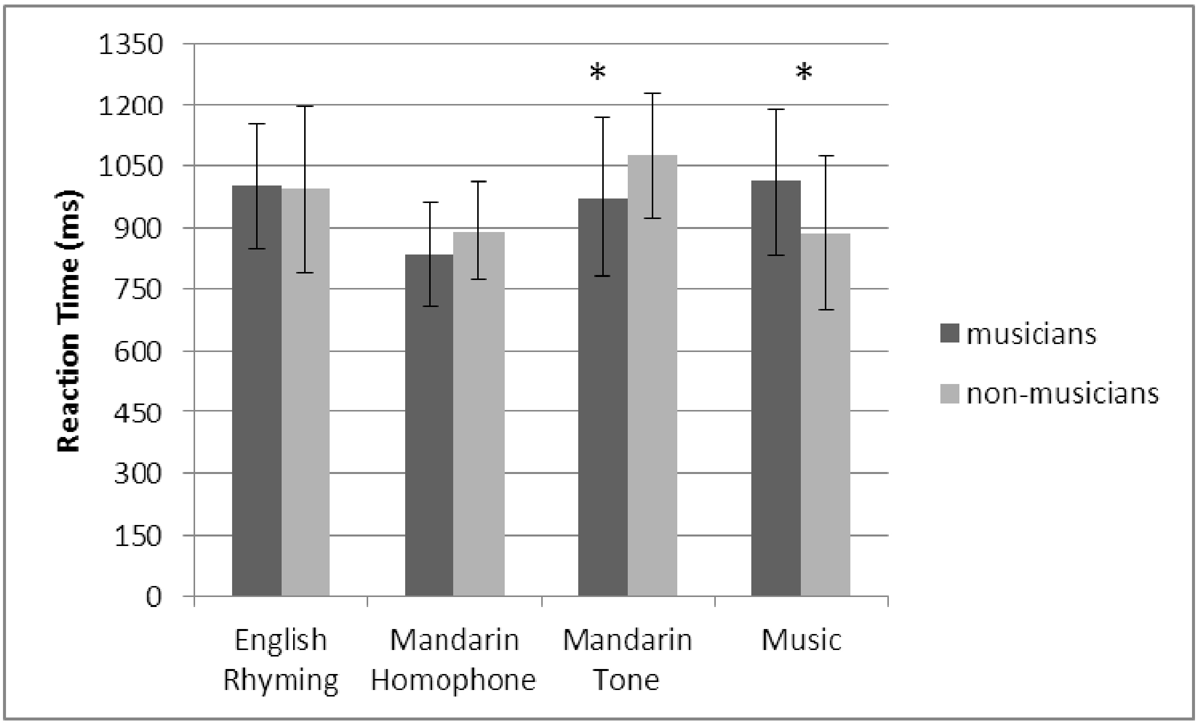


Figure 4.5 Reaction times of musicians and non-musicians for the 1-back tasks.

Within the musician group, a one-way repeated measures ANOVA was used to examine reaction time for the 1-back tasks. There was a significant difference in the reaction times of musicians across 1-back tasks [F (3, 87) = 50.81, p < .0005], with the fastest speed in the Mandarin homophone task (M = 735, SD = 83), followed by the Mandarin tone task (M = 876, SD = 131), the music task (M = 901, SD = 134), and the English rhyming task (M = 939, SD = 120). Within the non-musician group, reaction time data across 1-back tasks were also examined using a one-way repeated measures ANOVA. A significant difference in reaction times across all 1-back tasks was observed [F (3, 87) = 14.76, p < .0005] in that non-musicians showed that the fastest speed in the Mandarin homophone task (M = 779, SD = 94), followed by the English rhyming task (M = 919, SD = 130), the Mandarin tone task (M = 934, SD = 133), and the music task (*M* = 943, *SD* = 159). Mean reaction times for the 1-back tasks are shown above in Table 4.4.

The speed results of the 2-back tasks were examined separately from the 1-back tasks. The 2-back task accuracy data were also subjected to a 2 (Group) x 4 (Task) mixed model ANOVA. There were one significant main effects and one interaction effect. There was no significant main effect of group [F (1, 55) = .04, p = .836] in speed between musicians (*M* = 955, *SD* = 138) and non-musicians (*M* = 964, *SD* = 131) across all 2-back tasks. There was a significant main effect of task [F (3, 165) = 24.30, p < .0005], in that all participants showed the fastest reaction time in the 2-back Mandarin homophone task (*M* = 863, *SD* = 126), followed by the 2-back music task (*M* = 954, *SD* = 192), the 2-back English rhyming task (*M* = 997, *SD* = 178), and the 2-back Mandarin tone task (*M* = 1024, *SD* = 181). There was a significant interaction between group and task in reaction time [F (3, 165) = 10.65, p < .0005]. The sources of this interaction will be described below and in Figure 4.6.



*$p < .05$

Figure 4.6 Reaction times of musicians and non-musicians for the 2-back tasks.

Among two groups, four one-way ANOVA were conducted. Significant differences in reaction time were observed between the musicians and non-musicians in the 2-back Mandarin tone and the 2-back music tasks. In the Mandarin tone task, the reaction times of musicians were shorter (i.e., faster) than non-musicians in the 2-back condition [F (1, 58) = 5.34, p = .024]. In the music task, the reaction time of musicians was longer (i.e., slower) than the non-musicians in the 2-back task [F (1, 58) = 6.64, p = .013].

Within the musician group, reaction time data in the 2-back tasks were examined using a one-way repeated measures ANOVA. There was a significant difference in reaction time across tasks [F (3, 87) = 18.65, p < .0005], with musicians showing the fastest reaction times in the Mandarin homophone task (*M* = 835, *SD* = 128), followed by the Mandarin tone task (*M* = 972, *SD* = 193), the English rhyming task (*M* = 1002, *SD* = 151), and the music task (*M* = 1013, *SD* = 178). Within the non-musician group, reaction times in the 2-back tasks also were subjected to a one-way repeated measures ANOVA. Significant differences in reaction time across all 2-back tasks were observed [F (3, 78) = 16.24, p < .0005] in that non-musicians showed that the fastest speed in the Mandarin homophone task (*M* = 891, *SD* = 120), followed by the music task (*M* = 888, *SD* = 188), the English rhyming task (*M* = 993, *SD* = 204), and the Mandarin tone task (*M* = 1076, *SD* = 154). Mean reaction times for the 2-back tasks are shown above in Table 4.5.

Comparing the Effects of Visually Similar Stimuli.

Research Question 3: How is accuracy of performance in musicians and non-musicians in visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians in the visual recognition tasks with increased visual similarity of task stimuli.

The visual similarity of target stimuli was controlled in three of the four recognition tasks, including the English rhyming task, the Mandarin homophone task, and the music notation task. As described in Chapter Three, the English rhyming task and the Mandarin homophone task each included four different categories of target stimuli: orthographically similar/ phonologically similar (OSPS), orthographically similar/ phonologically dissimilar (OSPD), orthographically dissimilar/ phonologically similar (ODPS), and orthographically dissimilar/ phonologically dissimilar (ODPD). The music task included four different categories of target stimuli: the same visual location with the same pitch names (LSPS), the same visual location with different pitch names (LSPD), different visual locations with the same pitch names (LDPS), and different visual locations with different pitch names (LDPD). The accuracy rates of the musician group and the non-musician group for each stimulus category for these three visual recognition tasks are listed in Table 4.7. These accuracy rates are based on the 1-back tasks only.

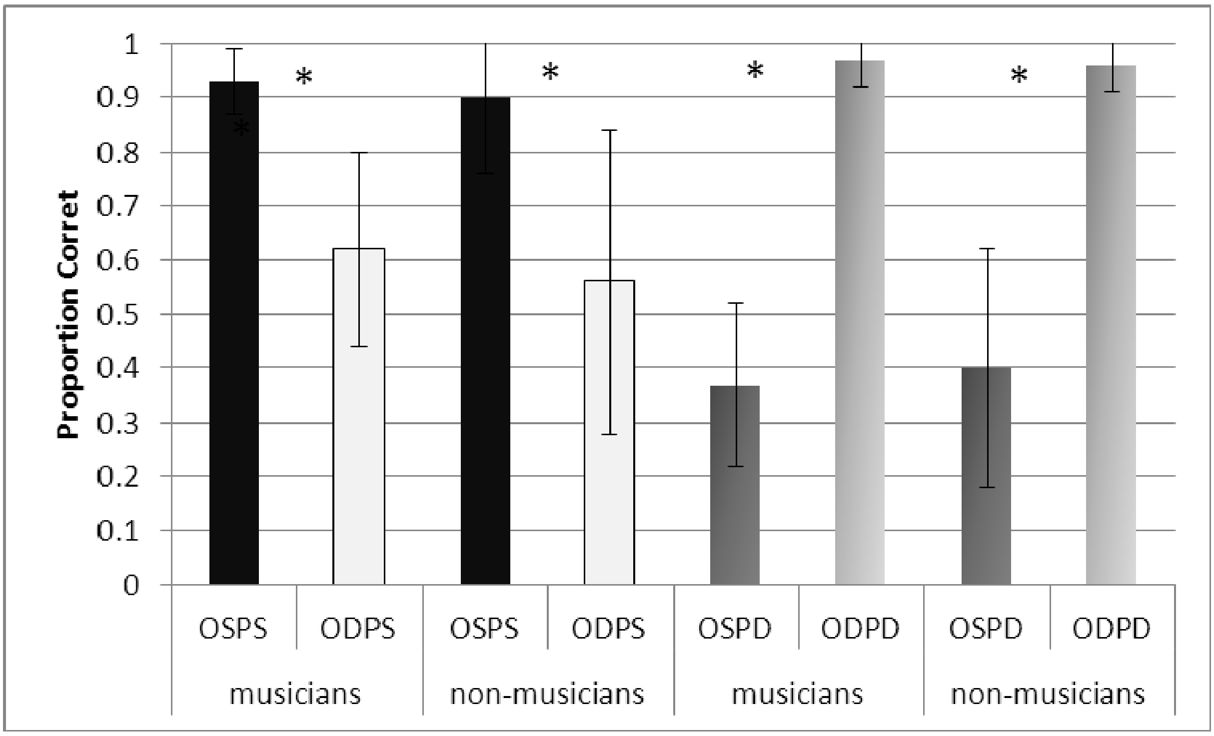
Table 4.7

Accuracy for Four Categories of Target Stimuli by Group

		Musicians [a]		Non-musicians [b]	
		Mean	*SD*	*Mean*	*SD*
English Rhyming	OSPS	.93	.06	.90	.14
	ODPS	.62	.18	.56	.28
	OSPD	.37	.15	.40	.22
	ODPD	.97	.05	.96	.05
Mandarin Homophone	OSPS	.91	.05	.87	.10
	ODPS	.95	.06	.91	.07
	OSPD	.99	.02	.98	.04
	ODPD	.99	.02	.99	.01
Music	LSPS	.91	.09	.77	.15
	LDPS	.90	.07	.28	.26
	LSPD	.95	.06	.83	.24
	LDPD	.99	.04	.91	.07

[a] n = 30 [b] n = 30.

The English Rhyming Task: Effects of Visual Similarity on Accuracy. To examine the effects of visual similarity on accuracy rates in the rhyming task, a $2_{between}$ (group: musicians versus non-musicians) x 4_{within} (categories: ODPS versus OSPD versus OSPS versus ODPD) two-way ANOVA was conducted. Accuracy rates were not significantly different for the stimulus categories between musicians and non-musicians in the rhyming task [F (3, 174) = 1.11, p = .345]. However, results from within group comparisons showed significantly different accuracy rates of the four categories in the rhyming task for each group. Among musicians, the accuracy rates of the four categories were significantly different [F (3, 87) = 187.36, p < .0005]. Musicians showed the most accurate performance for ODPD (M = .97, SD = .05), followed by OSPS (M = .93, SD = .06), ODPS (M = .62, SD = .18), and OSPD (M = .36, SD = .15) categories in the rhyming task. Among non-musicians, the accuracy rates of the four categories were significantly different [F (3, 87) = 60.01, p < .0005]. Non-musicians showed the best performance at ODPD (M = .96, SD = .05), followed by OSPS (M = .90, SD = .14), ODPS (M = .56, SD = .28), and OSPD (M = .40, SD = .22) categories in the rhyming task. Accuracy rates for each group for the four categories of target stimuli in the English rhyming task are shown in Figure 4.7.



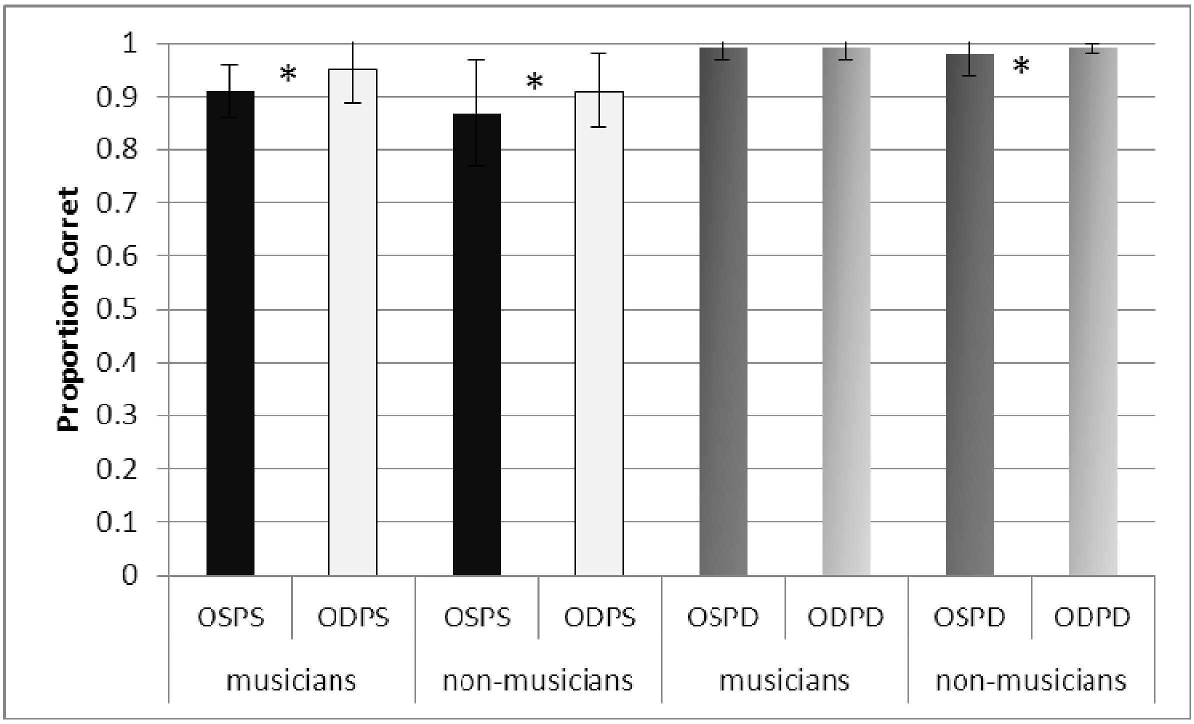
*p < .05

Figure 4.7 Accuracy rates for each stimulus category for English rhyming task in both groups.

To examine further the effect of visual similarity in the rhyming task, the accuracy rates for the four categories of target stimuli were compared within each group. Among musicians, performance for the OSPS category was significantly more accurate than for the ODPS category [t (29) = 10.37, $p < .0005$]. In other words, musicians were more likely to accept a correct match (i.e. phonologically similar match) when the target word looked similar to the probe than when the target word looked different from the probe. Also, musicians were significantly more accurate for the ODPD category as compared to the OSPD category [t (29) = 21.70, $p < .0005$] in the English rhyming task. In other words, musicians were more likely to incorrectly accept a phonologically dissimilar item (i.e., false positive response) when the target word looked similar to the probe than when the target word looked different from the probe.

Similarly, among non-musicians, performance for the OSPS category was significantly more accurate than the ODPS category [t (29) = 8.09, $p < .0005$] and performance for the ODPD category was significantly more accurate than the OSPD category [t (29) = 13.93, $p < .0005$] in the English rhyming task. In other words, for both musicians and non-musicians performance in the rhyming task was driven by the visual similarity of the target stimuli.

The Mandarin Homophone Task: Effects of Visual Similarity on Accuracy. To examine the effects of visual similarity on accuracy rates in the Mandarin homophone task, a $2_{between}$ (group: musicians versus non-musicians) x 4_{within} (categories: ODPS versus OSPD versus OSPS versus ODPD) two-way ANOVA was conducted. Accuracy rates were not significantly different for the four stimulus categories between musicians and non-musicians in the Mandarin homophone task [F (3, 174) = 2.35, $p = .074$]. However, results from within group comparisons showed significantly different accuracy rates for the four categories in the homophone task for each group. Results of a one-way ANOVA of the musician group showed that accuracy rates for the four categories were significantly different [F (3, 87) = 28.29, $p < .0005$]. Musicians showed the most accurate performance at ODPD (*M* = .99, *SD* = .02) and OSPD (*M* = .99, *SD* = .02), followed by ODPS (*M* = .95, *SD* = .06), and OSPS (*M* = .91, *SD* = .05) categories in homophone task. Results of a one-way ANOVA of the non-musician group showed that accuracy rates for the four categories were significantly different [F (3, 87) = 27.19, $p < .0005$]. Non-musicians showed the same pattern as the musicians in that they showed the most accurate performance for ODPD (*M* = .99, *SD* = .01), followed by OSPD (*M* = .97, *SD* = .04), ODPS (*M* = .91, *SD* = .07), and OSPS (*M* = .87, *SD* = .09) categories in the homophone task. Accuracy rates for each group for the four categories of target stimuli in the Mandarin homophone task are shown in Figure 4.8.



*$p < .05$

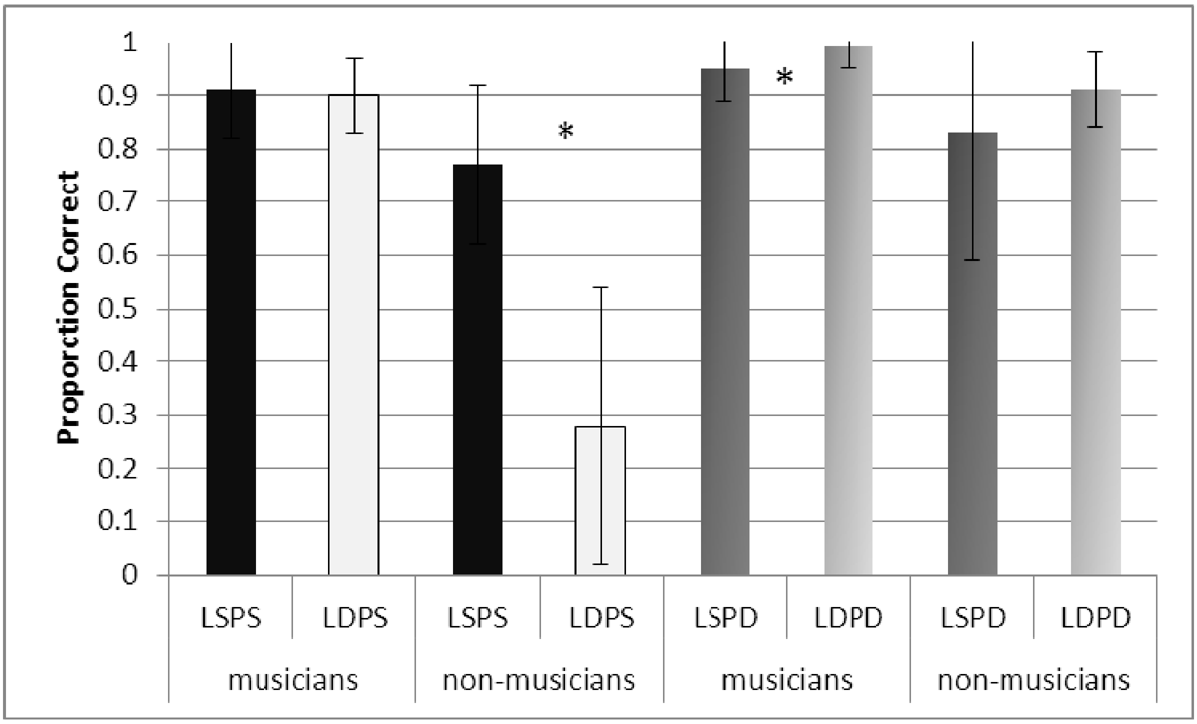
Figure 4.8 Accuracy rates for each stimulus category for Mandarin homophone task in both groups.

To further compare the performance of musicians versus non-musicians within each category, a one-way ANOVA was conducted for each category. Results from within group comparisons showed significantly different accuracy rates for the ODPS category in the homophone task [F (1, 58) = 5.34, p = .024] between musicians (M = .95, SD = .06) and non-musicians (M = .91, SD = .07). These analyses showed no other significant differences in accuracy rates for the other stimulus categories between the two groups in the Mandarin homophone task.

Among musicians, accuracy for the ODPS category was significantly more accurate than for the OSPS category [t (29) = -2.82, p = .009] in the homophone task. There was no significant difference in accuracy between the ODPD category and the OSPD category [t (29) = .92, p = .365] for musicians in the homophone task. Both of these results show that the performance of musicians in the homophone task was not driven by the visual similarity of target items. Among non-musicians, performance for the ODPS category was significantly more accurate than the OSPS category [t (29) = -2.451, p = .020] and performance in the ODPD category was significantly more accurate than the OSPD category [t (29) = 3.126, p = .004] in the homophone task. These results for both groups are inconsistent in terms of visual similarity in that visual similarity did not result in more correct acceptances of phonologically similar (PS) items (in fact, the reverse was true), but visual similarity did result in more incorrect acceptances of phonologically different (PD) items for non-musicians.

The Music Task: Effects of Visual Similarity on Accuracy. To interpret the effects of visual similarity on the accuracy rates in the music task, a $2_{between}$ (group: musicians versus non-musicians) x 4_{within} (categories: LDPS versus LSPD versus LSPS versus LDPD) two-way ANOVA was conducted. Accuracy rates were not significantly different for the four stimulus categories between musicians and non-musicians in the music task [F (3, 174) = 30.30, p

< .0005]. Results from within group comparisons showed significantly different accuracy rates for the four categories in the music task for each group. Results of a one-way ANOVA of the musician group showed that accuracy rates for the four categories were significantly different [F (3, 87) = 18.24, p < .0005]. Musicians showed the most accurate performance for LDPD (M = .99, SD = .04), followed by LSPD (M = .95, SD = .06), LSPS (M = .91, SD = .09), and LDPS (M = .90, SD = .07) categories in the music task. Results of a one-way ANOVA of the non-musician group showed that accuracy rates for the four categories were significantly different [F (3, 87) = 66.50, p < .0005]. Non-musicians showed the most accurate performance for the LDPD (M = .91, SD = .07), followed by LSPD (M = .83, SD = .24), LSPS (M = .77, SD = .15), and LDPS (M = .28, SD = .26) categories in the music task. Accuracy rates for each group for the four categories of target stimuli in the music task are shown in Figure 4.9.



*p < .05

Figure 4.9 Accuracy rates for each stimulus category for the music task in both groups.

To further compare the performance of musicians versus non-musicians within each category in the music task, a one-way ANOVA was conducted for each category. Results from these comparisons showed that accuracy rates of musicians were significantly more accurate than non-musicians for the LDPS category [F (1, 58) = 164.06, p < .0005], the LSPD category [F (1, 58) = 7.339, p = .009], the LSPS category [F (1, 58) = 20.20, p < .0005], and the LDPD category [F (1, 58) = 26.43, p < .0005] in the music task.

To examine further the effect of visual similarity in the music task, the accuracy rates for correct 'yes' (phonologically similar, PS) targets were compared (LDPS versus LSPS) and the accuracy rates for correct 'no' (phonologically dissimilar, PD) targets were compared (LSPD versus LDPD) for each group. Among musicians, there was no significant difference in accuracy for the LDPS and LSPS categories [t (29) = -.19, p = .851] in the music task. Performance in the LDPD category was significantly more accurate than for the LSPD category [t (29) = -4.40, p < .0005] for musicians. Thus, visual similarity did not influence performance of musicians given the PS (correct 'yes') targets, but it did influence their performance given the PD (correct 'no') targets in that they were more likely to incorrectly accept a PD target if it looked similar to the probe.

Among non-musicians, accuracy of performance for the LSPS category was significantly more accurate than the LDPS category [t (29) = 9.00, p < .0005]. In other words, non-musicians were more likely to accept a correct match as being phonologically similar if it also looked similar as compared to when it looked different from the probe. However, there was no significant difference in accuracy for the LDPD category and the LSPD categories [t (29) = 1.789, p = .084] in the music task. Thus, overall in the music task the influence of visual similarity on performance accuracy was inconsistent.

Research Question 4: How is speed of performance in musicians and non-musicians on visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians on visual recognition tasks with increased visual similarity of task stimuli.

As described above, target stimuli in the English rhyming task, the Mandarin homophone task, and the music task were controlled to allow the possible influence of visual similarity to be assessed. The reaction times of the musician group and the non-musician group for each stimulus category for these three visual recognition tasks are listed in Table 4.8. These reaction times (in milliseconds) are based on correct responses in the 1-back tasks only.

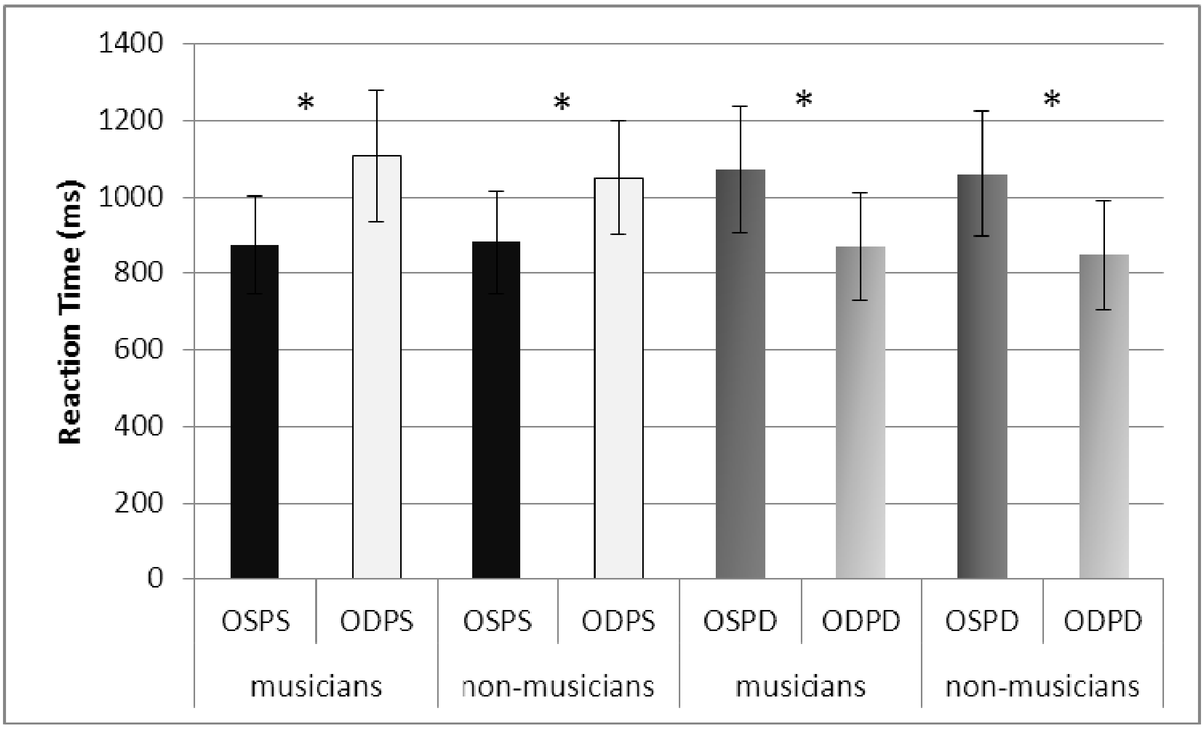
Table 4.8

Reaction Time for Four Categories of Target Stimuli by Group

		Musicians [a]		Non-musicians [b]	
		Mean	*SD*	*Mean*	*SD*
English Rhyming	OSPS	873	128	880	135
	ODPS	1105	172	1049	151
	OSPD	1071	167	1060	164
	ODPD	869	140	847	145
Mandarin Homophone	OSPS	751	88	797	86
	ODPS	723	78	778	108
	OSPD	775	95	821	112
	ODPD	691	95	726	117
Music	LSPS	859	113	905	131
	LDPS	945	142	1093	189
	LSPD	952	159	960	168
	LDPD	855	156	946	193

[a] $n = 30$ [b] $n = 30$.

The English Rhyming Task: Effects of Visual Similarity on Reaction Time. To interpret the effects of visual similarity on response speed in the English rhyming task, a $2_{between}$ (group: musicians versus non-musicians) x 4_{within} (categories: ODPS versus OSPD versus OSPS versus ODPD) two-way ANOVA was conducted. There was no significant difference for the four categories between musicians and non-musicians in the rhyming task [F (3, 174) = 1.19, p = .315]. However, results from within group comparisons showed significantly different reaction time for the four categories of target stimuli in the rhyming task for each group. Results of a one-way ANOVA of the musician group showed that accuracy rates for the four categories were significantly different [F (3, 87) = 38.23, $p < .0005$]. Musicians showed the fastest reaction time for ODPD (M = 869, SD = 140), followed by OSPS (M = 873, SD = 128), OSPD (M = 1071, SD = 167), and ODPS (M = 1105, SD = 151) categories in the rhyming task. Results of a one-way ANOVA of the non-musician group showed that accuracy rates for the four stimulus categories were significantly different [F (3, 87) = 44.84, $p < .0005$]. Non-musicians showed the fastest reaction time for ODPD (M = 847, SD = 145), followed by OSPS (M = 880, SD = 135), ODPS (M = 1049, SD = 151), and OSPD (M = 1060, SD = 164) categories in the rhyming task. Reaction times for each category in the English rhyming task are presented in Figure 4.10.



*$p < .05$

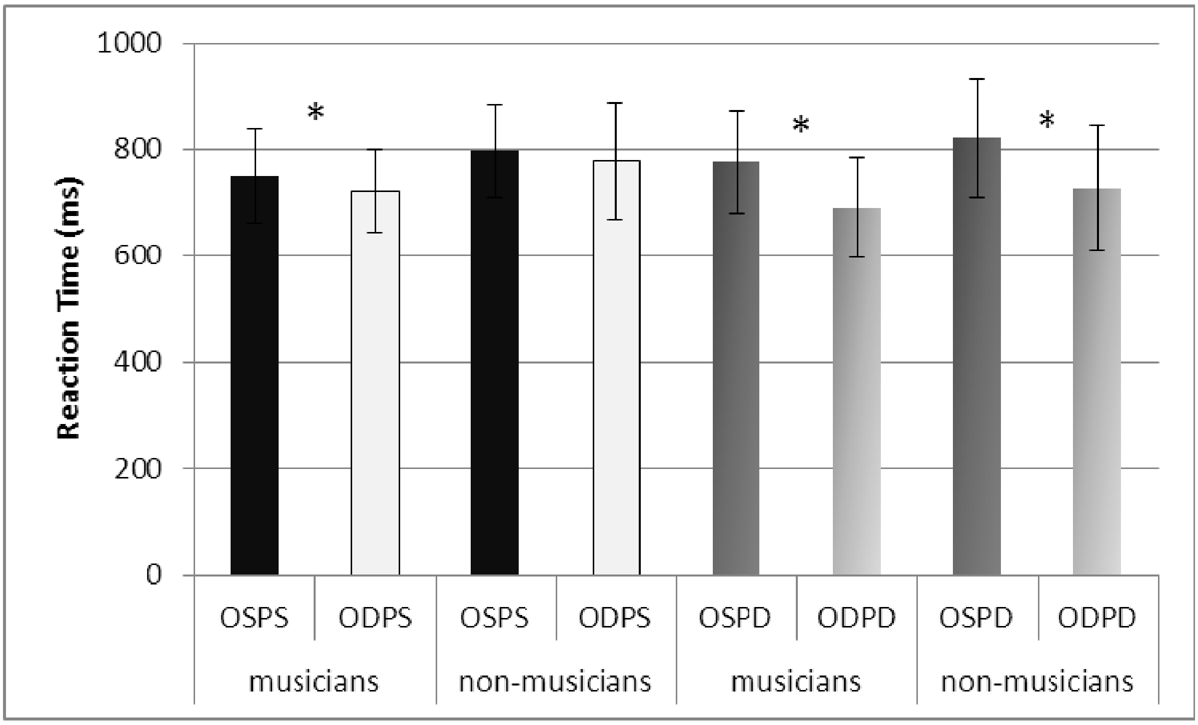
Figure 4.10 Reaction time for each stimulus category for the English rhyming task in both groups.

To further compare the performance of musicians versus non-musicians within each category in the music task, a one-way ANOVA was conducted for each category. Results from these comparisons showed no significant differences in reaction times across the four stimulus categories in the rhyming task.

Among musicians, reaction times for the OSPS category were significantly faster than the ODPS category [t (29) = 9.15, p < .0005] and reaction times for the ODPD category were also significantly faster than the OSPD category [t (29) = 7.49, p < .0005] in the rhyming task. Non-musicians showed the same pattern as musicians in that their reaction times for the OSPS category were significantly faster than the ODPS category [t (28) = 8.15, p < .0005] and reaction time for the ODPD category was significantly faster than the OSPD category [t (29) = 9.14, p < .0005] in the rhyming task. Thus, visual similarity of stimuli influenced reaction times for both groups in that all participants were faster in accepting a correct (PS) match if the target looked like the probe (OSPS) as compared to when the target did not look like the probe (ODPS). Also, all participants were faster in incorrectly accepting a phonologically dissimilar item (i.e., false positive response) when the target word looked similar to the probe (OSPD) than when the target word looked different from the probe (ODPD).

The Mandarin Homophone Task: Effects of Visual Similarity on Reaction Time. To interpret the effects of visual similarity on reaction times in the Mandarin homophone task, a 2_{between} (group: musicians versus non-musicians) x 4_{within} (categories: ODPS versus OSPD versus OSPS versus ODPD) two-way ANOVA was conducted. There was no significant interaction effect of the four stimulus categories and group in the Mandarin homophone task [F (3, 174) = .40, p = .753]. Results from within group comparisons showed significantly different reaction times for the four categories in the homophone task for each group. Results of a one-way

ANOVA of the musician group showed that reaction times for the four categories were significantly different [F (3, 87) = 25.80, p < .0005]. Musicians showed the fastest reaction time for ODPD (M = 691, SD = 95), followed by ODPS (M = 723, SD = 78), OSPS (M = 751, SD = 88), and OSPD (M = 775, SD = 95) categories in the homophone task. Results of a one-way ANOVA of the non-musician group showed that reaction times for the four categories were significantly different [F (3, 87) = 14.40, p < .0005]. Non-musicians showed the fastest reaction times for ODPD (M = 726, SD = 117), followed by ODPS (M = 778, SD = 108), OSPS (M = 797, SD = 86), and OSPD (M = 821, SD = 112) categories in the homophone task. Reaction times for each group for the four categories of target stimuli in the Mandarin homophone task are shown in Figure 4.11.



*p < .05

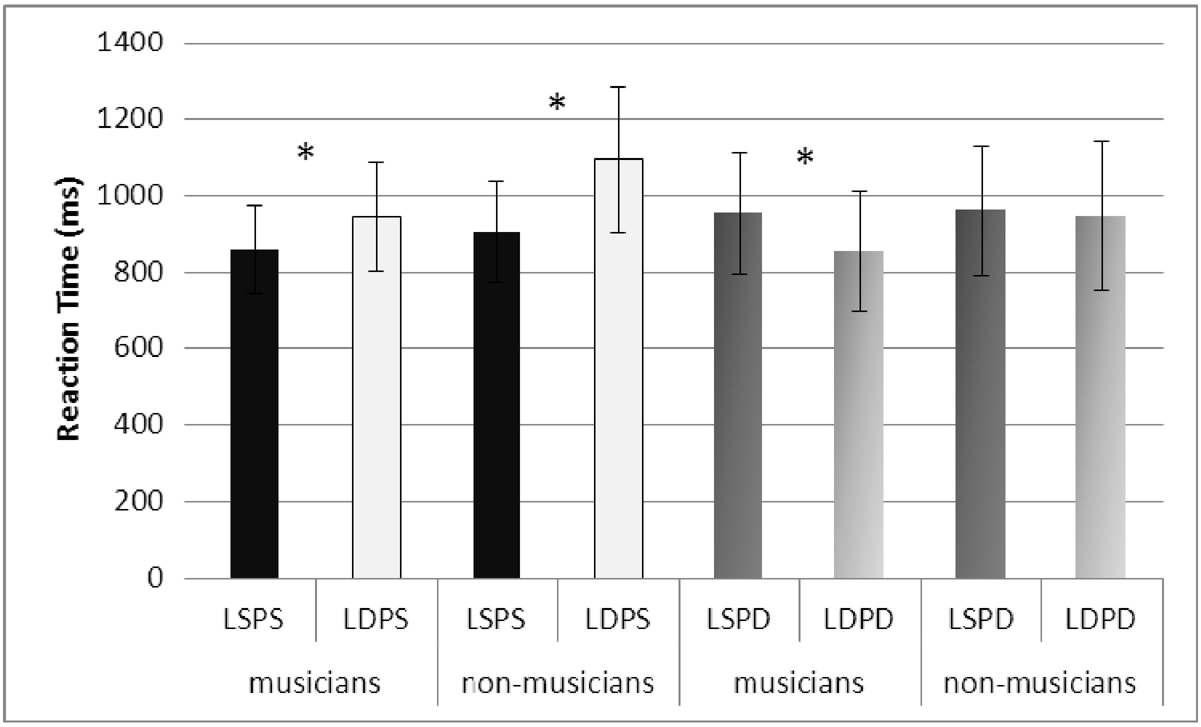
Figure 4.11 Reaction time for each stimulus category for the Mandarin homophone task in both groups.

To further compare the performance of musicians versus non-musicians within each category in the music task, a one-way ANOVA was conducted for each category. These comparisons showed that there was a significant difference between musicians and non-musicians in reaction time for the ODPS category [F (1, 58) = 5.12, p = .027] and for the OSPS category [F (1, 58) = 4.17, p = .046] in the homophone task. These analyses showed no other significant differences in reaction times for the other stimulus categories between the two groups in the homophone task.

Among musicians, reaction times for the ODPS category were significantly faster than for the OSPS category [t (29) = 3.78, p = .001], and reaction times for the ODPD category were faster than reaction times for the OSPD category [t (29) = 9.21, p < .0005] in the homophone task. Among non-musicians, reaction times for the OSPS category and ODPS category showed no significant difference [t (29) = 1.47, p = .153], whereas reaction times for the ODPD category were faster than the OSPD category [t (29) = 8.41, p < .0005]. Thus, the effects of visual similarity on reaction time for both the musician and the non-musician groups in the homophone task were inconsistent.

The Music Task: Effects of Visual Similarity on Reaction Time. To interpret the effects of visual similarity on reaction times in the music task, a 2_{between} (group: musicians versus non-musicians) x 4_{within} (categories: LDPS versus LSPD versus LSPS versus LDPD) two-way ANOVA was conducted. There was no significant difference in reaction times for the four stimulus categories between musicians and non-musicians in the music task [F (3, 174) = .40, p = .753]. Results from within group comparisons showed significantly different reaction times for the four categories in the music task for each group. Results of a one-way ANOVA of the musician group showed that reaction times for the four categories were significantly different [F

(3, 87) = 19.40, $p < .0005$]. Musicians showed the fastest reaction times for LDPD ($M = 855$, $SD = 156$), followed by LSPS ($M = 859$, $SD = 113$), LDPS ($M = 945$, $SD = 142$), and LSPD ($M = 952$, $SD = 159$) categories in the music task. Results of a one-way ANOVA of the non-musician group showed that reaction times for the four categories were significantly different [F (3, 87) = 12.71, $p < .0005$]. Non-musicians showed the fastest reaction time for LSPS ($M = 905$, $SD = 131$), followed by LDPD ($M = 946$, $SD = 193$), LSPD ($M = 960$, $SD = 168$), and LDPS ($M = 1093$, $SD = 189$) categories in the music task. Reaction times for each group for the four categories of target stimuli in the music task are shown in Figure 4.12.



*$p < .05$

Figure 4.12 Reaction time for each stimulus category for the music task in both groups.

To further compare the performance of musicians versus non-musicians within each category in the music task, a one-way ANOVA was conducted for each category. Results from these comparisons showed that showed that reaction times of musicians were significantly faster than non-musicians for the LDPS category [*F* (1, 54) = 10.98, p = .002] and for the LDPD category [*F* (1, 58) = 4.05, p = .049] in the music task. These analyses showed no other significant differences in reaction times for the other stimulus categories between the two groups in the music task.

Results of a one-way ANOVA of the musician group showed that reaction times for the LDPD category were significantly faster than the LSPD category [t (29) = 8.86, $p < .0005$], and reaction time for the LSPS category was significantly faster than for the LDPS category [t (29) = 5.93, $p < .0005$] in the music task. Results of a one-way ANOVA of the non-musician group showed that reaction times for the LSPS category were significantly faster than for the LDPS category [t (25) = 5.72, $p < .0005$]; whereas LSPD category and LDPD categories showed no difference [t (29) = .592, p = .558] in speed in the music task. Thus, the effects of visual similarity on reaction time for both the musician and the non-musician groups in the music task were inconsistent.

The results presented in this chapter were presented in relation to the research questions stated above in Chapter Two. These similarities and differences between musicians and non-musicians on the experimental tasks will be discussed in the next chapter.

CHAPTER 5

DISCUSSION

The present study addressed three areas of interest: 1) comparison of print to sound translation of English versus Mandarin Chinese versus musical notation; 2) comparison of how performance on these reading tasks may relate to musical training; and 3) how increased working memory load may affect performance across these reading tasks. The effect of visual similarity of target stimuli on performance was also examined. The four reading tasks involved a range of visual stimuli that correspond to phonological information (i.e., written English), combined phonological and tone information in a task emphasizing phonological information (i.e., written Chinese), combined phonological and tone information in a task emphasizing tone information (i.e., written Chinese), and pitch name information (i.e., Western musical notation system).

The Effect of Musical Training on Verbal and Tonal Working Memory

Two of the research questions posed in Chapter Two (Questions 1 and 2) addressed the issue of how increased demands on working memory may affect task performance in this study. It was assumed that the 2-back tasks were more demanding of working memory than the 1-back tasks, and that this increased task difficulty would be reflected in lower accuracy scores and longer reaction times.

Research Question 1: How is accuracy of performance in musicians and non-musicians in four visual recognition tasks influenced by increased task demands on working memory? Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians on four visual recognition tasks with increased task demands on working memory.

Research Question 2: How is speed of performance in musicians and non-musicians in

four visual recognition tasks influenced by increased task demands on working memory?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians in four visual recognition tasks with increased task demands on working memory.

The results of this study support the assumption that the 2-back task is more difficult than the 1-back task. Accuracy rates in the four visual recognition tasks after increased task demands were significantly different in that the accuracy rate of 1-back tasks was higher than 2-back tasks. Presumably, this difference relates to working memory in that the 2-back task requires comparison of a target to a probe presented two items back in a list, whereas the 1-back task requires only a comparison of the target to the item immediately preceding it in the list.

In the current study, musicians were significantly more accurate than non-musicians in three 1-back tasks (Mandarin homophone, Mandarin tone, and music tasks) and in two 2-back tasks (Mandarin tone and music tasks). Reaction times did not differ across groups in the 1-back tasks, but the musicians were faster to respond in the 2-back Mandarin tone task and they were slower to respond in the 2-back music task as compared to the non-musicians. These accuracy and reaction time results presumably reflect the effects of musical training. Three explanations have been proposed to account for superior performance of musicians in verbal tasks as compared to non-musicians: 1) musical training results in expanded working memory capacity; 2) musical training provides LTM support for normal working memory; and, 3) musical training results in improved executive function.

As noted in Chapter Two, Gudnubdsdittur (2010) claimed that sight reading of the musical notation system requires general mental capacities such as working memory and mental speed. Lehman (2007) also suggested that the task of music reading is demanding of working memory. Skilled performers such as musicians and chess players appear to have an expanded working memory capacity (Chase & Simon, 1973; Stigler, 1984). Thus, one explanation for the better performance of musicians as compared to non-musicians in verbal tasks in the current study is that musical training resulted in increased working memory capacity in the musicians, making them less vulnerable to high working memory demands.

The notion that working memory capacity can be expanded by musical training may reflect the operation of a "single acoustic loop" (Schendel, 2006). Based on evidence from articulatory suppression experiments and irrelevant sound effect experiments using verbal versus music stimuli (described in Chapter Two), Schendel (2006) proposed that a single acoustic loop with a single rehearsal mechanism is the most parsimonious explanation for similarities in language and music results. He suggested that using one acoustic loop for both phonological and tonal information may allow musicians to conserve cognitive resources. Interestingly, in the current study the performance of musicians in the 2-back Mandarin tone task was not significantly different in accuracy or speed from their performance in the 2-back music task. These results could be interpreted as evidence that working memory for language and working memory for music are both governed by a single acoustic loop. On this account, potential improvements in tonal working memory as a result of musical training could expand the capacity of this working memory loop and consequently improve verbal working memory.

Superior performance of musicians to non-musicians was also observed by Schulze and colleagues (2011) in a study comparing processing of verbal syllables versus sine wave tones in the two groups; however, they did not compare linguistic tone and musical tone. Schulze and colleagues (2011) proposed that two working memory systems may exist in musicians: a phonological loop supporting maintenance of phonological information, and a tonal loop supporting maintenance of pitch information. More research is needed to test this hypothesis of two working memory systems in musicians as opposed to the single acoustic loop hypothesized by Schendel (2006). However, similar performance of musicians in the Mandarin tone and music tasks in the current study appears to support the hypothesis of a single acoustic loop.

The second explanation for superior performance of musicians in verbal tasks as compared to non-musicians noted above is that musical training provides LTM support for normal working memory. On this account, long-term memory for musical information could interact with working memory. In Figure 2.2 above (Vallar, 2006), connections from verbal LTM to the phonological short-term store and the rehearsal process illustrate the support of LTM systems in aspects of immediate retention. Presumably, LTM for musical information could function in the same way as verbal LTM to support working memory. In the current study, the musicians have expertise in musical tone and this may be why they performed significantly more accurately and faster than non-musicians in recoding print into linguistic tone in the Mandarin tone task. Similar performance of musicians in the Mandarin tone and music tasks in the current study suggests that these forms of tonal information may exert a similar influence on working memory.

Long-term memory does appear to have influenced performance in verbal tasks in the current study in that both musicians and non-musicians performed worse in the English rhyming task as compared to the Chinese tasks. One possible interpretation of these results is that they reflect long-term memory and its effects on verbal working memory. As noted above, language is a learning process influenced by top-down knowledge, such that high-level information (e.g., stored phonological, lexical and semantic representations) guides the search for lower-level information (e.g., from letters to words, character to meaning, or note to melody) (Medin, Ross & Markman, 2004). In the current study, Chinese is a native language to all participants. Thus, long-term memory may help participants to be less vulnerable under heavy working memory load in Chinese as compared to English.

Within the Chinese tasks in the current study, higher accuracy rates and faster speed were observed in the homophone task as compared to the Mandarin tone task. This pattern was observed in non-musicians for both for the 1-back and 2-back tasks. Musicians also performed faster in the homophone task as compared to the Mandarin tone task (for both 1-back and 2-back) but only demonstrated this difference in accuracy for the 2-back homophone and Mandarin tone tasks. Thus, for the musicians the effects of expert ability (i.e., musical training) diminish under conditions of increased working memory load and the differences in difficulty between the homophone and the Mandarin tone tasks are revealed. Taken together, these accuracy and reaction time for both groups suggest that it is easier to extract phonological information from Chinese script than tonal information. An alternative explanation is that it is more difficult to maintain linguistic tone information in working memory as compared to phonological information.

A suprasegmental phonological feature, such as a linguistic tone, may be harder to retrieve than phonological segments in Mandarin Chinese because tone cues are embedded in phonological information (Leong, 2002; Taft & Chen, 1992). A benefit of the current study design is that direct comparison of phonological and tonal recoding of Mandarin script was possible, allowing the greater difficulty of extracting tone information from Chinese as compared to phonological information to become evident. In line with the conclusions of Li and Ho (2011) who found that tone processing is harder than phonological processing in Chinese reading for dyslexic children, in the current study this pattern was observed in normal skilled readers under time pressure and with increased working memory load.

Interestingly, even though both groups performed best in the Mandarin homophone task among all the experimental tasks in the current study, the non-musicians were less accurate than the musicians in this task. Results of the current study show that musicians are better at determining the similarities and differences between Mandarin homophone word pairs than non-musicians, as reflected in the musicians' superior accuracy in the Mandarin homophone task. This homophone task requires phonological judgments, not tone judgments, and yet musical training makes it easier to perform.

The third explanation noted above for superior performance of musicians in verbal tasks as compared to non-musicians is that musical training results in improved executive function. Studies addressing the relationship between central executive function and musical working memory have focused on the musical training effect (Degé, Kubicek & Schwarzer, 2011; Hargreaves & Aksentijevic, 2011; Moreno, Bialystok, Barac, Schellenberg, Cepeda & Chau, 2011; Schellenberg, 2011). Moreno and colleagues (2011) found short-term musical training enhanced verbal intelligence and executive function. They concluded that musical training improved not only musical skills, but also transferred to improved verbal ability because cognitive processing of music overlaps with cognitive mechanisms used in language (specifically, executive function). In the current study, specific tests of executive function were not employed. In any case, it is difficult to disentangle the effects of executive function and working memory in musical training given the range of behaviors included in musical training and the need to engage executive function and other cognitive resources in demanding tasks of expertise such as reading musical scores or playing chess.

Both the musician and non-musician groups in the current study showed the same pattern of accuracy across visual recognition tasks in that they were most accurate in the Mandarin homophone task, followed by the Mandarin tone task, the music task and then the English rhyming task. It is not surprising that the musicians performed more accurately than the non-musicians on the music reading task; however, both groups performed very poorly on the English rhyming task. English is not the native language of the participants, and the time constraints of the task do not allow effortful articulation or grapheme to phoneme conversion strategies during phonological judgments. As discussed below, participants relied on the visual similarity of targets and probes in attempting to perform the rhyming task.

The Effect of Visual Similarity of Target Stimuli

Two of the research questions in Chapter Two (Questions 3 and 4) addressed the issue of how performance on the visual recognition tasks may relate to the visual characteristics of the target stimuli. Few verbal working memory experiments have been conducted using visual stimuli although it has been noted that orthographic similarity can affect performance in phonological working memory experiments (Tree et al., 2011).

Research Question 3: How is accuracy of performance by musicians and non-musicians in visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in accuracy of performance of musicians and non-musicians in visual recognition tasks with increased visual similarity of task stimuli.

Research Question 4: How is speed of performance in musicians and non-musicians in the visual recognition tasks influenced by increased visual similarity of task stimuli?

Null Hypothesis: There are no significant differences in speed of performance of musicians and non-musicians in visual recognition tasks with increased visual similarity of task stimuli.

The visual features and corresponding sound (i.e., phonological or pitch) characteristics of stimuli were carefully controlled in three of the four recognition tasks: English rhyming, Mandarin homophone, and musical notation. The results of each task will be discussed separately.

In the English rhyming task, the performance of musicians and non-musicians was driven by the visual similarity of the target stimuli, a similar pattern to that reported by Tree and colleagues (2011). These results suggest that, given the time constraints of the task, participants were unable to transcode print to sound and instead relied on visual similarity to judge whether target stimuli rhymed with the probes. This pattern of reliance on visual information was evident both in accuracy and reaction time results. Within the conceptual framework of Figure 2.2 (Vallar, 2006), a visual stimulus can be held briefly in a visual "short-term store" (i.e., iconic memory); thus, the *n*-back task can be performed using this visual information more rapidly than using phonological information, which is normally obtained later during the phonological recoding or phonological buffer stages.

In the Mandarin homophone task, accuracy results were inconsistent in regard to the visual characteristics of target stimuli. For both groups, visual similarity did not increase correct acceptances of phonologically similar (PS) items (in fact, the reverse was true), but visual similarity did increase incorrect acceptances of phonologically different (PD) items in the non-musician group. Also, the effects of visual similarity on reaction time for both the musician and the non-musician groups in the homophone task were inconsistent. The fact that both groups performed best in the Mandarin homophone task relative to the other experimental tasks suggests that participants were attempting to perform the task by recoding print into sound and were not relying merely on visual characteristics of the stimuli to perform the task.

In the music task, the performance of the musician group is emphasized here because the non-musicians were not trained to read musical notation. Regarding the effect of visually similar target items, results were inconsistent overall. Non-musicians were more likely to accept a correct match as being phonologically similar if it also looked similar as compared to when it looked different from the probe. Musicians' significantly slower reaction times overall in this task, compared to the non-musician group, suggests that the musicians likely attempted to transcode the notes into sound. In regard to the visual characteristics of target stimuli, the performance of musicians was inconsistent. Visual similarity did not influence accuracy performance of musicians given the PS (correct 'yes') targets, but it did influence their performance given the PD (correct 'no') targets in that they were more likely to incorrectly accept a PD target if it looked similar to the probe. Visual similarity did influence reaction times of the musicians in the music task in that they were faster to accept a correct match when it looked similar to the probe than when it did not look like the probe, and they were faster to correctly reject mismatches when they looked different from the probe than when they looked similar to the probe.

Conclusion

In conclusion, for the tasks and stimuli created for this study, participants performed more accurately across tasks in the 1-back condition as compared to the 2-back condition, presumably as a result of increased demands on working memory in the 2-back condition. Interestingly, in the Mandarin homophone and Mandarin tone tasks, musicians and non-musicians showed significantly different accuracy rates. Analysis of the effects of visual similarity on task performance provided more detailed information than previous studies about the influence of visual versus phonological information in working memory tasks involving visual stimuli.

Although the tonal information in Mandarin is embedded in phonological information, the current study provided evidence that musicians more readily (in terms of both accuracy and speed) extract this tonal information from print than are non-musicians. A striking result of the current study is that even in the Mandarin homophone task, which requires phonological judgments of print, the musicians demonstrated superior performance.

More accurate performance by musicians as compared to non-musicians in verbal working memory tasks may reflect several processes including an expanded working memory capacity, the influence of long-term memory on working memory, and/or improved executive function. The results of the current study do not rule out any of these hypotheses; however, similar performance in musicians for the Mandarin tone task and the music task in the current study suggest that linguistic tone and musical tone may involve the same or overlapping processes of working memory and/or long-term memory. Moreover, the current study provides evidence that musical training facilitates phonological language processing.

Limitations and Future Studies

Several limitations of the current study can be identified that may influence the design of future studies. First, although the study participants had studied English for an average of more than 11 years, the English rhyming task proved to be too difficult for these native speakers of Mandarin even in the 1-back condition. Their performance in the rhyming task provided a clear example of how the visual similarity of task stimuli can drive performance when participants are unable to convert print to sound under time constraints; however, poor performance in this task did not allow for further comparisons of English reading versus Mandarin reading in the current study.

In future studies, native English speakers could be included to compare performance of musicians versus non-musicians in the English rhyming task versus the music task. Also, future studies with native Mandarin speakers could incorporate an *n*-back task involving stimuli from the Zhu-Yin-Fu-Hao phonetic system. Zhu-Yin-Fu-Hao is used in Taiwan to represent pronunciations of Chinese characters in a manner similar to the phonetic system of English. In a future study, performance on the English rhyming task could be compared to a homophone task that uses stimuli from the Zhu-Yin-Fu-Hao system. Although the results may show better accuracy rates for the homophone task compared to the rhyming task, reaction times may show no difference between the two tasks because Zhu-Yin-Fu-Hao may require more time to make a judgment as compared to reading the Chinese characters directly. The time constraints of the English rhyming task could also be modified so that native Mandarin speakers could perform the task more successfully.

A second limitation of the current study as it relates to the study of print to sound translation in Mandarin is that the stimuli used did not represent the full range of written Mandarin categories, which include pictographs, ideographs, compound ideographs, loan characters, analogous characters, and phonetic compounds. Attempting to examine all of these categories of written Mandarin in one study would have been too ambitious; however, systematic analysis of print to sound translation of these different features of Mandarin in future studies would provide the basis for more detailed theoretical models of Mandarin reading.

A third limitation of the current study is that the Mandarin tone stimuli were not controlled for visual similarity, unlike stimuli in the other experimental tasks. The Mandarin tone stimuli were controlled for the two variables of phonological similarity and tone similarity, and an additional subtest contrasting visual similarity with tone similarity was not included to limit the overall time required for each participant to complete the experiment. However, in a future experiment, this additional subtest could be included to provide a more direct comparison of visual and tonal features as was completed in the music task.

A fourth limitation noted in this study is that there were more stimuli in the *n*-back task for which the correct response was ‘no’ as compared to ‘yes’. This reflected the many probe and filler items that were included in addition to the target items (for which half were correct ‘yes’ responses and half were correct ‘no’ responses). In this context, participants may have become biased toward responding ‘no’ on the target items. However, the task design was the same across all 1-back and 2-back tasks, allowing for the differential effects of English versus Mandarin versus music stimuli to be observed in musicians versus non-musicians.

Evidence from the current study suggests that musical training can facilitate language processing. Although the differences between musicians and non-musicians in the current study were often significant but fairly small in terms of proportion correct, the effect of musical training on language is likely to be modest in healthy adults with normal reading ability. However, for children learning to read, particularly dyslexic children, even a fairly small facilitation effect of musical training may have a strong real-world impact in acquisition of reading skills.

In summary, this is the first study to compare similarities and differences between verbal and musical working memory using written English, Mandarin Chinese and musical notation in musicians and non-musicians. This study provides evidence that music and language share cognitive processes related to working memory.

APPENDIX

ALL STIMULI- English Rhyming Task- Phonological Similar

English Rhyming Task- OSPS				English Rhyming Task- ODPS			
probe	*f*	target	*f*	probe	*f*	target	*f*
mark	83	dark	185	head	424	bed	127
arm	94	farm	140	pool	111	rule	76
damp	16	lamp	18	float	3	quote	17
man	1207	can	1771	wait	88	late	179
fence	30	hence	58	cope	21	soap	22
yield	16	shield	8	clue	15	zoo	9
cold	171	told	413	comb	6	roam	6
spark	12	bark	14	more	2214	four	360
smile	58	mile	48	you	3277	two	1398
win	55	skin	47	fix	14	kicks	3
soy	1	toy	6	line	298	sign	94
skip	5	chip	17	chew	2	shoe	14
coil	6	boil	12	leap	14	peep	2
hip	10	flip	4	heat	97	meet	149
peer	8	deer	13	goal	60	hole	58
nest	20	pest	4	boot	13	flute	1
cane	12	pane	3	bait	2	skate	1
cat	23	hat	56	bone	33	thrown	40
lock	24	rock	75	bead	1	deed	1
fool	37	cool	62	dual	1	jewel	1
noon	25	moon	60	fail	37	pale	58
fill	50	kill	63	low	173	go	626
boat	73	coat	43	he	9547	key	86
nine	81	fine	162	break	88	take	610
beat	67	seat	54	pair	50	bear	57
zip	1	lip	18	route	43	root	30
speed	83	feed	122	mail	47	scale	59
bat	18	mat	5	tail	24	dale	5
cake	13	fake	10	phone	54	loan	43
fear	127	year	656	spoon	6	tune	10

English Rhyming Task- Phonological dissimilar

English Rhyming Task- OSPD				English Rhyming Task- ODPD			
probe	*f*	target	*f*	probe	*f*	target	*f*
cow	29	slow	60	lunch	33	maid	31
paid	145	said	1962	meat	45	chair	66
put	438	but	4383	foam	37	gate	37
poor	113	floor	158	beer	35	fat	60
youth	82	south	239	ball	110	hand	431
five	289	give	389	deck	23	bowl	23
deaf	12	leaf	13	couch	12	berry	9
ski	5	hi	6	horse	117	news	102
crow	2	brow	6	town	212	son	165
love	232	move	171	part	499	help	311
sown	3	crown	19	bug	4	cart	5
soot	1	hoot	9	mash	1	frost	7
does	485	goes	89	fuel	17	bow	15
blown	9	clown	3	seed	40	tool	44
lost	173	most	1159	tale	21	pack	25
gasp	3	wasp	2	neck	81	king	88
peach	3	poach	1	nut	15	pill	15
pour	9	sour	3	kit	2	bland	3
wear	31	ear	25	yard	35	van	32
your	923	hour	145	boy	244	soon	199
blind	47	wind	63	dear	54	foot	70
drive	104	live	177	cave	9	bush	14
card	26	ward	25	mood	37	beach	61
gross	65	cross	54	paw	3	noose	3
lose	58	nose	60	pole	18	juice	11
word	265	lord	93	pat	35	sand	28
cheat	3	sweat	23	hope	172	play	200
blood	121	food	147	mint	7	stair	2
stove	5	glove	9	fall	147	size	141
hint	9	pint	13	shine	5	brake	2

Mandarin Homophone Task- Homophone

Mandarin Homophone Task- OSPS				Mandarin Homophone Task- ODPS			
probe	***f***	**target**	***f***	**probe**	***f***	**target**	***f***
狐	**85**	弧	**42**	頒	**983**	班	**5168**
柱	**354**	注	**3904**	財	**9377**	裁	**2586**
淑	**1039**	叔	**258**	酬	**532**	愁	**148**
精	**5319**	睛	**353**	城	**3340**	程	**16133**
青	**4933**	清	**11820**	持	**15542**	遲	**1243**
聶	**49**	躡	**11**	膜	**299**	磨	**448**
濃	**1202**	農	**12696**	距	**1580**	句	**973**
維	**6033**	唯	**2049**	輝	**2651**	灰	**592**
扭	**443**	紐	**1043**	叛	**487**	盼	**941**
眉	**481**	楣	**117**	其	**32547**	齊	**1235**
裡	**4617**	理	**36671**	陽	**2492**	洋	**2646**
鈔	**478**	抄	**192**	風	**9058**	豐	**4272**
玫	**148**	枚	**397**	服	**8150**	浮	**774**
啤	**136**	脾	**119**	共	**17387**	貢	**624**
洛	**1230**	絡	**1550**	悅	**416**	月	**33109**
捕	**3229**	補	**6426**	脹	**356**	仗	**209**
橋	**3358**	僑	**1185**	震	**2118**	鎮	**11808**
蚊	**167**	紋	**365**	植	**1319**	執	**8177**
椅	**370**	倚	**107**	價	**18350**	架	**2028**
祇	**624**	紙	**2210**	劇	**3082**	聚	**1688**
棉	**421**	綿	**240**	練	**3141**	戀	**477**
紅	**3632**	虹	**216**	烈	**3768**	獵	**224**
豬	**1723**	諸	**1417**	龍	**5681**	隆	**5047**
笛	**138**	迪	**535**	欺	**595**	妻	**1881**
冒	**1112**	帽	**276**	攤	**2653**	坍	**179**
河	**3006**	荷	**725**	桃	**3427**	淘	**334**
池	**1341**	馳	**335**	晚	**5699**	碗	**183**
錶	**413**	表	**44788**	續	**9537**	序	**2790**
沙	**2481**	砂	**1341**	惠	**2174**	匯	**2828**
托	**306**	託	**2112**	飯	**1772**	犯	**6095**

Mandarin Homophone Task- Non-homophone

Mandarin Homophone Task- OSPD				Mandarin Homophone Task- ODPD			
probe	***f***	**target**	***f***	**probe**	***f***	**target**	***f***
柏	**2317**	怕	**2068**	爸	**505**	傻	**116**
課	**4938**	裸	**225**	抹	**203**	硬	**1514**
彬	**369**	淋	**211**	肌	**207**	玲	**778**
鍋	**394**	禍	**1187**	杯	**1161**	山	**14240**
貓	**410**	描	**406**	俊	**2033**	籠	**475**
落	**5979**	客	**8584**	騎	**1184**	律	**3793**
疏	**1086**	流	**8690**	罷	**1226**	燒	**1566**
槍	**6293**	創	**3585**	訊	**8963**	殼	**349**
村	**8077**	材	**2617**	畢	**2193**	鬆	**1020**
促	**3321**	捉	**381**	隻	**1383**	梅	**1607**
貧	**527**	貪	**444**	送	**9174**	扮	**774**
哀	**369**	衷	**284**	貨	**7474**	知	**10443**
狼	**178**	狠	**199**	末	**934**	紀	**3821**
宮	**1850**	官	**10246**	斬	**189**	樣	**5328**
旱	**166**	早	**5623**	博	**2268**	稍	**1211**
喘	**117**	端	**2522**	菌	**296**	鳥	**712**
倉	**624**	創	**3585**	家	**34730**	領	**6553**
掌	**2321**	賞	**1578**	飲	**1685**	綠	**1220**
住	**8832**	汪	**471**	伴	**590**	順	**4313**
司	**26120**	可	**42120**	漢	**1818**	只	**14311**
斯	**6509**	期	**22594**	營	**14537**	碑	**327**
般	**4367**	船	**6311**	火	**7553**	指	**25436**
席	**6424**	度	**20312**	悲	**780**	扇	**141**
砰	**19**	枰	**63**	涼	**319**	島	**2032**
巴	**3200**	已	**33836**	必	**10482**	押	**1725**
谷	**910**	俗	**1398**	玻	**614**	抨	**647**
亦	**7901**	示	**34047**	瞬	**117**	泊	**305**
諒	**446**	京	**2417**	功	**5385**	訓	**3592**
登	**4502**	澄	**609**	轟	**553**	佔	**3090**
易	**9396**	踢	**113**	牙	**941**	牌	**3330**

Mandarin Tone Task- The same tone

Homophone with the same tone				Non-homophone with the same tone			
probe	*f*	target	*f*	probe	*f*	target	*f*
金	24968	今	19780	肉	2006	熱	4853
蜜	314	密	3656	琴	919	旗	1522
鼠	326	屬	7260	倍	1370	畢	2193
份	11062	憤	788	波	6113	撥	1398
書	9328	舒	374	氣	10424	去	16165
麵	544	面	24564	賭	2212	底	7095
移	5705	宜	5600	繞	436	讓	7625
地	55300	弟	2060	誠	1895	沉	316
堡	294	保	20560	雞	1208	居	5995
畫	9145	話	8922	嘗	382	廚	387
球	8110	求	14712	坦	1672	躺	157
秋	1730	丘	299	皮	1988	葡	294
錢	4668	前	50655	昌	2014	初	6269
欣	1394	心	25445	賣	5751	慢	1152
欲	2376	玉	3475	穿	1560	春	4073
汐	213	夕	663	簡	3496	緊	3636
祖	1398	組	15363	笛	138	讀	2162
扶	566	伏	781	短	4098	黨	21650
陵	257	菱	161	湯	498	他	49438
塵	506	臣	305	品	17493	敏	1438
庫	2739	酷	189	車	28603	稱	6637
協	11837	鞋	705	集	9984	局	30181
僅	5973	錦	1817	批	4075	拍	2048
規	13668	歸	2156	母	3193	米	1845
權	14201	泉	1870	串	1663	創	3585
隱	777	引	9066	庭	3995	糖	1237
帆	189	繁	1452	海	14160	喊	1237
瓜	914	刮	503	最	21422	造	12164
鬥	1140	豆	793	果	15486	古	2898
分	46884	芬	709	慎	1438	上	59278

Mandarin Tone Task- Different tones

Homophone with different tones				Non-Homophone with different tones			
probe	*f*	target	*f*	probe	*f*	target	*f*
販	**3230**	反	**14983**	抹	**203**	玲	**778**
國	**87186**	過	**32398**	領	**6553**	知	**10443**
回	**13702**	匯	**2828**	功	**5385**	紀	**3821**
師	**8974**	史	**2956**	殼	**349**	押	**1725**
窗	**822**	闖	**480**	估	**3090**	營	**14357**
划	**225**	花	**7458**	悲	**780**	湖	**2968**
牛	**2119**	扭	**443**	隻	**1383**	末	**934**
土	**9679**	兔	**102**	籠	**475**	價	**18350**
水	**23359**	稅	**6938**	家	**34730**	柏	**2317**
腿	**556**	退	**4768**	轟	**553**	送	**9174**
信	**11240**	辛	**1071**	鬆	**1020**	傻	**116**
促	**3321**	粗	**403**	貨	**7474**	排	**5726**
猜	**246**	財	**9377**	扇	**141**	牙	**941**
八	**29615**	壩	**147**	伴	**590**	可	**42120**
必	**10482**	鼻	**398**	緣	**1220**	台	**57187**
星	**4335**	行	**60532**	山	**14240**	訓	**3592**
喝	**1108**	合	**24693**	漢	**1818**	博	**2268**
洗	**1361**	西	**8364**	俊	**2033**	燒	**1566**
麻	**1590**	罵	**529**	牌	**3330**	杯	**1161**
四	**38338**	司	**26120**	順	**4313**	常	**11207**
高	**36004**	告	**10675**	溪	**3966**	全	**27249**
晚	**5699**	完	**10112**	袋	**1071**	娶	**150**
有	**106582**	由	**39598**	火	**7553**	罷	**1226**
是	**90378**	十	**76679**	只	**14311**	硬	**1514**
打	**10559**	大	**83354**	節	**8033**	扮	**772**
美	**23111**	妹	**335**	鐘	**2010**	菌	**296**
把	**6299**	拔	**870**	享	**1578**	民	**53007**
媽	**1518**	馬	**6467**	場	**36564**	刷	**514**
任	**19008**	人	**109366**	共	**17387**	百	**28979**
泛	**554**	煩	**525**	輛	**3062**	石	**5765**

Music Reading Task – the same pitch names

The same location with the same pitch name				Different location with the same pitch name			
file name	probe	file name	target	file name	probe	file name	target
11.jpg	d	11.jpg	d	192.jpg	e flat	185.jpg	E flat
35.jpg	a'	35.jpg	a'	173.jpg	f	511.jpg	f
51.jpg	d	51.jpg	d	149.jpg	b	142.jpg	B
70.jpg	c sharp	70.jpg	c sharp	127.jpg	G	532.jpg	G
81.jpg	A	81.jpg	A	104.jpg	D	111.jpg	d
101.jpg	A	101.jpg	A	86.jpg	F	551.jpg	F
126.jpg	F	126.jpg	F	63.jpg	C sharp	70.jpg	c sharp
152.jpg	e	152.jpg	e	48.jpg	a	573.jpg	a
175.jpg	a' flat	175.jpg	a' flat	22.jpg	B	29.jpg	b
186.jpg	F	186.jpg	F	14.jpg	g	585.jpg	g
505.jpg	G	505.jpg	G	51.jpg	d	44.jpg	D
523.jpg	E	523.jpg	E	95.jpg	a'	606.jpg	a'
548.jpg	c	548.jpg	c	125.jpg	E flat	132.jpg	e flat
565.jpg	G	565.jpg	G	170.jpg	c	628.jpg	c
594.jpg	b flat	594.jpg	b flat	181.jpg	A flat	188.jpg	a flat
610.jpg	e	610.jpg	e	690.jpg	e flat	165.jpg	e flat
627.jpg	B flat	627.jpg	B flat	673.jpg	a flat	666.jpg	A flat
653.jpg	a	653.jpg	a	629.jpg	d	151.jpg	d
661.jpg	C	661.jpg	C	644.jpg	F sharp	651.jpg	f sharp
692.jpg	g	692.jpg	g	603.jpg	E	145.jpg	E
4.jpg	D	4.jpg	D	595.jpg	c'	588.jpg	c
34.jpg	g	34.jpg	g	571.jpg	f sharp	113.jpg	f sharp
42.jpg	B flat	42.jpg	B flat	541.jpg	C	555.jpg	c'
73.jpg	f sharp	73.jpg	f sharp	527.jpg	B	22.jpg	B
93.jpg	f	93.jpg	f	512.jpg	g	505.jpg	G
513.jpg	a	513.jpg	a	514.jpg	b	69.jpg	b
522.jpg	D	522.jpg	D	542.jpg	D	549.jpg	d
554.jpg	b flat	554.jpg	b flat	581.jpg	C	130.jpg	C
563.jpg	E	563.jpg	E	626.jpg	A flat	633.jpg	a flat
581.jpg	C	581.jpg	C	665.jpg	G	87.jpg	G

Music Reading Task- different pitch names

The same location with different pitch names				Different location with different pitch names			
file name	probe	file name	target	file name	probe	file name	target
5.jpg	E	505.jpg	G	81.jpg	A	46.jpg	F
526.jpg	A	26.jpg	F sharp	115.jpg	a'	7.jpg	G
52.jpg	e	552.jpg	g	90.jpg	C	514.jpg	b
563.jpg	E	63.jpg	C	525.jpg	G	588.jpg	C
84.jpg	D	584.jpg	F	625.jpg	G	21.jpg	A
608.jpg	c	108.jpg	a	681.jpg	C	47.jpg	G
130.jpg	c	630.jpg	e	1.jpg	A	63.jpg	c sharp
651.jpg	f sharp	151.jpg	d sharp	22.jpg	B	143.jpg	c sharp
169.jpg	b flat	669.jpg	d	155.jpg	a'	163.jpg	C
690.jpg	e flat	190.jpg	c	144.jpg	D sharp	114.jpg	g sharp
31.jpg	d	531.jpg	F sharp	6.jpg	F	32.jpg	e
574.jpg	b	74.jpg	g	7.jpg	G	69.jpg	B
13.jpg	f	513.jpg	a	502.jpg	D	530.jpg	e
606.jpg	A	106.jpg	F sharp	510.jpg	e	588.jpg	c
147.jpg	G sharp	647.jpg	B	127.jpg	G	44.jpg	D
683.jpg	e flat	183.jpg	C	514.jpg	b	523.jpg	E
164.jpg	D flat	664.jpg	F	67.jpg	G	551.jpg	f
622.jpg	D	122.jpg	B flat	88.jpg	A	114.jpg	G sharp
95.jpg	a'	595.jpg	c'	1.jpg	A	22.jpg	B
547.jpg	B flat	47.jpg	G	43.jpg	C	64.jpg	D
65.jpg	E	565.jpg	G	87.jpg	G	105.jpg	E
581.jpg	C	81.jpg	A	126.jpg	F	148.jpg	a
109.jpg	b	609.jpg	d	501.jpg	C	522.jpg	D
628.jpg	c	128.jpg	a flat	552.jpg	G	573.jpg	A
174.jpg	g	674.jpg	B flat	601.jpg	C sharp	583.jpg	E flat
642.jpg	D	142.jpg	B	195.jpg	a' flat	174.jpg	g
181.jpg	A flat	681.jpg	C	153.jpg	f sharp	89.jpg	b flat
513.jpg	a	13.jpg	f	3.jpg	C	27.jpg	G
28.jpg	a	528.jpg	c	64.jpg	D	89.jpg	b flat
652.jpg	g sharp	152.jpg	e	108.jpg	a	132.jpg	e flat

REFERENCES

Acheson, D. J., & MacDonald, M. C. (2009a). Verbal working memory and language production: common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135 (1),* 50-68.

Acheson, D. J., & MacDonald, M. C. (2009b). Twisting tongues and memories: explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language, 60,* 329-350.

Alley, T. R., & Greene, M. E. (2008). The relative and perceived impact of irrelevant speech, vocal music and non-vocal music on working memory. *Current Psychology, 27,* 277-289.

Awh, E., Jonides, J., Smith, E., Schumacher, E. H., Koeppe, R. A., & Katz, S. (1996). Dissociation of storage and rehearsal in verbal working memory: Evidence from positron emission tomography. *Psychological Science, 7(1),* 25-31.

Baddeley, A. D., Chincotta, D., Stafford, L., & Turk, D. (2002). Is the word length effect in STM entirely attributable to output delay? Evidence from serial recognition. *The Quarterly Journal of Experimental Psychology, 55A (2),* 353-369.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed), *Recent advances in learning and motivation* (Vols. 8, pp. 47-90). New York: Academic Press.

Baddeley, A. D. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4(10),* pp.829-839.

Baddeley, A. D., & Larsen, J. D. (2007). The phonological loop unmasked? A comment on the evidence for a "perceptual-gestural" alternative. *The Quarterly Journal of Experimental Psychology, 60(4),* p.497-504.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D. (2010). Working memory. *Current biology, 20 (4),* R136-140.

Bao, M., Li, Z. H., Chen, X. C., Zhang, D., R. (2006). Backward inhibition in a task of switching attention within verbal working memory. *Brain Research Bulletin, 69,* 214-221.

Beeson, P. M., & Hillis, A. E. (2001). Comprehension and production of written words. In R. Chapey (4^{th} Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (pp. 572-595). New York, NY: Lippincott Williams & Wilkins.

Behrmann, M., Plaut, D. C., & Nelson, J. (1998). A literature review and new data supporting an interactive account of LBL reading. *Cognitive Neuropsychology, 15 (1/2),* 7-51.

Berz, W. L. (1995). Working memory in music: A theoretical model. *Music Perception, 12(3),* 353-364.

Bi, Y., C. (2006). *The representation and access of homophone in speech production*. Harvard University, Cambridge, Massachusetts.

Binder, K. & Borecki, C. (2008). The use of phonological, orthographic, and contextual information during reading: a comparison of adults who are learning to read and skilled adult readers. *Reading and Writing, 21,* 843-858.

Blumstein, S. E. (1998). Phonological aspects of aphasia. In M. T. Sarno (3^{rd} Ed.), *Acquired Aphasia* (pp.157-185), New York: Academic Press.

Buchsbaum, B. R., & D'Esposito, M. (2008). The search for the phonological store: from loop to convolution. *Journal of Cognitive Neuroscience, 20(5),* 762-778.

Campoy, G., & Baddeley, A. D. (2008). Phonological and semantic strategies in immediate serial recall. *Memory, 16(4),* 329-340.

Caplan, D., Rochon, E., & Water, G. S. (1992). Articulatory and phonological determinates of word length effects in span tasks. *Quarterly Journal of Experimental Psychology: Human*

Experimental Psychology, 45 (A), 177-192.

Caplan, D., & Water, G. S. (1994). Articulatory length and phonological similarity in span tasks: A reply to Baddeley and Andrade. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 47 (A),* 1055-1062.

Chase, W. G. & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing,* (pp. 215-281), New York, Academic Press.

Chen, S. H. A., & Desmond, J. E. (2005). Cerebrocerebellar networks during articulatory rehearsal and verbal working memory tasks. *Neuroimage, 24(2),* 332-338.

Chen, Z. Y., Cowell, P. E., Varley, R., & Wang, Y. C. (2009). A cross-language study of verbal and visuospatial working memory span. *Journal of Clinical and Experimental Neuropsychology, 31(4),* 385-391.

Chmiel, N. (1984). Phonological recoding for reading: the effect of concurrent articulation in a Stroop task. *British Journal of Psychology, 75,* 213-220.

Chung, K. K. H., & Chang, C. M. B. (2011). Executive functioning skills uniquely predict Chinese word reading. *Journal of Educational Psychology, 103(4),* 909-921.

Chung, K. K. H., Ho, C. S. H., Chan, D. W., Tsang, S. M., & Lee, S. H. (2011). Cognitive skills and literacy performance of Chinese adolescents with and without dyslexia. *Reading and Writing, 24(7),* 835-859.

Chialant, D., & Caramazza, A. (1998). Perceptual and lexical factors in a case of letter-by-letter reading. *Cognitive Neuropsychology, 15 (1/2),* 167-201.

Chincotta, D, & Underwood, G. (1997). Digit span and articulatory suppression: A cross-linguistic comparison. *European Journal of Cognitive Psychology, 9(1),* 89-96.

Christensen, S. C., & Wright, H. H. (2010). Verbal and non-verbal working memory in aphasia:

what three n-back tasks reveal. *Aphasiology, 24 (6-8): 752-762.*

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior, 15,* 17-31.

Coltheart, V. (1993). Effects of phonological similarity and concurrent irrelevant articulation on short-term-memory recall of repeated and novel word lists. *Memory & Cognition, 21(4),* 539-545.

Coltheart, M., Curtis, B., Atkins, P., & Haller, M. (1993). Models of reading aloud: dual-route and parallel-distributed processing approaches. *Psychological Review, 100,* 589-608.

Crosson, B. (2000). Systems that support language processes: verbal working memory. In S. E. Nadeau, L. J. G. Rothi, & B. Crosson (Eds.), *Aphasia and language: Theory to practice* (pp.399-418)*,* New York: Guilford Press.

Crosson, B., Rao, S. M., Woodley, S. J., Rosen, A. C., Bobholz, J. A., Mayer, A., … Stein, E. A. (1999). Mapping of semantic, phonological, and orthographic verbal working memory in normal adults with Functional Magnetic Resonance Imaging. *Neuropsychology, 13(2), 171-187.*

Davidson, I. (2011). Phonetic, phonemic, and phonological factors in cross-language discrimination of phonotactic contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 37(1),* 270-282.

Degé, F., Kubicek, C., & Schwarzer, G. (2011). Music lessons and intelligence: A relation mediated by executive functions. *Music Perception, 29(2),* 195-201.

Deutsch, D. (1999). *The psychology of music* (2nd Ed.). San Diego, CA, US: Academic Press.

Ericsson, K. A., & Walter, K. (1995). Long-term working memory. *Psychological Review, 102(2),* 211-245.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). Mini-mental State Examination: A practical guide for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189-198.

Francis, W. N., & Kučera, H. (1956). Frequency analysis of English usage: Lexicon and Grammar information mediate access to meaning? *Journal of Memory and Language, 37,* 41-57.

Franklin, M. (2008). The effects of musical training on verbal memory. *Psychology of Music, 36(3),* 353-365

Gathercole, S.E., & Baddeley, A. D (1993). *Working memory and language.* Hove, UK: Lawrence Erlbaum Associates Ltd.

Greenwald, M., & Berndt, R. S. (1999). Impaired encoding of abstract letter order: severe alexia in a mildly aphasic patient. *Cognitive Neuropsychology, 16 (6),* 513-556.

Greenwald, M. (2004). “Blocking” Lexical competitors in severe global agraphia: A treatment of reading and spelling. *Neurocase, 10(2),* 156-174.

Goldrick, M. & Rapp, B. (2007). Lexical and post-lexical phonological representations in spoken production. *Cognition, 102,* 219-260.

Gottardo, A., Chiappe, P., Yan, B., Siegel, L., & Gu, Y. (2006). Relationships between first and second language phonological processing skills and reading in Chinese-English speakers living in English-speaking contexts. *Educational Psychology, 26(3),* 367-393.

Grainger, J., Muneaux, M., Farioli, F., & Ziegler, J. C. (2005). Effects of phonological and orthographic neighborhood density interact in visual word recognition. *The Quarterly Journal of Experimental Psychology, 58A (6),* 981-998.

Gudnubdsdittur, H. R. (2010). Pitch error analysis of young piano students’ music reading performance. *International Journal of Music Education, 28(1),* 61-70.

Gupta, P., & MacWhinney, B. (1995). Is the articulatory loop articulatory or auditory? Reexamining the effects of concurrent articulation on immediate serial recall. *Journal of Memory and Language, 34,* 63-88.

Hallé, P. A., Chang, Y. C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Mandarin Chinese vs. French listeners. *Journal of Phonetics, 32,* 395-421.

Hargreaves, D. J., & Aksentijevic, A. (2011). Music, IQ, and executive function. *British Journal of Psychology, 102(3),* 306-308.

Hedden, T., & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20(5),* 511-528.

Hickok, G., Buschsbaum, B., Humphrise, C., & Muftuler, T. (2003). Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *Journal of Cognitive Neuroscience, 15 (5),* 673-682.

Hoosain, R. (1984). Experiments on digit spans in the Chinese and English languages. In H.S.R. Kao & R. Hossain (Eds.), *Psychological studies of the Chinese language,* pp. 23-38. Hong Kong: The Chinese Language Society of Hong Kong.

Jacobs, A. M., & Grainger, J. (1992). Testing a semistochastic variant of the interactive activation model in different word recognition experiments. *Journal of Experimental Psychology: Human Perception and Performance, 18(4),* 1174-1188.

Jing, Y., & Lu, H. (2009). The effects of self-explaining and working memory in Chinese reading comprehension. *Psychological Science, 32(4),* 1009-1011.

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: implications for phonological coding in working memory. *Journal of Experimental*

Psychology: Learning, Memory, and Cognition, 19, 369-381.

Jones, D. M., Hughes, R. W., & Macken, W. J. (2007). Commentary on Baddeley and Larsen (2007) The phonological store abandoned. *The Quarterly Journal of Experimental Psychology, 60 (4),* 505-511.

Kay, J., Lesser, R., & Coltheart, M. (1997). *Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction*. East Sussex: Psychology Press.

Kielar, A., & Joanisse, M. F. (2009). Graded effects of regularity in language revealed by N400 indices of morphological priming. *Journal of Cognitive Neuroscience, 22(7),* 1373-1398.

Kim, J. J., Kim, M. S., Lee, J. S., Lee, D. S., Lee, M. C., & Kwon, J. S. (2002). Dissociation of WM processing associated with native and second language: PET investigation. *NeuroImage, 15,* 879-891.

Koelsch, S., Schulze, K., Sammler, D., Fritz, T., Müller, K., & Gruber, O. (2009). Functional Architecture of verbal and tonal working memory: An fMRI study. *Human Brain Mapping, 30,* 859-873.

Krumhansl, C. K., & Toiviainen, P. (2003). Tonal Cognition. In I. Peretz & R. Zatorre (Eds.), *The cognitive neuroscience of music* (pp.95-108). New York: Oxford University Press.

Lee, C. Y., Tao, L., & Bond, Z. S. (2008). Identification of acoustically modified Mandarin tones by native listeners. *Journal of Phonetics, 36,* 537-563.

Lee, Y. S., Lu, M. J., & Ko, H. P. (2007). Effects of skill training on working memory capacity. *Learning and Instruction, 17:* 336-344.

Lehman, A. C. (2007). Book Review of Component skills involved in sight reading music. *Psychomusicology: A Journal of Research in Music Cognition, 19 (2),* 91-94.

Leong, C. K., Tse, S. K., Loh, K. Y., & Hau, K. T. (2008). Text comprehension in Chinese children: relative contribution of verbal working memory, pseudoword reading, rapid automatized naming, and onset-rime phonological segmentation. *Journal of Educational Psychology, 100(1),* 135-149.

Leong, C. K. (2002). Segmental analysis and reading in Chinese. In H. S. R. Kao, C. K. Leong, D. G. Gao (Eds.), *Cognitive neuroscience studies of the Chinese language* (Vols. 8 pp. 227-246). Hong Kong: Hong Kong University Press.

Leong, C. K., Cheng, P. W., & Tan, L. H. (2005). The role of sensitivity to rhymes, phonemes and tones in reading English and Chinese pseudowords. *Reading and Writing, 18,* 1-26.

Leshowitz, B., & Hanzi, R. (1974). Serial position effects for tonal stimuli. *Memory & cognition, 2 (1-a),* 112-116.

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22,* 1-75.

Lieberman, M. D., & Rosenthal, R. (2001). Why introverts can't always tell who likes them: multitasking and nonverbal decoding. *Journal of Personality and Social Psychology, 80 (2), 294-310.*

Li, W. S., & Ho, C. S. H. (2011). Lexical tone awareness among Chinese children with developmental dyslexia. *Journal of Child Language, 38,* 793-808.

Lin, C. H., Wu, C. H., Ting, P. Y., & Wang, H. M. (1996). Frameworks for recognition of Mandarin syllables with tones using sub-syllabic units. *Speech Communication 18, 175-190.*

Lobley, K. J., Baddeley, A. D., & Gathercole, S. E. (2005). Phonological similarity effects in verbal complex span. *The Quarterly Journal of Experimental Psychology, 58A (8),* 1462-1478.

Logie, R. H., Della Sala, S., Wynn, V., & Baddeley, A. D. (2000). Visual similarity effects in immediate verbal serial recall. *The Quarterly Journal of Experimental Psychology, 53 A (3),* 626-646.

Lovatt, P., Avons, S. E., & Masterson, J. (2000). The word-length effect and disyllabic words. *The Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 53 (A),* 1-22.

Lu, C. I., Greenwald, M., Bowyer, S. (2010, June). *Brain activation during letter sounding and silent transposing of musical notation: A MEG study.* Poster session presented at the 16th Annual Meeting of the Organization for Human Brain Mapping, Barcelona, Spain.

Lu, Z., & Zhang, Y. (2007). The effects of phonological loop of the working memory in Chinese reading comprehension. *Acta Psychologica Sinica, 39 (5),* 768-776.

Malins, J. G., & Joanisse, M. F. (2010). The roles of tonal and segmental information in Mandarin spoken word recognition: An eyetracking study. *Journal of Memory and Language, 62,* 409-420.

Marshall, J., & Newcombe, F. (1973). Patterns of paralexia: a psycholinguistic approach. *Journal of Psycholinguistic Research, 2,* 175-199.

Medin, D. L., Ross, B. H., & Markman, A. B. (2004). *Cognitive Psychology* (4th ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Miller, P., & Kupfermann, A. (2009). The role of visual and phonological representations in the processing of written words by readers with diagnosed dyslexia: evidence from a working memory task. *Annual of Dyslexia, 59,* 12-33.

Miozzo, M., & Caramazza, A. (1998). Varieties of pure dyslexia: the case of failure to access graphemic representation. *Cognitive Neuropsychology, 15 (1/2),* 203-238.

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: Evidence from a learning paradigm. *Cognitive Science, 35,* 184-197.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., & Howerter, A. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: A latent variable analysis. *Cognitive Psychology, 41,* 49-100.

Moreno, S., Bialystok, E., Barac, R., Schellenberg, E. G., Cepeda, N. J., & Chau, T. (2011). Short-term music training enhances verbal intelligence and executive function. *Psychological science* (0956-7976), 22 (11), 1425 -1433.

Mueller, S. T., Seymour, T. L., Kieras, D. E., & Meyer, D. E. (2003). Theoretical Implications of Articulatory Duration, Phonological Similarity, and Phonological Complexity in verbal working memory. *Journal of Experimental Psychology, 29 (6),* 1353-1380.

Neath, I., Bireta, T. J., & Suprenant, A. M. (2003). The time-based word length effect and stimulus set specificity. *Psychonomic Bulletin & Review, 10(2),* 430-434.

Oberauer, K., & Lange, E. B. (2009). Activation and binding in verbal working memory: A dual-process model for the recognition of nonwords. *Cognitive Psychology, 58,* 102-136.

Ockelford, A. (2007). A music module in working memory? Evidence from the performance of a prodigious musical savant. *Musicae Scientiae, Special Issue,* 5-36.

Oldfiled, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9 (1),* 97-113.

Patel, A. D. (2008). Sound elements: pitch and timbre. In A. D. Patel (Ed.), *Music, language, and the brain* (9-94). NY: Oxford University Press.

Patterson, K. E., & Hodges, J.R. (1992). Deterioration of word meaning: Implications for reading. *Neuropsychologia, 30(12),* 1025-1040.

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

Plaut, D.C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired reading: computational principles in quasi-regular domains. *Psychological Review, 103,* 56-115.

Pollatsek, A., Tan, L. H., & Rayner, K. (2000). The role of phonological codes in integrating information across saccadic eye movements in Chinese character identification. *Journal of Experimental Psychology: Human Perception and Performance, 26(2),* 607-633.

Rapp, B., & Goldrick, M. (2000). Discreteness and interactivity in spoken word production. *Psychological Review, 107 (3),* 460-499.

Raymer, A. M., & Rothi, L. J. G. (2001). Cognitive approaches to impairments of word comprehension and production. In R. Chapey (4th Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (pp. 524-550). New York, NY: Lippincott Williams & Wilkins.

Reinarz, L. E. (1997). *Effects of orthographic/phonological conflict, phonological coding skills, and response modality on repetition blindness.* Hempstead, NY: Hofstra University.

Repovs, G., & Baddeley, A. D. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience, 139 (1),* 5-21.

Rey, A., Dufau, S., Massol, S., & Grainger, J. (2009). Testing computational models of letter perception with item-level event-related potentials. *Cognitive Neuropsychology, 26(1),* 7-22.

Rosazza, C., Appollonio, I., Isella, V., & Shallice, T. (2007). Qualitatively different forms of pure dyslexia. *Cognitive Neuropsychology, 24(4),* 393-418.

Rueckl, J. G., Mikolinskil, M., Raveh, M., Miner, C. S., & Mars, F. (1997). Morphological priming, fragment completion, and connectionist networks. *Journal of Memory and Language, 36,* 382-405.

Rudner, M., & Rönnberg, J. (2008). The role of the episodic buffer in working memory for language processing. *Cognitive Process, 9,* 19-28.

Saito, S., & Ishio, A. (1998). Rhythmic information in working memory: effects of concurrent articulation on reproduction of rhythms. *Japanese Psychological Research, 40 (1),* 10-18.

Schellenberg, E. G. (2011). Examining the association between music lessons and intelligence. *British Journal of Psychology, 102,* 283-302.

Schendel, Z. A. (2006). The irrelevant sound effect: similarity of content or similarity of process? *Dissertation Abstracts International Section B: The Science and Engineering, 67(6-b),* 3477.

Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *NeuroReport, 13(17),* 2285-2289.

Schulze, K., Zysset, S., Mueller, K., Friederici, A. D., & Koelsch, S. (2011). Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. *Human Brain Mapping, 32,* 771-783.

Schwartz, A. I., Kroll, J. F., & Diaz, M. (2006). Reading words in Spanish and English: mapping orthography to phonology in two languages. *Language and Cognitive Processes, 22(1),* 106-129.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). Fractionating the reading process in dementia: evidence for word-specific print-to-sound associations. In Coltheart M., Patterson, KE, Marshall, JC (Eds.), *Deep Dyslexia* (pp. 259-269). London: Routledge and Kegan Paul.

Shriberg, L. D., & Kent, R. D. (2003). *Clinical Phonetics* (3rd ed.). Boston, MA: Pearson Education, Inc.

Shu, H., Meng, X., Chen, X., Luan, H., & Cao, F. (2005). The subtypes of developmental dyslexia in Chinese: evidence from three cases. *Dyslexia (11),* 311-329.

Silverman, M. J. (2007). The effect of paired pitch, rhythm, and speech on working memory as measured by sequential digit recall. *Journal of Music Therapy, 44(4),* 415-427.

Simmons, M. R. (2000). *The central executive and working memory: A dual-task investigation of the n-back task*. Gainesville, Florida: University of Florida.

Siok, W. T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Developmental Psychology, 37 (6),* 886-899.

So, D., & Siegel, L. S. (1997). Learning to read Chinese: Semantic, syntactic, phonological and working memory skills in normally achieving and poor Chinese readers. *Reading and Writing: An Interdisciplinary Journal, 9,* 1-21.

Song, X., & Cheng, L. (2006). Language Learner Strategy Use and Test Performance of Chinese Learners of English. *Language Assessment Quarterly. 3(3),* 243-266.

Soto, D., & Humphreys, G. W. (2007). Automatic Guidance of Visual Attention from Verbal Working Memory. *Journal of Experimental Psychology: Human, Perception and Performance, 33(3),* 730-757.

Sparks, R. W. (2001). Melodic Intonation Therapy. In R. Chapey (Eds.), *Language intervention strategies in Aphasia and related neurogenic communication disorders* (pp.703-717), New York: Lippincott Williams & Wilkins.

Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: the role of phonological information. *Cognition, 76,* B1- B11.

Stewart, L., Vincent, W., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66(2),* 183-195.

Stigler, J. W. (1984). "Mental abacus": The effect of abacus training on Chinese children's mental calculation. *Cognitive Psychology, 16,* 145-176.

Surprenant, A. M., Pitt, M. A., & Crowder, R. G. (1993). Auditory recency in immediate memory. *Quarterly Journal of Experimental Psychology, 46A,* 193-223.

Sweet, L. H., Paskavitz, J. F., Haley, A. P., Gunstad, J. J., Mulligan, R. C., Nyalakanti, P. K., & Cohen, R. A. (2008). Imaging phonological similarity effects on verbal working memory. *Neuropsychologia, 46,* 1114-1123.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Pearson/Allyn & Bacon.

Taft, M., & Chen, H. C. (1992). Judging homophone in Chinese: the influence of tones. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese. Advances in psychology* (Vol. 7, pp. 151-172). Oxford, England: North-Holland.

Tan, L. H., Hoosain, R., & Siok, W. W. T. (1996). Activation of phonological codes before access to character meaning in written Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22(4),* 865-882.

Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language, 37,* 41-57.

Tong, Y., Francis, A. L., & Gandour, J. T. (2007). Processing dependencies between segmental and suprasegmental features. *Language and Cognitive Processes, 23(5),* 689-708.

Tree, J. J., Longmore, C., & Besner, D. (2011). Orthography, phonology, short-term memory and the effects of concurrent articulation on rhyme and homophone judgments. *Acta*

Psychologica, 136, 11-19.

Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human perception and performance, 35(2),* 480-498.

Vallar, G. (2006). Memory systems: the case of phonological short-term memory. A festschrift for cognitive neuropsychology. *Cognitive Neuropsychology, 23,* 135-155.

Wang, M. (2001). The development of spelling and its relationship to decoding and phonological processing in Chinese ESL children. *Dissertation Abstracts International Section A: Humanities and Social Science, 61(11-A),* 4289.

Wang, Y., Jongman, A., & Sereno, J. A. (2001). Dichotic perception of Mandarin tones by Chinese and American listeners. *Brain and Language, 78,* 332-348.

Wang, M., Anderson, A., Cheng, C., Park, Y., & Thomson, J. (2008). General auditory processing, Chinese tones processing, English phonemic processing and English reading skill: a comparison between Chinese-English and Korean-English bilingual children. *Read and Writing, 21(6),* 627-644.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third edition.* San Antonio, TX: The Psychological Corporation.

Weekes, B. S., Chen, M. J., & Gang, Y. W. (1997). Anomia without dyslexia in Chinese. *Neurocase, 3,* 51-60.

Williamson, V. J., Baddeley, A. D., & Hitch, G. J. (2010). Musician's and nonmusicians' short-term memory for verbal and musical sequences: comparing phonological similarity and pitch proximity. *Memory & Cognition, 38 (2),* 163-175.

Xu, Y., Pollatsek, A., & Potter, M. C. (1999). The activation of phonology during silent Chinese word reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (4),* 838-857.

Xu, Z., & Li, B. Q. (2009). The mechanism of reverse word length effect of Chinese in working memory. *Acta Psychologica Sinica, 41 (9),* 802-811.

Yeung, P. S. (2007). Learning to read and spell in English among Chinese English-as-a-second-language learners in Hong Kong. *Dissertation Abstracts International: Section B: The Sciences and Engineering, 67(8-B),* 4742.

Zhou, X., & Marslen-Wilson, W. (1999a). Phonology, orthography, and semantic activation in reading Chinese. *Journal of Memory and Language, 41,* 579-606.

Zhou, X., & Marslen-Wilson, W. (1999b). The nature of sublexical processing in reading Chinese Characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25 (4),* 819-837.

ABSTRACT

THE EFFECT OF MUSICAL TRAINING ON VERBAL AND TONAL WORKING MEMORY

by

CHING-I LU

August 2012

Advisor: Dr. Margaret Greenwald

Major: Communication Sciences and Disorders

Degree: Doctor of Philosophy

This dissertation explored the translation from print to sound of the tonal language Mandarin versus English versus musical notation in healthy volunteers. The performance of musicians and non-musicians was compared across a variety of reading tasks in an attempt to examine whether musical training can facilitate Mandarin tone or phonological processing. The effects of increasing working memory load on reading performance across tasks were also examined. Results showed that increasing demands on working memory in visual recognition tasks significantly decreased performance accuracy for both musicians and non-musicians across tasks. Significant differences in accuracy rates were observed between musicians and non-musicians. Although the tonal information in Mandarin is embedded in phonological information, the current study provided evidence that musicians are better able to extract this tonal information from print than are non-musicians, or to maintain it in working memory. Even in the Mandarin homophone task, which requires phonological judgments of print, the musicians demonstrated superior performance. The current study provides evidence that musical training facilitates phonological language processing.

AUTOBIOGRAPHICAL STATEMENT

Ching-I Lu received her Bachelor of Arts degree in Music from the National Taiwan Normal University in 1992. She received her Master of Arts degree in Music from the National Taiwan Normal University in June 1996, and she received her Master of Science degree in Psychology from the National Taiwan University in June 2004. She has worked as a senior high school music teacher since 1992 in Taipei, Taiwan. Her research interests include studies in reading, dyslexia, music, and neuroimaging, including fMRI, TMS, and MEG. Upon completion of the doctorate, Ms. Lu plans to pursue a career in research and teaching at the university level.