5-1-2021

THE L2 ACQUISITION OF MANDARIN ALVEOLO-PALATAL PHONEMES BY NATIVE ENGLISH SPEAKERS

Brooke Sherrell

THE L2 ACQUISITION OF MANDARIN ALVEOLO-PALATAL PHONEMES BY NATIVE ENGLISH SPEAKERS

by

BROOKE SHERRELL

Presented to the Faculty of the Honors College of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

HONORS BACHELOR OF ARTS IN

CRITICAL LANGUAGES AND INTERNATIONAL STUDIES

THE UNIVERSITY OF TEXAS AT ARLINGTON

May 2021

ACKNOWLEDGMENTS

I would like to thank Dr. Szu-Yen Liang for the guidance provided during this research, along with continuing to inspire my interest in the Chinese language. I would like to thank Dr. Ivy Hauser for providing the opportunity to observe the LING 5322 course in order to learn auditory analysis. I would like to thank Assistant Director of Undergraduate Research Bobbie Brown for her tremendous support throughout the completion of this Honors Senior Project.

April 23, 2021

ABSTRACT

THE L2 ACQUISITION OF MANDARIN ALVEOLO-PALATAL PHONEMES BY NATIVE ENGLISH SPEAKERS

Student Name, B.A. Critical Languages and International Studies

The University of Texas at Arlington, 2021

Faculty Mentor: Szu-Yen Liang

This study investigated the acquisition of Mandarin alveolo-palatal phonemes [tɕ, tɕʰ, ɕ] by English L1 learners of Mandarin. The acquisition of these phonemes is often difficult for English speakers as the alveolo-palatal place of articulation is not found in English, but this acquisition has not been extensively studied. This study consisted of two perception tasks and one production task. The first perception task was an AX discrimination task to determine participants' ability to distinguish between Mandarin and English phonemes. This task was followed by an identification task. The production task collected auditory data, which was analyzed with Praat and compared to two native Mandarin speakers' production. The results of the perception tasks showed learners had difficulty identifying the [tɕʰ] phoneme but not the [tɕ] and [ɕ] phonemes, confirming a previous study. The acoustic analysis confirmed the hypothesis that learners' production of the alveolo-palatal phonemes was not native-like.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Research Overview

This study investigates the degree to which adult English-speaking learners of Mandarin are able to perceive and produce the alveolo-palatal phonemes found in Mandarin; /ɕ/, /tɕ/, and /tɕʰ/. (written in pinyin as x, j, q respectively) These phonemes share the alveolo-palatal place of articulation (POA), a POA not found in the English consonant inventory. Many learners of a second language strive to achieve native-like abilities in their L2, however, the possibility of an adult learner reaching this level of acquisition is still up for debate (Hyltenstam & Abrahamsson, 2000). Researchers have extensively investigated other phonetic and phonological aspects of Mandarin that are not found in English, such as retroflex phoneme acquisition (Ou & Guo, 2014) or the acquisition of tones at the suprasegmental level (Wiener et al., 2020), however, this cannot be said for the alveolo-palatal phonemes. This research aims to explore the ability English Mandarin L2 learners have to accurately perceive these three alveolo-palatal phonemes, as well as determining how their production compares to native speakers' production through acoustic analysis.

1.2 Second Language Acquisition Theory

1.2.1 Critical Period Hypothesis

The Critical Period Hypothesis (CPH) was proposed by Lenneberg (1967) to explain what effect age had on second language acquisition (SLA). This theory posits

there is an ideal period when the brain is most receptive to learning language (typically the prepubescent period), and after this time passes, it becomes increasingly difficult to truly acquire language. This theory has been debated by many modern scholars. Modern theory argues while there may be a sensitive period where the brain is most receptive to linguistic input, there is no evidence suggesting age is a determining factor in successful language acquisition (Bialystok, 1997). Environmental factors such as L2 input via language immersion or solely in a classroom setting have been deemed to be more of a determining factor of acquisition than age, which may explain some of the trends suggesting young children learn languages quicker than teenagers or adults, since many older language learners learn the language in the classroom setting rather than through immersion (Singleton, 2003). For the purposes of this study, all participants are those who have studied Mandarin in a classroom environment, which should be kept in account when determining the feasibility of English native speakers fully acquiring the alveolo-palatal phonemes.

1.2.2 Contrastive Analysis Hypothesis

Lado (1957) postulated that the more different a sound was in an L2 compared to the learners' L1, the more difficult the sound would be to acquire, known as the Contrastive Analysis Hypothesis (CAH). Through positive transfer of L1 knowledge to L2 knowledge (Stockwell et al., 1965), language learners are able to more easily acquire the sounds in an L2 that are similar to their L1. Negative transfer occurs when a given structure of an L2 more highly contrasts with that of an L1, causing difficulties and allowing one to predict the errors that may be made when learning a second language.

1.2.3 Speech Learning Model

The Speech Learning Model (SLM) proposed by Fledge (1995) argues against the CAH by proposing that the sounds and structures are more unlike those found in a learners' L1 are more likely to be noticed by the learner, and thus the learner will put more attention to these sounds and more easily create new phonetic categories for them. For sounds similar to those found in a learners' L1, it is argued these will not be perceived as different and will instead be categorized as the same sound, leading to mistakes in perception and production.

1.3 Mandarin Sound Inventory

1.3.1 Mandarin Consonant Inventory

The following table outlines the consonant inventory of Mandarin. This study examines the L2 perception and production of the affricates and the fricative in the alveolo-palatal POA, /ɕ/, /tɕ/, and /tɕʰ/. As shown in Table 1.1, Mandarin contains consonants in the alveolar, retroflex, and alveolo-palatal POA. The retroflex and alveolo-palatal POA are not found in English, but the alveolar and postalveolar POA are found.

Table 1.1: Mandarin Consonant Inventory (based on Kim, 2018)

	Bilabial	Labiodental	Alveolar	Retroflex	Alveolo-palatal	Velar
Plosive	p pʰ		t tʰ			k kʰ
Affricate			ts tsʰ	ʈʂ ʈʂʰ	tɕ tɕʰ	
Fricative		f	s	ʂ	ɕ	x
Nasal	m		n			ŋ
Approximant				ɻ	j	
Lateral			l			

A key feature of Mandarin phonology that differs from English is the contrastive distribution between plosives and affricates in Mandarin is found through aspiration and unaspiration (Kim, 2018). English, on the other hand, contains voiced consonants and finds contrast between voiced and unvoiced phonemes. While unvoiced aspirated consonants are found in English phonology, the unvoiced aspirated consonants in Mandarin are more heavily aspirated than the ones found in English (Chen et al., 2008).

1.3.2 Pinyin

During the language reformation performed by the PRC during the mid-19th century, the Chinese government created a system to standardize a romanization of Mandarin Chinese. In 1958, a romanization called Hanyu Pinyin (汉语拼音 Hàn yǔ pīn yīn), or pinyin, was published by the Chinese government and added to school curriculums. Pinyin was later accepted as the standard romanization of Mandarin Chinese internationally in the 1980s (Zhou, 2004). Pinyin is used by a majority of Mandarin L2 learners to learn how to pronounce the Chinese orthography. Tables 1.2 and 1.3 outline pinyin and the corresponding IPA pronunciation of each letter, as found in Chen, 2016. All participants of this study have studied and understand how to pronounce pinyin, and thus pinyin will be used during the examination.

Table 1.2: Vowel Pinyin Chart with IPA in Parenthesis (based on Chen, 2016)

	Front	Central	Back
High	i, ü/y (/i/, /y/)	i (/ɨ/)	u (/u/)
Mid	e (/ɛ/)	e (/ə/, /ɤ/)	o (/o/)
Low		a (/a/)	

Table 1.3: Consonant Pinyin Chart with IPA in Parenthesis (based on Chen, 2016)

Labial	b (/p/)	p (/pʰ/)	m (/m/)	f (/f/)
Dental	d (/t/)	t (/tʰ/)	n (/n/)	l (/l/)
Velar	g (/k/)	k (/kʰ/)	h (/x/)	-ng (/ŋ/)
Alveolar	z (/ts/)	c (/tsʰ/)	s (/s/)	
Alveolo-palatal	j (/tɕ/)	q (/tɕʰ/)	x (/ɕ/)	
Retroflex	zh (/tʂ/)	ch (/tʂʰ/)	sh (/ʂ/)	zh (/ɻ/)

1.4 Hypotheses and Predictions

The prediction for this study is that the results will be reflect the claims of the SLM, thus due to the closeness of alveolar phonemes found in English and the alveolo-palatal phonemes found in Mandarin, the following hypotheses are made:

1. Non-nativelike production of the alveolo-palatal phonemes.
2. Alveolo-palatal phonemes will be perceived as alveolar phonemes.
3. Alveolo-palatal phonemes will not be accurately identified.

CHAPTER 2

LITERATURE REVIEW

Research surrounding the L2 acquisition of Mandarin phonemes has been conducted in the past, but not with the same grouping of phonemes. Other places of articulation have been the focus of studies, such as the study of the acquisition of retroflex sounds (Ou & Guo, 2014), or phonemes have been grouped by the manner of articulation, such as the study of Mandarin voiceless fricatives (Shih & Kong, 2011) or affricates (Yang & Yu, 2019). In all of these studies, it can be concluded learners of an L2 generally perform better with perception tasks than production tasks.

2.1 Previous Perception Studies

In Shih & Kong (2011), participants listened to two Mandarin and/or Taiwanese stimuli and were tasked with judging the similarity between the sounds based on a five-point Likert scale (1-the same, 2-very similar, 3-similar, 4-different, 5-very different). By conducting a perception test using this format, the researchers were able to analyze the data based on a mean rating of the stimuli pairs and by plotting the perceptual distance of various phonemes on a one-dimensional chart. Through this research, they concluded the "sa-sha pair" (stimuli written in pinyin) was the most difficult fricative for a bilingual of Mandarin and Taiwanese. This methodology will be duplicated in the present study to determine which sound pairs speakers of English and Mandarin most struggle with.

In Yang & Yu (2019), an identification task was completed as part of the perceptual analysis of the acquisition of Mandarin affricates by American L2 learners. In this study,

participants were tasked with listening to target syllables and identifying them as one of six answer choices, all listed using the Mandarin romanization pinyin. Through this analysis, it was determined these participants had a lower identification accuracy with the aspirated alveolo-palatal q /tɕʰ/ and the unaspirated retroflex affricate zh /tʂ/ compared with the other Mandarin affricates. This study analyzes how English speakers are able to acquire the Mandarin affricates, a manner of articulation, whereas the present study analyzes the acquisition of a single place of articulation, the alveolo-palatal phonemes. An identification task will be conducted in the present study to determine if difficulties with the aspirated alveolo-palatal affricate q /tɕʰ/ are also found.

2.2 Previous Production Studies

Previous studies that conducted an acoustic analysis of Mandarin affricates and sibilants by L2 speakers have concluded that learners had not fully acquired these sounds. Yang & Yu (2019) concluded that English L1 learners of Mandarin had not created new categories for the Mandarin affricates not found in English and instead were assimilating these sounds to the English alveolar place of articulation. In Kin (2016), it was concluded Cantonese L1 speakers of Mandarin were not completely assimilating Mandarin sibilants to the Cantonese place of articulation, but had instead created a "new form" residing in the middle of the two sets of sounds. Both studies analyzed beginner to intermediate learners of Mandarin, or those who have three months to two years of Mandarin experience in a classroom environment. The present study analyzes those who also have had at least three months and at most three years of Mandarin classroom instruction.

CHAPTER 3

METHODOLOGY

3.1 Participants

For the two perception tasks, thirteen participants, aged 18-23, were recruited. All the participants have taken or were enrolled in Mandarin courses at the University of Texas at Arlington, thus have an understanding of the Pinyin romanization system, and were native English speakers. Six of the thirteen participants opted to also provide auditory samples for the production task. Two native speakers of Mandarin recorded samples for the auditory analysis and for the perception tasks. One native speaker of English recorded samples for the distinguishment perception task.

3.2 Stimuli

Table 3.1 presents the set of stimuli used for the perception tasks. The six monosyllabic CV words consisted of syllables with the alveolo-palatal POA [tɕ, tɕʰ, ɕ] initials followed by the [iɑ] vowel final and retroflex POA [tʂ, tʂʰ, ʂ] initials followed by the [ɑ] vowel final. The [ɑ] vowel final was chosen because this is the only vowel in Mandarin which can be followed by both alveolo-palatal and retroflex consonants due to phonological constraints, but due to a vowel fronting rule following Mandarin alveolo-palatal phonemes, the [ɑ] vowel is realized as [iɑ] (Hauser, 2020). Each of the monosyllabic words are first-tone segments to keep the suprasegmental variance at a minimum. Three disyllabic words were used for the perception tasks as well, each containing at least one of the target alveolo-palatal phonemes.

Table 3.2 presents the set of stimuli recorded for the production task. Pinyin and characters were provided. Each word contains at least one of the target alveolo-palatal phonemes, all of which were analyzed during acoustic analysis.

Table 3.1: Perception Test Stimuli

Alveolo-palatal	xiā (虾) [ɕiɑ]	jiā (家) [tɕiɑ]	qiā (掐) [tɕʰiɑ]
Retroflex	shā (沙) [ʂɑ]	chā (差) [tʂʰɑ]	zhā (扎) [tʂɑ]
Disyllabic	xiū xi (休息) [ɕiɤʊ ɕi]	shí jiān (时间) [ʂʅ tɕiɛn]	qǐ chuáng (起床) [tɕʰi tʂʰu̯ɑŋ]
English Stimuli	[ʃɑ]	[dʒ]	[tʃ]

Table 3.2: Production Test Stimuli

Chinese Character	Pinyin	IPA
休息	xiū xi	[ɕiɤʊ ɕi]
时间	shí jiān	[ʂʅ tɕiɛn]
起床	qǐ chuáng	[tɕʰi tʂʰu̯ɑŋ]
喜欢	xǐ huan	[ɕi xu̯an]
家	jiā	[tɕiɑ]
夏天	xià tiān	[ɕiɑ tiɛn]
清楚	qīng chu	[tɕʰiŋ tʂʰu]
契机	qì jī	[tɕʰi tɕi]

3.3 Procedures

The first of two perception tasks were a distinguishment task. Participants listened to two stimuli and rated the similarity of the two words based on a five-point Likert scale

(same, very similar, similar, different, very different). Response time was not recorded, and participants could listen to the audio more than once. A total of 156 responses (12 stimuli pairs x 13 subjects) were recorded.

The second perception task was an identification task. Participants first studied Table 3.3 for as long as needed, then listened to a single stimulus and identified it out of 4 possible choices. The choices were either Pinyin for the Mandarin syllables or IPA for the English phonemes, as shown in Table 3.3. Response time was not recorded, participants could listen to the audio more than once, and could refer back to the table if needed. A total of 156 responses (12 stimuli x 13 subjects) were recorded.

Table 3.3: Table found in Identification Task

'x'	Pinyin, like in the word "xi huan" (喜欢)
'j'	Pinyin, like in the word "ji" (鸡)
'q'	Pinyin, like in the word "qu" (去)
'sh'	Pinyin, like in the word "shi" (是)
'zh'	Pinyin, like in the word "zhi" (只)
'ch'	Pinyin, like in the word "chi" (吃)
' ʃ '	IPA, like in the English word "<u>sh</u>e"
' tʃ '	IPA, like in the English word "<u>ch</u>eese"
' dʒ '	IPA, like in the English word "<u>j</u>uice"

The participants who opted to participate in the production section were tasked with recording 8 audio samples. Table (3.2) shows the 8 stimuli used. Pinyin and characters were provided. A total of 48 responses (8 stimuli x 6 subjects) were recorded.

3.4 Data Analysis

For the perception tasks, the means of both the ratings and accuracy of responses for the distinguishment and identification tasks respectively were recorded. For the production task, four metrics were recorded and compared with the two native speakers to determine the degree of native-like pronunciation of the alveolo-palatal phonemes. The initial vowel F1 and F2 values were recorded to determine the level of vowel fronting, which is a phonological process that occurs in Mandarin following alveolo-palatal phonemes (Hauser, 2020). The peak amplitude frequency of the consonants were measured to determine consonant raising (Edwards & Beckman, 2008). Finally, the consonant length of the fricative [ɕ] and VOT of the affricates [tɕ] and [tɕʰ] were recorded. The mean of these values was recorded and compared to the native speakers' production of the same set of stimuli. The spectral moments are not analyzed in this paper.

CHAPTER 4

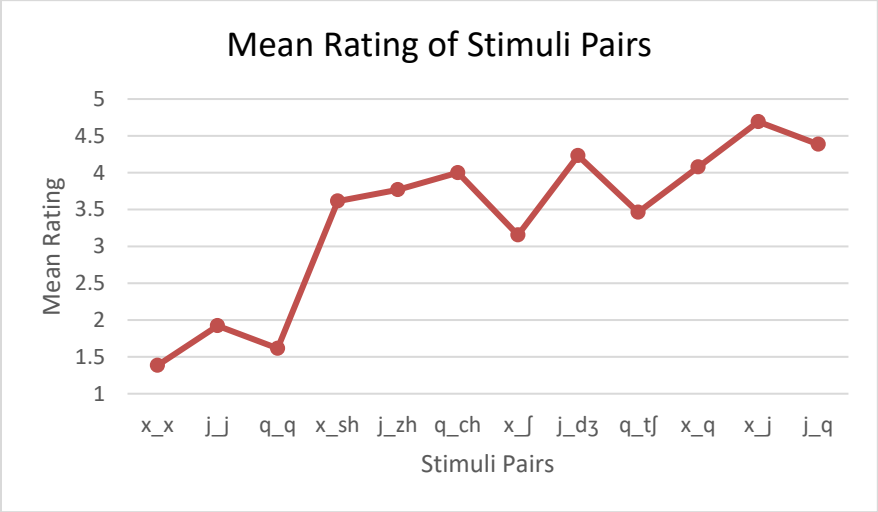
RESULTS

4.1 Perception Tasks

4.1.1 Distinguishment Task

Figure 4.1 graphs the results of the distinguishment task. Each point on the five-point Likert scale has been assigned a value as follows: same: 1, very similar: 2, similar: 3, different: 4, very different: 5. The averages of each value are graphed along the y-axis with the stimuli pairs on the x-axis. Pinyin is used to represent the Mandarin stimuli and IPA for the English stimuli.

Figure 4.1: Mean Rating of Stimuli Pairs



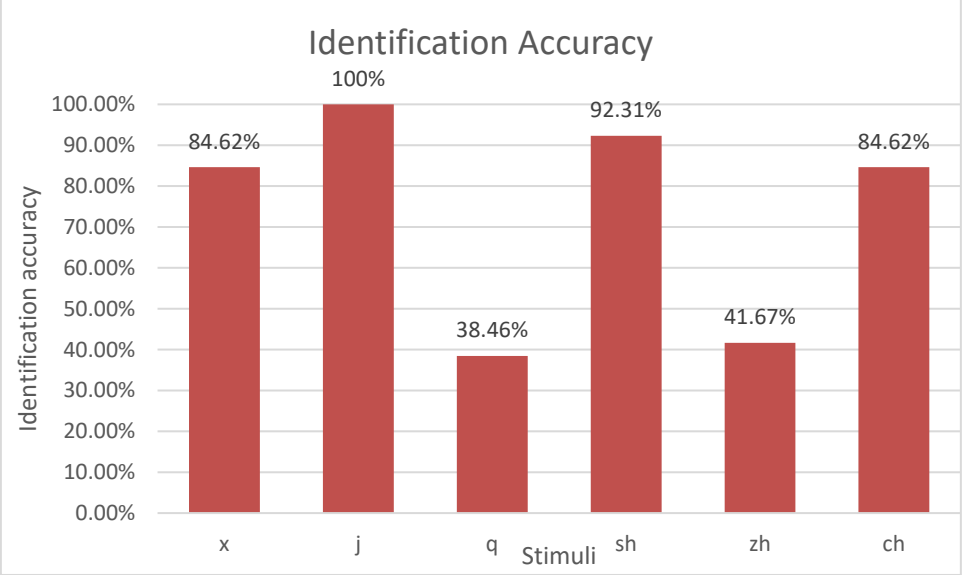
The results were divided into two groups; the first contains the three stimuli pairs that are the same (e.g. x_x), and the second group contains the nine stimuli pairs that are different (e.g. x_sh). A one-way ANOVA test was completed with the second group in order to determine if the lower ratings found in the x_ʃ and q_tʃ pairs were significant. The

test found there is a significant variation in the means of the second group ($p<0.04$). Pair-wise comparisons using the Tukey HSD test indicated the average rating of the x_ʃ pair was significantly lower than the other ratings ($p<0.04$), but not for the q_tʃ pair ($p>0.05$).

4.1.2 Identification Task

Figure 4.2 graphs the results of the identification task. The x-axis shows each stimulus, and the y-axis shows the accuracy of identifying the stimuli after listening to the audio clip.

Figure 4.2: Identification Accuracy



A t-test for dependent means was performed to determine if the lower accuracy of the aspirated palatal q /tɕʰ/ and the unaspirated retroflex zh /tʂ/ was significant.

Both accuracies were significant with $p<0.01$. The test was again performed three more times to determine if aspiration, place of articulation, or manner of articulation significantly correlated with lower accuracy scores. No significance was found in the accuracy of aspirated or unaspirated phonemes ($p>0.05$) or between the retroflex or alveolo-palatal place of articulation ($p>0.05$), but significance was found between the accuracies of affricates and fricatives ($p<0.01$).

4.2 Production Task

Four metrics were collected from each of the alveolo-palatal phonemes produced by both the two native speakers and the 6 participants. For x /ɕ/, q /tɕ^h/, and j /tɕ/, the onset vowel F1 and F2 values were recorded to determine vowel raising and fronting. The peak amplitude frequency of the consonants was also recorded. For the fricative x /ɕ/, consonant length was recorded. For the affricates q /tɕh/ and j /tɕ/, VOT was recorded. The values for these metrics are shown in Tables 4.1, 4.2, and 4.3, along with the mean values for the native speaker group, mean values for the participant group, and p-values comparing the two groups found using a t-test for independent means.

Acoustic analysis of phoneme x /ɕ/ found that fricative length was significantly shorter with the participants compared to the native speakers ($p<0.01$). The mean fricative length for the native speakers was 0.37 seconds while the mean length for the participants was 0.172 seconds. Analysis of the phoneme q /tɕ^h/ found that VOT was also significantly shorter with the participants compared to the native speakers ($p<0.01$). The mean VOT for the native speakers was 0.183 seconds while the mean VOT for the participants was 0.121 seconds. Analysis of the phoneme j /tɕ/ found that the peak amplitude frequency was significantly shorter with the participants compared to the native speakers ($p<0.05$). The mean peak amplitude frequency for the native speakers was 6167 Hz compared to 4549 Hz with the native speakers. No significant difference in vowel quality was determined following any of the three target phonemes.

Table 4.1: Acoustic Analysis of Phoneme x /ɕ/

	NS1	NS2	P1	P2	P3	P4	P5	P6	NS Mean	P Mean	P-value
Vowel F1 (Hz)	403	324	495	424	332	346	359	368	364	387	0.3234
Vowel F2 (Hz)	2309	2028	1619	2055	2519	2650	2323	2542	2169	2285	0.3543
Peak Amplitude Frequency (Hz)	4467	5442	2955	4466	3168	5783	3345	4686	4955	4067	0.1688
Fricative Length (s)	0.26	0.48	0.16	0.14	0.17	0.26	0.15	0.15	0.37	0.17	0.0090

Table 4.2: Acoustic Analysis of Phoneme q /tɕ^{h}/

	NS1	NS2	P1	P2	P3	P4	P5	P6	NS Mean	P Mean	P-value
Vowel F1 (Hz)	401	355	313	281	314	318	359	380	378	328	0.064
Vowel F2 (Hz)	2599	2395	2023	2312	2019	2636	2325	2658	2498	2329	0.2312
Peak Amplitude Frequency (Hz)	6905	4857	3296	4411	4014	5737	4515	5089	5881	4010	0.0674
VOT (s)	0.185	0.181	0.09	0.116	0.15	0.15	0.09	0.12	0.183	0.121	0.008

Table 4.3: Acoustic Analysis of Phoneme j /tɕ/

	NS1	NS2	P1	P2	P3	P4	P5	P6	NS Mean	P Mean	P-value
Vowel F1 (Hz)	440	536	386	357	477	398	393	494	488	418	.0909
Vowel F2 (Hz)	1756	2352	2065	2134	2536	2550	2226	2478	2054	2332	.1205
Peak Amplitude Frequency (Hz)	6661	5672	3053	4731	5403	5344	3587	5176	6167	4549	.0413
VOT (s)	0.082	0.092	0.027	0.04	0.122	0.096	0.077	0.104	0.087	0.077	.3749

CHAPTER 5

DISCUSSION

5.1 Interpretation of Results

5.1.1 Perception Tasks

Results from the perception tasks suggest a partial acquisition of the alveolo-palatal phonemes. Hypothesis two predicted that alveolo-palatal phonemes will be perceived as alveolar phonemes by the English L1 participants. The stimuli pair x_ʃ was perceived to be significantly more similar compared to the other stimuli pair, with an average rating of 3.154, but it was still perceived as "similar" to each other, so participants could distinguish the sounds. Consequently, hypothesis two is rejected. Hypothesis three is partially confirmed by the results of the identification task. Participants had very high identification rates of j /tɕ/ (100% accuracy) and x /ɕ/ (84.62% accuracy), but a significantly lower rate of identification of the phoneme q /tɕ^{h}/ (38.46% accuracy). Thus, participants only struggled to identify one of the three alveolo-palatal phonemes found in Mandarin.

5.1.2 Production Tasks

Table 5.1 displays the three significant results found during the auditory analysis. VOT was not native-like during the production of q /tɕh/, but the difference in VOT between native speakers and the participants for the production of j /tɕ/ was not significant. A cross-language study comparing aspiration between Mandarin and English performed by Chen (2008) found the VOT of unaspirated phonemes in English and Mandarin are similar, whereas the aspirated phonemes in Mandarin are more heavily aspirated compared

to English aspirated phonemes. The findings regarding participants VOT in q /tɕʰ/ and j /tɕ/ in this study can be explained by this difference. The English L1 participants did not produce heavy enough aspiration on the aspirated q /tɕʰ/ because English aspirated phonemes are not as heavily aspirated.

The lower mean peak amplitude frequency of the phoneme j /tɕʰ/ found with the participants compared to the native speakers suggests the place of articulation was more anterior than the native speakers, thus closer to the English alveolar phoneme (Edwards & Beckman, 2008).

Since each of the three alveolo-palatal phonemes was produced in a non-native-like manner, either due to shorter VOT, shorter fricative length, or lower peak amplitude frequency, hypothesis one is confirmed.

Table 5.1: Significant Findings of Each Phoneme after Auditory Analysis

Significant Findings of Each Phoneme	Native Speakers Mean Value	Participants Mean Value	P-Value
x (/ɕ/): Fricative Length	0.37 seconds	0.172 seconds	$p<0.01$
q (/tɕʰ/): Voice Onset Time	0.183 seconds	0.121 seconds	$p<0.01$
j (/tɕ/): Peak Amplitude Frequency	6167 Hz	4549 Hz	$p<0.05$

5.2 Comparison with a Previous Study: Yang & Yu (2019)

In 2019, Chunsheng Yang and Alan C. L. Yu of The University of Connecticut and The University of Chicago respectively completed a study titled “The Acquisition of Mandarin Affricates by American Second Language Learners.” This study completed

perception and production tests for sixteen American English L1 learners of Mandarin. The stimuli in this study were the following Mandarin affricates: [ts], [ts^{h}], [tʂ], [tʂ^{h}], [tɕ], [tɕ^{h}]. Because that study analyzed the acquisition of affricates while the present study analyzed the acquisition of alveolo-palatal phonemes, there was some overlap in the research.

The results found in the identification task of Yang & Yu (2019) were the identification accuracy of the phonemes zh /tʂ/ and q /tɕ^{h}/ were significantly lower than the other phonemes. The results in the present study mirror these results, as the identification accuracy of zh /tʂ/ was 41.67% and the accuracy of q /tɕh/was 38.46%. Both results were significant at $p<0.01$. Both studies also found that the aspiration counterparts of these two phonemes (j /tɕ/ and ch /tʂ^{h}/) had a high identification accuracy, so the difficulty in the identification of zh /tʂ/ and q /tɕh/ cannot be linked with aspiration. As the present study analyzed a place of articulation rather than a manner of articulation, comparisons were able to be made between the identification accuracy of affricates and fricatives. Significance was found between the accuracy of identification of affricates and fricatives ($p<0.01$), at a 66.19% accuracy rate of affricates and an 88.47% accuracy rate of fricatives. It is still difficult to interpret the link between zh /tʂ/ and q /tɕ^{h}/, given that the aspiration and place of articulation of these two phonemes are different, but through the present study the trend has that American English L1 learners Mandarin significantly identify these phonemes at a lower accuracy rate has been reduplicated. These results indicate the need for future studies to identify why native American English speakers struggle to identify these phonemes.

5.3 Implications for Mandarin Pedagogy

The results of this study provide implications for future studies of Mandarin pedagogy, specifically for the creation of new phonetic categories. As previously mentioned, the Speech Learning Model (Fledge, 1995) proposes phonetic categories which are the same or different in an L1 compared to an L2 will be easily acquired, but phonetic categories that are similar will be less easily acquired because the slight differences in those phonetic groups will not be noticed by the learners. This present study supports this theory, as the alveolo-palatal phonemes were not fully acquired, and in fact, participants struggled to distinguish between the Mandarin x /ɕ/ phoneme and the English /ʃ/ phoneme. Given this difficulty in acquiring similar phonemes, further research should be completed to investigate the best teaching methodology to teach learners to notice the differences in these phonemes, thus creating new phonetic categories.

There is currently conflicting research on the topic of explicit phonetic instruction and its effects on phonetic acquisition in an L2. Kissling (2013) investigated the effects of explicit versus implicit phonetic instruction with English L1 Spanish learners but found that both sets of learners improve in pronunciation equally. This study concluded that "input, practice, and/or feedback" may be the more important factors in learning L2 pronunciation rather than a specific instruction methodology. Another study investigated explicit phonetic instruction with Mandarin tones (Wiener, 2020). Results of this study found that English L1 learners of Mandarin noticeably improved their production of tones after explicit instruction and more exposure to input. Both studies suggest improvements in the acquisition of new phonemes and phonetic information after explicit instruction, but more research should be completed to better understand the benefits of different types of

instruction, or if simply active listening and more exposure to variable input are equally as effective as any particular instruction methodology.

5.4 Research Limitations

Major limitations of the present study include a small sample size and low recording quality for the production task. The sample size for the perception tasks was thirteen, which included both beginner and intermediate learners of Mandarin. A larger sample size would not only render results that more accurately represent the population but would also allow for comparison between levels of Mandarin knowledge. This would allow an analysis of how perception and production abilities increase with more exposure and knowledge of Mandarin, which would better inform future studies regarding Mandarin pedagogy. For the duration of this study, access to a phonology lab with recording equipment was not available, thus all recordings were made on participants' individual recording devices. While the recordings were screened to ensure there was no major background noise or that the recordings were too low quality to analyze, higher quality recordings would increase the accuracy of the auditory analysis.

CHAPTER 6

CONCLUSION

The present study investigated the degree of acquisition of the Mandarin alveolo-palatal phonemes [tɕ, tɕʰ, ɕ] by English L1 learners of Mandarin. Two tests were performed to analyze perception skills: a distinguishment task and an identification task. Acoustic analysis using four metrics was performed after a production task and compared to native speakers to determine the level of native-like pronunciation. The results indicated participants had some difficulties distinguishing the x /ɕ/ phoneme from the English /ʃ/ phoneme but were still able to distinguish the sounds. Participants had a significantly low identification rate of the j /tɕ/ phoneme. Finally, the production of the alveolo-palatal phonemes was not native-like based on the VOT, fricative length, and peak amplitude frequency. These results suggest a need for future studies to analyze the effects of various phonetic instruction on the acquisition of Mandarin phonemes.

REFERENCES

Bialystok, E. (1997). The structure of age: in search of barriers to second language acquisition. Second Language Research, 13(2), 116–137.

Chen L, et al. (2008). A Cross-Language Study of Stop Aspiration: English and Mandarin Chinese. *2008 Tenth IEEE International Symposium on Multimedia*, 556-561.

Chen, L., Zhou, X., & Wang, Y. (2016). The effectiveness and mechanism of pinyin in Chinese learning. 5, 715–720.

Edwards, J.R., & Beckman, M. (2008). Methodological questions in studying consonant acquisition. *Clinical Linguistics & Phonetics, 22*, 937 - 956.

Flege, J. E. (1995). Second Language Speech Learning Theory, Findings, and Problems. *Speech Perception and Linguistic Experience: Issues in Cross-Language Research*

García Mayo, María del Pilar & Garcia Lecumberri, Maria. (2003). Age and the acquisition of English as a foreign language.

Hauser, I. (2020. Coarticulation with alveopalatal sibilants in Mandarin and Polish: Phonetics or phonology? The University of Minnesota.

Hyltenstam, K. & Abrahamsson, N. (2000), Who can become native-like in a second language? All, some, or none?. *Studia Linguistica*, 54: 150-166.

Kim, T. E. (2018). Comparative study of Mandarin phonology and English phonology. *Mandarin Loanwords*. 29–57.

Kin, W. (2016). Production and Perception of Mandarin Sibilants by Cantonese Speakers. *Department of Linguistics and Translation,* City University of Hong Kong.

Kissling, E. (2013) Teaching pronunciation: Is explicit phonetics instruction beneficial for FL learners? *Modern Language Journal*, 97: 720-744.

Lado, R. (1957). Linguistics across Cultures. University of Michigan Press, Ann Arbor.

Lenneberg EH (1967) Biological foundations of language. New York: Wiley.

Ou, S., & Guo, Z. (2014). Mandarin Retroflex Sounds Perceived by Non-native Speakers. 39-76.

Shih, Y., & Eunjong K. (2011). Perception of Mandarin Fricatives by Native Speakers of Taiwan Mandarin and Taiwanese. *North American Conference on Chinese Linguistics*, 110–119.

Stockwell, R., Brown, J. D., & Martin, J. D. (1965). The grammatical structure of English and Spanish. *Chicago: University of Chicago Press*.

Wiener, S. et al. (2020). Do Explicit Instruction and High Variability Phonetic Training Improve Nonnative Speakers' Mandarin Tone Productions? *The Modern Language Journal*, 104: 152-168.

Yang, C., & Yu, A. (2019) The Acquisition of Mandarin Affricates by American Second Language Learners. *Taiwan Journal of Linguistics,* 17.2, 91-122.

Zhou, M., & Hongkai, S. (2004). Language policy in the Peoples Republic of China: theory and practice since 1949. Boston, MA: Kluwer Academic.

BIOGRAPHICAL INFORMATION

Brooke Sherrell is a Senior Honors student at the University of Texas at Arlington. She majors in Critical Language and International Studies with a Concentration in Chinese, minors in Linguistics, minors in Chinese Localization and Translation, and is earning a TESOL Certification through UTA. Throughout her academic career, she has shown strong interest towards the Chinese language through serving as the President of the Chinese Culture and Language Association and working as a Chinese tutor in the Department of Modern Languages. Post-graduation, she plans to live in China to teach English in order to continue her Chinese studies and to gain valuable teaching experience. In the future, due to her interest in the Chinese language and general language pedagogy, she plans to work as a language teacher.