Shared Orthography: Do shared written symbols influence the perception of native-nonnative sound contrasts?

by

Carolyn Pytlyk
B.A., University of Saskatchewan, 1996

A Thesis Submitted in Partial Fulfillment of the Requirements of the Degree of

MASTER OF ARTS

in the Department of Linguistics

SUPERVISORY COMMITTEE

Shared Orthography:
Do shared written symbols influence the perception of native-nonnative sound contrasts?

by

Carolyn Pytlyk
B.A., University of Saskatchewan, 1996

Supervisory Committee

Dr Hua Lin (Department of Linguistics)
Supervisor

Dr John Esling (Department of Linguistics)
Departmental Member

Dr Sonya Bird (Department of Linguistics)
Additional Member

Dr Richard King (Department of Pacific and Asian Studies)
Outside Member

ABSTRACT

Supervisory Committee

Dr Hua Lin (Department of Linguistics)
Supervisor
Dr John Esling (Department of Linguistics)
Departmental Member
Dr Sonya Bird (Department of Linguistics)
Additional Member
Dr Richard King (Department of Pacific and Asian Studies)
Outside Member

The research presented in this thesis investigated whether second language learners who learn via a familiar orthography (i.e. Pinyin) differed from second language learners who learned via a non-familiar orthography (i.e. Zhuyin) in their perception of English-Mandarin sound contrasts. The assumption was that significant differences in perceptual performances between experimental groups could be attributed to the orthography's influence in the categorization of Mandarin sounds. Also investigated in this research was the degree of confusability of particular Mandarin sounds based on their relative similarities to their English counterparts.

The data were obtained from thirty-two native Canadian English speaking participants through a series of experimental tasks – pre-test > training phase > post-test. The pre- and post-tests assessed the participants' sensitivities to English-Mandarin contrasts in an oddity discrimination task. Between the two tests, the participants underwent a short training phase where they learnt Mandarin via the orthographic medium assigned to their group. Perceptual performance of the participants was measured in terms of error rates (ER), a-prime (A') and response times (RT).

The hypotheses concerning orthographic influence were not supported by the results. Three-way repeated measures ANOVAs revealed that there were no significant differences in the perceptual performances of the three groups for any of the three measures of the dependent variable. The lack of significant differences is discussed in terms of: 1) the strength of the established L1 orthographic system, 2) the cognitive load placed on the participants, and 3) the insufficient time given for the development of new symbol-sound associations within the new L2 orthographic system. The hypotheses concerning the degree of confusability were confirmed. The statistical analyses revealed three groups of perceptual sensitivity; 1) a highly sensitive group, 2) a slightly sensitive group, and 3) an insensitive group. These results are discussed in terms of acoustic saliency and models of speech processing.

This is the first systematic study to investigate the potential influence of the L1 orthographic code on second language speech perception. Two major conclusions were drawn from the results. First, Mandarin instruction via Pinyin appears to have slight advantage over instruction via Zhuyin as the conflict between the two orthographic systems appears to neutralize any potential benefits. Second, participants exhibit varying degrees of perceptual sensitivity to L1-L2 sound contrasts due to the type of differences between the native and nonnative sounds.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

The completion of this project would not have been possible if not for the support and guidance of many wonderful individuals. First, I would like to thank my supervisory committee. From the conception of the research idea through to the final draft, Dr Richard King and Dr John Esling have been very helpful with their insightful feedback and suggestions. A special thank you to Dr Sonya Bird for her phonetics expertise, unlimited patience and without whom I could not have made it through the stimuli creation process. I am especially grateful to have had the honour to work with Dr Hua Lin, my supervisor. Her support, guidance and enthusiasm for my research have been an invaluable source of motivation and inspiration. Also, thank you to Dr Daniel Bryant for kindly agreeing to be my external examiner.

I wish to extend my sincere gratitude to all the people involved in the creation, collection and analysis of the data. Thank you Abbey Bell, Laura Hawkes, Allison Benner, Wang Qian, Tian Jun, and Zhu Yinxin for being readily available to record and judge the experimental stimuli. Thank you to the native English speakers who gave up their valuable time to participate in this project. Thank you to Huang Shumin for her energy in and dedication to teaching the Mandarin classes. Also, thank you to Dr Natasha Warner from the University of Arizona who generously agreed to help me navigate the murky waters of statistics.

The Department of Linguistics milieu has been an excellent place for the development of my academic career. I cannot have imagined a kinder, more supportive group of faculty, staff and graduate students In particular, I would like to thank my friend and colleague, Janet Leonard, who was always there to listen and save me from myself.

I would also like to acknowledge the institutions that provided me with financial support over the course of my M.A. program. Thank you to the Department of Linguistics for employing me as a Teaching Assistant (2005-2007). Thank you to the University of Victoria for an M.A. Fellowship (2005-2006) and an internal research grant awarded to Dr Hua Lin. Thank you to Henry and Michiko Warkentyne for a Graduate Scholarship (2006).

Finally, I am eternally grateful for my extraordinary friends and family for their unfailing patience, encouragement and love.

DEDICATION

for Siobhan and Magdalena

Chapter One

INTRODUCTION

"Interlingual identification often amounts to fitting square pegs into round holes." (Flege, 1991, p.702)

1.1 Background

Current research leaves little doubt that first languages (L1) play a pivotal role in the acquisition of second languages (L2). In the words of Major (2001), "our NL [native language] has an inescapable influence on our L2" (p. 36). Indeed, it was often thought that L1 influence was one of the major barriers that prevented language learners (especially adults) from achieving native-like mastery of the L2. What has yet to be determined is specifically how, why and in what way the L1 exerts influence over subsequent language learning. Researchers like Kuhl (1993, 2000), Flege (1987, 1995) and Best (1995, 2001) argue that L2 sounds are perceived in relation to the L1 phonetic categories that already exist within the learners' phonetic systems. In the area of psychophysiology, research has demonstrated that a visual symbol strongly leads listeners to anticipate a sound that has already been associated to that particular symbol (Hallé et al., 2000; Widmann et al., 2004). Also, psychological research on orthography has shown that spelling-to-sound and sound-to-spelling inconsistencies increase response latencies and error rates in word recognition for native speakers (e.g., Glushko, 1979; Ziegler & Ferrand, 1998; Stone et al., 1997; Ziegler et al., 2004; etc.). From these areas of research, it seems natural to hypothesize that shared symbols used to represent L1 *and* L2 sound contrasts would have a profound effect on how second language learners would perceive and, ultimately, classify the L2 sounds. If shared orthography does prove to interfere with perception of L1 and L2 sound contrasts, it may be necessary to revise teaching methodologies to incorporate this knowledge and inform language teachers.

When learning Mandarin Chinese, foreigners are traditionally introduced to the Mandarin sound system via an alphabetic system of representation. The most widely used system is Pinyin. One benefit to using such a system is that learning Mandarin becomes more accessible to those foreign learners who are accustomed to Latin alphabet. While representing and/or teaching Mandarin via Pinyin has numerous benefits, this study is interested in exploring if it also has any drawbacks. Particularly, does using an L2 orthographic medium that is strongly linked to that of the L1 affect the learners' abilities to perceive closely related sound contrasts between the L1 and the L2? In short, this study will examine whether the orthographic medium can be considered a contributing factor in the perception (or lack thereof) of sound contrasts between English and Mandarin. Therefore, for the purposes of this study, the term *orthographic medium* refers to the writing system employed to represent the sounds of a language and *perceptual sensitivity* refers to learners' (listeners') abilities to distinguish between English-Mandarin sound pairs.

1.2 Research statement

This study examines whether there is a relationship between learners' abilities to distinguish native-nonnative sound contrasts and the orthographic medium with which the contrasts are presented to learners; specifically as it relates to native English speakers learning Mandarin. This study also examines whether learners exhibit varying degrees of sensitivity to the particular contrasts based on the relative similarities between the two sounds in any given contrast pair.

1.3 Research questions

For the purposes of this study, several research questions have been articulated. They have been organized into two groups of questions. The first group poses the research questions regarding orthographic influence and the second group poses the research questions regarding varying degrees of perceptual sensitivity.

1.3.1 Orthographic influence

The primary research question is as follows:

1. Does the type of orthography used for L2 instruction affect the second language learners' perception of L1 and L2 sound contrasts?

More specifically,

2. Are learners able to distinguish between L1 and L2 sound contrasts that are represented by the same orthographic symbol (grapheme) in the L1 and the L2?
3. Are learners able to distinguish between L1 and L2 sound contrasts that are represented by separate orthographic symbols in the L1 and the L2?
4. Are learners more successful at distinguishing L1 and L2 sound contrasts when the L2 orthographic symbols do not also represent L1 sounds?

1.3.2 Degrees of perceptual sensitivity

1. Are learners more successful at distinguishing some L1-L2 contrasts than others?
2. Do the target contrasts differ in terms of perceptual saliency?
3. How are the L2 sounds mapped in relation to their English counterparts?

1.4 Research hypotheses

Again, this section is divided into two parts; hypotheses that mirror the research questions posed in Section 1.3 have been formulated for 'orthographic influence' and 'degrees of perceptual sensitivity'.

1.4.1 Orthographic influence

1. Shared orthographic symbols that represent separate sounds in the L1 and L2 will **negatively impact** learners' abilities to distinguish between the L1 and L2 sounds. Therefore, learners who learn Mandarin via Pinyin, which shares orthographic symbols with English, would be led to equate L2 phonemes with their English counterparts.
2. Non-shared orthographic symbols that represent separate sounds in the L1 and L2 will **positively impact** learners' abilities to distinguish between the L1 and L2 sounds. Therefore, learners who learn Mandarin via Zhuyin, an entirely separate orthographic system, would be better able to make distinctions between similar sounding L2 phonemes and their English counterparts.
3. Absence of orthographic symbols will not affect learners' abilities to distinguish between L1 and L2 sounds. Therefore, learners who learn Mandarin via no orthography would fall between the performances of the Pinyin and Zhuyin groups.

1.4.2 Degrees of perceptual sensitivity

1. Learners will exhibit different levels of discrimination performance based on the perceptual saliency of the acoustic cues in the speech signal.
2. Discrimination will be poor for those contrasts where learners perceive the L2 sound to belong to the same phonetic category as the L1 contrast counterpart.

3. Discrimination will be good for those contrasts where learners perceive the L2 sound to belong to a different phonetic category from the L1 contrast counterpart.

1.5 Thesis outline

As outlined above, the purpose of this research is to discover whether shared orthography affects learners' perception of L2 sounds and whether learners demonstrate different levels of perceptual sensitivity to the target contrasts. This thesis is presented in six chapters. Chapter One provides an introduction by outlining the research questions and hypotheses on which the research was founded. Chapter Two reviews the pertinent literature on English and Mandarin consonant inventories, romanization of Mandarin, theories of first language influence, models of speech processing, and orthographic influence. Chapter Three provides a comprehensive description of the methodology employed in the research. The results of the data collected are provided in Chapter Four. The thesis continues with an in-depth discussion of the results in Chapter Five by returning to the research questions and hypotheses. Finally, Chapter Six concludes the research project by summarizing the results, discussing some of the limitations, highlighting its contributions and offering suggestions for future research.

Chapter Two

LITERATURE REVIEW

"The effect that L1 transfer has in acquisition of L2 phonology is one that imposes a mismatch between the L2 input and what is actually perceived by, and therefore, accessible to the learner as L2 intake." (Matthews and Brown, 2004, p. 6)

The primary goal of the research is to investigate the potential existence of L1 influence on native-nonnative sound contrasts due to shared orthographic representation, specifically with regards to native English speakers learning Mandarin. Areas of previous research that shed light on the issue at hand include: (1) a comparison of the English and Mandarin consonant inventories, (2) the romanization of Mandarin, (3) theories of L1 influence, (4) models of speech processing, and (5) orthographic influence on word and sound recognition. All areas are reviewed in this chapter. Finally, this chapter concludes with the last section identifying the significance of the problem under investigation.

2.1 English and Mandarin consonants

Prior to any discussion of shared orthographic interference, it is first important to become familiar with the sound systems of English and Mandarin. Figure 2-1 is a slightly revised version of a handout from LING 378 (2005) taught by Dr Hua Lin at the University of Victoria (see also Lin, 2001; Duanmu, 2000 and Rogers 2000) that has been created to compare the consonant inventories of English and Mandarin. The highlighted columns offset the Mandarin system from the English system.

			LABIAL		DENTAL		ALVEOLAR		ALVEO-PALATAL		PALATAL		VELAR		GLOTTAL	
			E	M	E	M	E	M	E	M	E	M	E	M	E	M
STOPS	-vc	+asp	pʰ	pʰ			tʰ	tʰ					kʰ	kʰ	(ʔ)*	
		-asp	p	p			t	t					k	k		
	+vc		b				d						g			
AFFRI-CATES	-vc	+asp				tsʰ			tʃ	tʂʰ		(tɕʰ)				
		-asp				ts				tʂ		(tɕ)				
	+vc								dʒ							
FRICA-TIVES	-vc		f	f	θ	s̪	s		ʃ	ʂ		(ɕ)		x	h	
	+vc		v		ð		z		ʒ	ʐ						
NASALS			m	m			n	n					ŋ	ŋ		
LIQUIDS							l/ɫ	l								
							ɹ	ɻ								
GLIDES			w	(w)							j	(j)				

* The sounds in brackets () are not considered to be phonemic.

Note: The **}** notation (as suggested by Dr J. Esling) for the English voiceless stops indicates that the two consonants belong to the same phonemic category where the underlying phoneme is not clear.

FIGURE 2-1 English and Mandarin consonant inventories

The above table demonstrates that while Mandarin and English do share some sounds, there are also many sounds that differ between the two languages. If the goal is to employ the same alphabetic script, an obvious problem arises as to how to represent the sound systems of both English and Mandarin with a limited amount of letters. One possible solution is to have the different L1 and L2 sounds share the same orthographic grapheme(s) which is indeed the case with most romanization systems of Mandarin (see below for more details).

2.2 Romanization of Mandarin

Romanizing Chinese is not a new idea. The first alphabetical system was designed in 1605 by an Italian missionary named Matteo Ricci. After the Opium War (1840-2), according to Duanmu (2000), the "proponents of language reform believed that alphabetical writing was a key to the strength of a modern nation" (p. 6). As a result, in the decades between 1890 and 1930 approximately 30 different designs were proposed. After the Communist Revolution ended in 1949, the new government revived the interest in language and writing reform. According to Zhou (2001), the goal of the communist regime was linguistic and ethnic assimilation where, based on the advice of Soviet specialists, one standard language and one standard writing system would make the assimilation process smoother. With this mandate, the new government proposed the *Hanyu Pinyin* ('Chinese Spelling System') "to further the political objective of promoting a less elitist form of the educated Peking dialect into a national vernacular" (Killingly, 1998, p. 2). In 1958, almost a decade after it was first proposed, Pinyin (as the system became known) was adopted as the official spelling system by the Chinese government.

The original plan intended for Pinyin to eventually supplant the use of Chinese characters. Ironically, it was the government that sanctioned the adoption of Pinyin as the official romanization system, yet the high level intellectual officials oppose the promotion of the system and Pinyin has been forced to play a secondary role (DeFrancis, 2006). Still, Pinyin has garnered widespread national and international popularity and acceptance. For example, Pinyin has the advantage of the governmental endorsement as the official transcription system in China. Chinese names have been standardized on the basis of the Pinyin orthography. The educational system requires the instruction of Pinyin for all children in grade one. Its use has been extended to dictionaries, library catalogues, and other reference materials. Pinyin is used on street signs, building names, and product labels. In addition, entities which defer to the existence of the People's Republic of China (PRC), such as the mass media and the United Nations, have accepted Pinyin (DeFrancis, 1990). So although Pinyin has not replaced Chinese characters, its position as a romanized transcription system seems quite secure. Its supporters "have succeeded in consolidating Pinyin as THE [original emphasis] system of representing Chinese in an alphabetic script" (DeFrancis, 1990, p. 10) which makes Pinyin "the most important Chinese romanization system in modern usage" (Killingly, 1998, p. 5).

As for the system itself, the Pinyin designers borrowed graphemes that either had a basis in western orthographies or linguistic notation (DeFrancis, 1990) to represent the individual Mandarin phonemes. Table 2-1 below presents the Pinyin letter-sound relationship for the 24 Mandarin consonants (Lin, 2001).

labial	**b-** [p]	**p-** [p^h]	**m-** [m]	**f-** [f]
alveolar	**d-** [t]	**t-** [t^h]	**n-** [n]	**l-** [l]
velar	**g-** [k]	**k-** [k^h]	**-ng** [ŋ]	**h-** [x]
dental	**z-** [ts]	**c-** [ts^h]	**s-** [s̪]	
alveo-palatal	**zh-** [tʂ]	**ch-** [tʂ^h]	**sh-** [ʂ]	**r-** [ʐ]
palatal	**j-** [tɕ]	**q-** [tɕ^h]	**x-** [ɕ]	
glides	**y-** [j]	**w-** [w]		

Note: The IPA equivalents are provided in square brackets.

TABLE 2-1 Pinyin consonant representations

As is evident in Table 2-1, the 'romanization' of Mandarin or more specifically, Pinyin (although not the only motivation for its creation) makes the orthographic representation of Mandarin more similar to languages that also employ an alphabetic system of writing, such as English.

2.3 First language influence

If China's adoption of an alphabet makes English and Mandarin more similar orthographically, the question then becomes would orthographic similarity facilitate (or hinder) the acquisition of certain Mandarin elements such as the sound and spelling systems? Generally speaking, does L1-L2 orthographic similarity prevent or foster L1 influence on learning an L2? Contrastive Analysis Hypothesis (CAH) (Lado, 1957) was one of the most influential theories that attempted to explain the effect of L1 on learning an L2. CAH maintained that interference from the learner's first language was the major barrier to the acquisition of a second language. Learning (or non-learning) was contingent on the notion of transfer. Archibald (1998) defines transfer as "the process whereby a feature or rule from a learner's first language is carried over to the IL [interlanguage – the developing second language system] grammar" (p. 3). It is important to note that, according to CAH, there are two types of transfer. Positive transfer refers to when an L1 feature is carried over into the interlanguage and **facilitates** the learning of the L2. Negative transfer, on the other hand, refers to when an L1 feature is carried over into the interlanguage but **hinders** the learning of the L2. For proponents of CAH, all errors in the L2 could be explained by negative transfer from the L1.

CAH claimed that by systematically comparing the target language and the native language, it could predict where the learners would have difficulty and where they would not

(Lado in Major, 2001). It predicted that similar elements,[1] between the L1 and L2, would be easy for the learner to acquire while different elements would be difficult. Simply put, similarity equaled ease and dissimilarity equaled difficulty. While the predictive powers of CAH were initially very appealing, it soon became apparent that there were some serious shortcomings with the hypothesis. Most seriously, CAH predicted some errors that did not occur but failed to predict all the errors that were made by language learners. For example, one type of error that CAH could not account for was what would become known as 'developmental'. A second major criticism, as Brown (2000) notes, was that CAH also failed to account for why learners with different L1s would substitute different L1 sounds for the same L2 sound. For example, why would Japanese speakers substitute [s] for [θ] but Russian speakers substitute [t] for the same sound when [t] and [s] exist in both Japanese and Russian (Hancin-Bhatt, 1994).

In response to this heavy criticism, Oller and Ziahosseiny (1970) sought to remedy the shortcomings of CAH by proposing a *moderate version* of the hypothesis that considered degrees of similarity between L1 and L2 elements. In their study of English spelling errors, they discovered that those learners whose native language did not use the Roman alphabet made fewer mistakes than those whose native language did. Oller and Ziahosseiny claimed that learning similar 'sounds, sequences and meanings' would cause more difficulty for language learners than dissimilar ones because "whenever patterns are minimally distinct in form in one or more systems, confusion may result" (p. 186). Their speculations were a complete reversal of the original tenets of CAH.

In light of Oller and Ziahosseiny's study and the criticisms against the traditional CAH, researchers gradually shifted from thinking that dissimilar elements would be the most

---

[1] 'Elements' is a general term that refers to features, structures and/or rules inherent to a given language. These elements can be either phonetic, phonological, morphological, syntactic, semantic or pragmatic.

difficult to thinking that, perhaps, they would be the easiest to acquire. Dissimilar elements are those elements that have no corresponding structure in the L1. As a result, researchers posit that dissimilar elements are less likely influenced by transfer; and therefore, are more likely to be learned (Wode, 1983; Major, 2001). In contrast, where similar elements exist between the two systems, transfer is likely to dominate and learning is less likely to occur.

The question that remains is 'why similar elements should be more difficult to learn?'. Why would transfer hinder rather than facilitate, as predicted by CAH, learning a second language? Major (2001) suggests the reason has to do with perceptual saliency. That is, the greater the difference between structures, the easier the learner will be able to perceive that difference and learn it. Conversely, the more similar the elements the less likely the learner will actually perceive the minute differences. Major offers the following explanation with regards to phonological similarity:

> The psycholinguistic reason why similar sounds tend to be more difficult than dissimilar sounds seems to be that gross differences are more often noticed, due to perceptual saliency, whereas minimal differences are less likely to be noticed, resulting in non-learning. (p. 37)

Intuitively, this appears to be a reasonable assumption. No doubt, when differences are not perceptually salient, learners will perceive the L2 element as the same as the L1 element. However, it is important to point out that, unfortunately, the concept of what constitutes 'similarity' has yet to be satisfactorily defined and a methodology for measuring such a concept still eludes researchers.

2.4 Models of speech processing

The claim that similar elements are more difficult to acquire than dissimilar ones has sparked much research in the area of second language research. Specifically, research in the area of the acquisition of L2 phonology has attempted to characterize the relationship and interaction between the existing L1 phonological system and the L2 phonology to be acquired. Recent models suggest that native sound experience gives learners an 'organizing perceptual framework' with which to discriminate and classify nonnative sounds. (Best, 2001, p. 776). In short, many researchers have sought to demonstrate "how L2 sounds are mapped onto L1 sounds" (Brown, 2000, p. 8).

One such model of speech perception is Kuhl's (1993, 2000) *Native Language Magnet* (NLM). NLM holds that innate factors and linguistic experience influence speech perception. It addresses, 1) how native language categories are created, and 2) how L2 sounds interact with L1 sounds. According to NLM, "infants respond to speech in a linguistically relevant way" (Kuhl, 1993, p. 125). This predisposition manifests itself into a "general auditory processing mechanism" that allows infants to "perceptually partition a series of speech sounds at the places where the world's languages divide the series of stimuli into phonetic categories, rather than at arbitrary places" (Kuhl, 1993, p. 125). That is, infants are able to group sounds into gross universal categories based on the acoustic features of the sounds. However, NLM claims that from birth, as infants gain more experience and input from the ambient language (L1), they start to reconfigure the gross category boundaries and create mental maps of speech sounds that "warps the acoustic dimensions underlying speech, producing a complex network, or filter, through which language is perceived" (Kuhl, 2000, p. 11854). In short, NLM holds that speech perception is altered by linguistic experience.

During the reconfiguration process of phonetic boundaries, language specific prototypes are also established. It is these prototypes, NLM maintains, that act like 'perceptual magnets' that distort the phonetic space and reduce the perceptual distance between the prototype and the stimuli (Kuhl, 1993). Magnets attract nearby sounds to make them more similar to the category prototype. This has serious implications for second language perception. According to NLM, foreign sounds are more difficult to discriminate when they closely resemble a native magnet. Therefore, the closer the new sound is to an L1 magnet the more likely it will be assimilated to and indistinguishable from the L1 prototype.

Another model that tries to account for the role that the L1 plays in the perception of nonnative contrasts is Best's *Perceptual Assimilation Model* (PAM) (Best, 1995, 2001). Like NLM, PAM maintains that learners are heavily influenced by their knowledge of their established native categories. As the name suggests, PAM predicts that learners will assimilate L2 sounds to native phones "whenever possible based on detection of commonalities in the articulators, constriction locations and/or constriction degrees used" (Best, 2001). When it comes to assimilation of nonnative sounds, according to PAM, learners categorize nonnative sounds in one of three ways. A sound can be categorized either as: 1) part of a native category where it may range from an excellent to poor exemplar of the phoneme, 2) an uncategorizable speech sound where the phone falls somewhere between the native phonetic categories, or 3) an unassimilatable non-speech sound that bears no detectable similarity to language.

In addition to accounting for native-nonnative contrast perception, PAM also considers nonnative-nonnative contrast perception. PAM predicts that there are a number of pair-wise assimilation types. For example, it is possible for two nonnative sounds to be assimilated to two separate categories. In this case, PAM predicts that discrimination would be very good. In contrast, it is also possible for two nonnative sounds to be assimilated to the same category. Here, discrimination would be poor. Because of the many pair-wise types, Best (2001) argues for levels of discrimination where native phonology aids perception when contrasts are separated by a phonological boundary but hinders when both sounds are assimilated to the same category.

Unfortunately, both NLM and PAM are concerned strictly with the perception sound contrasts and neither model makes any provisions for the possible *production* of nonnative contrasts. Also, neither model addresses the question of whether it is possible for learners to create new phonemic categories for new or uncategorizable speech sounds and/or unassimilatable non-speech sounds. It is important to recognize that, while an integral part of language learning, *perception* of nonnative sounds is not the only factor to influence acquisition. An equally important factor is the *production* of the nonnative sounds. Flege's (1987, 1991, 1995, 2004) *Speech Learning Model* (SLM) makes the connection between perception and production. Like NLM and PAM, SLM predicts that similar sounds will be more challenging for learners due to the learners' tendencies to equate a similar nonnative sound with an already existing one. However, unlike the first two models, SLM goes further and makes predictions about learners' abilities to produce nonnative sounds.

SLM distinguishes two kinds of sounds: *new* and *similar*. New sounds are defined as those sounds that have no counterpart in the L1 (Flege, 1987). Where there is no L1 counterpart, dissimilarity exists. SLM predicts that "the greater the perceived dissimilarity between an L2 sound and the closest L1 sound, the more likely it is that phonetic differences between the sounds will be discerned" (Flege, 1995, p. 239). Since the learners are more likely to discern the differences with *new* sounds, SLM further predicts that learners are also more likely to perceive and produce new sounds authentically. Supporting this claim are the results from Bohn and Flege's (1992) study of German-speaking learners of English. Part of this study focused on German speakers' acquisition of the English vowel [æ] which has no counterpart in German. Bohn and Flege discovered that given enough time and exposure, German speakers were successful at perceiving and authentically producing the English phoneme [æ]. While the results of this study indicate that learners may need considerable exposure to a new L2 contrast, Bohn and Flege claim that it is possible for adult learners to establish new phonemic categories for sounds that do not exist in their native language.

Similar sounds, according to SLM, are a greater challenge for language learners. Flege (1987) maintains that similar sounds are those sounds that "differ systematically from an easily identifiable L1 counterpart" (p. 48). The differences between the L2 sounds and the closest L1 phones are much more subtle; and therefore, the differences are much more difficult to discern. SLM attributes this difficulty to *equivalence classification* where:

> Category formation for an L2 sound may be blocked by the mechanism of equivalence classification. When this happens, a single phonetic category will be used to process perceptually linked L1 and L2 sounds (diaphones). Eventually, the diaphones will resemble one another in production. (Flege, 1995, p. 239)

In short, equivalence classification of sounds hinders or prevents the learner's ability to create new phonetic categories for similar sounds. Unfortunately, in SLM, Flege makes no specific provisions for how sounds are categorized as either 'new' or 'similar'. For example, at what point are L1 sounds considered 'counterparts'?

The message from the above theories is clear; learners' acquisition of L2 phonology is greatly influenced by their L1 (although not necessarily the only influencing factor) and that, generally speaking, L2 sounds are perceived in relation to the existing L1 categories. With these two assumptions in mind, the remaining sections of this chapter will focus on the potential difficulties of second language learning when an orthographic influence is added to the equation. More specifically, it will focus on isolating some potential problems that may arise when teaching Mandarin via the romanized system Pinyin.

2.5 Orthographic influence on word and sound recognition

As demonstrated in the previous section, NLM, PAM and SLM all argue that existing native phonetic categories ultimately determine how L2 sounds are categorized. But it is unlikely that these are the only influences on how L2 sounds become categorized. Treiman and Cassar (1997) claimed that once children begin to read (in alphabetic systems at least) they are no longer able to completely separate sound and spelling. If true, then logically, orthography must exert some influence on listeners' perceptions of the sounds in their native languages. Orthographic research has focused on many aspects of influence by taking into account issues such as consistency and frequency. For instance, research into 'consistency effects' has shown that word pronunciation and recognition is affected by knowledge of similarly spelt words. This means that where there are spelling-to-sound or sound-to-spelling inconsistencies, processing times and error rates will be affected (Glushko, 1979; Jared et al., 1990; Stone et al., 1997; Lacruz and Folk, 2004; Ziegler et al., 2004).

2.5.1 Visual recognition consistency effects

The first researcher to consider consistency effects was Glushko (1979). In his study, he classified words as either 'regular' or 'exception' as well as consistent or inconsistent. Regular words are words whose pronunciation can be derived from a regular orthographic rule. This type of word can be further classified as either 1) 'regular consistent' where the rimes do not have any alternate pronunciations [ex. H**EAP**] or 2) 'regular inconsistent' where the rime has an alternate pronunciation(s) [ex. G**AVE**]. In other words, both regular consistent and regular inconsistent words have predictable pronunciations based on an orthographic rule. Exception words, in contrast, have irregular spelling-to-sound correspondence [ex. H**AVE**] and are always inconsistent. For example, both the words *gave* and *have* contain the rime –AVE. In the case of *gave*, the pronunciation is predicable based on the A—E orthographic rule, whereas the rime in *have* has an alternate irregular pronunciation. As a result, since the rime –AVE has more than one pronunciation, words such as *gave* and *wave* are considered regular inconsistent and words like *have* are considered exceptions. Figure 2-2 illustrates these spelling-to-sound relationships.

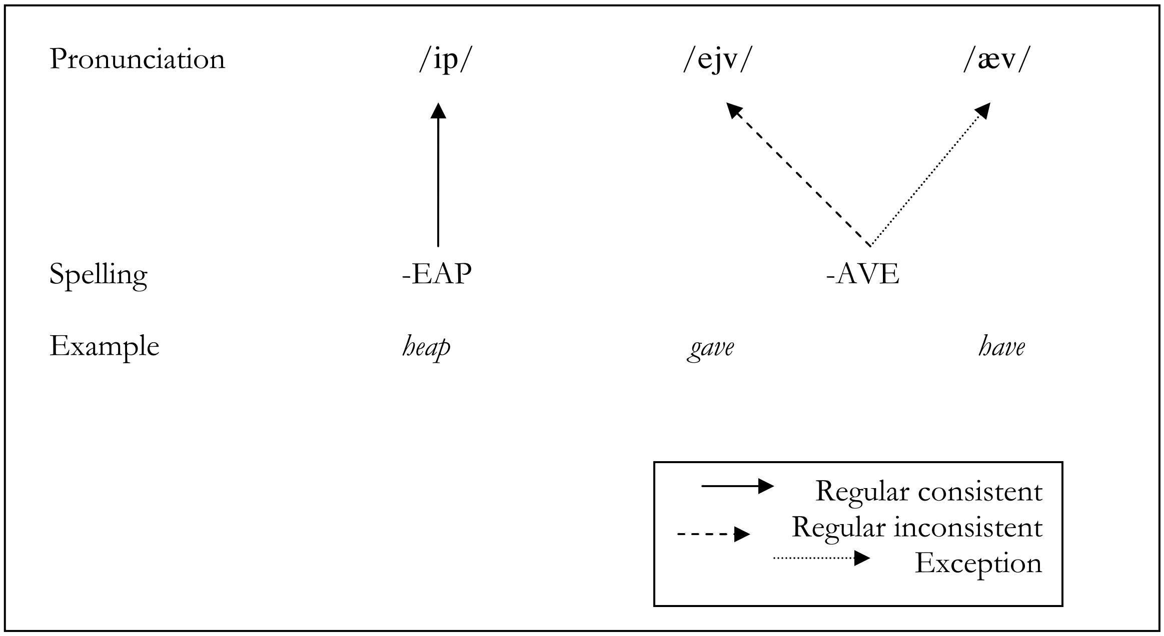


FIGURE 2-2 Glushko's spelling-to-sound relationships

Glushko discovered that exception and regular inconsistent words produced longer naming latencies and higher error rates than consistent words. As a result, Glushko claimed that word pronunciations were influenced by knowledge of similarly spelt words.

In the twenty-five years since Glushko's seminal work, research has attempted to support his claims on consistency effects. In 1990, Jared, McRae and Seidenberg attempted to address why the research into consistency effects up to that point, were not very robust. For these researchers, consistency effects were only part of the equation. According to them, a word's 'neighbourhood' was also an important factor. *Neighbours* are words that share the same rime and *neighbourhoods* are sets of words that share the same rime. Within neighbourhoods it is possible to have 'friends' (neighbours that are pronounced the same) and 'enemies' (neighbours that are pronounced differently). Figure 2-3 demonstrates the notion of a *word neighbourhood* with the rime –OWN.

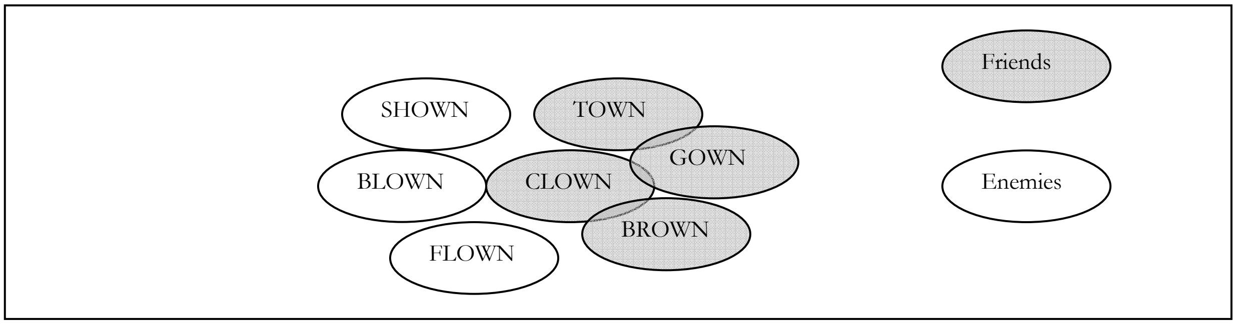


FIGURE 2-3 –OWN neighbourhood

Jared et al. hypothesized that consistency effects are contingent on the word frequency as well as on the frequency of inconsistent neighbours ('enemies'). Their results indicated that high frequency words are rarely affected by consistency effects as the enemy influence is 'washed out' due to the large degree of exposure of the target word. Low frequency words, on the other hand, are more susceptible to influence especially if their rival pronunciations have a high frequency.

Also attempting to investigate the lack of robust results in previous research were Stone, Vanhoy and Van Orden (1997). They suggested the problem arose because the previous research focused on only one direction of influence, from spelling-to-sound. This direction of influence was termed 'feedforward consistency' where inconsistent words have multiple pronunciations for one spelling (see Figure 2-2). The feedforward theory predicts that alternate inputs *cannot* influence performance. This means that even if the pronunciation has multiple spellings, they will not interfere with word recognition. Stone et al. however argued that this is not the case. Rather, they claimed that alternative spellings *can* influence performance. Thus, they posited a 'bidirectional flow of activation' where feedback consistency (from sound-to-spelling) also needs to be taken into account. This bidirectional flow can be seen in Figure 2-4.

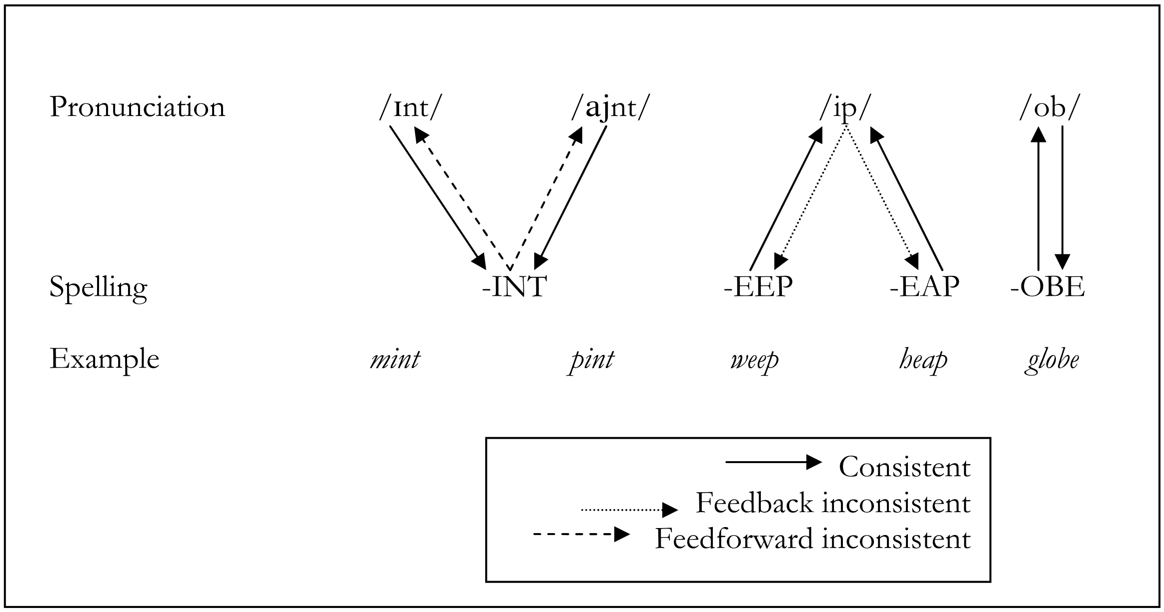


FIGURE 2-4 Bi-directional activation between orthography and phonology

Stone et al.'s results demonstrated that, in addition to feedforward inconsistencies, feedback inconsistencies also slow processing time of visual word recognition. Therefore, they concluded that the lack of control for feedback inconsistencies may explain the lack of robust results in the previous research.

Building on the work of Stone et al. (1997), Lacruz and Folk (2004) investigated feedforward and feedback consistency with regards to high- and low- frequency words. Their research yielded two major findings. First, they were able to confirm Stone et al.'s results that both feedforward and feedback consistency had an impact on the processing of monosyllabic words. According to Lacruz and Folk, "there is an interactivity between a word's sound and spelling – that information does not just flow from spelling to sound in visual word recognition tasks, but also from sound to spelling" (p. 1275). More importantly, and contrary to Jared et al.'s (1990) results, these researchers discovered that both feedforward and feedback inconsistent words were processed more slowly than consistent words regardless of word frequency.

2.5.2 Auditory recognition consistency effects

Up until 1998, all the research on consistency effects had been done in terms of visual word recognition performance. Ziegler and Ferrand (1998), however, were interested in investigating auditory performance. They believed that consistency effects in auditory word recognition should operate in the same way as visual word recognition. To test this theory, they asked native French speakers to determine whether auditorily presented French words and pseudowords were real or not. They discovered the words with phonological rimes that could be spelt more than one way produced longer auditory lexical decision latencies and error rates. These results prompted Ziegler and Ferrand to question the traditional assumption that speech processing is independent from written language because even without visual orthographic information, orthographic knowledge appeared to influence the processing of the spoken words.

In another study on auditory word recognition, Ziegler, Ferrand and Montant (2004) were interested in 'degrees of inconsistency.' For their purposes, they identified pairs of feedback inconsistent words with *dominant* and *subdominant* phonological mappings onto the spellings. Dominant words were identified as those where there is a greater probability of spelling (ex. –INE) and subdominant words were identified as those where the spelling occurs in fewer words (ex. -IGN). Like the previous research, Ziegler et al. were able to confirm that feedback consistent words were processed faster than inconsistent words in auditory word recognition. Their research also demonstrated that inconsistent words with dominant spellings were processed faster than subdominant spellings.

Here, for all intents and purposes, dominant and subdominant spellings could be equated with spelling frequency since dominant spellings by their very definition occur more often than subdominant spellings. Therefore, what Ziegler et al. appear to have demonstrated is that spelling frequency does not appear to influence the auditory latency divisions between consistent and inconsistent words but does appear to influence auditory latencies between different types of inconsistencies; namely, high-frequency and low-frequency spellings. As a possible explanation for why, when speech is primary, orthography affects auditory word recognition, Ziegler et al. suggest that stored words contain both orthographic and phonological information.

2.5.3 Phonemic consistency effects

While intriguing, the above research has been limited in scope. The multitude of previously mentioned studies were all interested in native speakers' perception and word recognition in their native languages. However, only a small body of research exists that goes below the word level to investigate whether consistency effects operate at the phonemic level. Although sparse, research has shown that orthography does have a role in perception of phonemes. Research into phoneme awareness has shown that alphabetic experience allows listeners to break down words into their component sounds (Read, Zhang, Nie, and Ding, 1986; Chueng and Chen, 2004; Carroll, 2004). In the words of Chueng and Chen (2004):

> The idea is that whereas an awareness of syllables, onsets and rimes emerges spontaneously with speech development, the ability to analyse spoken language into phonemes, or phonemic awareness, requires support from alphabetic reading. This is because the identity of the phoneme is made explicit only in alphabets (p. 3).

Based on their study of Cantonese-Chinese participants, Chueng and Chen argued that alphabets 'sensitize' readers to the phonemic level, which allows them to segment syllables into phonemes. The existing research agrees that alphabetic orthographies "radically alter the way listeners analyze speech into sound" (Hallé et al., 2000, p. 619) and that learning to read and write as children dictates the functional organization of the brain (Castro-Caldas, 1998).

This conclusion is in line with Treiman and Cassar's (1997) position that phonemic awareness arises from both: (1) the phonetic properties of the sounds, and (2) how the sounds are represented in print. Their results indicated both beginning readers and experienced readers of alphabetic languages were, for the most part, able to separate phonemes and spelling except when a combination of phonemes also sounds like the name of the letter (ex. /ɑr/ and /ɛm/). For these combinations, listeners were more likely to report the existence of one phoneme rather than two. Furthermore, because there was no significant difference between the speech processing abilities of beginning and experienced readers, Treiman and Cassar argued that complete separation of sound and spelling is no longer possible once individuals begin learning to read alphabetic letters. Again it is apparent that orthographic knowledge has an inescapable influence on the perception of sounds.

Similar to the studies on visual and auditory word recognition, other research has demonstrated that sounds that can be spelt in more than one way take longer to process than those sounds that have only one possible spelling. For example, Frauenfelder, Segui and Dijkstra (1990) found that /k/ took longer for French participants to detect in spoken words than /p/. These researchers attributed this effect to the fact that /k/ has many more possible spellings than /p/, which only has one. In another study, Dijkstra, Roelofs and Fieuws (1995) found that because /k/ in Dutch can also be spelt in more than one way, the detection of the /k/ in the subdominant 'c' spelling words took longer to process. These two studies suggest that orthographic consistency as well as frequency influence perception below the word level.

Perhaps one of the most enlightening research endeavours is Hallé, Chéreau and Segui's (2000) investigation of the possible contention between phonetic code (the surface phonology) and the linguistic code (the representations in print). This study compared the perceived stops in French words like 'absurde' [apsyrd]. Their first experiment confirmed that because of the process of French voice assimilation, the stop produced was /p/. The participants reported hearing /p/ when the stop was segmented out of the target words and presented in isolation. However, the second experiment showed that when *auditorily* given the entire word, participants had a tendency to report hearing a /b/ instead of /p/. These results led Hallé et al. to conclude that even in auditory tasks, orthography interferes with the perception of surface phonology. Apparently, in case where the phonetic code and linguistic code do not match, the linguistic code overrides the phonetic code and thereby causes listeners to misperceive the sound actually produced. These results suggest that orthographic influence is indeed a prevalent contributor to native sound perception. If this is the case then intuitively, it would be reasonable to assume that it would also be a prevalent contributor to nonnative sound perception.

2.5.4 Orthography and second language acquisition

Even smaller still is the body of research investigating L1 orthography's effect on nonnative sounds and second language acquisition. Erdener and Burnham's (2005) results indicated the existence of orthographic interference on the production of nonnative sounds (Irish and Spanish) by their Turkish participants. For example, orthographic 'j' exists in both

Turkish and Spanish. But the sounds this letter represents are not the same; it represents [ʒ] in Turkish and [x] in Spanish. Interestingly, when given only auditory information, the participants had 0% error rates for reproduction of the nonnative Spanish words. But when given both auditory and orthographic information, the Turkish participants' error rates jumped to 46%. The researchers attributed this phenomenon to the orthographic representation overriding the auditory representation of the word. As a result, the participants substituted their L1 sound, [ʒ], associated with the letter, 'j', they saw. Although this was not a study of Spanish and Irish second language learners, these results offer valuable insight into just how powerful a listener's L1 orthography can be.

Finally, some research that explores orthographic influence on second language learning are the studies conducted on bilinguals' activation of phonological representations (Jared & Kroll, 2001; Jared & Szucs, 2002). These authors claimed that when two languages share an alphabetic writing system but use some or all of the letters to represent different sounds, it may be possible to have conflicting pronunciations active at the same time (Jared & Szucs, 2002). This activation of conflicting pronunciations would in turn slow processing times as the bilinguals take the time to reconcile the conflict. In their study of interlingual homographs (words that are spelt the same but pronounced differently in both languages), Jared and Szucs found that two variables influenced the simultaneous activation of non-target and target representations. The first variable was whether the bilinguals were reading in their L1 or L2. According to the researchers, non-target activation is rarely strong enough to influence reading/naming in the L1. Conversely, when reading in their L2, readers demonstrated a clear non-target activation influence as reflected in longer response times and more errors. The second variable was whether the readers had recently named words in the non-target language. The results showed that readers were significantly slower at naming homographs after reading the non-target language words. In short, conflicting pronunciations were more likely to be activated when reading in the L2 and after reading non-target L1 words. Of course, the existence of conflicting pronunciations would have serious repercussions for second language acquisition.

2.5.5 Summary

In sum, from the literature reviewed above, the important tendencies of orthographic influence can be succinctly summarized into the following six points. First, both spelling-to-sound and sound-to-spelling inconsistencies influence reader/listener performance as manifested by longer latencies and higher error rates (Glushko, 1979; Jared et al., 1990; Stone et al., 1997; Lacruz & Folk, 2004). Second, even in the absence of visual stimuli, orthographic knowledge influences speech processing in literate listeners (Ziegler & Ferrand, 1998; Ziegler, et al., 2003; Ziegler et al., 2004). Third, orthographic information exerts influence at the phonemic as well as the word level (Frauenfelder et al., 1990; Dijkstra et al., 1995). Fourth, frequency and dominance of orthography determine the degree of influence (Jared et al., 1990; Ziegler et al., 2004). Fifth, from the early stages, alphabetic experience creates strong bonds between sound and spelling such that learners are unable to completely separate the two (Treiman & Cassar, 1997; Jared & Szucs, 2002). Finally and perhaps most importantly, in some cases, orthographic information 'suggests' sounds and overrides the actual phonologic information (Hallé et al., 2000; Erdener & Burnham, 2005).

2.6 Statement and significance of the problem

The traditional assumption is that shared orthography facilitates language learning by capitalizing on positive transfer. Indeed, it is one of the reasons that alphabetic scripts have been adopted to teach languages like Mandarin and Japanese. Presumably, alphabetic systems make languages that use logographic orthographies (i.e. characters) more accessible to those learners who are more accustomed to a phonemic representation of sounds. Although this may be true, it would be a mistake to assume that sharing orthographic systems would be without its dangers. Clearly, negative transfer needs to be considered as well. As the above research demonstrates, orthographic information exerts an inescapable influence on native speakers' perception of their own systems. It is not difficult then to imagine that because of strong L1 sound-grapheme associations, the L1 orthographic information may exert an even greater amount of influence on L2 learning when a single set of graphemes is meant to also represent two sound systems. Therefore, it is important to question if shared orthography hinders /interferes with second language acquisition. Does L1 orthographic knowledge negatively influence L2 speech processing when shared alphabetic letters represent different sounds? If so, does negative transfer from L1 to L2 magnify orthographic influence? More specifically, are L1 phonological representations automatically activated when using a shared alphabet?

In her book introducing beginners to the romanization systems of Mandarin, Killingly (1998) offers this caveat:

> It is well worth the extra time and effort spent on following the detailed instructions on how to produce sounds which you might otherwise *mechanically* [emphasis added] substitute with sounds related to the spelling system of your own language (p. 7).

Sound advice, no doubt, that could be passed on to students learning any L2 that shares a writing system with their L1. The point here is that where separate L1 and L2 sounds share the same orthographic symbol, learners must disassociate the symbols from the L1 sound they have already attached to it. That is, in order to perceive (and ultimately produce) the L2 sounds accurately, language learners must be able to transcend the symbol-sound associations they have already made in their L1. But the important question remains: Are language learners able to achieve this transcendence?

The research would suggest otherwise. Recall that Treiman and Cassar (1997) found that once children begin to read, they could not fully separate sound and spelling. Jared and Szucs (2002) argue that native language "phonological representations were automatically activated and could not be effectively deactivated or inhibited" (p. 236). Furthermore, Landerl, Frith and Wimmer (1996) claim that the automatic orthographic and phonological co-activation makes individuals more "susceptible to unwanted interference" (p. 12). In short, research appears to suggest that complete disassociation of L1 sound-symbol associations is not entirely possible.

Closely related to the idea of strong symbol-sound associations is the idea of shared phonology. Of course, along with shared orthography comes the assumption that the phonology represented by those shared letters would also be shared. The important question here is whether learners do assume that shared letters also represent shared sounds. Often there is a positive transfer of L1 sound-symbol to L2 sound-symbol. But what about instances where the sounds are similar but not exactly the same? Research has demonstrated that orthographic information can and does override phonetic information and lead listeners to either disregard or misperceive what is heard (Hallé et al., 2000; Erdener & Burnham, 2005). Therefore, is it not conceivable that shared orthography may also lead learners to misperceive an L2 sound? This misperception may in turn lead to Flege's (1987, 1995) notion of equivalence classification whereby the learners equate the L1 and L2 sounds as one. Thus, shared orthography may have the potential to provoke equivalence classification tendencies in language learners so that they are unable to perceive the differences between the L1 and L2 sounds.

Another consideration is one of frequency and dominance. It seems plausible that orthographic influence would have more of an impact if a shared symbol represents a high frequency L1 sound and a low frequency L2 sound and less of an impact if the shared symbol represents a low frequency L1 sound and a high frequency L2 sound. This brings many questions to the fore. How are relative frequencies determined across languages? All other things being equal, would L1 sounds be considered more frequent than L2 sounds simply because they are entrenched and the learner has had much more experience with them? In other words, is it possible to consider an L1 sound as the dominant sound and the L2 sound as the subdominant?

Consider Table 2-2. This table demonstrates that 13 of the shared English and Pinyin consonantal letters represent separate sounds in each language.

Shared Letter(s)	*English*		*Mandarin*	
	sound	*example*	*sound*	*example*
b	/b/	(**b**oy)	/p/	把(bǎ)
d	/d/	(**d**og)	/t/	讀 (dú)
g	/g/	(**g**oat)	/k/	狗 (gǒu)
h	/h/	(**h**at)	/x/	好 (hǎo)
ch	/tʃ/	(**ch**art)	/tʂ^{h}/	臭(chòu)
sh	/ʃ/	(**sh**are)	/ʂ/	上(shàng)
z	/z/	(**z**oo)	/ts/	走(zǒu)
s	/s/	(**s**ee)	/s̪/	色 (sè)
c	/s/	(**c**ycle)	/ts^{h}/	菜(cài)
r	/ɹ/	(**r**aw)	/ʐ/	人 (rén)
j	/dʒ/	(**j**udge)	[tɕ]	叫(jiào)
q	/kw/	(**qu**iet)	[tɕ^{h}]	錢(qián)
x	/ks/	(x-ray)	[ɕ]	學(xué)

TABLE 2-2 Shared orthographic symbols and the separate sounds they represent in English and Mandarin

From this table it is clear that many Mandarin sounds have a closely related (though not exactly the same) counterpart in English, excepting the final two letters, 'q' and 'x'. But each pair of sounds is represented by the *same* letter(s). For example, 'sh' represents [ʃ] in English but [ʂ] in Mandarin. Thus, it is not unreasonable to speculate that this method of sound representation is likely to contribute to difficulties in the learners' abilities to distinguish between the L1 and L2 sounds in each contrast. Therefore, the question becomes whether Pinyin further reduces the perceptual saliency of some Mandarin sounds.

However, it is important to keep in mind that shared orthography is not enough of a basis for confusion between L1 and L2 sounds. For example, in a discrimination task, it is unlikely that listeners would confuse [ɕ] with [ks] or [tɕʰ] with [kw]. There must be another connection between the sounds on which confusion between the two will be founded. That is, for L1 and L2 sounds to be confusable, they must share many of the same acoustic properties. As the discussion in Section 2.4 pointed out, the differences between L1 and L2 elements (in this case, phonemes) must be perceptually salient for learners to be able to distinguish the L2 from the L1, and therefore, learn the L2 element. In this research, the differences between the acoustic properties of the native-nonnative contrasts (i.e. the English and Mandarin sounds) must be salient in order for participants to make accurate L1-L2 discriminations. Therefore, varying degrees of similarity (or dissimilarity) between L1 and L2 sounds will inversely correspond to learners' degrees of perceptual sensitivities to those sound contrasts. Thus, the greater the dissimilarity the greater the perceptual sensitivity and vise versa.

Regardless of whether one believes in the necessity of a native-like mastery of a second language, or simply being understood, it is important to understand what constitutes the barriers that make L2 learning a challenging task for most learners. This research has the potential to identify and characterize one such barrier, shared orthography. It will attempt to shed some light on how much influence (if any) orthography exerts on second language learners' perceptions of Mandarin sounds and thereby will contribute to the sparse body of research on orthography's relationship to L2 sound perception. Further, while it is highly unlikely that Pinyin would be abandoned as a medium of instruction, the results from this study may suggest Mandarin language instructors adjust their approaches to using Pinyin in their classrooms. That is, the results may indicate that language instructors need to be more explicit by highlighting the differences between the L1 and L2 sounds that share alphabetic symbols.

Chapter Three

METHODOLOGY

"Perception is a two-way street." (Stone et al., 1997, p. 337)

For the purposes of this research, the experiment focused on native Canadian English speakers who had no previous experience with any Chinese language. The three-part experiment (which included pre- and post-instruction perception tests and instruction phases) was designed to determine whether L2 sound categorization, and ultimately sound discrimination, would be influenced by the orthographic medium (or lack thereof) with which the participants had been instructed. This chapter describes in detail the methodology employed for the collection and analysis of the data. The first section provides the background information of the participants and the recruitment process. The second and third sections describe the creation of the stimuli and the experimental materials used in the perception tests, respectively. The fourth section outlines the data collection procedure. And finally, the fifth section describes the statistical analysis of the data.

3.1 Participants

Thirty-two participants were recruited from the University of Victoria. The participants were all native speakers of Canadian English[2]. The recruitment process involved posting recruitment posters around the university. Additional recruitment also involved the researcher, some of the participants and some of the researcher's colleagues sending out emails to persons who might be interested free Chinese lessons and the project. Not all

---

[2] The consonant inventory given in Figure 2-1 is based on North American English. The underlying assumption here is that there are few differences between the *consonants* of different dialects of English (Rogers, 2000). That is, there are no observable differences between the consonants of the American and Canadian dialects.

participants were university students, as a number of participants were informed of the project by email and word of mouth. Participants contacted the researcher via email and the classes and test appointments were all organized through this electronic medium. None of the participants had previous language instruction (either formal or informal) in any Chinese language. Only those volunteers who reported no auditory or visual impairments were accepted for the study. Table 3-1 provides a synopsis of the participants' characteristics.

Group	Number of Participants	Mean Age	Nationality	Other languages learned
PN **(Pinyin)**	n = 11	23.5 years	Canadian	▪ French, Italian, Dutch, Japanese, Spanish, Russian, German, Greek, Arabic
ZN **(Zhuyin)**	n = 11	29.2 years	Canadian	▪ German, French, Italian, Punjabi, Hindi, Japanese, Spanish
CL **(Control)**	n = 10	23.5 years	Canadian	▪ Japanese, French, Spanish, Portuguese, German, Latin

TABLE 3-1 Participant demographics

3.2 Stimuli

The target contrasts investigated in this research were limited to those separate sounds from the L1 and L2 that share an orthographic representation between the English alphabet and Pinyin, which is the most common orthography used in teaching Mandarin. That is, the English-Mandarin (native-nonnative) target pairs investigated are represented by the following seven orthographic symbols given in Table 3-2.

Orthographic symbol	Sound contrast English	Mandarin
c[3]	[s]	[tsʰ]
z	[z]	[ts]
s	[s]	[s̺]
ch	[tʃ]	[tʂʰ]
sh	[ʃ]	[ʂ]
r	[ɹ]	[ʐ]
h	[h]	[x]

TABLE 3-2 The 7 target English-Mandarin contrasts and the orthographic symbol(s) they share

Other native-nonnative contrasts were not investigated for a number of reasons. Vowel contrasts were excluded because both English and Mandarin vowels are very complex and would warrant a much more complicated and lengthy investigation. In addition to the vowels, two groups of separate consonants that share orthographic representation were also excluded. The stops that share 'b', 'd' and 'g' graphemes were not considered because they are cued significantly by formant transitions within the vowel, and hence are extremely difficult to extract from the vowel context (a more detailed description of the stimuli creation process is given below). The second group of sounds that was excluded was the one containing the sounds that share 'q', 'j' and 'x'. These sounds were not studied because the Mandarin counterparts are considered to be allophones of the retroflex alveo-palatal consonants and only occur in limited environments (Lin, 2001), namely, before the high front vowels [i] and [y].

[3] The pronunciation of orthographic 'c' for the purposes of this study are limited to those words where 'c' is a soft 'c' in words such as *cycle, conceive*, *cite*, and so on. However, since the /s/ pronunciation of orthographic 'c' does not occur before a low back vowel, it was necessary to use an 's' word (i.e. *somber*) to gather the /s/ target stimuli for orthographic 'c' (see Appendix A; also see below for the rationale behind the use of the target stimuli followed by a low back vowel.).

The target consonants were collected from English and Mandarin words (see Appendix A for the complete list of words) that were recorded by three native speakers of each language. These speakers were all female graduate students at the University of Victoria. The rationale behind the use of multiple speakers was to minimize the participants' abilities to make perceptual decisions based on individual speakers' voice qualities alone[4]. Each English and Mandarin word was recorded five times to ensure the recording process yielded clear native-like exemplars of each sound. The words were recorded using the software Praat (version 4.4.13) at a sampling frequency of 44.1 kHz with an M-Audio Luna Large-Diaphragm Condenser microphone on a Windows XP workstation connected to M-Audio Firewire 4/10 Preamp Interface. The resulting sounds files were saved as *wav*. files.

For the perception tests, the desired stimuli were simple consonant-vowel (CV) syllables. The motivation to study the native-nonnative contrasts in the onset position was because Mandarin only allows nasals and /ɻ/ in coda position. In addition to the onset consonants, a single exemplar of the vowel [a] was used, which occurs in both English and Mandarin vowel inventories. The stimuli creation process was as follows:

1. All target sounds were recorded in either stressed CV or CVC syllables, where the vowel in the syllable was a low back vowel (see Appendix A).
2. The target sounds were identified in their respective words using the software Audacity (version 1.2.4).
3. The low back vowel (and the coda) in the original word was removed and replaced a single low back vowel, [a].

---

[4] The term 'speakers' in this paper refers to the native subjects from whom the sounds were gathered to create the project stimuli while the term 'participants' refers to the native English subjects who participated in the Mandarin classes and the perceptual pre- and post-tests.

From the five recorded exemplars of each sound, five new CV syllables were created using the above procedure. This process yielded 80 sound stimuli (16 sounds x 5 exemplars).[5]

The decision to segment the target consonants from the rest of the word was made for two reasons. First, not all the target sounds could be found in English-Mandarin homophone pairs. Second, vocalic information was a concern. In order to minimize the vocalic information available to the participants, a single vowel, [a], was used in the creation of the CV stimuli. One native English speaker recorded numerous tokens of [a]. By considering native-likeness and duration, the best vowel token was then determined by a different native English speaker (the researcher). By using low back vowels both in the original words and in the manipulated CV syllables, the consonant to vowel transitions in the manipulated CV stimuli were kept as natural as possible. The five tokens for each sound were created in the same *wav*.file and were separated by 1 second of silence. Finally, each sound file was normalized to a peak intensity of -3dB.

Also, in the case of voiced consonants, the pitch contours of the targets and the vowel were examined to ensure that they formed a natural contour. When unnatural, Praat was used to adjust the contour of the target into the vowel. Figure 3-1 demonstrates this adjustment process. The 'natural contour' was based on the pitch contour from target consonant into the vowel in the original sound recording.

[5] Two additional sounds, English [f] and [v], were also recorded and segmented for use in the practice trial section of the pre- and post-tests.

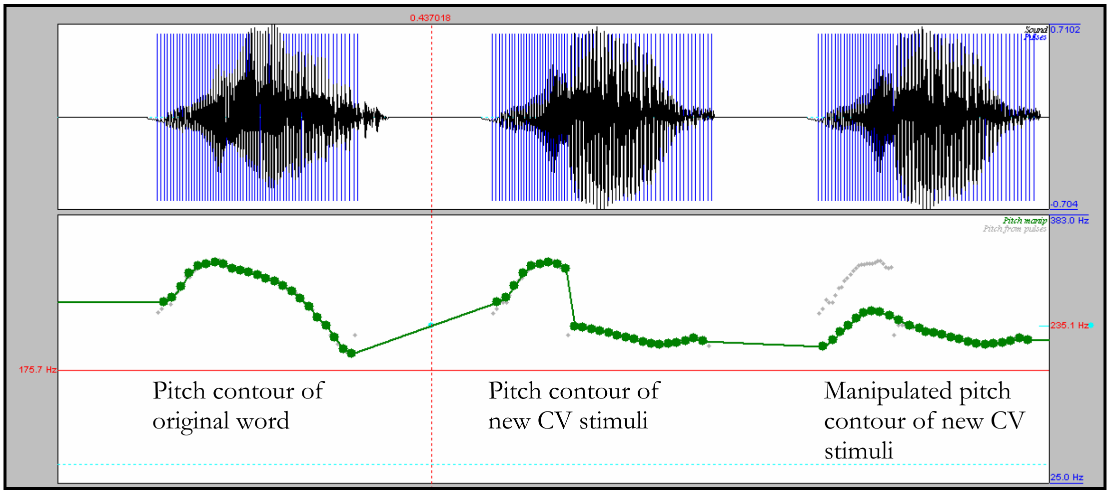


FIGURE 3-1 Sample pitch contour adjustment for new /ẓa/ CV stimuli

The same contour pattern was used for each of the five tokens for each adjusted consonant.

Once the 5 new CV tokens were created for each target sound, native English (NE) and native Chinese (NC) judges (3 for each language) evaluated the tokens. More specifically, the NE judges judged the new English CV stimuli while the NC judges judged the Mandarin CV stimuli. Each judge completed three tasks. First, they identified the target sound by writing the corresponding orthographic symbol(s) for the target. This means that the NEs were instructed to identify the target sounds according to standard English spelling conventions and the NCs were instructed to identify the target sounds according to standard Pinyin spelling conventions. This task was designed to confirm that the target sounds were indeed orthographically represented by their presumed letters. Second, since the stimuli were created by segmenting and pasting, it was imperative that the native judges assessed the 'naturalness' of the stimuli to ensure they were as native-like as possible. And finally, the native judges indicated the tokens they felt 'best' represented the target sound.

Only those tokens on which all three judges agreed were good instances of the intended category were used for the creation of the stimuli triad sets. Table 3-3 provides the complete list of the acceptable CV syllables (as determined by the 6 native speaker judges) used in the triad creation. Here, NS refers to the native speaker from which the tokens were gathered and the numbers (1 through 5) indicate those CV stimuli that all three native judges agreed were acceptable exemplars of that particular target sound.

Symbol	ENGLISH				MANDARIN			
		Acceptable tokens				Acceptable tokens		
	sound	NS_1	NS_2	NS_3	sound	NS_4	NS_5	NS_6
c	[s]	2,3,4,5	3,5	4,5	[tsʰ]	2,3,4,5	1,2,3,4,5	1,2,3,4,5
z	[z]	1,2,3,5	3,4,5	4,5	[ts]	1,2,3,5	4,5	1,2,4,5
s	[s]	2,3,5	1,2,3	4,5	[s̺]	1,2,3,4,5	1,2,3,4,5	2,4,5
ch	[tʃ]	1,2,3,4,5	1,2,3	1,2,3,4,5	[tʂʰ]	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5
sh	[ʃ]	1,5	1,4,5	2,3,5	[ʂ]	1,2,4,5	1,2,3,4	3,4,5
r	[ɹ]	1,2,3,4,5	4,5	1,4	[ʐ]	3,4,5	2,3,4,5	2,4,5
h	[h]	3,4,5	1,2	1,4	[x]	1,2,3,4	1,2,3,4,5	1,2,3,4,5
f	[f]	1,2,3,4,5	1,5	1,2,3,4	--	--	--	--
v	[v]	1,2,3,4,5	4,5	1,2	--	--	--	--

TABLE 3-3 Stimuli tokens deemed good exemplars by native speaker judges for each target sound

For example, in the case of English speaker 3's (NS_3) tokens of 'sh', the NE_1 judge felt tokens 2, 3, 4, and 5 were all good exemplars of 'sh', NE_2 also felt that tokens 2, 3, 4, and 5 were good exemplars, but NE_3 felt 1, 2, 3, and 5 were good exemplars. In other words, only tokens 2, 3 and 5 of 'sh' were agreed upon by all 3 NE judges to be good exemplars of English 'sh' (the corresponding cell in Table 3-3 is highlighted above). All undesirable exemplars were deleted from the sound files. As shown in Table 3-3, the number of acceptable stimuli varied for each sound. For the creation of the token triads (see section 3.3), the stimuli used were randomly chosen from the acceptable stimuli. This means that not all the acceptable stimuli were used. Three stimuli from each speaker for each sound were chosen; however; in the cases where there were only two acceptable tokens, one stimulus was used twice.

3.3 Experiment materials

Based on the methodologies employed in recent studies by Flege and McKay (2004), Guion, Flege, Akahane-Yamada, and Pruit (2000), and Baker, Trofimovich, Mack and Flege (2002), the perception task chosen for this experiment was an 'oddity' discrimination task. This type of task is similar to an AXB discrimination task except instead of choosing which stimuli (either 'A', the first, or 'B', the last) 'X' (the middle) matches, participants are required to choose the 'odd' item out. These odd stimuli can be in any one of the 3 positions. For the purposes of an 'oddity' task, two types of triads were created. The first type of triads contained 'catch' triads (ex. [h_1a] [h_2a] [h_3a])[6] where each of the three CV stimuli was produced by a different speaker and thereby including in each catch triad physically different exemplars of the *same* sound category. The purpose of the 'catch' trials is to test "participants' ability [sic] to ignore audible but phonetically irrelevant within-category variation" (Flege & MacKay, 2004, p. 9). The other set of triads were 'odd' triads (ex. [h_2a] [h_1a] [x_4a]) where each of the CV stimuli were again produced by different speakers but included an 'odd' item representing a different sound category from the other two tokens. Both the catch and odd triads were created from those stimuli deemed acceptable by the native judges. In order to ensure that each triad contained tokens from different speakers,

[6] The subscript indicates from which speaker the target sound was collected.

the triads were manually created in Audacity with 1500ms[7] of silence between the stimuli in each triad. The triads were then saved as single *wav.*files.

The appropriate length of silence to be inserted between the presentations of each stimulus was one extremely important aspect that needed careful consideration. This is what the literature refers to as the Interstimulus Interval (ISI). ISI research (ex. Pisoni, 1973; Werker and Tees, 1984; Werker and Logan, 1985; Matthews and Brown, 2004) has established a strong inverse relationship between the length of the ISI and the amount of information available to the listener. Simply put, the longer the ISI, the greater the amount of short-term memory decay. This means that the type of perceptual discrimination possible is contingent on how much time there is between the presentations of the stimuli.

Werker and her colleagues (1984, 1985) argue for three levels of processing that are dependent on ISI length. *Acoustic processing* is considered to encompass non-linguistic information and is only accessible with an ISI of less than 500ms. This type of processing allows listeners to demonstrate sensitivity to acoustic differences between stimuli rather than phonological differences. For example, in less than 500ms, listeners are able to distinguish between different utterances of the same sound from the same speaker.

Phonetic processing, which is accessible in ISI durations between 500 to 1500ms, carries less detail (due to memory decay) than the fine detail associated with acoustic processing. Within this level, listeners demonstrate sensitivity to allophonic variations whereby listeners can distinguish between exemplars of the same category (within-category distinctions). This means that listeners are able to demonstrate a sensitivity to phonetic distinctions employed in other (nonnative) languages.

---

[7] The judged CV syllables were separated by 1000ms of silence while the CV syllables in the perception triads where separated by 1500ms of silence.

The third type of processing is *phonemic processing*, which is activated with ISI durations of over 1500ms and retains the least amount of linguistic information. Here, sensitivity is restricted to only those stimuli that are separated by native-language phonological boundaries (between-category distinctions) (Werker and Logan, 1985). For example, at the acoustic level, the fine detail enables listeners to discriminate between different tokens of the same sound provided by the same speaker. At the phonetic level, listeners are no longer able to distinguish between same speaker tokens but they can still distinguish between different speaker tokens of the same sound. And finally, at the phonemic level, the lack of acoustic and phonetic detail prevents listeners from distinguishing between different speakers' productions of the same sound.

Werker and Tees (1984) maintain that during ***phonetic*** processing, "sounds may be perceived as synthesized (non-meaningful) percepts, and thus may be retained in memory for a short duration, but have a more rapid decay period" (p. 1876). In contrast, they claim that during ***phonemic*** processing, "stimuli are treated as meaningful events that can be efficiently encoded and represented in memory for a long duration" (p. 1876). It seems then that although longer ISIs reduce the amount of detailed information accessible to listeners, the information that does remain can withstand decay and can be stored in long-term memory. In the words of Matthews and Brown (2004), "presenting pairs of stimuli with a substantial ISI (eg. 1500ms) precludes subjects from relying on acoustic representation in short-term memory to perform a discrimination task" (p. 11).

With these perceptual processing considerations in mind, the importance of appropriate ISI durations is readily apparent. Since the goal of this research is to investigate whether orthography influences the perception, and ultimately, the categorization of Mandarin sounds in relation to the English phonemic system, it was necessary to activate ***phonemic***

processing in the discrimination task. If phonemic processing only demonstrates sensitivity to between-category distinctions, the expectation then would be that NE learners of Mandarin would be able to successfully distinguish the 'odd item out' only for those sounds that have been stored as belonging to separate sound categories. Similarly, unsuccessful attempts would suggest that the sounds are stored as exemplars of the same category[8]. Therefore, to activate phonemic processing and determine whether NE Mandarin learners can make phonemic distinctions between separate English and Mandarin sounds that share orthographic symbols, an ISI of 1500ms was inserted between the three stimuli in each triad.

Another important consideration was the number of stimuli to incorporate in the perception tests. Each symbol/contrast pair that exists between Mandarin and English was represented by 12 triad sets. The number 12 was chosen for four reasons: to ensure that (1) the participants would have multiple opportunities to distinguish between the target contrasts, (2) there were not too many tokens pairs as to make the test too long and tedious, (3) there were an equal number of 'catch' and 'odd' triads, and finally 4) there were an equal number of '1', '2', and '3' odd out triads. For example, the 12 token triads for the orthographic symbol 'h' are 3 [h] catch triads, 3 [x] catch triads, 3 [x] odd triads, and 3 [h] odd triads. Table 3-4 demonstrates the distribution of triads.

[8] In her discussion of Exemplar Dynamics, Pierrehumbert (2001, 2002) argues that phonemic categories are represented by "a large cloud of remembered tokens" that form a cognitive map where "memories of highly similar instances are close to each other and memories of dissimilar instances are far apart" (Pierrehumbert, 2001, p. 140).

CATCH TRIADS		ODD TRIADS	
English *catch*	Mandarin *catch*	English *odd*	Mandarin *odd*
$[h_2a\ldots h_3a\ldots h_1a]$*	$[x_5a\ldots x_6a\ldots x_4a]$	$[\mathbf{h_2a}\ldots x_6a\ldots x_4a]$	$[\mathbf{x_5a}\ldots h_3a\ldots h_1a]$
$[h_3a\ldots h_2a\ldots h_1a]$	$[x_6a\ldots x_5a\ldots x_4a]$	$[x_6a\ldots \mathbf{h_3a}\ldots x_4a]$	$[h_3a\ldots \mathbf{x_6a}\ldots h_1a]$
$[h_1a\ldots h_2a\ldots h_3a]$	$[x_4a\ldots x_5a\ldots x_6a]$	$[x_4a\ldots x_5a\ldots \mathbf{h_1a}]$	$[h_1a\ldots h_2a\ldots \mathbf{x_4a}]$

* each subscript denotes from which speaker the consonant token was gathered and the […] indicate that each token was separated by 1500ms of silence.

TABLE 3-4 12 token triads for the [h-x] contrast represented by orthographic 'h'

The triad groups for each one of the English-Mandarin sound contrasts investigated was done in this manner. This means that there were a total of 84 token triads for both the pre- and post-tests [7 contrasts x 3 English 'catch' x 3 Mandarin 'catch' x 3 English 'odd' x 3 Mandarin 'odd']. There were also 12 token triads for the [f-v] contrast used in the practice trials.

For the presentation of the stimuli, the software SuperLab Pro 4 was used to create the perception test. In the test, the 84 stimuli triads were manually organized into three blocks of 28 triads each. This manual organization was motivated by the need to ensure that each block contained an equal number of 'catch' and 'odd' triads as well as an equal distribution of 1, 2, and 3 answers for the 'odd' trials. Further, within each block, the triads were also put into four trials where each trial contained one triad of each sound contrast. This was done to minimize the likelihood of having same sound triads following each other. During the presentation of the stimuli, the triads were randomized within each trial and the trials were randomized within each block.

For the presentation of each triad, a sound cue followed by 25ms of silence introduced the triad. This sound cue was added to prepare the participants for each stimuli triad. During the course of the test, the SuperLab program recorded (in a text file) the participant responses for each triad and whether these responses were either 'correct' (C) or 'incorrect'

(E) (see Section 3.4 for a more in-depth description of the task). Response times (RTs) in milliseconds were also recorded. RTs were recorded from the offset of the triad presentation to the depression of any key on the keyboard. However, the program was set to timeout if it had not received a response from the participant within 5 seconds of the offset of the triad. In this event, the response (or lack thereof) was recorded as a 'no response' (NR) with a response time of 0ms and the program moved on to the next triad. Once a block was completed, the program enforced a mandatory 2 minute break. The participants were free to take a longer break but they could not take a shorter one. The total duration of this process was approximately 20 minutes depending on whether the participant chose to take more than the mandatory 2 minute break between each block.

3.4 Procedure

This section outlines the procedure utilized in the three-part experiment devised for this research. The first section describes the organization of participants into their respective experimental groups. The second section describes the procedure for the pre-test. The third section describes the training phase of the project and the last section describes the post-test procedure.

3.4.1 Group organization

The participants were assigned to one of the three experimental groups by virtue of their availability and when they contacted the researcher. That is, because it took several weeks to find enough participants to fill all three groups, it was a concern that those who volunteered very early would no longer be available if they had to wait until all the participants had been recruited. Therefore, the first ten volunteers who were available for the same days and times were put into the Pinyin (PN) group. The next ten were put into the Zhuyin (ZN) group and the final ten were put into the Control (CL) group (see Section 3.4.3 for a detailed description of the training process). Once a group was full, the classes were organized and conducted for that group. Each group was organized to run for a week and all three groups were completed in three consecutive weeks. As mentioned above, the first week's group was the PN, the second was ZN and the final week's group was CL. Due to a high attrition rate, it was necessary to conduct an additional PN experimental group and ZN experimental group. The data from these second sets of lessons was compiled with the first sets to give the final numbers for the PN and ZN groups. There was no attrition in the CL group, and therefore, there was no need to conduct a second set of classes for this group.

3.4.2 Pre-test

Prior to the commencement of each group's Mandarin lessons, each participant met individually with the researcher on the day before or the day of the first class. During this meeting, the participant, filled out the consent form and background questionnaire (see Appendices C and D, respectively) as well as took the perception pre-test. These individual perception tests were conducted in a sound treated room in the Phonetics Laboratory of the Department of Linguistics at the University of Victoria. At a comfortable listening level and using MDR-V600 Sony headphones, the stimuli were presented to the participants via the software SuperLab 4 on an Acer Aspire 3000 personal laptop. All instructions were given verbally by the researcher in the participants' native language, English.

For the pre-test, each participant was told that the goals of the test were to investigate listeners' abilities to distinguish Mandarin sounds from English and to establish a base line with which to compare the post-test results. The researcher explained that the participants would hear 84 triads containing three syllables (hence the name *triad*). They were informed that half of the triads contained three examples of the same sound category. However, they were warned that although each of the syllables was an example of the same sound category, each syllable was spoken by a different speaker. They were told that this meant they would not be able to base any of their judgements on individual speakers' voice qualities so they would have to listen carefully to the sounds themselves. In the cases when the participants thought all three syllables to be examples of the same sound, the participants were asked to indicated their response by pressing the button marked 'N'.

For the other half of the triads, the participants were told that one of the three syllables was an example from a different sound category and again each was spoken by a different speaker. When they thought they discerned an 'odd' syllable out, the participants were asked to indicate their response by pressing the number that corresponded to the position of the odd syllable in the sequence of three (either '1', '2', or '3') The response keys were clearly marked on the keyboard with yellow tabs labeled as either '1', '2', '3', or 'N'. The participants were encouraged to answer as quickly but as accurately as possible and in the event that they did not know they were asked to make their best guess.

In addition, the researcher explained that the 84 triads were broken up into three blocks of 28 triads and that there was a mandatory 2 minute break between each of the blocks. The participants were also warned that the program would wait for a response from them after each triad was presented but that if they took longer than 5 seconds to answer, the program would automatically move on to the next triad. The program randomly presented each of the triads for each participant.

Before the start of the pre-test, the participants had an opportunity to familiarize themselves with the test procedure by doing a practice trial. In this trial the participants practiced discriminating between [f] and [v]. The results from the practice trial were not recorded nor was the [f]-[v] sound contrast used in the data collection process. During the practice trial, the researcher remained in the room with the participant to answer any questions that might arise about the test. After the practice trial, the researcher left the sound booth before the participant started the actual perception test. Both the correct and incorrect responses and the response times (RTs) were recorded and saved in text files. After the participants finished the test, the researcher then confirmed that the data were successfully recorded and saved on the computer.

3.4.3 Language training

During the training phase (i.e. the Mandarin classes) of the study, the PN group was taught Mandarin phonemes through the alphabetic system, Pinyin. The ZN group was taught the same Mandarin phonemes via the phonetic system called Zhuyin (see Appendix B for a comparison of the Pinyin and Zhuyin orthographies) and the CL group was taught the Mandarin phonemes with no orthographic medium. In addition to learning the phonemes, each group was taught how to count to 100 and useful everyday Mandarin phrases such as 'hello,' 'goodbye,' 'my name is ____' and so on (see Appendix E for a sample activity). The participants received four and a half hours of Mandarin instruction via the orthographic medium assigned to their particular group. Four and a half hours of instruction was chosen so that participants would gain repeated exposure to and practice of the Mandarin sounds but also to make the instruction load of the instructor manageable for this project. These hours of instruction were broken up into three lessons (which were taught on three separate days within the space of a week) of one and a half hours each.

To control for 'teacher effect', all three groups had the same Mandarin instructor. Since it was necessary to have an instructor who was intimately familiar with both Pinyin and

Zhuyin, the instructor was from Taiwan where Zhuyin is taught[9]. While many Taiwanese Chinese (including the instructor in this research) have mastered Pinyin, few Mainland Chinese know Zhuyin. At no point during the treatment stage did the instructor draw the participants' attentions to contrasts between the pairs of L1 and L2 sounds.

While it is virtually impossible to prevent the ZN and CL participants from making Mandarin / English phoneme associations in their minds, they were prevented from making any visual associations. That is, for the ZN and CL groups, the participants were not allowed to make any notes about the sounds with alphabetic representation. In other words, if they thought that a Mandarin phoneme sounded like an English phoneme, they were not permitted to write the alphabetic symbol to help them remember what it sounded like. For example, if they thought that Mandarin [ʂ] sounded like English [ʃ], they could not use a written 'sh' to help them remember it. The participants, however, were allowed to make notes on how the sounds were pronounced. For instance, in the case of the retroflex phonemes, the participants could write that their tongues needed to 'curl back in their mouths' to pronounce this group of sounds correctly.

3.4.4 Post-test

The day after the completion of the Mandarin classes, the participants took their perception post-tests. As with the pre-tests, each post-test was done individually in the same sound-treated room as the pre-test, in the Phonetics Laboratory of the Department of Linguistics at the University of Victoria. The participants were told that they would be doing the same test as the one they took before the Mandarin classes. They were given the same instructions and did the practice trial to re-familiarize themselves with the task. Once the

---

[9] This instructor was also one of the native Mandarin speakers who provided the target consonants that were used in the CV stimuli. Her tokens were judged by mainland Mandarin speakers to be 'nativelike'.

test was finished, the researcher asked each participant for any feedback s/he had about the project.

3.5 Analyses

This section identifies and defines the independent and dependent variables as well as discusses the analyses of the data.

3.5.1 Independent variables

For the purposes of this study, there were three independent variables. The first independent variable was ***orthographic medium*** and was operationally defined as *the system of visual symbols that are assigned to represent and used to teach the sounds of any given language*. The orthographies used in this study were restricted to two types: (1) ***Pinyin***, which was modeled on western alphabetic orthographies and developed by the Chinese government for the purpose of representing Chinese phonemes, and (2) ***Zhuyin***, which is a syllabary orthography that also represents Chinese sounds (see Appendix B). There was also a control group where Mandarin was learnt without the use of an orthographic medium. The control group was used to gage how much perceptual sensitivity could be attributed to auditory learning alone. In short, each experimental group (i.e. Pinyin, Zhuyin, and Control) represents a level of the independent variable – orthographic medium – therefore, this independent variable will henceforth be referred to as 'group'. Since it is not possible to quantify orthography, this independent variable was placed on a *nominal scale*, where each orthographic group was assigned a number (Pinyin = 1, Zhuyin = 2 and Control = 3). The second independent variable was ***perception test***. Again, because it is not possible to quantify this variable, the pre- and post-tests were also assigned numbers and placed on a nominal scale (pre-test = 1 and post-test = 2). And finally, the third independent variable was the ***sound contrast***. As mentioned above, seven sound contrasts were investigated. These contrasts were also given a number and placed on a nominal scale (/s-tsʰ/ = 1, /tʃ-tʂʰ/ = 2, /h-x/ = 3, /ɹ-ʐ/ = 4, /s-ʂ/ =5, /ʃ-ʂ/ = 6, and /z-ts/ = 7). Other important variables were also recorded for possible consideration in the results. These variables include: 1) gender, 2) number of second languages learned, 3) the block of stimuli presentation, and 4) the position of the correct answer.

3.5.2 Dependent variable and measurements

The single dependent variable used to measure the effect of orthography was the learners' ***perceptual sensitivity*** to native-nonnative sound contrasts. In this study, the operational definition of perceptual sensitivity is *a learner's ability to distinguish between an L1 and an L2 phoneme* that was measured in terms of error rates, a-prime and response times (all three measures are described in more detail below). This study employed a pre-test and a post-test to assess the participants' perceptual sensitivities to the seven English-Mandarin sound contrasts. In the pre- and post-tests, the participants were asked to identify whether groups (triads) of three CV syllables were examples of the same sound category or whether the triads contained an 'odd' item out. These answers were recorded and classified as either correct or incorrect.

The first measurement of perceptual sensitivity was error rates (ER). ERs were calculated as the percentage of incorrect responses relative to the total number of responses. The error rates were calculated for each: 1) participant, 2) sound, 3) group, and 4) test. They were used to determine whether a group's overall performance improved or worsened from the pre-test to the post-test. Perceptual sensitivity was also measured in terms of response times (RT). As mentioned above, RTs were measured in milliseconds and recorded as the time between the offset of the sound stimuli and the depression of any key on the computer. RTs were of interest to assess the processing times required for particular contrasts and groups.

The final measurement used was A-prime. In order to calculate perceptual sensitivity, A-prime (A') scores were determined from the participants' correct and incorrect responses for each of the 7 consonant sound contrasts investigated (Werker and Tees, 1984; Guion et al., 2000; Baker et al., 2002; Flege and MacKay, 2004). According to Guion et al. (2000):

> The A' scores provide an unbiased measure of perceptual sensitivity by taking into account the responses to the different trial and the catch trials. An A' score of 1.0 indicates perfect discrimination of a contrast and an A' score of 0.5 or lower indicates insensitivity to a contrast. (p. 2718)

The calculation of A' values (Snodgrass and Corwin, 1988) are determined by 'hits' and 'false alarms'. The 'hit rate' (H) is the proportion of odd triads where the odd item was **correctly** selected (Triads that were determined 'odd' but had the wrong token chosen were considered incorrect and not 'hits'.). The 'false alarm rate' (FA) is the proportion of catch triads where an odd item was **incorrectly** selected. Table 3-5 demonstrates the classification of participant responses with regard to hits and false alarms.

Correct response	Possible Response	Classification
1 odd	**1 odd**	HIT
	2 odd	error
	3 odd	error
	catch	error
2 odd	1 odd	error
	2 odd	HIT
	3 odd	error
	catch	error
3 odd	1 odd	error
	2 odd	error
	3 odd	HIT
	catch	error
catch	**1 odd**	FALSE ALARM
	2 odd	FALSE ALARM
	3 odd	FALSE ALARM
	catch	correct

TABLE 3-5 Classification of possible participant responses in terms of hits and false alarms for the calculation of A' scores

Once the H and FA rates were calculated, their numeric relations were considered and then put in the appropriate A' formula. The A' formulae are as follows (Snodgrass and Corwin, 1988; Guion et al., 2000):

1. If $H>FA$ then $A' = 0.5 + [(H-FA)(1+H-FA)] / [4H(1-FA)]$
2. If $H=FA$ then $A' = 0.5$
3. If $H<FA$ then $A' = 0.5 - [(FA-H)(1+FA-H)] / [4FA(1-H)]$

For example, from the data given in Table 3-6 it is possible to derive PN01's pre-test A' score for the English-Mandarin [s-ts^h] contrast that is represented by the orthographic symbol 'c'.

GROUP	PART.	SOUND	PART. RESPONSE	CORRECT RESPONSE	ERROR CODE	REACTION TIME
pinyin pre-test	PN01	C*	1 odd	1 odd	C	834
pinyin pre-test	PN01	C	1 odd	1 odd	C	776
pinyin pre-test	PN01	C	2 odd	2 odd	C	789
pinyin pre-test	PN01	C	catch	2 odd	E	2494
pinyin pre-test	PN01	C	3 odd	3 odd	C	1056
pinyin pre-test	PN01	C	3 odd	3 odd	C	908
pinyin pre-test	PN01	C	catch	catch	C	976
pinyin pre-test	PN01	C	catch	catch	C	892
pinyin pre-test	PN01	C	catch	catch	C	988
pinyin pre-test	PN01	C	catch	catch	C	2347
pinyin pre-test	PN01	C	1 odd	catch	E	1070
pinyin pre-test	PN01	C	2 odd	catch	E	966

6 'odd' triads for the [s]-[ts^h]contrast
H = #of C/6

6 'catch' triads for the [s]-[ts^h] contrast
FA = #of E/6

* The shared orthographic symbol is used in the spreadsheet to denote the English-Mandarin contrast.

TABLE 3-6 Sample data spreadsheet for the pre-test responses of Participant 1 in the PN group (PN01) for the [s-ts^h] contrast

Recall (Table 3-4) that each contrast was assessed with 12 stimuli triads: 6 'odd' triads and 6 'catch' triads. According to the sample spreadsheet in Table 3-6, PN01 correctly identified 5 of the 6 odd triads which yielded a 'hit rate' of 0.83 (5÷6)[10], and PN01 incorrectly identified 2 of the 6 catch triads which yielded a 'false alarm rate' of 0.33 (2÷6). Therefore, H = 0.83 and FA = 0.33. Since H>FA, then A' is calculated according to the first formula (A' = 0.5 + [(H-FA)(1+H-FA)] / [4H(1-FA)]). In short, PN01's A' score for the [s-ts^h] contrast was 0.84. This A' score is well above the 0.5 sensitivity threshold and means that PN01 was highly sensitive to this contrast.

---

[10] The 'hit rate' is calculated by dividing the number of correct selections of odd items by the total number of odd triads for the contrast. Conversely, the 'false alarm rate' is calculated by dividing the number of incorrect selections of odd items for the catch triads by the total number of catch triads for the contrast.

In sum, because this dependent variable was evaluated in terms of the pre- and post-test ERs, RTs and A' scores, it was necessary to place perceptual sensitivity on an ***interval scale*** that calculates the mean ERs, RTs, and A' scores in relation to the 3 independent variables (orthographic medium, sound contrast, and perception test).

3.5.3 Statistical Analyses

Since the experimental design incorporated three independent variables (group, sound contrast, and test) and one dependent variable (perceptual sensitivity), 3-way repeated measures ANOVAs were conducted to calculate the significance. More specifically, 'group' was a between-subjects variable while 'sound contrast' and 'test' were within-subjects variables. Since 'sound contrast' and 'test' were within-subjects variables, the repeated measures were performed on these two independent variables where the *sound contrast* factor had seven levels (/s-tsʰ/, /z-ts/, /s-s̺/, /tʃ-tʂʰ/, /ʃ-ʂ/, /ɹ-z̢/, /h-x/) and the *test* factor had two levels (pre-test and post-test). Since the dependent variable was measured in terms of A', ER and RT, separate 3-way repeated measures ANOVAs were carried out for each measurement. The significance level was set at 0.05 such that any *p-value* less than 0.05 was considered statistically significant.

Chapter Four

RESULTS AND ANALYSES

"The human mind has developed into a remarkable seeker of patterns."
(Borden et al., 2003, p. 151)

Up to this point, the first half of this thesis, namely Chapters 1, 2 and 3, has been dedicated to providing the necessary theoretical and methodological backgrounds from which this current research has grown. The goal of the second half of this thesis is to report on and discuss the findings of the research. As the title suggests, Chapter 4's central focus is to report on results and the statistical analyses of the data. This chapter is broken down into two major sections. The first section provides the overall results of each orthographic group's performance in terms of error rates (ER), A-prime (A'), and response times (RT) with respect to the 7 target sound contrasts investigated in the perception tests. The second section delves deeper into the relationship between the sound contrasts themselves. In this section, the sound contrast data are presented and assessed in terms of the varying degrees of discrimination difficulty in the performance on individual contrasts.

4.1 Orthographic medium

Recall that the fundamental question this research seeks to answer is whether the orthographic medium within which learners were instructed played a role in the learners' perceptual sensitivities to native-nonnative sound contrasts. Specifically, did shared orthographic symbols facilitate or hinder perception of L1-L2 sound contrasts? From this question stemmed three hypotheses about the results of the data. The first hypothesis predicted that learners who learnt Mandarin via Pinyin, which shares orthographic symbols with English, would be led to equate a Mandarin (L2) phoneme with its English counterpart.

This equivalence classification (Flege, 1987, 1995) would be demonstrated by the PN participants having greater difficulty distinguishing the English-Mandarin sound contrasts. The second hypothesis predicted that learners who learnt Mandarin via Zhuyin, which is an entirely separate orthographic system, would not be thus affected and would be better able to make distinctions between a similar sounding L2 phoneme and its English counterpart. This ability to make distinctions would be demonstrated by the ZN participants being more successful at distinguishing the Mandarin from the English sounds. The third and final hypothesis predicted that learners who learnt Mandarin via no orthography would fall somewhere between the performance of the PN and ZN groups (see Chapter 1, Section 1-4).

Pre- and post-test error rates (ER), A-prime (A') scores, and response times (RT) were employed to test the three hypotheses. Based on the experimental design outlined in Chapter 3, 3-way repeated measures ANOVAs (3 orthographic groups x 7 sound contrasts x 2 tests) were conducted to determine the existence of any statistical significance. The repeated measures were performed on the final two independent variables. Significance (or lack thereof) would indicate if the type of orthographic instruction had a bearing on the classification of the new Mandarin sounds, and ultimately, the perceptual sensitivity necessary to distinguish between these L2 sounds and their entrenched English counterparts. The following sections provide the results and statistical analyses of the data.

4.1.1 Error rates (ER)

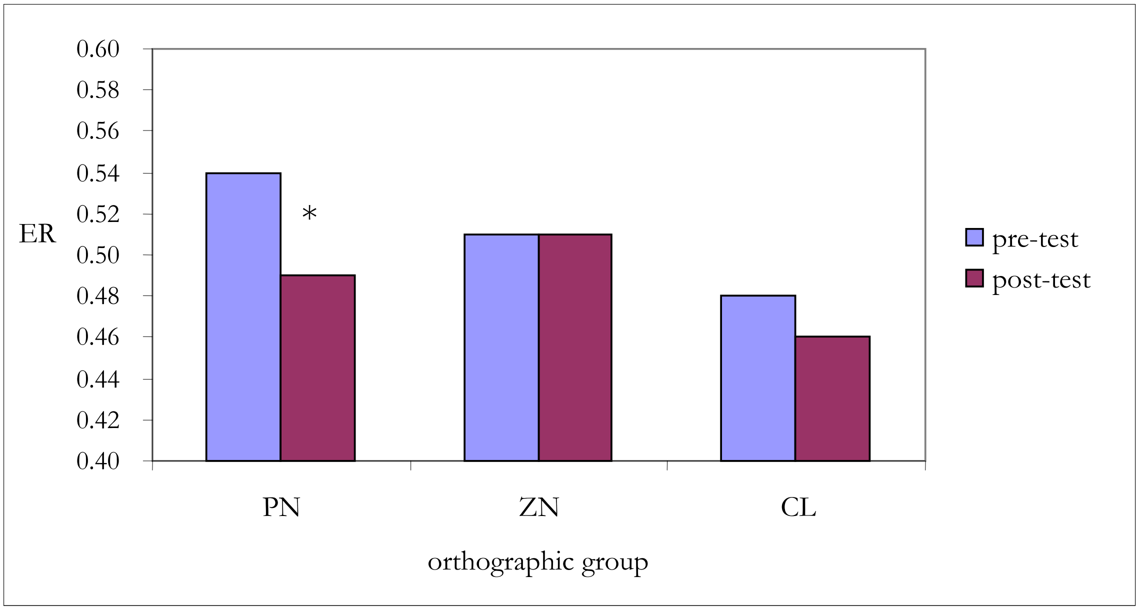
In Table 4-1, the highlighted column draws attention to the overall performance of each of the three experimental groups in terms of their ERs for both the pre- and post-tests. Also provided in this table are the mean ERs and standard deviations (SD) for each of the target sound contrasts.

Group	Test	Error Rates (SD) s-tsʰ	z-ts	s-ʂ	tʃ-tʂʰ	ʃ-ʂ	ɹ-ʐ	h-x	**AVERAGE**
PN	pre-	0.27(0.18)	0.27(0.14)	0.71(0.10)	0.73(0.14)	0.64(0.14)	0.53(0.18)	0.64(0.12)	**0.54**(0.06)
	post-	0.26(0.13)	0.18(0.15)	0.72(0.12)	0.60(0.17)	0.61(0.13)	0.45(0.12)	0.58(0.10)	**0.49**(0.08)
	difference	0.01*	0.09	-0.01	0.13	0.03	0.08	0.06	**0.05**
ZN	pre-	0.29(0.14)	0.25(0.12)	0.70(0.11)	0.59(0.14)	0.63(0.10)	0.55(0.16)	0.57(0.09)	**0.51**(0.05)
	post-	0.27(0.13)	0.29(0.11)	0.69(0.12)	0.67(0.11)	0.63(0.11)	0.51(0.13)	0.54(0.11)	**0.51**(0.06)
	difference	0.02	-0.04	0.01	-0.08	0	0.04	0.03	**0.00**
CL	pre-	0.27(0.15)	0.23(0.20)	0.65(0.15)	0.66(0.17)	0.52(018)	0.45(0.10)	0.60(0.12)	**0.48**(0.09)
	post-	0.23(0.14)	0.16(0.18)	0.65(0.17)	0.62(0.12)	0.53(0.18)	0.49(0.09)	0.56(0.17)	**0.46**(0.09)
	difference	0.04	0.07	0	0.04	-0.01	-0.04	0.04	**0.02**

* The difference was calculated by subtracting the post-test ERs from the pre-test ERs. Positive numbers indicate an improvement while negative numbers indicate a worsening from the pre- to the post-test.

TABLE 4-1 Mean ERs for Pinyin (PN), Zhuyin (ZN), and Control (CL) pre- and post-tests across each of the 7 sound contrasts

In addition to the numeric data given in Table 4-1, Figure 4-1 below offers a more visually assessable representation of the three groups' overall performances with respect to their pre- and post-tests as well as to one another. The height of the bar in the graph indicates the percentage of errors for each group and each test.



* indicates that the difference between the pre- and post-test scores is statistically significant at p<0.05

FIGURE 4-1 Mean pre- and post-test ERs for all three orthographic groups

From Table 4-1 and Figure 4-1, two observations are apparent. First, the pre-test error bars in Figure 4-1 indicate that the three experimental groups did not start out on an equal footing. Rather, each group was characterized by a different mean ER in the pre-test. The fact that each group started out with a different overall ER means that there was no standard base line against which all three groups' post-test scores could be measured. Therefore, post-training performance could only be measured by the difference between the pre- and post-test scores for each of the three experimental groups. This comparison between the pre- and post-tests leads to the second observation that can be gleaned from the above data. Contrary to the first hypothesis, the Pinyin (PN) group improved the most from 54% to 49% between the pre- and post-tests, a difference of 5%, and contrary to the second hypothesis, the Zhuyin (ZN) group did not improve at all from the pre-test to the post-test. The only hypothesis that appears to be partially supported is the third hypothesis that predicted the Control group performance would be between the PN and ZN groups. The Control (CL) improved slightly from 48% to 46%, a difference of 2%.

While the descriptive data in Table 4-1 and Figure 4-1 suggest that there is a difference between the groups with regards to perceptual performance, the 3-way repeated measures analysis of variance demonstrated otherwise. The relationship between the orthographic group and the test scores was not statistically significant, F (2, 29) = 2.52, p = 0.098. In other words, the statistical analysis revealed that there was no statistically significant difference between the groups' overall perceptual performances. The analysis yielded a main effect for test (F (1, 29) = 5.07, p = 0.03) indicating that the average ER in the post-test condition (M = 0.49, SD = 0.08) was significantly lower than in the pre-test condition (M = 0.51, SD = 0.07). The interaction effect of orthographic group was not significant, F (2, 29) = 1.171, p = 0.324. Individual 2-way repeated measures ANOVAs (one for each orthographic group) revealed that only the difference between the pre- and post-test ERs for the PN group to be statistically significant (F (1, 10) = 6.631, p = 0.028) indicating that PN was the only group to improve significantly from the pre- to the post-test.

As well as assessing the overall perceptual performance of each orthographic group, it was important to investigate the ERs of each of the 7 target contrasts. The three graphs in Figure 4-2 represent the differences between the pre- and post-test ERs for each of the experimental groups where: a) are the pre- and post-test ERs for the Pinyin group, b) are the pre- and post-test ERs for the Zhuyin group, and c) are the pre- and post-test ERs for the Control group.

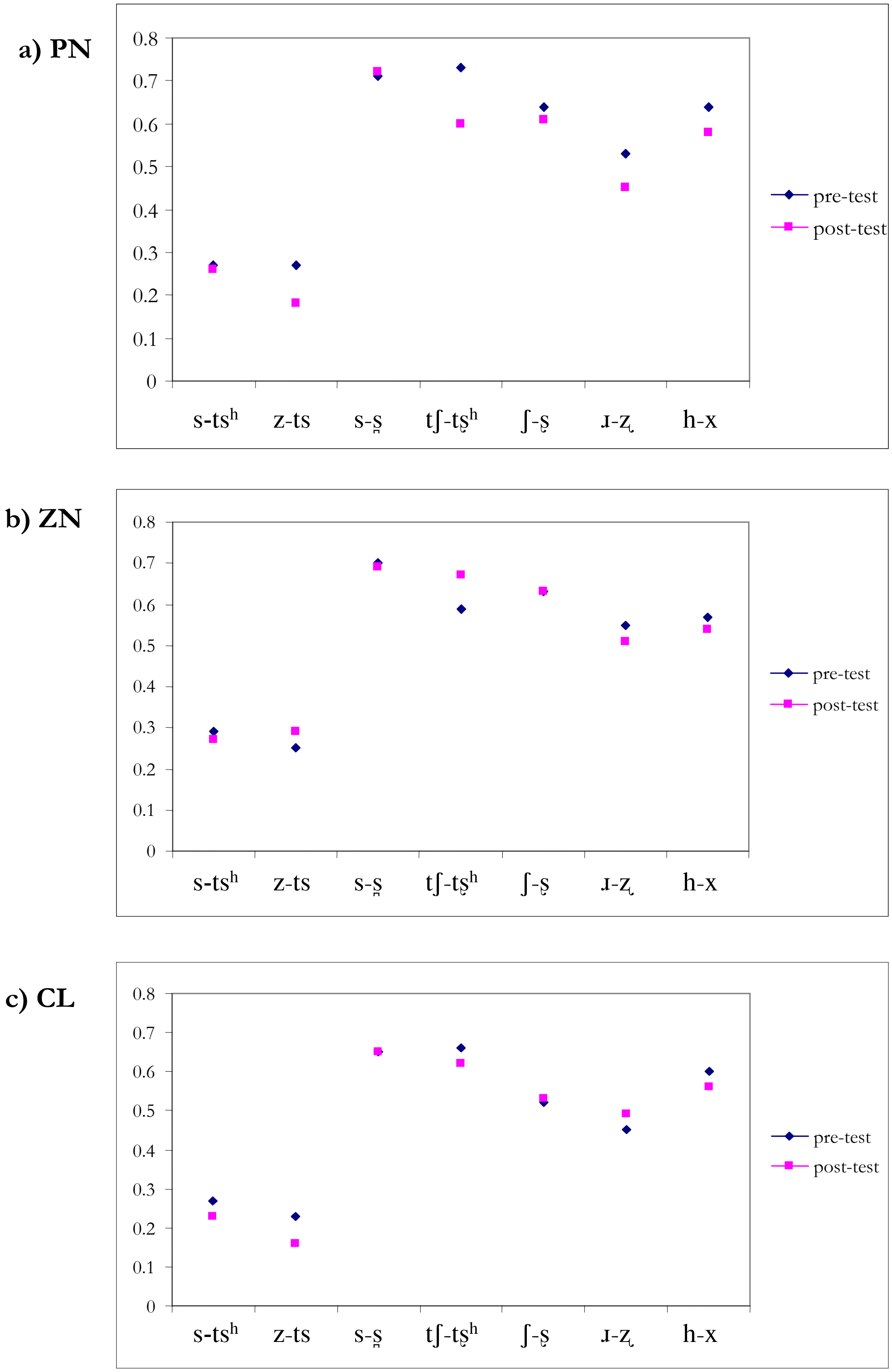


FIGURE 4-2 Mean pre- and post-test ERs across all 7 sound contrasts for the PN, ZN, and CL groups

Figure 4-2 demonstrates that all groups had similar pre- and post-test ER patterns. This figure also shows that the PN group improved in accuracy from the pre- to the post-test – by 1, 9, 0, 12, 2, 9 and 6 percent (see Table 4-1) respectively – in all but one of the sound contrasts /s-s̪/. ZN only improved in accuracy for four of the contrasts; /s-tsʰ/, / s-s̪/, /ɹ-ʐ/, and /h-x/ – by 2, 1, 4, and 3 percent, respectively. Similarly, CL only improved in four of the 7 contrasts; /s-tsʰ/, /z-ts/, /tʃ-tʂʰ/, and /h-x/ – by 5, 7, 4, and 4 percent, respectively. As shown in Table 4-1, the numbers demonstrate that although ZN did improve on some of the contrasts, their increase in accuracy (i.e. lower ERs) was smaller than the increases for PN and CL for each of the target contrasts. The only sound contrasts that all three groups improved in were /s-tsʰ/ and /h-x/. In light of the statistical results (i.e. the lack of significance) for the 'orthographic group x test' analysis, it is not surprising then that there was also no significance found for 'orthographic group x sound x test' (F (12, 174) = 1.05, p = 0.405). However, there was a significant main effect for sound contrast (F (6, 174) = 108.78, p<0.001) indicating that there was significant difference in the mean ERs for at least some of the target sound contrasts (see section 4.2).

4.1.2 A-prime (A')

The orthographic groups' ERs were investigated for the native-nonnative contrasts because ERs are the most commonly (and traditionally) used measure of participant performance. However, previous research has argued that ERs may not, in fact, be accurate measures for perceptual sensitivity as they often do not take into account participant bias. Therefore, recent research projects (e.g. Flege and McKay, 2004; Guion et al., 2000; Baker et al., 2002) rejected ERs as a performance measure in favour of a more accurate measure called A-prime (or A'; see Chapter 3, section 3.5 for a more in-depth discussion of A') in order to account for participant bias. With this concern in mind, A' scores were also calculated for each group with regards to each of the 7 target contrasts. Table 4-2 provides the pre- and post-test A' scores for the 7 contrasts for PN, ZN and CL. The standard deviations are provided in brackets. The highlighted cells isolate the contrasts on which each group increased in sensitivity.

		A-prime (SD)						
Group		s-ts^h	z-ts	s-ʂ	tʃ-tʂ^h	ʃ-ʂ	ɹ-ʐ	h-x
PN	pre-	**0.77**(0.19)◆	**0.80**(0.17)	0.22(0.13)	0.19(0.13)	0.33(0.15)	0.44(0.21)	0.34(0.16)
	post-	**0.77**(0.17)	**0.85**(0.14)	0.21(0.15)	0.40(0.23)	0.37(0.19)	**0.55**(0.15)	0.38(0.15)
	difference‡	0	0.05	-0.01	0.21	0.04	0.11	0.04
ZN	pre-	**0.79**(0.19)	**0.84**(0.11)	0.23(0.14)	0.38(0.18)	0.28(0.14)	0.43(0.20)	0.41(0.13)
	post-	**0.77**(0.16)	**0.77**(0.12)	0.25(0.14)	0.27(0.14)	0.30(0.14)	0.48(0.15)	0.45(0.14)
	difference	-0.02	-0.07	0.02	-0.11	0.02	0.05	0.07
CL	pre-	**0.81**(0.17)	**0.80**(0.25)	0.31(0.18)	0.31(0.24)	0.47(0.22)	**0.56**(0.14)	0.35(0.18)
	post-	**0.83**(0.15)	**0.89**(0.16)	0.32(0.21)	0.35(0.18)	0.46(0.24)	0.49(0.13)	0.45(0.22)
	difference	0.02	0.09	0.01	0.04	-0.01	-0.07	0.09

◆ Boldface type highlights the values that demonstrate a degree of sensitivity to a sound contrast.
‡ The difference was calculated by subtracting the pre-test A' value from the post-test A' value. Positive numbers indicate an improvement while negative numbers indicate a worsening from the pre- to the post-test.

TABLE 4-2 Mean A' scores for Pinyin (PN), Zhuyin (ZN), and Control (CL) pre- and post-tests across each of the 7 sound contrasts

Since A' is a pair-wise measure that assesses perceptual sensitivity of a single contrast, it is not appropriate to give an overall mean A' for each group as sensitivity for each contrast varies significantly (see Section 4.2.1).

Figure 4-3, like Figure 4-2, represents the differences between the pre- and post-test. In this case, the values are represented in terms of A'. Again, it is possible to see that each group had similar perceptual sensitivity patterns across all 7 target contrasts, patterns that inversely mirror the ER patterns for the orthographic groups.

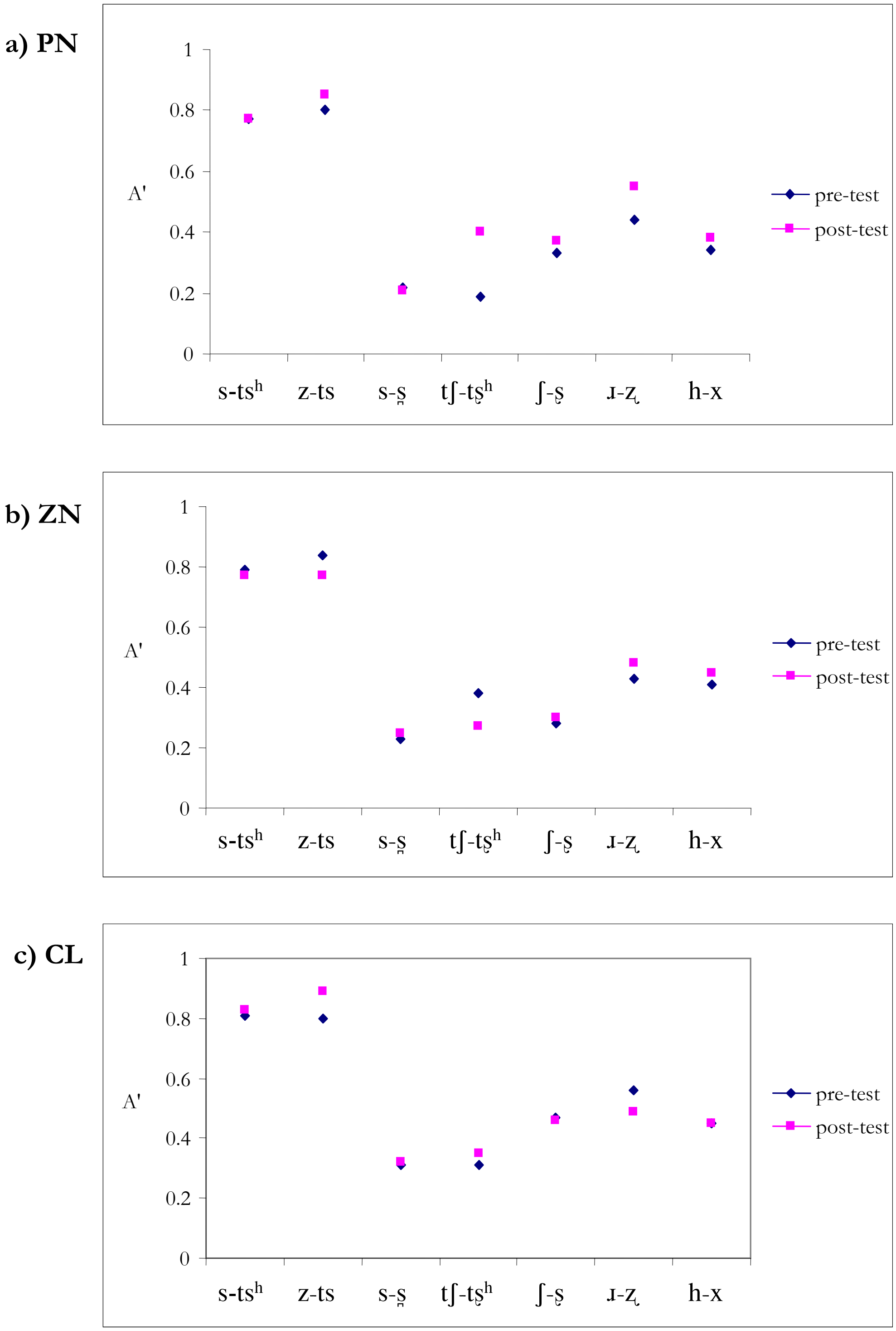


FIGURE 4-3 Mean pre- and post-test A' scores across all 7 sound contrasts for the PN, ZN, and CL groups

As shown in Table 4-2 and Figure 4-3, on the whole, the PN group appears to have increased its perceptual sensitivity for each contrast to a greater degree than the ZN and CL groups. These A' scores were also subjected to a 3-way repeated measures ANOVA that revealed a non-significant relationship between the orthographic group, the sound contrast and the test, F (12, 174) = 1.532, p = 0.117. Unlike the main effect of test for ERs, the main effect of test for A' merely approached significance (F (1, 29) = 3.62, p = 0.067). The main effect for sound contrast was significant (F (6, 174) = 106.94, p<0.001) such that there was a significant difference between the mean A' scores of at least some of the contrasts (see section 4.2). The interaction effect of orthographic group was non-significant, F (2, 29) = 1.55, p = 0.23.

4.1.3 Response times (RT)

Response Times (RTs) provided a third opportunity to observe possible performance differences between the three orthographic groups. Figure 4-4 highlights the mean pre- and post-test RTs for each group.

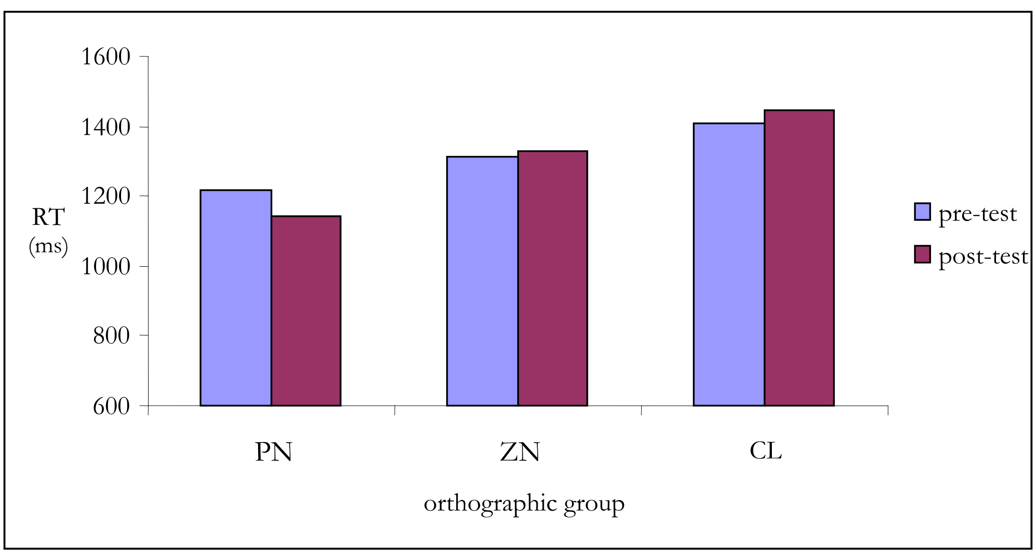


FIGURE 4-4 Mean pre- and post-test RTs for all three orthographic groups

This figure shows that, on the whole, the PN group's RT were faster in both test conditions. ZN was the second fastest group and CL was the slowest. As in the cases with the ERs and A', the RTs were subjected to the 3-way repeated measures ANOVA. Neither the 'orthographic group x test' analysis (F (2, 29) = 0.858, p = 0.434) nor the 'orthographic group x sound x test' (F (12, 174) = 0.889, p = 0.560) were found to be statistically significant. Also, both the main effect of test (F (1, 29) < 0.001, p = 0.990) and the interaction effect of orthographic group (F (2, 29) = 1.70, p = 0.200) were not significant.

4.1.4 Summary

Three 3-way repeated measures ANOVAs were conducted to assess the perceptual performances of the PN, ZN, CL groups in terms of three measures – ERs, A', and RT – of the dependent variable. As a whole, the data yielded four findings:

1. The perceptual performances of three groups did NOT differ significantly from each other in any of the three measures of the dependent variable.
2. All three dependent measures indicated that the interaction effect of orthographic group was non-significant.
3. The analysis of ERs indicated a main effect for test, while the analyses for A' and RT did not.
4. All three dependent measures suggest significant differences exist between at least some of the 7 target contrasts themselves (see section 4.2).

In sum, the above results do not support the hypotheses formulated for this research. In fact, these findings suggest that the type of orthographic instruction may have little influence perceptual sensitivity to native-nonnative sound contrasts (at least not during the very beginning stages of Mandarin language instruction).

However, it is important to note that although there were no statistically significant differences between the groups, a performance trend did emerge. From the data presented in the previous tables and figures, it does appear that the PN group's improved performance from the pre- to the post-test is generally better than the other two groups. That is, the PN participants improved their accuracy and sensitivity to the sound contrasts as well as increased their speed in responding to the token stimuli. By contrast, the ZN participants failed to improve their accuracy (0% change in performance) and sensitivity (They only improved on 4 of the 7 contrasts.) and were slightly slower in responding in the post test. Also, the CL participants did improve in terms of their accuracy (2%) and sensitivity (5 of the 7 contrasts) – although not a much as the PN group – but they were slower in responding to the stimuli in the post-test. In short, the Pinyin group was the only group to consistently improve on all three measures.

4.2 Sound contrast

The data and results in the previous section have demonstrated that there were no significant differences between each of the three experimental orthographic groups in perceptual performance in either ERs, A' or RTs. Because of this result, it is possible to collapse all three groups into one group and from there look deeper into the differences between the sound contrasts themselves. As with the previous section, the perceptual sensitivities to specific contrasts have been assessed according to the 3 measures of ER, A' and RT.

4.2.1 Error rates (ER)

Now that the data from each of the orthographic groups has been merged, this subsection presents the participants' perceptual performances for each of the individual sound contrasts in terms of the first measure, error rates. The error rates gage the participants' levels of accuracy in distinguishing the L2 sound (Mandarin) from the L1 sounds (English). Figure 4-5 presents the pre- and post-test ER data for each contrast.

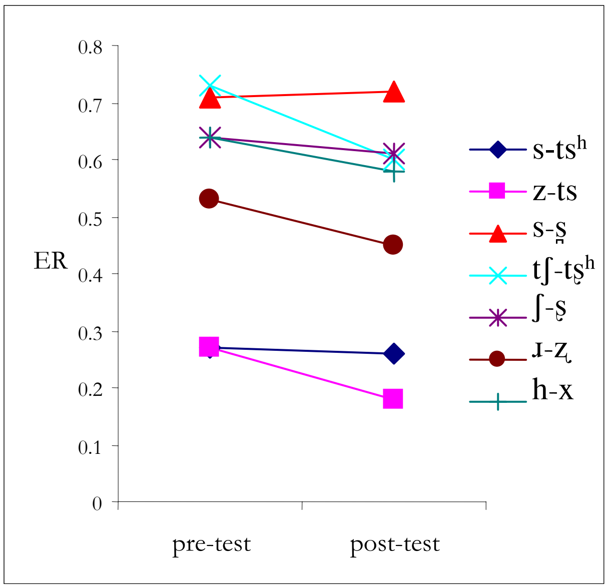


FIGURE 4-5 Mean pre- and post-test ERs for each of the target contrasts

The data here show that the participants' accuracy in distinguishing the L2 from the L1 varied across the sound contrasts. Figure 4-5 clearly illustrates that some contrasts were much more easily distinguished than others. Specifically, the contrasts /s-tsʰ/ and /z-ts/ were the most accurately distinguished. The ERs for these two were lower than 30%. The participants performed around the 50% level for the contrast /ɹ-ʐ/. Finally, the participants had difficulty in distinguishing the other four contrasts. The ERs were higher than 60%.

4.2.2 A-prime (A')

As mentioned above, since there were no significant differences between the orthographic groups, the three experimental groups were collapsed into one group thereby removing one of the three independent variables, orthographic group. The statistical analysis also revealed that there was no statistically significant relationship between the A' difference of the sound contrasts and the pre- and post-test (F (6, 174) = 0.311, p = 0.931). Again, this result allowed the merging of the test results to remove another independent variable, test. The remaining independent variable to be analysed was sound contrast. This subsection presents the data and analyses for the A' values of the individual sound contrasts. In Figure 4-6, the overall mean A' scores are plotted for each of the 7 target contrasts.

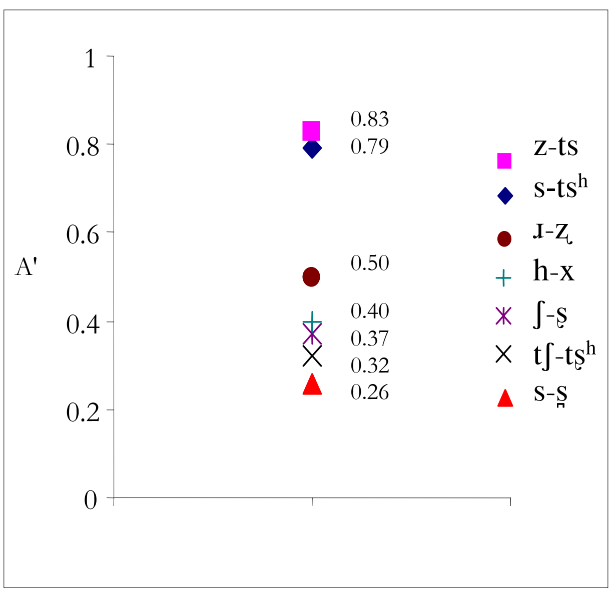


FIGURE 4-6 Mean A' scores for each of the 7 target sound contrasts

Like the ER results, this figure demonstrates that the participants were not equally sensitive (or insensitive) to all of the contrasts. Some contrasts were clearly more easily distinguishable than others. As shown, the contrasts fall along a 'sensitivity' continuum where participants were the most sensitive to the /z-ts/ contrast (0.83), the most insensitive to the /s-s̺/ contrast (0.26), and the other 5 contrasts fell somewhere in between.

With these glaring descriptive differences in perceptual sensitivity between contrasts, the obvious question is: How to they statistically compare to one another? Table 4-3 provides the differences in A' scores between each contrast where a single asterisk represents a significant difference ($p<0.05$) and double asterisks represent a highly significant difference ($p<0.001$). The highlighted cells illustrate those contrasts that differ significantly from one another.

	s-tsʰ	**z-ts**	**s-s̺**	**tʃ-tʂʰ**	**ʃ-ʂ**	**ɹ-ʐ**	**h-x**
s-tsʰ		-0.04	0.53**	0.47**	0.42**	0.29**	0.39**
z-ts	0.04		0.57**	0.51**	0.46**	0.33**	0.43**
s-s̺	-0.53**	-0.57**		-0.06	-0.11*	-0.27**	-0.14*
tʃ-tʂʰ	-0.47**	-0.51**	0.06		-0.06	-0.18**	-0.08
ʃ-ʂ	-.0.42**	-0.46**	0.11*	0.06		-0.13*	-0.03
ɹ-ʐ	-0.29**	-0.33**	0.27**	0.18**	0.13*		0.10
h-x	-0.39**	-0.43**	0.14*	0.08	0.03	-0.10	

*p<0.05
**p<0.001

TABLE 4-3 Mean A' differences between each of the 7 target sound contrasts

This table can be summed up as follows: the A' value(s) of

- the contrasts **/s-tsʰ/** and **/z-ts/** exhibit no statistically significant difference from each other but the two contrasts do exhibit highly significant differences from the other contrasts.
- the contrast **/ɹ-ʐ/** does not exhibit a significant difference from /**h-x**/. It does, however, exhibit a significant difference from /ʃ-ʂ/ and highly significant differences from all the other contrasts.

- the contrast **/h-x/** does not exhibit significant differences from the contrasts /tʃ-tʂʰ/, /ʃ-ʂ/, and /**ɹ-ʐ**/. It does exhibit a significant difference from /**s-s̺**/ and highly significant differences from /s-tsʰ/ and /z-ts/.
- the contrast **/s-s̺/** does not exhibit a significant difference from /tʃ-tʂʰ/ but does exhibit significant differences from /ʃ-ʂ/ and /**h-x**/ and highly significant differences from /**ɹ-ʐ**/, /s-tsʰ/ and /z-ts/.
- the contrast **/tʃ-tʂʰ/** does not exhibit significant differences from /ʃ-ʂ/, /s-s̺/, and /**h-x**/ but does exhibit highly significant differences from /**ɹ-ʐ**/, /s-tsʰ/ and /z-ts/.
- the contrast **/ʃ-ʂ/** does not exhibit significant differences from /tʃ-tʂʰ/ and /**h-x**/ but does exhibit significant differences from /s-s̺/ and /**ɹ-ʐ**/ as well as highly significant differences from /s-tsʰ/ and /z-ts/.

The analysis of the results points to the existence of three significantly different groups of perceptual sensitivity within which the contrasts fall. In order for a contrast to fall within a sensitivity category, it must adhere to two criteria. First, the contrast must be above (or below) the identified A' value. Second, it must be significantly different from the contrasts in the other sensitivity categories. For illustrative purposes, these three statistically identified groups have been defined below:

1. ***Highly Sensitive Contrasts*** – A' values are higher than 0.75.
2. ***Slightly Sensitive Contrasts*** – A' values are around the 0.50 mark.
3. ***Insensitive Contrasts*** – A' values are lower than 0.50[11].

[11] According to the A' literature (ex. Guion et al., 2000), participants are considered to be insensitive to contrasts that fall below 0.5. But the above significant differences between the insensitive contrasts /s-s̺/ and /ʃ-ʂ/ suggest that there may also be varying degrees of insensitivity. This question, however, is beyond the scope of this study and should be addressed in future research.

Figure 4-7 offers a representation of the statistically identified sensitivity groups along the sensitivity continuum and the categorization of the contrasts.

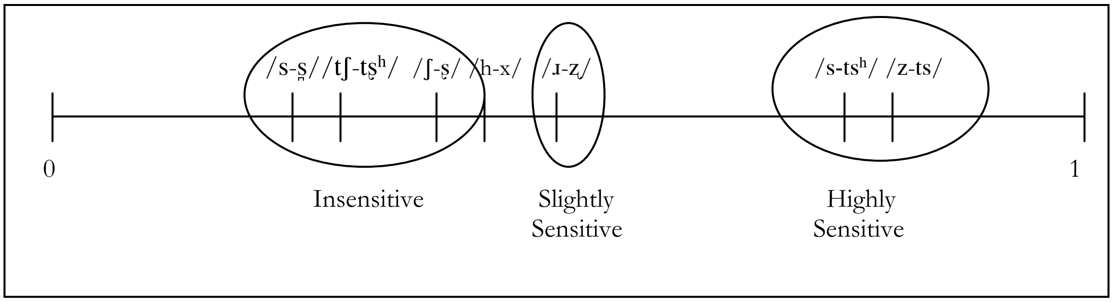


FIGURE 4-7 Three sensitivity groups identified according to statistically significant differences among the 7 target contrasts

All the contrasts fall into categories of sensitivity that are significantly different from one another; the only exception being the /h-x/ contrast. Interestingly, this contrast appears to hover on the cusp between the 'slightly sensitive' group and the 'insensitive' group as the A' differences between it and the two nearest contrasts (on either side) are not statistically significant. While the contrast /h-x/ appears to pattern with the insensitive group, it only satisfies one of the 2 aforementioned criteria, namely, having an A' score below 0.5. However, it is not statistically different from the slightly sensitive contrast /ɹ-ʐ/. This is why /h-x/ has not explicitly been categorized with the insensitive group.

4.2.3 Response times (RT)

The third and final measure of perceptual performance was response times. RTs assess the speed with which the participants responded to the stimuli. Like the results with A', the lack of a statistically significant relationship between the sound contrasts and the pre- and post-test (F (6, 174) = 1.493, p = 0.183) allowed for the merger of the pre- and post-test RTs. This leaves only one independent variable ('sound contrast') to be analysed in terms of response times. Figure 4-8 below demonstrates the speed with which the participants distinguished the contrast pairs.

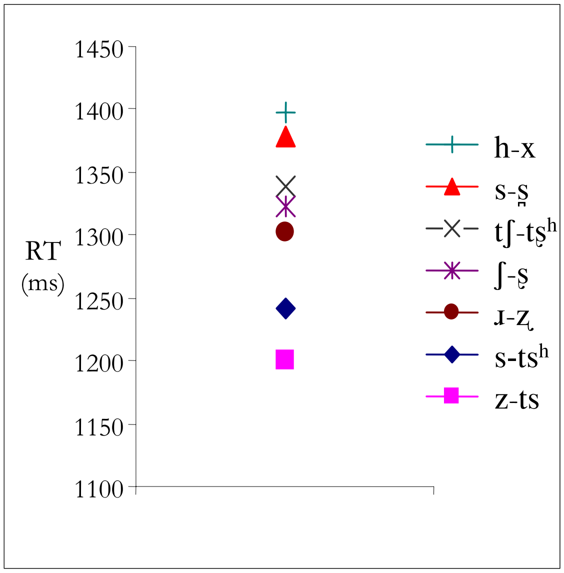


FIGURE 4-8 Mean RTs for each of the target contrasts

Again, it is possible to see that the participants performed differently according to the contrast. They were obviously faster at distinguishing some contrasts than others. Specifically, participants were the quickest at distinguishing the contrast /z-ts/ (1201 ms) and the slowest at distinguishing /h-x/ (1398 ms). All other contrasts fell somewhere along a continuum between these two speeds.

Table 4-4 provides the statistically relevant pair-wise comparisons of contrasts. Those contrasts that differ significantly in RT are highlighted.

	s-tsʰ	z-ts	s-s̺	tʃ-ts̺ʰ	ʃ-s̺	ɹ-z̺	h-x
s-tsʰ		39.6	-157.0*	-98.1	-82.0	-61.2	-157.5*
z-ts	-39.6		-196.6*	-137.8	-121.6	-100.9	-197.2*
s-s̺	157.0*	196.6*		58.8	75.0	95.8	-0.5
tʃ-ts̺ʰ	98.1	137.8	-58.8		16.2	36.9	-59.3
ʃ-s̺	82.0	121.6	-75.0	-16.2		20.8	-75.5
ɹ-z̺	61.2	100.9	-95.8	-36.9	-20.8		-96.2
h-x	157.5*	197.2*	0.5	59.3	75.5	96.2	

*$p<0.05$

TABLE 4-4 Mean RT (ms) differences between each of the 7 target sound contrasts

Unlike the statistically defined sensitivity groups found for A', the RTs were less well defined. The two contrasts /s-tsʰ/ and /z-ts/ are statistically significant from the two contrasts /h-x/ and /s-s̺/ while the RTs for /ɹ-z̺/, /tʃ-ts̺ʰ/, and /ʃ-s̺/ are not significantly different from any of the other contrasts. Generally one would expect to observe the response times decrease as the A' scores increase but there is no clear correlation between A' scores and response times. The only correlation with the RT data appears to be with the 2 highly sensitive contrasts where they are distinguished much faster than the more confusable sounds. As for the other five contrasts there is no distinct relationship between their A' sensitivities and their response times.

4.2.4 *Summary*

In sum, this section explored the participants' sensitivities to the target contrast in terms of the relationship between each of the contrasts in the perception test triads. The data and statistical analyses support the hypotheses generated in Chapter One, Section 1.4.2. The hypotheses were tested in terms of accuracy (ER), sensitivity (A') and speed (RT). That is, not all contrasts were equally (un)distinguishable. For the most part, the contrasts appear to fall into one of three sensitivity categories: 1) the highly sensitive category contains the contrasts /z-ts/ and /s-ts^{h}/, 2) the slightly sensitive category contains the contrast /ɹ-ʐ/, and 3) the insensitive category contains the contrasts /s-ʂ/, /ʃ-ʂ/, and /tʃ-tʂ^{h}/. Only the contrast /h-x/ appears difficult to categorize as it borders on the cusp of the insensitive and slightly sensitive group.

4.3 General summary

The goal of Chapter 4 has been to present and analyse the data collected for this research project that was designed to investigate the effect of shared orthography on sound perception in L2 acquisition. Three-way repeated measures ANOVAs were conducted to analyse this data. The results did not support the initial hypotheses formulated with respect to orthographic influence as outlined in Chapter 1. Rather, the results indicated that there were no significant differences in the overall perceptual performances of the three experimental groups. The results did, however, point to an observable trend where the Pinyin group did appear to improve somewhat more than the Zhuyin and Control groups in accuracy (ERs), sensitivity (A') and speed (RTs). But since the statistical analyses revealed a non-significant effect between experimental groups, the trend observed for the PN group is inconclusive and requires further study.

While the analyses did not establish the existence of statistically significant differences between the perceptual performances of the experimental groups, they did reveal significant differences for the participants' sensitivities among contrasts. These results do support the hypotheses formulated with respect to the degrees of perceptual sensitivity. Between contrasts, the participants varied in degree of sensitivity to each of the target contrasts where, based on the A' scores, the contrasts could be categorized as either; highly sensitive, slightly sensitive or insensitive. As the ERs show, the participants' accuracies in distinguishing contrasts inversely mirror their sensitivities to the contrasts. While the data for the RTs is not as clear as the ER and A' data, the RTs do demonstrate that the participants' responses to the highly sensitive group of contrasts were faster than for the other two groups of contrasts.

The data and results reported in this chapter were unexpected but it does not follow that they must then also be unwelcome. Instead, they have sparked numerous interesting questions about the nature of experimental design, data collection, language training, performance measures and acoustic saliency. Some of these questions include:

1. Is shared orthography facilitative at the beginning stages of learning?
2. How long do learners need to form sound-symbol connections?
3. What is the minimum amount of time required for experimental training to be effective?
4. Why are some of the contrasts easier to distinguish than others?
5. How are the Mandarin sounds mapped onto the native English phonological system?

These questions as well as other issues will be explored and discussed at length in the next chapter.

Chapter Five

DISCUSSION

"One might liken the possession of an alphabetic code to a virus. This virus infects all speech processing […] Language is never the same again." (Frith, 1998, p. 1011)

In Chapter Four, the data and analyses isolated a number of interesting findings. These findings, in turn, raise many interesting questions about the effect of shared orthography on second language acquisition and the perception of individual native-nonnative sound contrasts. This chapter is organized into two major sections that address and discuss these questions. The first section explores why there were no significant differences in perceptual performances of the three experimental groups from the stand points of the L1 and L2 orthographic codes. The second section discusses the varying degrees of sensitivity to the target contrasts by considering the acoustic saliency of the differences between the sounds in the contrast pairs. This section also discusses the how the participants may have mapped the target sounds onto their L1 phonological system in reference to the current models of speech processing, namely Flege's Speech Learning Model (1987, 1995) and Best's Perceptual Assimilation Model (1995, 2001).

5.1 Perceptual performance of orthographic training groups

This study began with four research questions designed to investigate the primary purpose of this study, namely, whether learning a second language with a shared orthography (i.e. the Pinyin group) differed from learning Mandarin with non-shared orthography (i.e. Zhuyin and Control groups) in the perception of English-Mandarin sound contrasts. The assumption was that if the results demonstrated performance differences between the three experimental groups, these differences could be attributed to an orthographic influence on the learners' categorizations of the L2 phonemes in relation to the L1 system. That is, the L2 categorizations may be either facilitated or hampered by the orthographic instruction medium.

In order to examine this fundamental issue, native Canadian English speakers, who wanted to learn Mandarin, were assigned to one of the three experimental groups (either PN, ZN, or CL). The data were obtained from the participants through a series of experimental tasks – pre-test > training phase > post-test. In the pre- and post-tests, the participants were tested on their abilities to distinguish Mandarin phonemes from English phonemes in an oddity discrimination task. Between the two tests, the participants underwent a short training phase where they learnt Mandarin via the orthographic medium assigned to their group. Participant performance was determined by calculating the differences between the pre- and post-test results for the participants in each group and comparing these differences across the three groups. Perceptual performance was measured in terms of error rates (ER), a-prime (A') and response times (RT).

This study was developed based on the premise that at least two factors influence the categorization and perception of the new L2 sound system. The first factor is the nature of the L2 sounds themselves in relation to the entrenched L1 sound system and the second factor is the way the L2 sounds are represented in the orthographic instructional medium. According to the hypotheses outlined in Chapter One, the influence of the first factor on perceptual distinction for native-nonnative sound contrasts would be mediated by the influence of the second factor. That is, for the shared orthographic group, the influence of the second factor would be reflected by greater tendencies to equate the L2 sounds with the L1 sounds. On the other hand, the influence of the second factor on the first factor for the other two groups (the non-shared orthographic groups) would be reflected by greater tendencies to distinguish between the L2 and L1 sounds.

The results of this experiment did not confirm the hypotheses. Rather, the repeated measures ANOVA analyses of the results indicated that there were no significant differences in the perceptual performances of the three groups for any of the three measures for the dependent variable. One interpretation of these results could be that the choice of orthographic medium used in L2 instruction has little bearing on L2 sound categorization. In other words, the use of either the L1 or the L2 orthography does not matter in the acquisition of the L2 sound system. This interpretation, however, is far too simplistic.

A more plausible alternative interpretation is that potential facilitative effects of using a separate orthography in second language learning were neutralized by three important factors. First, the strength of the established L1 orthographic systems, at least during the beginning stages of learning, overrides the facilitative potential of a new orthographic system. Second, the cognitive demands placed on the participants were unequal between groups. Finally, the development of new symbol-sound associations requires a longer, more intensive training program than the training program employed in this project to effectively reduce the influence of the established L1 orthographic system. These three factors exist in a hierarchical relationship where the second factor is caused by the first and the third factor is, in turn, caused by the second and consequently the first. The following subsections address these factors more in-depth.

5.1.1 L1 orthographic code

Based on the existing literature reviewed in Section 2.6 of Chapter Two, it is not unreasonable to assume when an L2 phoneme is represented by a grapheme that is already associated with an L1 phoneme, the chances of a listener perceiving the L2 phoneme as an

L1 phoneme would be increased. Olson (1996) argues that "the writing system provides a model in the form of a set of categories in terms of which speech sounds are represented and thereby brings those aspects of speech into consciousness" (p. 95). Burnham's research (2003) suggests that once children learn an alphabetic script the "incoming sounds are best assigned to orthographically defined categories" where orthographic categorization suppresses the perception of nonnative sounds (p. 600). In other words, L2 sounds are further attenuated with orthographic knowledge. With this consideration in mind, it also seems logical then that representing different L1-L2 sounds with different graphemes may facilitate the perception of nonnative sounds rather than suppress them.

Yet the research here does not appear to support these assumptions. The statistical analyses revealed that the performances of the three orthographic groups did not differ significantly. The pressing question remains: Why? Part of the answer to this question may lie in the entrenchment of the orthographic code. That is, one possible factor leading to these results could be that regardless of the instructional medium present in the language classroom, native English learners learn Mandarin sounds via their native orthography. In other words, even without explicit instruction in the shared alphabetic system for the ZN and CL groups, it is entirely possible that the L1 orthographic code was still present in the language classroom.

Prior to a discussion on the ramifications of an entrenched L1 orthographic code on L2 learning, it is worthwhile to review the research that has established the strength of the native orthographic code. Like the L1 phonological system, after its establishment in childhood, the L1 orthographic system is equally pervasive. At the outset of orthographic learning very strong connections are formed between the graphemes and the phonemes that make it virtually impossible to completely separate them from there on (Treiman and Cassar, 1997). Much research has accumulated to support the idea that orthographic knowledge permanently alters the way people categorize and perceive spoken language. As the quotation from Frith (1998) at the beginning of the chapter creatively describes, orthographic knowledge permeates the phonological system like a virus that irrevocably changes speech processing such that language can never be the same after. The strength of the relationship between the orthographic and phonological codes is generally acknowledged by the research community. Burnham (2003) claims that from the beginning of childhood acquisition of an alphabet, the new orthographic knowledge forces the reorganization of the L1 phonological system. This reorganization creates an interdependence between the two systems.

In a similar vein, Landerl et al. (1996) argue that orthography 'intrudes' on phoneme awareness tasks. They suggest that in adult readers, orthography and phonology are so closely connected that they are automatically co-activated. Apparently, orthographic knowledge is so intimately linked and such a prominent feature, that readers cannot avoid thinking about the graphemes even when specifically instructed not to. Phonological awareness is more than a matter of implicit linguistic knowledge; it also involves sorting sounds into the categories provided by the orthography (Olson, 1996). Ziegler, Muneaux and Grainger's (2003) work also points to the intimate relations between the orthographic and phonological systems. They claim that orthographic knowledge "provides an additional constraint in driving segmental restructuring" (p. 790). Perception is affected because orthographic knowledge is automatically activated with print which in turn provides information to the phonological system. In short, research has suggested that by triggering phonological restructuring and reorganization, the two systems are not mutually exclusive.

The orthographic code becomes inseparably connected to the phonological system it has altered.

Up until the mid 1990's, research on word recognition was primarily concerned with the effect of orthographic knowledge on visual representation. For example, the presence of visual orthographic representation of the word allows listeners to hear speech that is embedded in noise (Frost, 1988). Other studies demonstrated that sound-to-spelling and spelling-to-sound inconsistencies resulted in higher error rates and latencies in visual word recognition tasks (e.g. Glushko, 1979; Jared et al., 1990; Stone et al., 1997). So far the research reviewed has demonstrated that with visual representation, orthographic knowledge is co-activated with and affects the phonological system, but what about in the absence of visual representation? Does orthographic knowledge affect auditory processing alone? Interestingly, research has demonstrated that in the absence of visual stimulation, orthographic knowledge still heavily influences how sounds are perceived in the mind (Hallé et al., 2000; Ziegler and Ferrand, 1998; Ziegler et al., 2004).

Ziegler and his colleagues (1998, 2004) speculated that spelling-to-sound and sound-to-spelling inconsistencies would influence auditory word recognition in the same way as they influence visual word recognition. They discovered that even when only given auditory information, listeners had higher error rates and longer response times for words with phonological rimes that could be spelt in more than one way. These results led Ziegler and Ferrand (1998) to conclude that speech processing is not independent of written language. They suggest instead that speech processing is affected by orthography as a result of a "coupling between orthography and phonology that is functional in both visual and auditory word perception" (p. 686).

While most of the research investigating orthographic influence on speech perception is couched in word recognition studies, there is one study on orthographic influence of individual phoneme perception that stands out. Hallé et al. (2000) claim that listeners' orthographic knowledge of how words are spelt tells them what they *should* hear rather than the sounds actually present in speech. In an auditory phoneme detection task, native French speakers were more likely to report hearing a /b/ in French words like *absurde* than a /p/ although the surface representation of the word is [apsyrd]. According to Hallé et al., hearing a /b/ rather than a /p/ implies that listeners were unable to ignore the orthographic code and thus suggests its dominant influence over the speech signal.

The point illustrated with the research above is the profound tenacity of the L1 orthographic code. As mentioned above, due to the pervasiveness of the L1 orthographic code, it is highly likely that L1 orthographic knowledge was indirectly present in the non-shared orthographic groups. Recall, the ZN and CL groups were not instructed in an alphabetic medium and were instructed not to physically write any English alphabetic graphemes that would help them remember or associate the new Mandarin sounds they were learning. All three groups were also asked to try and refrain from thinking about the Mandarin sounds in terms of English letters. Still, anecdotal evidence from the participants in the training sessions indicated that disassociating themselves from thinking in terms of English letters was virtually impossible. There were a few occasions when a participant would slip and say, for example, 'It sounds like a *k*.' or 'It is an *h*.' While these occurrences were relatively few, they did demonstrate that L1 orthography was present in the participants' minds during the Mandarin lessons regardless of whether there was visual representation of the English alphabet or not, and thus suggests that the participants most likely learnt the Mandarin sounds via their native orthography.

This section has argued that the lack of differences between the groups may be attributed to a lack of ability to completely disassociate from the L1 orthographic code as suggested by the research on the symbiotic relationship between orthography and phonology (Treiman and Cassar, 1997; Burnham, 2003; Landerl et al., 1996; Ziegler et al., 2003) and orthography's influence on auditory perception (Hallé et al., 2000; Ziegler and Ferrand, 1998; Ziegler et al., 2004). Consequently, the continued reference to the L1 orthographic code in the language classrooms influenced the effectiveness of the instructional medium, especially that of the nonnative orthography, thereby equalizing the perceptual performances of all three groups.

Yet, the influence of the L1 orthographic code does not fully explain the results of the study. If the participants learnt Mandarin via their native orthography (whether instructed in it or not), the perceptual performances of the three groups should have been equal. However, while there were no significant differences between the groups, there was a trend that indicated that the PN group consistently performed better than the other two groups. This trend indicates that, in addition to the influence of the established L1 orthographic code, other factors must have had a role in the overall performances of the three groups. These factors, addressed below, are: 1) the cognitive load placed on each group and 2) the length of language training.

5.1.2 L2 orthographic code

When considering the aforementioned factors, there are a number of important questions that should be asked here. Was the task load of the ZN and CL groups comparable to the PN group? In other words, by not having a familiar orthographic system (i.e. a Roman alphabetic system) were the ZN and CL groups asked to undertake a more difficult task during the training phase of the experiment than the PN group? Was the training period sufficient for the ZN group to create new sound-symbol associations based on the L2 orthographic and phonological systems? More specifically, how long a training period do learners need to form new (and influential) sound-symbol connections in their L2?

Again, it is important to remember that although there were no significant differences in group performances, the shared orthographic group (i.e. PN) did tend to outperform the other two groups (i.e. ZN and CL). This appears to suggest that shared orthographic representation may have a facilitative effect, at least in the very beginning stages of learning Mandarin. However, this facilitative effect may simply be a consequence of a smaller demand on Pinyin learners. Strictly speaking, the PN group had a slight advantage because they were asked to accomplish less. For this group, the orthographic medium was already in place as it is shared with the L1. The PN participants were only required to learn a new phonemic inventory and adjust the existing orthography accordingly. Since the alphabetic tools were already there, the participants did not need to allot significant time and cognitive resources to orthographic learning. Rather, they could concentrate almost all of their energy and time to learning the sounds and their grapheme associations.

Conversely, in retrospect, it appears that the ZN participants had considerably higher demands placed on them which in turn further diminished the facilitative potential of learning with a separate orthography. Not only did they learn a new phonemic inventory but they also learnt an entirely new orthographic system. It may be the case that as the ZN participants were struggling to learn both new L2 systems, the participants' learning abilities were seriously handicapped in that their cognitive processes were overextended (as compared to the PN participants) by having to learn two systems as opposed to one. These participants more than likely did not have any time to notice the abstract differences between the L1-L2 contrasts as they were too focused on trying to learn the new orthography and the sounds attached to each new grapheme. Thus the facilitative potential of learning with a different orthography may have been overruled by conflicting orthographies and thereby preventing any improvement from the pre-test to the post-test. A similar argument can be made for the CL group. Although they did not have another orthography to contend with, the lack of visual orthographic cues may have also made learning Mandarin phonemes more challenging as participants may have found remembering the new sounds more difficult without any visual stimuli.

If the L2 orthographic code is a hindrance at the beginning of learning because of the higher cognitive demand placed on learners, is there a point where the L2 orthographic code no longer hampers L2 learning? How much training, exposure and practice are necessary for an orthographic code to influence speech processing? Research suggests that orthography affects sound perception when children learn to read (Castro-Caldas et al., 1998, Olson, 1996; Luksaneeyanawin et al., 1997; Ziegler et al., 2004). Luksaneeyanawin et al. (1997) argue that perceptual sensitivities to nonnative sounds are attenuated at two important junctures; in early infancy and at the onset of reading. Similarly, for Olson (1996), "learning to read is learning to hear speech in a new way!" (p. 95). It seems then that reading in the target orthography is the key. While orthography starts to influence perception when the graphemes are learnt in isolation, it is not sufficient. It is only when children (and learners) become actively engaged in learning to read that orthographic knowledge really begins to take hold of speech perception.

Perhaps this explains, in part, why the ZN group was unable to make any improvement from the pre- to the post-tests. The participants spent their three lessons learning the new orthography and the new phonemic inventory. These lessons were specifically designed to focus on grapheme-sound association and recognition. However, the participants did not learn to read in the Zhuyin orthography per se. There were no activities where the participants practiced the new orthography through reading dialogues and/or texts written in Zhuyin. It is interesting to speculate that had the experimental groups had a longer, more intensive training program and the Zhuyin had more time and opportunity to solidify the new symbol-sounds associations through reading practice that the ZN group would have eventually translated the new orthographic knowledge into a facilitative tool to enhance the participants' perceptual sensitivities to the L1-L2 contrasts. No doubt, a very short training period in a new orthographic code (such as the training period in this research) would not be able to 'out influence' a lifetime of associations made in the L1 orthographic code. However, it is possible that given a longer training program that the non-shared orthographic groups may have surpassed their shared orthographic counterparts in perceptual performance.

5.1.3 Summary

The above sections have attempted to account for the lack of significant differences between the PN, ZN and CL groups in terms of the relative strengths of the L1 and L2 orthographic codes. If the L1 orthographic code is interconnected with (Olson, 1996; Landerl et al., 1996; Burnham, 2003; Ziegler et al., 2003) and inseparable from (Treiman and Cassar, 1997) the L1 phonological system and if learners perceive an L2 phonological system in reference to their L1 system (Kuhl, 1993, 2000; Flege, 1987, 1991, 1995; Best, 1995, 2001), then it is possible that all the participants in this study perceived the new L2 phonological system in reference to both the L1 phonological *and* the L1 orthographical systems. That is, as the research on auditory recognition suggests (Hallé et al., 2000; Ziegler and Ferrand, 1998; Ziegler et al., 2003), even in the absence of visual appearance, L1 orthographic knowledge was still a prevalent feature of the learning process for the non-shared orthographic groups. In short, the lack of significant differences between the three groups may in part be a function of language learning where learners perceive the L2 sound system in terms of their L1 orthography (at least for natives of an alphabetic orthography) regardless of the orthographic medium within which they are instructed.

In addition to the strength of the L1 orthographic code, two other factors were also identified as secondary contributors to the observed experimental results. First, the cognitive demands may have been higher for the ZN and CL groups. The ZN group was required to learn two new systems; a phonological and orthographic system. The CL group was required to learn only one new system but the participants may have been hampered by not being able to associate the new phonology to visual orthographic cues. Second, if, as the research suggests (Olson, 1996; Luksaneeyanawin et al., 1997; Ziegler et al., 2003), learning to read the orthography is what creates a strong relationship with the phonological system, then the ZN group did not have enough exposure and practice in reading the Zhuyin script to form strong new sound-grapheme connections.

In sum, it appears that the lack of significant differences between the experimental groups can be attributed to a hierarchy of the following three factors.

1. The strength of the entrenched L1 orthographic code which remains pervasive even without overt visual representation.
2. The unequal amount of cognitive demand placed on the learners.
3. The inadequate training program which did not allow sufficient time for the non-shared orthographic groups to reconcile themselves to the tasks and the new systems.

The combination of these three factors lends a slight advantage to using Pinyin during the beginning stages of learning Mandarin phonology as the potential advantages of employing a separate orthography, Zhuyin – or no orthography at all – appear to be neutralized.

However, whether Pinyin would retain this advantage over the course of extended learning or whether the benefits of learning via Zhuyin would surpass those of Pinyin in the long term are unattested and remain important questions for further research.

5.2 Sound contrast sensitivity

While the analyses of the results indicated that there were no significant differences in perceptual performance between the three experimental groups, the analyses did indicate that there were significant differences in perceptual sensitivities to each of the target contrasts. The 3-way repeated measures ANOVA isolated three sensitivity groups: 1) a highly sensitive group, 2) a slightly sensitive group, and 3) an insensitive group. Figure 5-1 revisits the sensitivity continuum illustrated in Chapter 4.

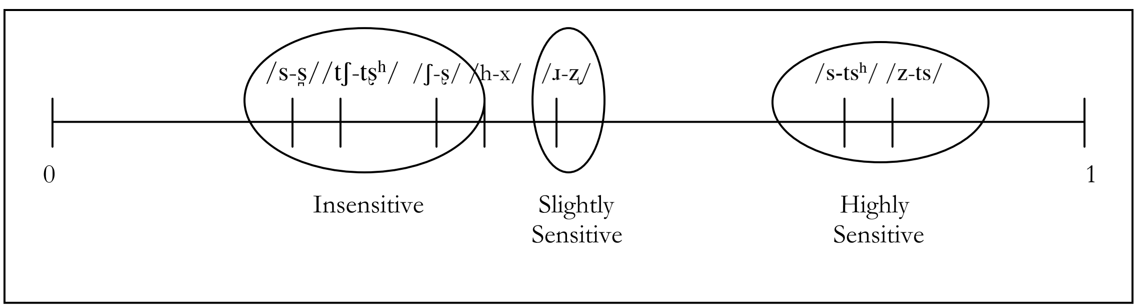


FIGURE 5-1 Three statistically significant sensitivity groups revisited

Based on the varying degrees of sensitivity, the purpose of this section is to discuss why the participants were more sensitive to some contrasts while they remained insensitive to others. The following subsections will discuss the contrast pairs in terms of their similarities, which potentially create confusion, and their differences, which potentially aid in perceptual discrimination. Taking into consideration these similarities and differences, this section will be rounded out with a discussion of L2 phoneme 'mapping' and models of speech processing.

5.2.1 Basis for potential confusability

Simply sharing orthographic representations is not a sufficient condition for the potential perceptual confusability between L1 and L2 phonemes. Take, for example, the grapheme 'j'; it exists in both the English and Spanish alphabets. Yet, the grapheme 'j' represents two completely different sounds in its respective languages. It represents the voiced alveo-palatal affricate [dʒ] in English and the voiceless glottal fricative [h] in Spanish. In a perceptual discrimination task, it is almost certain that listeners would easily distinguish between the two sounds since they are very different from each other in terms of articulatory and acoustic characteristics. Therefore, the potential contribution of shared orthography to perceptual confusability is only relevant when the sound pairs in each target contrast are closely related by exhibiting some important articulatory and /or acoustic similarities.

It was established in Chapters One and Two that the target contrasts chosen in this research share characteristics that make them more susceptible to perceptual confusion, although the degrees of similarity varies between each sound contrast. To review, Table 5-1 lays out the underlying similarities by providing the feature specifications for each of the 14 target sounds. It is important to note a number of details regarding Table 5-1. First, while there is still debate on some of the features, the assumption here is that the phonological features accurately represent the articulatory / acoustic properties of speech. Second, the feature specifications used to characterize the target sounds are based on Gussenhoven and Jacobs (2005). Third, the following table only presents those features where there are at least some differences between contrasting sounds. Finally, empty cells indicate that the sound is unspecified for that particular feature.

FEATURES		CONTRASTS													
		Highly sensitive				Slightly sensitive		Insensitive							
		z*	ts	s	ts^{h}	ɹ	ʐ	h	x	ʃ	ʂ	tʃ	tʂ^{h}	s	s̺
Major class	consonantal	+	+	+	+	+	+	-	+	+	+	+	+	+	+
	sonorant	-	-	-	-	+	-	-	-	-	-	-	-	-	-
	approximant	-	-	-	-	+	-		-	-	-	-	-	-	-
Manner	continuant	+	-	+	-	+	+		+	+	+	-	-	+	+
	strident	+	+	+	+		+	-	-	+	+	+	+	+	+
Voicing	voice	+	-	-	-	+	-		-	-	-	-	-	-	-
	spread	-	-	-	+	-	-	+	-	-	-	-	+	-	-
Place	CORONAL	√	√	√	√	√	√			√	√	√	√	√	√
	distributed	-	+	-	+	-	-			+	-	+	-	-	+
	anterior	+	+	+	+	+	-			-	-	-	-	-	+
	DORSAL								√						
# of different features		3		3		5		6		1		2		2	

Note: When one of the sounds is unspecified for a feature, this is considered a difference between contrast sounds.

* The contrasts above are given in order of their degree of sensitivity from the highest /z-ts/ to the lowest /s-s̺/.

TABLE 5-1 Feature specifications for the seven target contrast pairs

From Table 5-1 above, there is little doubt that the L1 and L2 contrast pairs under investigation do have many similarities. However, upon closer inspection it is also apparent that the degrees of similarity vary. These relative degrees of similarity as predicted in earlier chapters would result in different perceptual sensitivities to particular contrasts. But how are the degrees of similarity between sounds pairs to be established?

One possible method for assessing the relative similarity (or dissimilarity) between sounds would be to simply tally the number of featural differences between each set of sounds. The idea being that those sounds with fewer differences between them would be more similar than those with more differences. According to this method, the order of similarity from the least to the most similar would be:

/h-x/ > /ɹ-ʐ/ > /s-tʂʰ/ and /z-ts/ > /tʃ-tʂʰ/ and /s-s̺/ > /ʃ-ʂ/

The order, in turn, would predict an order of perceptual sensitivity where according to SLM (Flege, 1987, 1995), the more dissimilar two sounds are, the greater the chance that they will be perceived as distinct from one another. Therefore, if the sheer number of feature differences is responsible for the degree of perceptual sensitivity then the order of sensitivity should be the same as the order given above.

Yet, as discussed previously, the results do not reflect these predictions. In fact, the order of sensitivity discovered in the data is as follows:

/z-ts/ > /s-tʂʰ/ > /ɹ-ʐ/ > /h-x/ > /ʃ-ʂ/ > /tʃ-tʂʰ/ > /s-s̺/

This discrepancy between the number of differences and the degree of sensitivity to the sound pairs suggests that not all features are weighted equally. Rather, some features appear to carry more weight and are more acoustically salient than others which would influence the degree to which listeners are sensitive to a contrast. The following subsection addresses this issue.

5.2.2 Feature weight and acoustic saliency

Up to this point, it has been argued that the potential confusability between the target sound pairs lies in their similarity. But, it has also been established that the number of differences does not accurately predict how sensitive listeners will be to a contrast. So, to account for the order of sensitivity among the 7 target contrasts, the subtle manner and place differences between the L1 and L2 sounds in each contrast pair must be isolated. Manner and place are of interest here as the results demonstrated a correlation between the type of differences (either manner or place) and the contrast sensitivity where contrasts that differed in manner are more easily distinguished than contrasts that only differ in place. Table 5-2 breaks down where the sounds in each contrast differ in terms of these two categories. The darkly shaded rows indicate those two target contrasts to which the participants were highly sensitive. The lightly shaded row indicates the contrast to which the participants were slightly sensitive and the unshaded rows indicate those four contrasts to which the participants were insensitive.

Grapheme	Contrast		Manner			Place		
	English	Mandarin		English	Mandarin		English	Mandarin
c	**s**	**ts^h**	**different**	fricative	affricate	**different**	alveolar	dental
z	**z**	**ts**	**different**	fricative	affricate	**different**	alveolar	dental
r	**ɹ**	**ʐ**	**different**	liquid	fricative	**different**	alveolar	post-alveolar
h	**h**	**x**	**same**	fricative	fricative	**different**	glottal	velar
sh	**ʃ**	**ʂ**	**same**	fricative	fricative	**same***	post-alveolar	post-alveolar
ch	**tʃ**	**tʂ^h**	**same**	affricate	affricate	**same***	post-alveolar	post-alveolar
s	**s**	**s̪**	**same**	fricative	fricative	**different**	alveolar	dental

* Although there are no differences in the passive articulator (place), there are most likely differences in the grooving of the active articulator (i.e. the tongue) which results in the retroflex characteristic on the two Mandarin sounds.

TABLE 5-2 Comparisons of manner and place characteristics for the 7 target contrasts

In Tables 5-1 and 5-2, a distinct pattern emerges. That is, when the two sounds in a contrast differed in manner, listeners were at least somewhat sensitive to the contrast (discussed further below). This suggests that where sounds differ in manner, they are more likely to be distinguished from one another. No such pattern is evident for differences in place of articulation. For these contrasts, place differences alone do not appear to be sufficient for listeners distinguish between the sounds. For example, /s-s̪/ and /h-x/ differed in terms of place but were still confusable contrasts. Interestingly, even though /s-s̪/ differed in place and /tʃ-tʂ^h/ and /ʃ-ʂ/ did not (They only differ in the retroflex feature.), /tʃ-tʂ^h/ and /ʃ-ʂ/ were slightly less confusable than /s-s̪/. For native Mandarin speakers, Tse (2000) found retroflexion to be more perceptually salient than dentality. Perhaps this difference in perceptual saliency is also somewhat accessible to nonnative Mandarin speakers / learners. Of course, this potential hierarchy of perceptual sensitivity requires further research to confirm the degree of saliency for retroflexion. In short, the results suggest that manner differences between the sounds appear to be more acoustically salient and better indicators of perceptual distinguishability than place differences.

What Table 5-2 demonstrates is that perceptual sensitivity is not a matter of listeners being either sensitive or insensitive. Rather, as the results have shown, degrees of sensitivity exist and the degree to which listeners are sensitive to a contrast is contingent on whether the contrast pairs differ in manner of articulation and when they do differ what those differences in manner are. To further illustrate the varying degrees of similarities exhibited between the sounds in the contrast pairs, the following three figures offer example waveforms and spectrograms (one set for each sensitivity group) that visually represent the different acoustic properties.

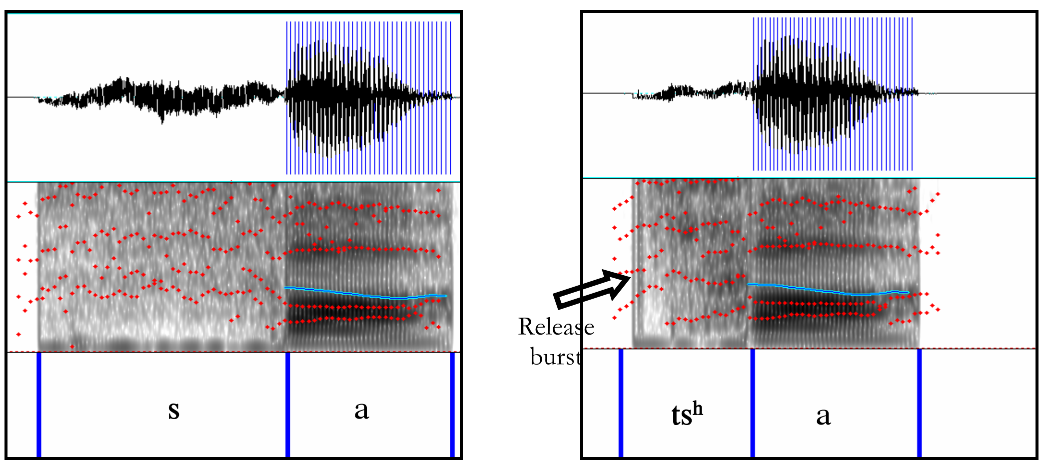


FIGURE 5-2 Waveforms and spectrograms of the *highly sensitive* contrast /sa/ and /tsʰa/

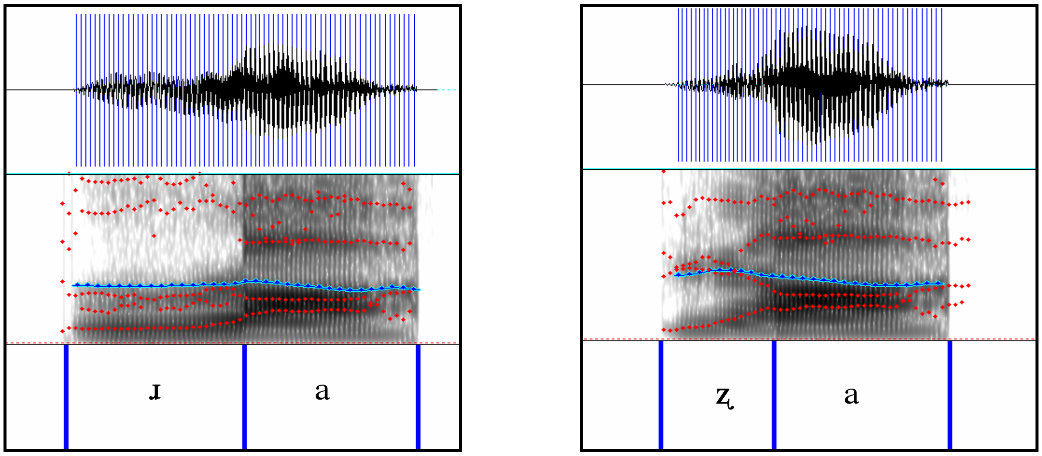


FIGURE 5-3 Waveforms and spectrograms of the *slightly sensitive* contrast /ɹa/ and /ʐa/

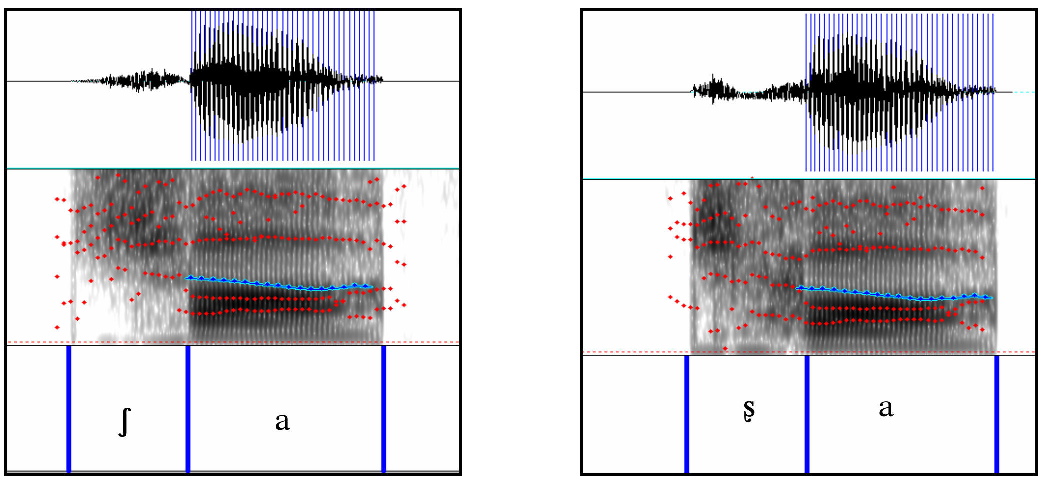


FIGURE 5-4 Waveforms and spectrograms of the *insensitive* contrast /ʃa/ and /ʂa/

Table 5-2 demonstrates that the fricative-affricate pairs (/s-ts^h/ and /z-ts/) are easier to distinguish than the liquid-fricative pair (/ɹ-ʐ/). This is not surprising considering the more salient acoustic differences between fricative and affricates than fricatives and liquids. According to Borden, Harris and Raphael (2003), the primary acoustic cue to the manner classes is the aperiodic component (i.e. the noise energy). Fricatives are characterized by a relatively long duration of frication and affricates are characterized by a sharp onset of noise (the release burst) and a shorter duration of frication. Research has suggested that frication duration is a sufficient acoustic cue for differentiating fricatives from affricates (Kluender and Walsh, 1992). The spectrograms in Figure 5-2 above show an obvious difference in frication duration between the fricative /s/ and the affricate /tsʰ/. In addition, the release burst characteristic of affricates is also clearly visible in the /tsʰ/ stimulus and no such burst exists in the /s/. While relatively long periods of frication are the principle acoustic cue for the perception of fricative manner, affricates are complex segments that contain acoustic cues for both stops and fricatives (Borden et al., 2003; Kent and Read, 2002). Apparently, the silence, the release burst, the rapid rise time, the frication duration and the formant transitions are all used to identify affricates (Borden et al., 2003). Indeed, sounds containing a release burst plus frication would be easier to differentiate from sounds only containing frication. The feature here that seems most responsible for the high degree of sensitivity appears to be [+/- continuant] where the fricative are [+ continuant] while the affricates are [- continuant].

For the /ɹ-ʐ/ contrast, the two sounds do differ in manner but the acoustic differences of this contrast appear to be less salient than the /s-tsʰ/ and /z-ts/ contrasts thereby resulting in only a slight sensitivity to the contrast. Both /ɹ/ and /ʐ/ are voiced and are characterized by an unobstructed degree of constriction in the vocal tract (properties of both fricatives and liquids). They are both [+continuant]. In this type of contrast, there is no silence or release burst to offset one sound from the other. Examples of the acoustic differences between fricative /ʐ/ and the liquid /ɹ/ can be seen in Figure 5-3 above. When compared with Figures 5-2 and 5-3, the spectrograms in Figure 5-4 demonstrate contrasts such as /ʃ-ʂ/ have fewer acoustic differences; and therefore, would be much less salient. In sum, when it comes to the varying degrees of sensitivity to the target contrast pairs, the results indicate that manner differences coupled with the [+/- continuant] feature differences seem to be more acoustically salient than others.

Although the noise segments are argued to be the primary cue to the manner class distinctions (Borden et al., 2003), they are not the only acoustic cue available to listeners. Formant transitions also aid in perceptual distinctions. In fact, for stops, fricatives, and affricates, F2 formant transitions to and from the neighbouring vowels are one of the prominent cues to place of articulation (Borden et al., 2003). Since each target consonant was extracted from its original word and added to a single low back vowel to create simple CV stimuli, the formant transitions from the consonant into the preceding vowel were incomplete. As a result, the listeners heard conflicting transition cues for the two segments that were pasted together in the CV stimuli. Therefore, it is very likely that the participants did not have access to the cue information necessary to distinguish the sounds that differed in place alone. This may explain in part why some of the place contrasts, which are cued by F transitions, were more confused than others; the saliency of the formant transition cues may have been reduced by the creation of the CV stimuli for this project. In sum, when it comes to manner and place differences between sounds in native-nonnative contrast pairs, the key differences that contribute to perceptual sensitivity are manner differences.

There is yet another plausible explanation for why the fricative-affricate pairs were the easiest contrasts to distinguish. The contrasts /s-ts^h/ and /z-ts/ may not have been perceived as differences between two single phonemes. That is, the Mandarin affricates /ts^h/ and /ts/ may have been perceived not as single phonemes but rather as exemplars of a consonant cluster. It is possible that the participants were making distinctions between single consonant onsets /s/ or /z/ and a 2-member cluster onsets /ts/. For example, in their native language, English speakers are adept at distinguishing between /s/ as in the word *mass* and /ts/ as in the word *mats*. Granted this distinction does not exist in the onset position in English; however, the Markedness Differential Hypothesis (MDH) (Eckman,

1985) predicts that if listeners (learners) have acquired the distinction in the more marked position, the coda, then they will have very little difficulty in applying it to a less marked position, the onset. The acoustic cues available for listeners will dictate how the sounds are categorized and ultimately perceived. Therefore, how the L2 phonemes are 'mapped' on to the native L1 phonological system becomes a vitally important question.

5.2.3 Models of speech processing

The underlying assumption on which all the current models of speech perception rests is that L2 sounds are perceived in reference and mapped on to the L1 phonological system. Yet, the debate on how exactly L2 sounds are perceived and mapped is far from resolved. For instance, the Native Language Magnet (NLM) theory (Kuhl, 1993, 2000) maintains that after the native language's phonological system is established in the first year of life, the L1 phonetic prototypes serve as magnets that 'warp' the phonological space and draw L2 sounds. The Speech Learning Model (SLM) (Flege, 1987, 1991, 1995) similarly maintains that L1 phonetic categories act as 'attractors' of L2 sounds but where sufficient differences exist between the L2 sounds and the closest L1 category, the establishment of new L2 phonetic categories is possible. According to SLM, perceptual sensitivity to nonnative sounds remains through adulthood. Finally, the Perceptual Assimilation Model (PAM) claims that L2 sounds are assimilated to L1 phonetic categories based on articulatory, acoustic and constriction commonalities (Best, 1995, 2001). Of particular interest here are the abilities of Flege's SLM and Best's PAM in accounting for the varying degrees of perceptual sensitivities observed in the results of this research.

However, prior to entering a discussion on SLM and PAM, some time must first be allotted to the clarification of the English-Mandarin sound correspondences made for the purposes of this research. Figure 5-5 (on page 100) is a reproduction of the English and

Mandarin consonant inventories from Chapter Two. The arrowed lines illustrate the correspondences (i.e. the contrast pairs) between the L1 and L2 sounds for each of the contrasts studied. The impetus behind investigating these correspondences was that the two sounds were connected by the same orthographic symbol(s). But as mentioned previously, shared orthography can sometimes be rather arbitrary, and therefore, shared orthography could not be the sole reason for a connection between two sounds in this study. The corresponding sounds must also share some phonetic / phonological properties that make the sounds similar to one another which, in turn, may or may not lead to confusion. In short, the correspondences (henceforth referred to as 'counterparts') illustrated in Figure 5-5 are products of the phonetic and phonological characteristics (as outlined in sections 5.2.1 and 5.2.2), as well as orthographic characteristics that are shared by the two sounds in each target contrast. It is important to remember that the L1-L2 correspondences created for this project do not necessarily mean that the L1 sounds are the phonetically / phonologically 'closest' relative to the L2 sound. For example, the broken arrow pointing from /ʐ/ to /ʒ/ suggests that /ʒ/ may be a closer relative to /ʐ/ than /ɹ/ (see below for a more detailed discussion).

			LABIAL		DENTAL		ALVEOLAR		ALVEO-PALATAL		PALATAL		VELAR		GLOTTAL	
			E	M	E	M	E	M	E	M	E	M	E	M	E	M
STOPS	-vc	+asp	p^h	p^h			t^h	t^h					k^h	k^h	(ʔ)*	
		-asp	p	p			t	t					k	k		
	+vc		b				d						g			
AFFRI-CATES	-vc	+asp				ts^h				tʂ^h		(tɕ^h)				
		-asp				ts			tʃ	tʂ		(tɕ)				
	+vc								dʒ							
FRICA-TIVES	-vc		f	f	θ	s̪	s		ʃ	ʂ		(ɕ)		x	h	
	+vc		v		ð		z		ʒ	ʐ						
NASALS			m	m			n	n					ŋ	ŋ		
LIQUIDS							l/ɫ	l								
							ɹ	ɻ								
GLIDES			w	(w)							j	(j)				

* The sounds in brackets () are not considered to be phonemic

Note: The } notation (as suggested by Dr J. Esling) for the English voiceless stops indicates that the two consonants belong to the same phonemic category where the underlying phoneme is not clear.

FIGURE 5-5 English and Mandarin consonant inventories revisited

The primary aim of SLM is to account for the attainment of L2 production (Flege, 1995) where perception of phonetic distinctiveness is a necessary condition for accurate production. Perception of L2 sounds is contingent on the perceptual similarity (or dissimilarity) of L2 sounds to the 'closest' L1 phonetic category. SLM argues that if the differences between the L2 sound and the closest L1 sound can be discerned then learners are more likely to create a new category for the L2 sounds which will ultimately lead to authentic production of the new L2 sound. If there are no discernable differences between the L2 and L1 sounds then learners are likely to equate the two sounds as one which will result in poor production of the L2 sound. (Flege, 1987, 1991, 1995).

From Figure 5-5 and the feature discussion in subsection 5.2.2, /tsʰ/ and /ts/ appear to be perceptually farther from their respective corresponding L1 counterpart, /s/ and /z/ than /s̺/, /ts̺ʰ/, /s̺/, and /x/ are from /s/, /tʃ/, /ʃ/, and /h/. SLM predicts that native English speakers would be more likely to create a new category(s) for /tsʰ/ and /ts/[12] while /s̺/, /ts̺ʰ/, /s̺/, and /x/ are more likely to be equated with their L1 counterparts. Indeed, these predictions are borne out in the results as demonstrated by the participants' high degree of sensitivity to the /s-tsʰ/ and /z-ts/ contrasts and the participants' insensitivity to the /s-s̺/, /tʃ-ts̺ʰ/, /ʃ-s̺/, and /h-x/ contrasts.

However, the problem with SLM becomes apparent with the final contrast in this research, /ɹ-ʐ/. Figure 5-5 shows that the closest L1 category to /ʐ/ may not be /ɹ/ but rather /ʒ/ based on more shared characteristics such as place (alveo-palatal), and as demonstrated in 5.2.2, most importantly, manner (fricative). SLM's preoccupation with production performance is only concerned with L1-L2 correspondences insofar as the L1 is

---

[12] It seems unlikely that native English speakers would categorize /tsʰ/ and /ts/ separately from each other as aspiration is not a contrastive feature in English. The point here is that /tsʰ/ and /ts/ would be categorized as separate from their L1 counterparts /s/ and /z/, respectively.

the phonetically closest sound to the L2. That is, SLM would only account for and consider the relationship between /ʐ/ and /ʒ/. Setting aside for a moment that the contrast under investigation is not this pair, SLM would predict that, like the other alveo-palatal contrast pairs, listeners would not be able to perceive the differences between them and the /ʐ/ would be equated with the /ʒ/. If indeed, /ʐ/'s closest L1 phoneme is /ʒ/, SLM makes no provisions for the perceptual confusion between /ʐ/ and other sounds such as /ɹ/.

Perhaps it could be argued that the only relationship of perceptual importance is the one between the L2 sound and the closest L1 phonetic category. However, the results suggest otherwise. If it was the only important relationship then, regardless of whether they are equated with each other or a new category is formed, when the L2 sound is contrasted with another phoneme perceptually farther away, they should be relatively easy to distinguish. In other words, even if /ʐ/ is equated with /ʒ/, participants should be relatively sensitive to the contrast /ɹ-ʐ/ as /ɹ/ and /ʒ/ are separate phonemes in English. But as the data indicated, the participants were only slightly sensitive (A' = 0.5) to this contrast. This indicates that there was a certain degree of confusability between /ɹ/ and /ʐ/ and suggests that the closest L1 phonetic category may not be the only phoneme to influence the perception of the L2 sounds.

Alternatively, /ʐ/ does share many important characteristics with /ɹ/; specifically, /ʐ/ and /ɹ/ are both apicial, [+ continuant] and [COR]. Indeed, /ʐ/ may occupy phonological space with both /ʒ/ and /ɹ/. Figure 5-6 visually speculates on the relationship between these three sounds.

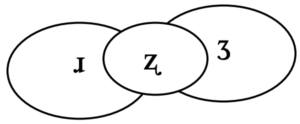


FIGURE 5-6 Diagram of possible shared phonological space between the Mandarin sound /ʐ/ and the English sounds /ʒ/ and /ɹ/

If true, then /ʐ/ would be confusable with both L1 phonemes. However, this question is beyond the scope of the current study and requires further research to ascertain whether /ʐ/ is closest L1 phoneme is either /ʒ/ or /ɹ/ and / or whether it shares phonological space with both L1 sounds.

The lack of clarity as to /ʐ/'s closest L1 phoneme highlights the fundamental flaw that plagues all the current models of speech processing. Models of speech perception hinge on one major assumption – L2 sounds are perceived in relation to the closest L1 phoneme. While intuitively this assumption makes sense, the notions 'closest' and 'similar' are ephemeral. All models talk about 'closeness' and 'similarity' but all have failed to adequate account for it. In other words, there is a distinct lack of well defined criterion for establishing the L1 and L2 similarity and/or proximity. Therefore, until there is an explicit method for determining degrees of 'closeness' between L1 and L2 sounds, models of speech perception will never be able to fully account for how L2 sounds are perceived.

At the outset, PAM (Best, 1995, 2001) appears similar to SLM where PAM suggests that listeners will "perceptually assimilate non-native phones to native phonemes wherever possible" (p. 777). In the cases where they cannot be assimilated to a native category, the L2 sounds will be classified as either an uncategorizable speech sound or an unassimilable non-speech sound. Best (2001) argues that perception of nonnative sounds "is *not* uniformly poor" [original emphasis] (p. 776); therefore, PAM attempts to account for poor to excellent discriminations of not only native-nonnative contrasts but also nonnative-nonnative contrasts. How each sound in a contrast pair is assimilated will dictate the level of discrimination possible, ranging from poor to excellent.

According to PAM, there are 6 pair-wise assimilation types for nonnative-nonnative contrasts.

1. **Single Category (SC)** assimilation where both sounds are assimilated to one native category.
2. **Two Category (TC)** assimilation where the 2 sounds are assimilated to separate categories.
3. **Category Goodness (CG)** assimilation where both sounds are assimilated to the same category but one sound is a better exemplar than the other.
4. **Uncategoried-Categorized (UC)** assimilation where one sound is assimilated to a native category and the other is recognized as a speech sound but remains uncategorized.
5. **Uncategorized-Uncategorized (UU)** assimilation where both sounds remain uncategorized.
6. **Unassimilable (NA)** assimilation where both sounds are perceived as non-speech sounds.

With these assimilation types, there is a hierarchy of discrimination success. PAM predicts that NA assimilations would yield good to excellent discrimination, TC and UC assimilations would yield good discrimination because of the existence of a category boundary, GC and SC would yield poorer discriminations because the sounds are in the same category and UU assimilations would yield fair to good discriminations (Best, 2001).

While the above pair-wise assimilations were formulated to account for nonnative-nonnative contrasts, they may also be useful for native-nonnative contrasts. As mentioned previously, to assume that the L2 sound can only be perceived in relation to the closest L1 phoneme effectively reduces the predictive power of the speech perception model. Therefore, extending the assimilation types to encompass native-nonnative contrasts as well avoids assuming that the L1 sound in the contrast is necessarily the 'closet' phoneme to the L2 sound. In this way, PAM can offer some insight into the relationship between the Mandarin and English sound contrasts investigated in the current research. That is, with

PAM's assimilation types as a guide, the varying degrees of perceptual sensitivity observed in the data suggest how the Mandarin sounds may have been mapped in relation to their L1 contrast counterpart. When L2 sounds are assimilated to native categories, PAM predicts a discrimination order of TC>CG>SC (Best, 1995, 2001). This order has been confirmed (Best, 2001) where TC assimilations exhibited excellent discriminations, CG assimilations exhibited fair discriminations and SC assimilations exhibited poor discriminations.

The three perceptual sensitivity groups discovered in this research mirror Best's excellent, fair and poor discriminations of contrasts. Based on the discrimination patterns of the English-Mandarin contrasts, within PAM, this suggests then that the Mandarin sounds were mapped in similar ways. Naturally, the English sounds are associated with their native categories. The question is then: Are the Mandarin sounds assimilated to the same native categories as their contrast counterparts or are they assimilated to separate categories. Ease of discrimination suggests that the Mandarin phone may have been assimilated to a separate category. Fair discrimination suggests the Mandarin phone may have been perceived as an aberrant variant of the English phoneme and poor discrimination suggests that the Mandarin phone may have been assimilated as a good variant of the English phoneme. Table 5-3 provides a summary of the discrimination performance for each contrast and the suggested assimilation pattern according to PAM.

Contrast	Discrimination	L2 sound probably assimilated…
s-ts^{h}	excellent	to a separate category (TC)
z-ts	excellent	to a separate category (TC)
s-ʂ	poor	to the same category (SC)
tʃ-tʂ^{h}	poor	to the same category (SC)
ʃ-ʂ	poor	to the same category (SC)
ɹ-ʐ	fair	as an aberrant exemplar of /ɹ/ (CG)
h-x	poor to fair	to the same category (SC) or as an aberrant exemplar of /h/ (CG)

TABLE 5-3 Interpretation of possible assimilation patterns for the English-Mandarin contrasts based on the levels of discrimination

In conclusion, of the recent models developed to describe second language speech processing, only PAM's gradient levels of discrimination reflect how the Mandarin L2 sounds are potentially perceived in relation to native English phonemes.

However, in order to truly test PAM in relation to the acquisition of Mandarin sounds, further research would need to devise a method that first externally identifies the 'right' assimilation categories for each contrast. That means that the procedure would need to have another component that had native speakers report on how the sounds are perceived. For example, native speaker judgements on L2 sounds compared with their L1 counterparts may report that /ts^{h}/ is a different sound from /s/ which would confirm that the /s-ts^{h}/ contrast falls within a Two Category assimilation. Similarly, native speakers' reports that /ʐ/ is a 'weird'/'poor' example of /ɹ/ would confirm that the /ɹ-ʐ/ contrast falls with in a Category Goodness assimilation.

5.2.4 Summary

The focus of this section has been to discuss why the participants demonstrated different perceptual sensitivities to the target contrasts. The first part of the section established that similarities existing between any given English-Mandarin sound pairs may result in confusion and an inability to distinguish between the two sounds. In the second part, the discussion on acoustic characteristics revealed a hierarchy of saliency.

1. Manner plus [+/- continuant] differences appear to be the most acoustically salient.
2. Manner differences appear to be less acoustically salient than manner plus [+/- continuant] differences but more acoustically salient than place differences alone.
3. Place differences do not appear to be acoustically salient which may in part be due to the creation of the CV stimuli.

Finally, in the third part, the evaluation of the current models of speech processing determined that Best's PAM provides the most satisfactory account of how the varying degrees of perceptual sensitivity illustrate the ways in which the Mandarin phones may be mapped onto the native English phonological system and thus perceived. Yet, this finding is far from conclusive and requires adjustments to the methodological procedure and further research.

Chapter Six

CONCLUSION

"By the time we reach adulthood […] language experience has had a profound effect on speech perception." (Kuhl, 1993, p. 125)

This final chapter presents four concluding sections that complete the investigation undertaken in this research project. The first section provides a summary of the project and the findings of the research. The second section considers its limitations. In light of the results and remaining questions, the third section briefly suggests some future research possibilities. The last section outlines the possible contributions that this research makes to the growing body of knowledge on second language acquisition.

6.1 Summary of research

The focus of this research was to determine whether the orthographic medium used for second language instruction had an effect on learners' abilities to accurately perceive L1-L2 sound contrasts. The study sought to determine if perceptual performance of native English speakers taught Mandarin via an alphabetic writing system (i.e. Pinyin) differed significantly from that of native English speakers taught Mandarin via a syllabary writing system (i.e. Zhuyin) and/or from native English speakers taught Mandarin without a writing system (i.e. Control). The three-part experiment was designed to test the prediction that when the L1 and L2 share an orthographic system where the same graphemes represent separate sounds, learners are less likely to perceive the difference between the two sounds. Also of interest was whether learners of Mandarin would demonstrate varying degrees of perceptual sensitivity to the sound contrasts based on the relative similarities between the two sounds in each contrast.

To investigate these important questions, thirty-two native Canadian English speakers, were assigned to one of the three experimental groups (either the PN, ZN, or CL group). The three-part experiment design encompassed: 1) a pre-test, 2) a training phase, and 3) a post-test. The pre- and post-tests used an oddity discrimination task (Flege and McKay, 2004; Guion et al., 2000; Baker et al., 2002) to test the participants' abilities to distinguish the L2 Mandarin phonemes from the L1 English phonemes. Between the two tests, the participants participated in a training phase (4 ½ hours) where they learnt Mandarin via the orthographic medium assigned to their group. From the data collected in the pre- and post-tests, participant performance was determined by calculating the differences between the pre- and post-test results for the participants in each group and comparing these differences across the three groups. Perceptual performance was measured in terms of accuracy (error rates), sensitivity (a-prime), and speed (response times).

The research was based on the premise that if the three experimental groups demonstrated different perceptual sensitivities to the target sound contrasts that these differences could be indicative of orthographic influence on how the learners categorize the Mandarin phonemes in relation to the native English phonological system. At least two factors influence the categorization and perception of the new L2 sound system: 1) the nature of the L2 sounds themselves in relation to the L1 phonological system (Kuhl, 1993, 2000; Flege, 1987, 1995; Best, 1995, 2001) and 2) the way the sounds are represented in the orthographic medium used for instruction (Trieman and Cassar, 1997; Landerl et al., 1996; Olson, 1996; Hallé et al., 2000; Burnham, 2003; Ziegler and Ferrand, 1998; Ziegler et al., 2003; Ziegler et al., 2004). According to the hypotheses outlined in Chapter One, the influence of the L2 sounds on perceptual distinction between native-nonnative sound contrasts would be mediated by the influence of the orthography. In short, the hypotheses predicted that for the PN group, the influence of the shared orthographic medium would be revealed in higher tendencies to equate the L2 sounds with the L1 sounds. For the ZN and CL groups, the influence of the different or nonexistent orthographic medium would be revealed in higher tendencies to distinguish between the L2 and L1 sounds.

The results of this experiment did not confirm the hypotheses regarding the effect of shared orthography on L2 sound perception. Rather, the 3-way repeated measures ANOVA analyses of the results indicated that there were no significant differences in the perceptual performances of the three groups for either error rates, a-prime or response times. A number of factors were discussed as to why there was a lack of significant results. The first of which was attributed to the pervasiveness of the L1 orthographic code. As previous research has suggested even without the explicit presence of the L1 orthographic code, it almost certainly remained indirectly present in the minds of the participants in the Zhuyin and Control groups. In fact, the results of this study suggest that second language learners learn a L2 system via their native orthography regardless of the medium with which they were taught and the facilitative potential of using a separate orthography are counteracted by the entrenched native orthography..

While there were no significant differences between any of the experimental groups, an observable trend did appear. The shared orthography group (PN) was the only group to consistently and to a greater extent improve in accuracy, sensitivity and speed from the pre-test to the post-test. This is possibly attributed to the cognitive demands placed on the participants. Higher cognitive demands were more than likely placed on the ZN and CL groups as the ZN group had to learn two new systems (as opposed to the single system the PN group was required to learn) and the CL group did not have the benefit of visual cues to aid in learning. Furthermore, another explanation is, as research suggests, it is not until the onset of reading that the orthographic code starts to influence speech perception (Castro-Caldas et al., 1998, Olson, 1996; Luksaneeyanawin et al., 1997; Ziegler et al., 2004). No doubt, the very short training period was insufficient to supplant a lifetime of L1 orthographic associations.

While the results did not confirm the hypotheses about orthographic influence, they did confirm the hypothesis that the participants would exhibit different levels of sensitivity to each of the seven target contrasts. That is, not all L1-L2 contrasts were equally easy (or difficult) to distinguish. In fact, the 2-way repeated measures ANOVA isolated three sensitivity groups; 1) highly sensitive contrasts, 2) slightly sensitive contrasts, or 3) insensitive contrasts. These varying degrees of sensitivity suggest then that there were varying degrees of similarity within the contrasts whereby the more dissimilar the sounds, the easier they were to discriminate (Flege, 1987, 1995; Best 1995, 2001). Furthermore, a closer inspection of the differences between the sounds in each contrast suggests that manner differences are more perceptually salient than place differences.

6.2 Limitations

No experimental design is infallible. While every aspect was carefully considered and there was a rationale behind every decision, the experimental design of this project is not without its limitations. The primary methodological limitation lies in the length of the training phase. The short time allotted for the Mandarin language training may have been insufficient for the participants to create new L2 associations through reading. A second limitation of this study was the creation of the CV stimuli. The process of identifying the target consonants and replacing the vowels (and codas where present) of the original words with the same single low back vowel may have provided conflicting formant transitions which in turn may have reduced the necessary place cues needed to distinguish some of the contrast sounds. Additionally, although approved by the native speaker judges, these manipulated CV stimuli may have retained an element of unnaturalness. Another limitation was the relatively small number (PN = 11, ZN = 11, and CL =10) of participants in each group.

6.3 Future Research

As this research is the first to explore the effect of orthography on the perception and categorization of L2 phonemes in second language acquisition, the results are preliminary. Future research that incorporates a more refined research design addressing the limitations mentioned above is necessary to confirm the lack of perceptual differences between alphabetic and non-alphabetic learners of Mandarin. In addition to the need for replication, the results, and subsequently, the issues raised in this work have opened the door to a wealth of research possibilities. Particularly, one interesting issue arose in the data that is currently inexplicable and demands further research to isolate the answers. Language base appears to influence the degree of perceptual sensitivity to some of the contrasts. That is, for the contrasts /ɹ-ʐ/ and /h-x/, the participants were significantly more successful at identifying the odd item out when the odd stimulus was a Mandarin phoneme that was offset by two stimuli from their native language (English base). However, the reverse was true for the contrasts /s-tsʰ/ and /s-ʂ/. In these cases, the participants were more successful when the odd item out was an English phoneme that was offset by two stimuli from their second language (Mandarin base). For the final three contrasts /z-ts/, /tʃ-tʂʰ/, and /ʃ-ʂ/the base did not influence discrimination. At present, there is no discernable pattern.

Another research possibility would be to integrate a priming element to the discrimination task to investigate how immediate orthographic cues influence L2 phoneme perception. Perception, however, is only one side of the coin. Future research is also necessary in the area of shared orthographic influence on L2 production. The following are two potential Mandarin production research projects:

1. A comparison of production accuracy for different groups of learners reading aloud in either Pinyin or Zhuyin. Do Pinyin readers produce Mandarin sounds less authentically than Zhuyin readers?
2. A comparison of production accuracy for the same learners reading aloud in Pinyin and Chinese characters. Do learners produce Mandarin sound more authentically when they read words in characters than when they read words in Pinyin?

These suggestions for future research are by no means exhaustive, merely some of the most intriguing.

6.4 Contributions

The research in this project explores an area of second language acquisition that up to now has largely been ignored. In other words, few studies have been conducted that explore whether there is a relationship between first language influence, orthography and learners' perceptions of sound contrasts. The research here suggests, contrary to the hypotheses, that shared orthography does not hinder but may, in fact, facilitate L2 sound perception during the beginning stages of learning a second language by reducing the cognitive demands placed on the learners. This knowledge is valuable both theoretically and pedagogically. Theoretically, research of this type is important because it contributes to the growing body of knowledge on the potential influences to L2 sound perception. Pedagogically, the results have serious implications for language instructors who may question the value of teaching Mandarin via a romanized orthography. Further, the results seem to caution instructors that the introduction of a new orthography at the beginning of language learning may result in slower progress (at least initially) although the long term effects of non-shared orthography are far from clear.

REFERENCES

Abercrombie, D. (1949). Teaching pronunciation. *English Language Teaching,* 3, 113-122.

Ager, S. (1998). Pu&to⇐nghua⇑ (Mandarin). *Omniglot: Writing languages of the world.* April 14, 2006, http://www.omniglot.com/writing/mandarin_pts.htm.

Arichbald, J. (1998). *Second Language Phonology.* Amsterdam: John Benjamins Publishing Company.

Baker, W., Tromfimovich, P., Mack, M., & Flege, J.E. (2002). The effect of perceive phonetic similarity on non-native sound learning by children and adults. *Proceedings of the Annual Boston University Conference on Language Development,* 36-47. Somerville, MA: Cascadilla Press.

Best, C.T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.) *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (pp. 171-204). Baltimore: York Press.

Best, C.T. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *Journal of the Acoustical Society of America*, 109 (2), 775-794.

Best, C.T. & McRoberts, G.W. (2003). Infant perception of non-native consonant contrasts that adults assimilate in different ways. *Language and Speech*, 46 (2-3), 183-216.

Bohn, O.S. & Flege, J.E. (1992). The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 14 (2), 131-158.

Borden, G.J., Harris, K.S., & Raphael, L.J. (2003). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech.* 4th Ed. Baltimore MD: Lippincott Williams & Wilkens.

Brown, C. (2000). The interrelation between speech perception and phonological acquisition from infant to adult. In J. Archibald (Ed.) *Second Language Acquisition and Linguistic Theory* (p. 4-63). Malden: Blackwell Publishers Ltd.

Brown, J.D. (1988). Chapter 5: Critiquing statistical studies. *Understanding Research in Second Language Learning: A Teacher's Guide to Statistics and Research Design*. (pp. 43-62) New York: Cambridge University Press.

Brown, J.D. (1991). Statistics as a foreign language – Part 1: What to look for in reading statistical language studies. *TESOL Quarterly*, 25 (4), 569-586.

Brown, J.D. (1992). Statistics as a foreign language – Part 2: More things to consider in reading statistical language studies. *TESOL Quarterly*, 26 (4), 629-664.

Burnham, D.K. (2003). Language specific speech perception and the onset of reading. *Reading and Writing: An Interdisciplinary Journal*, 16, 573-609.

Carroll, J.M. (2004). Letter knowledge precipitates phoneme segmentation, but no phoneme invariance. *Journal of Research in Reading*, 27 (3), 212-225.

Castro-Caldas, A., Petersson, K.M., Reis, A., Stone-Elander, S., & Ingvar, M. (1998). The illiterate brain: Learning to read and write during childhood influences the functional organization of the adult brain. *Brain*, 121, 1053-1063.

Ch'en, T.T., Link, P., Tai, Y.J., & Tang, H.T. (1994). *Chinese Primer: Lessons.* Princeton, NJ: Princeton University Press.

Ch'en, T.T, Link, P., Tai, Y.J., & Tang, H.T. (1994). *Chinese Primer: Notes and Exercises.* Princeton, NJ: Princeton University Press.

Chueng, H. & Chen, H.C. (2004). Early orthographic experience modifies both phonological awareness and on-line speech processing. *Language and Cognitive Processes,* 19 (1), 1-28.

DeFrancis, J. (1990). The why of Pinyin grapheme selection. *Chinese Language Teachers Association Journal*, 25 (3), 1-14.

DeFrancis, J. (1996). How efficient is the Chinese writing system? *Visible Language*, 30 (1), 6-44.

DeFrancis, J. (2006). The prospects for Chinese writing reform. *Sino-Platonic Papers*, 171 (June), 1-24.

Dijkstra, T., Roelofs, A., & Fieuws, S. (1995). Orthographic effects on phonemic monitoring. *Canadian Journal of Experimental Psychology,* 49 (2), 264-271.

Duanmu, S. (2000). *The Phonology of Standard Chinese.* Oxford: Oxford University Press.

Eckman, F.R. (1985). Some theoretical and pedagogical implications of the Markedness Differential Hypothesis. *Studies in Second Language Acquisition,* 7 (3), 289-307.

Ellis, N. (2002). Frequency effects in language processing. *Studies in Second Language Acquisition*, 24, 143-188.

Erdener, V.D. & Burnham, D.K. (2005). The Role of audiovisual speech and orthographic information in nonnative speech production. *Language Learning,* 55 (2) June, 191-228.

Everson, M.E. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal,* 82 (ii), 194-203.

Field, A. (2005). *Discovering Statistics Using SPSS*. 2nd Ed. London: SAGE Publications Ltd.

Flege, J.E. (1987). The production of new and similar phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.

Flege. J.E. (1991). The interlingual identification of Spanish and English vowels: orthographic evidence. *The Quarterly Journal of Experimental Psychology*, 43A (3), 701-731.

Flege, J.E. (1995). Second language speech learning theory, findings and problems. In W. Strange (Ed.) *Speech Perception and Linguistic Experience: Issues in Cross Language Research* (p. 233-277). Baltimore: York Press.

Flege, J.E. (2003). Assessing constraints on second-language segmental production and perception. In N.O. Schiller & A.S. Meyer (Eds.) *Phonetics and Phonology in Language Comprehension and Production: Differences and Similarities*. (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J.E. & MacKay, I.R.A. (2004). Perceiving vowels in a second language. *Studies in Second Language Acquisition*, 26, 1-34.

Frauenfelder, U. H., Segui, J., & Dijkstra, T. (1990). Lexical effects in phonemic processing: Facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance,* 16 (1), 77-91.

Frith, U. (1998). Editorial: Literally changing the brain. *Brain*, 121, 1011-1012.

Frost, R., Repp, B.H., & Katz, L. (1988). Can speech perception be influenced by simultaneous presentation of print? *Journal of Memory and Language*, 27, 741-755.

Glushko, R. (1979). The organization and activation of orthographic knowledge in reading aloud. *The Journal of Experimental Psychology: Human Perception and Performance,* 5 (4), 674-691.

Guion, S.G., Flege, J.E., Akahane-Yamada, R., & Pruitt, J.C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107 (5), 2711-2724.

Gussenhoven, C., & Jacobs, H. (2005). *Understanding Phonology*. 2nd Ed. London: Oxford University Press.

Hallé, P.A., Chéreau, C., & Segui, J. (2000). Where is the /b/ in "absurde"[apsyrd]? It is in French listeners' minds. *Journal of Memory and Language,* 43, 618-639.

Hancin-Bhatt, B. (1994). Segment transfer: a consequence of a dynamic system. *Second Language Research*, 10 (3), 241-269.

Hatch, E. & Farhady, H. (1982). Chapter 2: What is a variable? *Research Design and Statistics for Applied Linguistics.* (pp. 12-17). Rowley: Newbury House Publishers.

Iverson, P., Kuhl, P., Akahane-Yamada, R., Diesch, E., Tohkura, Y., Kettermann, A., & Siebert, C. (2002). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, 87, B47-B57.

Jared, D., McRae, K., & Seidenberg, M.S. (1990). The basis of consistency effects in word naming. *The Journal of Memory and Language,* 29, 687-715.

Jared, D. & Kroll, J.F. (2001). Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language*, 44, 2-31.

Jared, D. & Szucs, C. (2002). Phonological activation in bilinguals: Evidence from interlingual homograph naming. *Language and Cognition,* 5 (3), 225-239.

Kent, R.D. & Read, C. (2002). *Acoustic Analysis of Speech.* 2nd Ed. Albany, New York: Thomson Learning.

Kess, J. & Miyamoto, T. (2006). The issue of phonological activation in Chinese and Japanese Lexical Access. In Grace E. Wiebe, Gary Libbon, Tom Priestly, Ron Smyth, & H. Samuel Wang (Eds.). *Phonology, Morphology, and the Empirical Imperative: Papers in Honour of Bruce L. Derwing.* (pp. 457-482). Taipei: Crane Publishing Co., Ltd.

Kessler, B. & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language,* 44, 592-617.

Killingley, S.Y. (1998). *Learning to Read Pinyin Romanization and its Equivalent in Wade-Giles: A Practical Course for Students of Chinese.* Newcastle: LINCOM Europa.

Kluender, K.R. & Walsh, M.A. (1992). Amplitude rise time and the perception of the voiceless affricate/fricative distinction. *Perception and Psycholophysicis*, 51, 328-333.

Koda, K. (1989). Effects of L1 orthographic representation on L2 phonological coding strategies. *Journal of Psycholinguistic Research,* 18 (Nov. 2), 201-222.

Kuhl, P. K. (1993). Early linguistic experience and phonetic perception: implications for theories of developmental speech perception. *Journal of Phonetics*, 21, 125-139.

Kuhl, P.K. (2000). A new view of language acquisition. *Proceedings of the National Academy of Sciences of the United States of America*, 97 (22) October, 11850-11857.

Lacruz, I. & Folk, J.R. (2004). Feedforward and feedback consistency effects for high- and low-frequency words in lexical decision and naming. *The Quarterly Journal of Experimental Psychology,* 57A (7), 1261-1284.

Lado, R. (1957). *Linguistics Across Cultures.* Ann Arbor: University of Michigan Press.

Landerl, K., Frith, U., & Wimmer, H. (1996). Intrusion of orthographic knowledge on phoneme awareness: Strong in normal readers, weak in dyslexic readers. *Applied Psycholinguistics*, 17, 1-14.

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

Luksaneeyanawin, S., Burnham, D., Francis, E., & Pansottee, S. (1997). The role of L1 background and L2 instruction in the perception of fricative contrasts: Thai and English children and adults. *Asia Pacific Journal of Speech, Language and Hearing*, 2, 25-42.

MacKay, I.R.A., Meador, D., & Flege, J.E. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58, 103-125.

Major, R. (2001). *Foreign Accent: The Ontogeny and Phylogeny of Second Language Phonology.* Mahwah: Lawrence Erlbaum Associates, Inc.

Major, R. (2002). The phonology of the L2 user. In V. Cook (Ed.) *Portraits of the L2 User.* (p. 67- 92). Clevedon: Multilingual Matters Ltd.

Matthews, J. & Brown, C. (2004). When intake exceeds input: Language specific perceptual illusions induced by L1 prosodic constraints. *International Journal of Bilingualism*, 8 (1), 5-27.

Muneaux, M. & Ziegler, J.C. (2004). Locus of orthographic effects in spoken word recognition: Novel insights for the neighbour generation task. *Language and Cognitive Processes*, 19 (5), 641-660.

Oller, J.W. & Ziahosseiny, S.M. (1970). The contrastive analysis hypothesis and spelling errors. *Language Learning*, 20 (2), 183-189.

Olson, D.R. (1993). How writing represents speech. *Language & Communication*, 13 (1), 1-17.

Olson, D.R. (1996). Towards a psychology of literacy: on the relations between speech and writing. *Cognition*, 60, 83-104.

Pastore, R.E., Crawley, E.J., Berens, M.S., & Skelly, M.A. (2003). "Nonparametric" *A'* and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, 10 (3), 556-569.

Peng, T.H. (1984). *Hanyu Pinyin.* Singapore: Times Books International.

Pierrehumbert, J. (2001). Exemplar dynamics: Word frequency, lenition and contrast. In J. Bybee & P Hopper (eds), *Frequency and the Emergence of Linguistic Structure*, (p. 137-157) Amsterdam: Benjamins.

Pierrehumbert, J. (2002). Word-specific phonetics. In C. Gussenhoven & N. Warner (eds.) *Laboratory Phonology 7*, (p. 101-139) Berlin & New York: Mouton de Gruyter.

Pisoni, D.B. (1973). Auditory and phonetic codes in the discrimination of consonants and vowels. *Perception and Psychophysics*, 13, 253-260.

Read, C., Zhang, Y.F., Nie, H.Y., & Ding, B.Q. (1986). The ability to manipulate speech depends on knowing alphabetic writing. *Cognition*, 24, 31-44.

Rogers, H. (2000). *The Sounds of Language: An introduction to phonetics.* Essex: Pearson Education Limited.

Seidenberg, M.S., Waters, G.S., Barnes, M.A., & Tanenhaus, M.K. (1984). When does irregular spelling of pronunciation influence word recognition? *The Journal of Verbal Learning and Verbal Behavior,* 23, 383-440.

Snodgrass, J.G. & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117 (1), 34-50.

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

Treiman, R. & Cassar, M. (1997). Can children and adults focus on sound as opposed to spelling in a phoneme counting task? *Developmental Psychology,* 33 (5), 771-780.

Tse, J. K. (2000). The perceptual saliency of palatality, dentality, and retroflexion in Mandarin consonants: An experimental study. *Studies in English Literature and Linguistics*, 26 (June), 97-111.

Werker, J. & Tees, R. (1984). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America*, 75, 1866-1878.

Werker, J. & Logan, J. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, 37, 35-44.

Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schroger, E. (2004). From symbols to sounds: Visual symbolic information activates sound representations. *Psychophysiology,* 41, 709-715.

Wode, H. (1983). *Phonology in L2 acquisition.* In H. Wode (Ed.), *Papers on language acquisition, language learning and language teaching.* (p. 175-187). Heidelberg, Germany: Groos.

Zhou, M. (2001). Language policy and reforms of writing systems for minority languages in China. *Written Language & Literacy*, 4 (1), 31-65.

Ziegler, J.C. & Ferrand, L. (1998). Orthography shapes the perception of speech: The consistency effect in auditory word recognition. *Psychonomic Bulletin & Review,* 5 (4), 683-689.

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

Ziegler, J.C., Ferrand, L., & Montant, M. (2004). Visual phonology: The effects of orthographic consistency on different auditory word recognition tasks. *Memory and Cognition,* 32 (5), 732-741.

Appendix A

CITATION WORDS FOR TARGET CONSONANTS

English

1. raw
2. hop
3. shawl
4. chop
5. soggy
6. somber
7. za
8. fawn
9. vault

Since orthographic 'c' representing the sound /s/ does not occur before a low back vowel, it was necessary to use an 's' word to gather the target stimuli.

Used to create the stimuli for the practice trial phase of the perception test.

中文。。。

1. 讓 'rang'
2. 哈蟆 'hama'
3. 沙子 'shazi'
4. 叉子 'chazi'
5. 撒謊 'sahuang'
6. 擦 'ca'
7. 咱們 'zamen'

Appendix B

PINYIN AND ZHUYIN ORTHOGRAPHIES

Initials

PINYIN	ZHUYIN
b	ㄅ
p	ㄆ
m	ㄇ
f	ㄈ
d	ㄉ
t	ㄊ
n	ㄋ
l	ㄌ
g	ㄍ
k	ㄎ
h	ㄏ
j	ㄐ
q	ㄑ
x	ㄒ
zh	ㄓ
ch	ㄔ
sh	ㄕ
r	ㄖ
z	ㄗ
c	ㄘ

Finals

PINYIN	ZHUYIN
a	ㄚ
o	ㄛ
	ㄜ
e	ㄝ
ai	ㄞ
ei	ㄟ
ao	ㄠ
ou	ㄡ
an	ㄢ
en	ㄣ
ang	ㄤ
eng	ㄥ
er	ㄦ
i	ㄧ
u	ㄨ
ü	ㄩ

Appendix C

Participant Consent Form

Shared Orthography: Do shared written symbols influence the perception of native and nonnative sound contrasts?

You are invited to participate in a study entitled "Shared Orthography: Do shared written symbols influence the perception of native and nonnative sound contrasts?" that is being conducted by Hua Lin and Carolyn Pytlyk.

Dr. Hua Lin is an Associate Professor in the Department of Linguistics at the University of Victoria. Ms. Pytlyk is an MA student in the department of Linguistics at the University of Victoria. You may contact either or both if you have further questions by e-mail at pytlykca@uvic.ca.

This research is being funded by University of Victoria Internal SSHRC Research Grant.

Purpose and Objectives
The purpose of this study is to discover whether the orthography used for instruction has an effect on second language learners' abilities to accurately perceive sound contrasts between their first and second languages. The study plans to investigate if perceptual performance of native English speakers taught Mandarin via an alphabetic writing system differs significantly from that of native English speakers taught Mandarin via a character-like writing system. It will test the hypothesis that when the first and the second language share orthographic symbols representing separate sounds that resemble each other, learners are less likely to perceive the difference between the two sounds.

Importance of this Research
Research of this type is important because it intends to make a contribution to the growing body of knowledge on SECOND LANGUAGE accents and how they result. Within this area, there is a substantial amount of research regarding FIRST LANGUAGE interference, perception of sound contrasts, and symbol-sound associations. However, there is a lack of research investigating the possible connection between all three areas of research. In other words, few studies have been conducted to explore whether there is a relationship between first language influence, orthography and learners' perception of sound contrasts.

Participants Selection
You are being asked to participate in this study because you represent the population of native Canadian English speakers who learn Mandarin as either a second or foreign language and you have no previous experience with any Chinese language.

What is involved?
If you agree to voluntarily participate in this research, your participation will include (1) taking a pre-instruction perception test, (2) having 4 ½ hours of instruction in Mandarin Chinese, and (3) taking a post-instruction perception test. You will also be asked to complete a questionnaire to ascertain your language learning background.

Inconvenience
Participation in this study may cause some inconvenience to you, including approximately the 6 hours you will need to allot for participation.

Risks
There are no known or anticipated risks to you by participating in this research.

Benefits
The potential benefits of your participation in this research include gaining an introductory understanding and knowledge of Mandarin Chinese via the free Mandarin lessons offered. Your participation will also contribute to the investigation into whether written symbols play a pivotal role in the acquisition of a second language sound system.

Voluntary Participation
Your participation in this research must be completely voluntary. If you do decide to participate, you may withdraw at any time without any consequences or any explanation. If you do withdraw from the study the information you provide in the background questionnaire and your data will not be used for the analysis and will be destroyed.

On-going Consent
To make sure that you continue to consent to participate in this research, you will be reminded that your participation is voluntary and that you can withdraw at any time during the research without any consequences or explanation.

Anonymity
In terms of protecting your anonymity, all data collected will be kept completely anonymous. All information and the data collect will be arranged and stored according to your identification numbers. Any analysis and mentioning of the testing processes will be anonymous; no names or other defining characteristics will be revealed.

Confidentiality
The confidentiality of your participation will be limited due to the group nature of the language classroom context. This means that your participation in the study will known by all the other participants in your Chinese language class. However, the confidentiality of your data will be protected by ensuring that all your data and information is stored in password protected files and/or in a locked cabinet in a locked office.

Dissemination of Results
It is anticipated that the results of this study will be shared with others in the following ways; (1) as a thesis done by Ms. Carolyn Pytlyk, (2) a presentation at scholarly meetings, and possibly (3) a published article.

Disposal of Data
Data from this study will not be disposed of. There is a possibility that it may be of use for future studies by either the researcher or other researchers.

Contacts
Individuals that may be contacted regarding this study include Carolyn Pytlyk at pytlykca@uvic.ca and Dr. Hua Lin at luahin@uvic.ca.

In addition, you may verify the ethical approval of this study, or raise any concerns you might have, by contacting the Associate Vice-President, Research at the University of Victoria (250-472-4545).

Your signature below indicates that you understand the above conditions of participation in this study and that you have had the opportunity to have your questions answered by the researchers.

Name of Participant	*Signature*	*Date*

A copy of this consent will be left with you, and a copy will be taken by the researchers.

Appendix D

PARTICIPANT QUESTIONNAIRE

Dear Participant,
Please be advised that this is not a test. It is a questionnaire designed to identify your previous first and language learning experiences. The information provided by you below will be used in conjunction with your data collected from the study. In order to ensure your anonymity no names will be elicited. Instead, you will be assigned an identification code. All data will be stored according to these identification numbers. If at any time you decide to withdraw from this study, all the data collected (including the information provided here) will not be used for the analysis and will be destroyed. If you have any questions or concerns, Ms. Carolyn Pytlyk (pytlykca@uvic.ca) will be happy to discuss them with you. Thank you for your participation.

Background Information

Age ______________________________ Gender__________________________

Place of Birth _____________________ Nationality _____________________

Where did you attend elementary school? __________________________________

Where did you attend high school? _______________________________________

Language Learning Information

Have you studied any languages other than English? YES NO

If so, what language(s)? ________________________________

How long? ____

What is you level of proficiency (ex. beginner, intermediate, advanced)?

____

Do you consider yourself:

☐ MONOLINGUAL ☐ BILINGUAL ☐ MULTILINGUAL

Do you have any experience with a Chinese language? YES NO

If so, which Chinese language (ex. Cantonese, Mandarin etc.)? _______________

Please briefly describe your Chinese experience.

Identification Number ___________________

Appendix F

EXAMPLE DIALOGUE ACTIVITY[13]

A: nǐ hǎo *(Hello.)* B: nǐ hǎo *(Hello.)* A: nǐ jiào shénme? *(What is your name?)* B: wǒ jiào ______. *(My name is ____.)* A: nǐ jǐ sùi? *(How old are you?)* B: wǒ ____ sùi. *(I am ____ years old.)* A: xìe xie *(Thank you.)* B: měi guǎn xi *(You're welcome.)*

1. míngzi: ______________ sǔi: __________	2. míngzi: ______________ sǔi: __________	3. míngzi: ______________ sǔi: __________
4. míngzi: ______________ sǔi: __________	5. míngzi: ______________ sǔi: __________	6. míngzi: ______________ sǔi: __________
7. míngzi: ______________ sǔi: __________	8. míngzi: ______________ sǔi: __________	9. míngzi: ______________ sǔi: __________

[13] The same activity was conducted in the ZN and CL group classes where in the ZN group the Q/A prompts were written in Zhuyin graphemes and in the CL group, there were no Q/A written prompts.