國立交通大學

外國語文學系

外國文學與語言學碩士班

碩　士　論　文

上海四歲兒童在華語子音，母音及聲調上的聲學表現

Acoustic realization of consonants, vowels and tones in Mandarin by 4-year-old Shanghainese children

研究生: 林郡儀

指導教授: 賴郁雯　教授

中　華　民　國　一　〇　〇　年　五　月

上海四歲兒童在國語子音，母音及聲調上的聲學表現

Acoustic realization of consonants, vowels and tones in Mandarin by 4-year-old Shanghainese children

研 究 生：林郡儀 Student：Chun-Yi Lin

指導教授：賴郁雯 Advisor：Yuwen Lai

國 立 交 通 大 學

外國語文學系外國文學與語言學碩士班

碩 士 論 文

A Thesis

Submitted to Graduate Institute of Foreign Literatures & Linguistics

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Linguistics

May, 2011

Hsinchu, Taiwan, Republic of China

中華民國一○○年五月

上海四歲兒童在國語子音，母音及聲調上的聲學表現

學生：林郡儀 **指導教授：賴郁雯**

國立交通大學 外國語文學系外國文學與語言學 碩士班

中文摘要

本研究旨在探討上海四歲兒童對於華語之子音、母音及聲調之聲學表現，並討論第一語言對第二語音習得之可能影響。本研究採用了跨語音音韻遲緩八十個目標字詞包含了華語的各個音段及主要結構，在不同字詞出現至少兩次，不同聲調出現至少一次。所有發音由一位實驗者轉寫紀錄，另外聲學分析可確認兒童是否念出了目標字詞。聽音轉寫的結果包含了九種錯誤：（1）非捲舌音發成捲舌音（2）捲舌音發成非捲舌音(3) 以其他音代替正確的音（4）鼻音混淆，（5）/h/ 錯誤，（6）/tɕʰ, tɕ, ɕ/ 混淆，（7）母音混淆，（8）聲調錯誤，以及（9）其他少見之錯誤。 本研究的聲學分析只包含了錯誤最多的捲舌音、母音以聲調的表現。 實驗結果顯示，有聲捲舌音的 F3 在正確發音時，如預期的出現了下降的現象，但在錯誤發音時就沒有 F3 下降的情形。另外在無聲捲舌音的結果顯示，當捲舌音在單元音(/i/, /e/, /u/)及 VN (/en/)之前，F3 下降的情形也只會出現在正確的發音，錯誤發音時則無此情形。 但有趣的是，當捲舌音後面接了雙元音(/ua/, /ou/)或 VVN (/uaŋ/)時，F3 下降的情形會同時出現在正確及錯誤的發音。頻譜動差 M1-M4 的結果顯示捲舌音的 M1 和 M2 的值比非捲舌音的值還要低。而 M3 和 M4 的值在捲舌音的情況下高於非捲舌音。在母音空間的聲學表現上，上海兒童的母音空間比上海及華語成人的母音空間小。由於上海成人的母音空間比華語成人的母音空間還小，因此上海兒童較集中的母音空間被推測為是受到第一語言的影響。在聲調方面，Jongman (1989)的結果顯示，華語成人三聲和二聲的聲調時長是最長的，四聲最短，在此篇研究結果，上海兒童在聲調時長和 F0 曲線上的表現在華語成人一樣。總結來說，實驗結果顯示上海兒童的發音同時受到一般普遍性發展及第一語言的影響。

Acoustic realization of consonants, vowels and tones in Mandarin by 4-year-old Shanghainese children

Student: Chun Yi Lin **Advisor: Dr. Yuwen Lai**

Graduate Institute of Foreign Literatures & Linguistics
National Chiao Tung University

英文摘要
ABSTRACT

This study examined the phonetic realization of 4-year-old Shanghainese children in Mandarin tones, consonants and vowels and the possible effect of their first language (Shanghainese) on their second language (Mandarin). 80 target words contain each segment, major word structure at least twice in different words and tone sequences at least once. The production data were transcribed as misarticulated or not by one experimenter and the acoustic analysis was further conducted to make sure that the subjects read the target words. Results of transcription included nine errors: (1) Non-retroflex were produced as retroflex, (2) Deretroflexion, (3) Substitution, (4) Nasal confusion, (5) /h/ substitution, (6) /tɕʰ, tɕ, ɕ/ confusion, (7) Vowel confusion, (8) Tone confusion, and (9) Others. Only retroflex, vowels and tones were included in the acoustic analysis. The result of F3 in voiced retroflex consonants indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for three different rhymes (VVN, VV, V). Results indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for retroflexes precede V (/i/, /e/, /u/) and VN (/en/) rhymes. However, F3 dipping was found in correct as well as incorrect pronunciation of retroflex precedes VV (/ua/, /ou/) and VVN (/uaŋ/) rhymes. By measuring M1-M4 of retroflexes, the result of M1-M4, M1 and M2 value of correct pronunciation (retroflex) is significantly lower than that of incorrect pronunciation (nonretroflex). M3 and M4 of correct pronunciation is higher than that of incorrect pronunciation. The vowel space of Shanghainese children was found narrower than that of Shanghainese and Mandarin adults. Given the fact that Shanghainese vowel space is found narrower than that of Mandarin's, narrower vowel space of Shanghainese children is postulated that because of the effect of L1 transfer to L2. The adult norm for tone duration is T3/T2 longer than T1, and T4 is shortest (Jongman, 1989), and Shanghainese children show the similar pattern. These children managed to utilize F0 cue and duration cue. In conclusion, results indicated that children's production was influenced by universal constraint and L1 background.

誌　　謝

在我的碩士生涯裡，首先要感謝的人當然是我的指導教授賴郁雯老師。她的出現真的像是天上掉下來的天使，讓我重新找到了前進的力量。老師的幽默風趣也讓我在學習過程中，享受到了更多不同的樂趣。老師親切的常讓學生們到她的辦公室去聊天，也用心的分享她的一切，讓所上也多了很多笑聲。最重要的是，老師不辭辛勞的閱讀我的碩論，並且給予指導指正，讓我也學到了撰寫論文的方法，尤其是老師口中的「龍捲風」結構，由大範圍寫到小細節的寫作方式，讓我在統整自己思路時更有方向和系統。能有一個這樣的指導教授，我真的很幸運！還有我的口試委員：張顯達老師和何延光老師，謝謝他們願意當我的口試委員，也謝謝他們給了我很多建議，讓我學習到很多東西，也讓我知道自己還有很多東西需要改進。還要謝謝雅玲，我常常都有很多不知道的事情要問她，她都很親切的回答我，超有耐心的！系辦有她在，有很多疑難雜症都解決了！

另外也要謝謝同學的鼓勵，尤其是博任總是被我吵著問問題，他真的是我碩士生涯裡的大貴人，從碩一到畢業都一直幫助我、鼓勵我，能有他的幫忙真的是一件很幸運的事。另外晉廷也幫我找了好多我遍尋不著的文章，讓我在找參考資料時不會那麼無助。還有在同一個實驗室的米琪，她要辛苦的報帳，但都和我互相分擔一些助理的工作，在我口試當天，還幫我泡茶和調整投影片，真的救了我一命！鈺楨在碩一時給我的鼓勵和幫忙；皓志總是怕我心情不好，要跟我聊天；伊凡、婉淇也都會適時的出現，大家的鼓勵都讓我覺得很溫暖。

當然也要謝謝我的家人，一路以來一直支持我，給我很多力量和溫暖。謝謝姿菁和喻雁，一直容忍我不能陪她們。謝謝慧中讓我的住宿生涯充滿樂趣，陪我吃飯聊天，被我慫恿去吃消夜，還要聽我吵鬧，有一起瘋的室友讓我在寫碩論時減少了許多壓力。要謝的人真的太多，在交大的這三年，志謙、丙昇、權訓、文鈞、柏宏，謝謝你們幫我加油、陪我聊天散心、聽我抱怨，謝謝所有幫助我、鼓勵我、以及陪在我身邊的人！

Table of content

Table of Figures

Table

ABSTRACT

This study examined the phonetic realization of 4-year-old Shanghainese children in Mandarin tones, consonants and vowels and the possible effect of their first language (Shanghainese) on their second language (Mandarin). 80 target words contain each segment, major word structure at least twice in different words and tone sequences at least once. The production data were transcribed as misarticulated or not by one experimenter and the acoustic analysis was further conducted to make sure that the subjects read the target words. Results of transcription included nine errors: (1) Non-retroflex were produced as retroflex, (2) Deretroflexion, (3) Substitution, (4) Nasal confusion, (5) /h/ substitution, (6) /tɕʰ, tɕ, ɕ/ confusion, (7) Vowel confusion, (8) Tone confusion, and (9) Others. Only retroflex, vowels and tones were included in the acoustic analysis. The result of F3 in voiced retroflex consonants indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for three different rhymes (VVN, VV, V). Results indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for retroflexes precede V (/i/, /e/, /u/) and VN (/en/) rhymes. However, F3 dipping was found in correct as well as incorrect pronunciation of retroflex precedes VV (/ua/, /ou/) and VVN (/uaŋ/) rhymes. By measuring M1-M4 of retroflexes, the result of M1-M4, M1 and M2 value of correct pronunciation (retroflex) is significantly lower than that of incorrect pronunciation (nonretroflex). M3 and M4 of correct pronunciation is higher than that of incorrect pronunciation. The vowel space of Shanghainese children was found narrower than that of Shanghainese and Mandarin adults. Given the fact that Shanghainese vowel space is found narrower than that of Mandarin's, narrower vowel space of Shanghainese children is postulated that because of the effect of L1 transfer to L2. The adult norm for tone duration is T3/T2 longer than T1, and T4 is shortest (Jongman, 1989), and Shanghainese children show the similar pattern. These children managed to utilize F0 cue and duration cue. In conclusion, results indicated that children's production was influenced by universal constraint and L1 background.

Chapter 1. Literature Review

The Universal Grammar (UG) claims that human beings have principles and parameters that make up the abstract knowledge of language under all language varieties (Ingram, 1989). He further defined that the principles are grammatical rules which make human languages similar to one another. On the other hand, these parameters are different in certain restricted ways, thus showing language specification (Ingram, 1989). Similar to Ingram, Slobin (1985) also defined the common and general patterns in child language development as "developmental universals", and the other which can be seen patterns of acquisition of specific properties vary from language to language is called "language-specific patterns" or "particulars" (Slobin, 1985). By examining comparable acquisition data from children of different language backgrounds, researchers could identify the developmental universals. By also comparing different languages, researchers could pinpoint language-specific developmental patterns (Slobin, 1985).

1.1 Overview

1.1.1 Developmental Universals

Consonant

In terms of consonant acquisition, O'Grady et al. (1997) suggested children begin to experiment with and gain control over their vocal apparatus in babbling. Despite the fact that they are exposed to different languages, children from different linguistic communities exhibit significant similarities in their babbling. O'Grady et al (1997) provided the tendencies of occurred consonant frequency based on fifteen different languages (English, Thai, Japanese, Arabic, Hindi and Mayan, etc). The consonants most frequently found were: /p/, /b/, /m/, /t/, /d/, /n/, /k/, /g/, /s/, /h/, /w/ and /j/. However, the consonants most infrequently found were: /f/, /v/, /θ/, /ð/, /ʃ/, /ʒ/, /ʧ/, /ʤ/, /l/, /r/ and /ŋ/.

In general acquisition, *markedness* is the term used to refer to frequency facts. White (1889) suggested that *markedness* could be divided into marked (peripheral phenomena) and unmarked (core grammar). One could simply define markedness in terms of what is or is not present in UG. Any principles or parameters proposed in linguistic theory would be unmarked, whereas other linguistic structures would be marked. Chomsky & Halle (1968) suggested that the frequency fact is that most frequent sounds/features are maximally unmarked and would be acquired first. They argued that environmental lexical/sound frequency affects acquisition order. Therefore, the marked features (affricates and aspiration) are later acquired and easily replaced by unmarked features.

Locke (1983) suggested that children produce stops, nasals and [h] more frequently, whereas fricatives and liquids are generally avoided in the stage of babbling, no matter what language the children is exposed to. In general, children master simpler distinctions before more complex ones. Jakobson (1968) proposed that the distributions of cross-linguistical speech sounds and the order in which different sounds are acquired show regular relationships. Jakobson proposed that high frequent sounds like voiceless stops, nasals and the vowel [a], which are found in virtually all languages, were the first to be acquired.

Different acquisition patterns might due to frequency and markedness, as well as articulatory control (Clark, 2009). The articulatory controls, for example, the voiced stops are acquired earlier than voiceless stops in babbling which not coordinated with the concept of markedness. Voiceless stops are more frequent and acquired earlier than voiced stops. This may due to the insufficient articulatory control for producing a voiceless sound.

Vowels

Vowel development is generally considered to be easier and acquired earlier than consonants. Vowels also seem to be less misarticulated when being produced. However, Paschall (1983) reported that the mean accuracy of vowel production in 20 children was below 60 %. Pollock and Keiser (1990) investigated 15 phonologically disordered children and found that half of the subjects produced vowels with less than 80 % accuracy and 14 of subjects had diphthong errors. Here provide different aspect to defined vowel acquisition.

1.1.2 Language-specific developments

Languages differ in their sound inventory and phonotactics. For example, the lexical tones in Mandarin which could identify word meanings do not exist in English. Children have the capacity to get statistical distribution of auditory phonetic information in their linguistic input, and create phonetic categories when they are 6-8 months old (Hannahs et al, 1997).

1.2 Phonological development

1.2.1 Segment

Consonant

Voicing

Voiceless means that the vocal folds are held far enough apart to allow airflow through the glottis. In contrast, "voiced refers to normal vocal fold vibration occurring along or all of the length of the glottis" (Clark et al., 2007). In terms of voicing, VOT is defined as the time between the onset of glottal pulsing and the release of the initial stop consonants. Such as long lead (prevoiced), short lag (voiceless), and long lag (voiced) refer to the degree of VOT. Preston, Yeni-Komshian, and Stark (1967) suggested that voiceless stops tend to predominate in babbling, and voiced (aspirated) stops emerge later (age 2 or older). Children generally acquire voiced sounds before voiceless sounds (Bernhardt & Stemberger, 1998). They also suggested that children acquire unaspirated sounds earlier than aspirated ones. Stemberger (2008) suggested that voicing is acquired earlier than voiceless and unaspirated is acquired earlier than aspirated.

Manner of articulation

Manner of articulation refers to the degree or extent of a constriction and the way that the constriction is formed in the vocal tract. When it comes to the manner of articulation, it is generally known that the stops are acquired first, followed by nasals, fricatives, and lastly the affricates. When children speak, stops, nasals and the [h] sound are frequently appearing. However, the fricatives and liquids are constantly avoided (Goodluck, 1991). Su (1985) investigated two Mandarin-Taiwanese bilingual children. Su also proposed that the reason children replace affricates with stops rather than fricatives might be related to the high frequency of occurrence of affricates in Mandarin. The reason of this acquisition order might be, as mentioned above, the more frequent sounds are acquired earlier than less frequent sounds. Also the physiological factors also might influence the acquisition. Stops involve a complete blockage of air; Fricatives involve a loose occlusion. For affricates, there is a complete closure followed by a gradual fricative release. To produce segmental contrasts requires different kinds of combinations between different mechanisms. These varied movements of articulation might cause children to acquire affricates later than other manner of articulation.

Bernhardt and Stemberger (1998) proposed that stops, nasals and glides were found in early inventories and occur with a high frequency. Fricatives, affricates and liquids tend to be mastered later (up to age 8 or 9) which is also supported by Stemberger (2008)'s study. Stemberger also suggested that stops, nasals and [h] are frequently found. Stops occur most frequently, followed by nasals, then fricatives, and finally affricates.

Jakobson (1968) proposed that the occurrence of back stops is less frequent than front stops, and that fricatives occur less frequently than stops. Also, he argued that a language may have front stops without back stops, or have stops without having fricatives, but not vice versa. According to the frequency facts, he proposed that front stops are acquired before back stops and stops are acquired before fricatives.

Place of articulation

Place of articulation refers to that while producing a particular consonant, the constriction which is located at some point in the vocal tract. In terms of place of articulation, Stemberger (2008) suggested that dentals (not alveolar) is the most frequently found in babbling, followed by bilabials then velar. Labiodentals, interdentals, alveolars, retroflexes, palatoalveolars, palatals, uvulars, pharyngeals are rare or non-occur in babbling.

Bernhardt and Stemberger (1998) suggested segments containing major articulator features ([Labial], [Coronal], [Dorsal]) were found in very early development, and features ([anterior], [grooved], [distributed]) were established later. According Bernhardt and Stemberger, the tongue position will influence the acquiring order. Front position is acquired earlier than back position. It is also suggested that the insufficient motor control influences early child phonological development (Wode, 1997). The articulator is not well controlled by children. Vihman (1996) suggested that it is because the front position is easier to see in children, the acquisition of front position is earlier than back position. At an early stage (about four to six months olds), the front voiceless stops, nasals and the vowel [a] will appear. In addition, dentals (not alveolar) appear next, then bilabial and velar. However, labiodentals, interdentals, alveolar, retroflexes, palatoalveolars, palatals, uvular and pharyngeal are rare or non-occurring at such an early stage. Dorsals are the least likely default place feature to appear in babbling and early words to be produced (Clark, 2009).

One particularly place of articulation which attracts scholarly attention is retroflex. Hamann defined all retroflexes share three articulatory characteristics: raising of the tongue tip, a sublingual cavity, and the tongue body retraction. According to Hamann (2005), from the 317 languages of the UPSI database, retroflex occurs infrequently: voiceless retroflex stop (8.5%), voiced stop (7.3%), voiceless retroflex fricative (5.3%) voiced fricative (0.9%), and retroflex nasal (5.9%). There are two voiceless retroflex affricates [ʈʂ, ʈʂ^{h}], a voiceless retroflex fricative [ʂ], and a voiced retroflex fricative [ʐ] in Mandarin which are more infrequently. The articulatory characteristics of Mandarin retroflexes move the tongue tip toward the hard palate rather than curling. Jeng (2004) suggested that retroflex in Mandarin is difficult to acquire and is thus the latest acquired phoneme. Besides occurrence frequency, other researchers suggested that retroflexes are acquired later than other segments. Jeng (2004) studied 301 Mandarin-speaking children with ages ranging from 2.5 - 6. She suggested that the most difficult consonants were coronals (includes retroflexes). She also discovered that bilabials were acquired when the child reaches an age of 2.5, and that retroflexes were acquired after the age of six (Table. 1). The infrequently and the articulatory complexity of retroflex might be too difficult for children in early acquisition stages.

Table 1. Age of phoneme acquisition by Mandarin-speaking children.

Age/ researches	Wang et al. (1984)	Jeng (2004)
Before 3	p,tʰ,pʰ,h,kʰ,k,t,n,m	p
3	k, tɕʰ,ts	n,k, tɕ
3.5	ɕ, tsʰ, s, f	t, l, h
4	tɕ	m, tʰ, kʰ, tɕʰ
4.5		
5		pʰ
5.5		
6		f, ɕ
After 6	ʂ, ʐ, tʂ, tʂ ʰ	tʂ, tʂ ʰ,ʂ,ʐ, ts, tsʰ, s

Retroflex - voiced

According to Hamann (2003) a coronal articulations generally result in a raised F2. However, when the place of constriction occurs from a front position to a back position, results showed a lowered F3. The articulatory properties that can vary in retroflexes are: the size of the sublingual cavity, the degree of tongue retraction, and additional lip-rounding. The size of sublingual cavity is partly vowel-dependent. The back vowels will have a bigger cavity than front vowels. This sublingual cavity has two acoustic effects (Stevens, 1998), which are the introductions of a low-frequency resonance F_R and the introduction of a zero Z_o. This F_R is about 1800Hz and located between F2 and F3; therefore it is often associated with the F2 of the adjacent vowel. The Z_o is located around 2000Hz. The effect of Z_o is F3 amplitude weakening and higher formants **(Figure 1)**. F_R and Z_o merge at the release of the retroflex closure, and the amplitude of F3, F4 and F5 will increase.

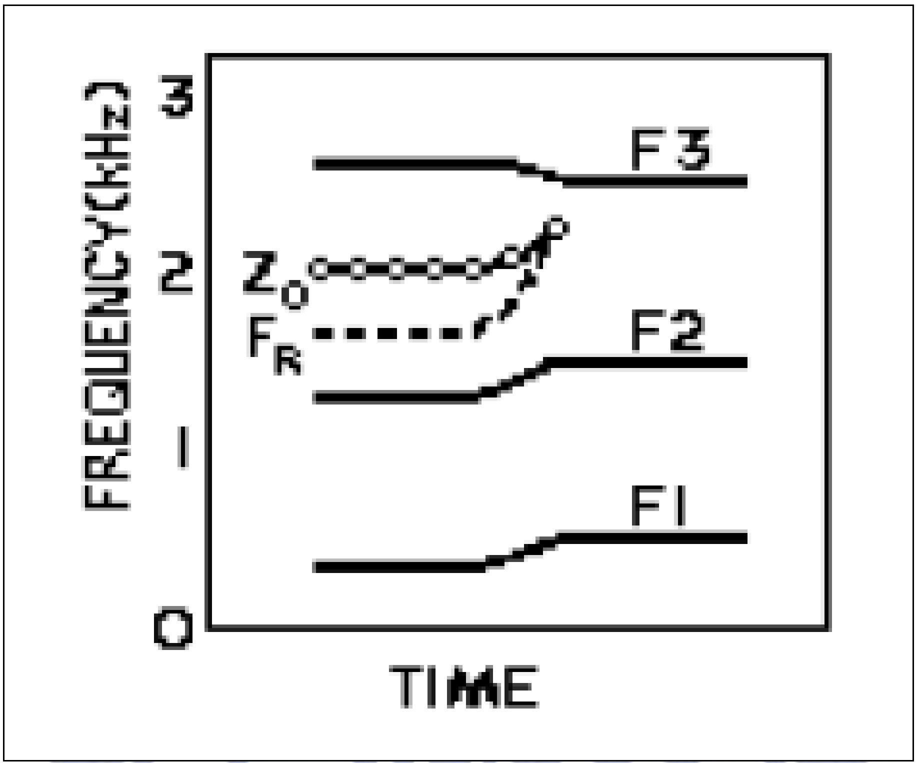


Figure 1 F1, F2 and F3 trajectories for a retroflex released into a vowel (Stevens 1998)

Thus, the additional formant and the zero introduced by the sublingual cavity are interpreted as a high F2, low F3 and higher formants during the consonant constriction. A retroflex consonant is expected to show a lowered F3 while the F2 value is depending on the vowel context. Thus retroflexes show a lowering from the vowel to the consonant and a raising from the consonant constriction to the vowel. The lowered trajectory of F3 is what actually distinguishes retroflexes from other coronals. However, for stops, parts of the formant trajectories are not observable, due to the closure silence and the stop burst. Thus in the present study, the examiner also measured F3 of the following vowels and the spectral moment (which will be introduced in the following section) of voiceless retroflex .

Hamann (2003) suggested that there is variation in degree of lowering F3, depending on the language specification and the vowel context. For example, a non-posterior retroflex has a less lowered F3, and the retroflex which follows by a back vowel shows a lower F3 than those in other context.

McGowen, Nittrouer and Manning (2004) investigated the development of /ɹ/ in eight English speaking children (14- 31 months). Subjects were recorded every 2 months interacting with their parents. In this study, Richard examined the prevocalic, postvocalic and medial syllabic /ɹ/. When investigating /ɹ/, F2 and F3 as well as F3-F2 in /ɹ/ and neighboring vowel should be measured. Here the only focus was on prevocalic /ɹ/. Results indicated that prevocalic /ɹ/ is difficult and are acquired later for children. To explain this result, Richard suggested that the prevocalic /ɹ/ requires two substantial tongue constrictions: in the oral cavity and in the pharynx. He provided Vorperian and Kent’s (1995) study to explain that children's tongue fills the oral cavity. Thus, articulating the prevocalic /ɹ/ properly is difficult to attain for children. The massive tongue body and a small pharyngeal cavity would make young children difficult to form both a palatal and pharyngeal constriction with the tongue.

Retroflex - voiceless

Jeng (2004) examined the spectral characteristics (spectral moment measurement) of retroflexed fricative and affricates in Taiwan Mandarin by using TF 32 acoustic analysis software developed by Milenkovic (2004) (Figure. 2). Jeng's study aims to compare moment parameters of retroflex and non retroflexes fricatives and affricates in Mandarin, and investigate the four moment parameters in distinguishing the retroflexion contrast in Taiwan Mandarin. 30 adults were recorded (20-23 years old), and four parameters measured: M1 (mean), M2 (variance), M3 (skewness), and M4 (kurtosis). M1 represents the mean of central gravity of spectrum, M2 is the energy dispersion (if the diversity is centralized, M2 value will be small), M3 is spectral tilt (if the power is located in low frequency, M3 is a positive value), and M4 is the peakedness of spectrum (if the distribution curve is sharp, M4 value would be big, which means that spectral power is centralized near the mean value). The spectral moments and the noise duration of fricatives and affricates were also measured. Speakers need to answer 21 quiz activities and read 42 disyllabic words. Results indicated that there were significant differences between the four spectral moments and noise duration, as well as between the retroflexed and the non-retroflexed. Compared to the non-retroflexed, the retroflexed fricatives and affricates have lower M1, and M4, and greater M2, and M3. Jeng suggested that retroflexes in Mandarin have lower central gravity frequency, low but weak energy and wide energy distribution.

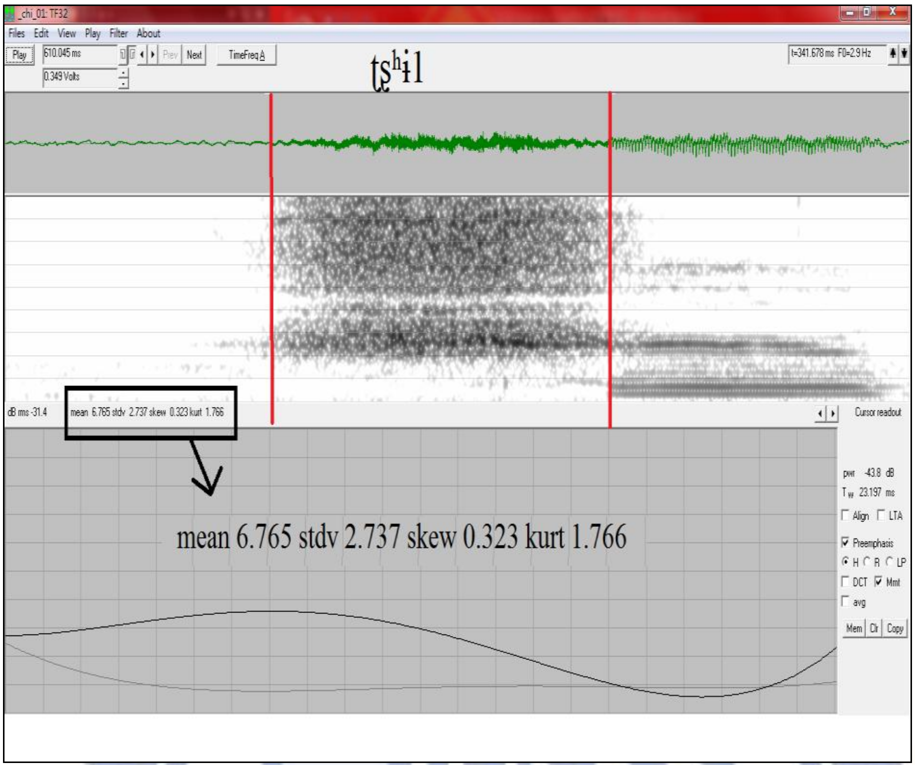


Figure 2 Spectral moment measurement and the four parameters of tʂʰɨ1 "eat". From top to down is wave form, spectrogram, then the spectral moment during the frication portion.

Vowels

Features: Front / Back ; High / Low; Round / Unrounded

Bernhardt and Stemberger (1998) summarized the results of previous studies in Germanic languages and indicated that the basic vowel triangle [i], [a], [u], mid back tense [oʊ] and mid central [ʌ] are the earliest mastered, followed by back and central lax vowels [ʊ], [ɔ], [ə] and low front [æ], front non-low vowels [eɪ], [ɛ], [ɪ] and rhotic vowels were mastered last. Selby, Robb and Gilbert (2000) investigated four English-speaking children aged from 15 to 36 months. Vowels [a], [i], [u], and [ʌ] were the first appear at age of 15 months. [ɔ] and [æ] were appeared at 18 months. [o] and [e] showed up until the age of 21 months. Levelt (1994) investigated 12 Dutch subjects and indicated a similar vowel development like English: The vowel triangle /i/, /a/, /u/ are early mastery, followed by [ɪ], [e], [o],[ɔ] , mid front vowel [ɛ] is mastered last.

On the other hand, some researches (Bernhardt & Stemberger, 1998) showed different patterns for vowel development. The low vowels [æ], [ɑ], [a], mid vowels [ʌ], [ɛ] and high vowel [i] are early mastered, back rounded vowel [u] is late mastery. Ito (1990) reported that the Japanese children acquire [a], [e] and [o] before the vowel [i]. Fee (1991) investigated Hungarian-learning and Spanish-learning children and suggested that vowel [a], [e] and [o] were first acquired and [u] was the last acquired vowel. Different studies show that the order of vowel development is hardly to be fully confirmed.

monophthong/ diphthong/ triphthong

Strokes and Wong (2002) studied Cantonese-speaking infants (10-27 months), and suggested that monophthongs are acquired first (at 15-18 months) followed by diphthongs (at 20-23 months), and finally trip-thongs (about 24-27 months). Chen (2004) investigated Mandarin-speaking children (from birth to 5 years old) by a longitudinal observation of vowel production. Vowel production was analyzed by transcription and acoustic measurement. Results indicated that monophthongs are first mastered by children (6-23 months). Diphthongs and triphthongs were found at 12 months but not well controlled, and which are more stable until 36 months old. The nasal-final (V+N, VVN) frequently occurs but is not well mastered during 48-60 months old. However, the articulatory control is more fluently.

Vowel space

The present study further examine vowel space development and the possible effect of young learners' L1 and L2. Peterson and Barney (1952) investigated general American English vowels produced by 61 adults and 15 children. Token lists 10 monosyllabic words: *heed, hid, head, had, hawed, hood, who'd, hud, heard.* Results indicated that the vowel space of children is higher than that of adults, but the vowel space of children is not wider than that of adults. **(Figure. 3)**.

On the other hand, Vorperian and Kent (2007) integrated 14 different English dialects published acoustic data on the development of vowel production and draw the vowel spaces. Subjects include 886 children (age 0.25 month to 18 years) 305 adults (19 years to 80 years). Results indicated that the formant frequency and F1-F2 area gradually reduce with age **(Figure 4-5)**. van der Stelt et al. (2006) investigated five Dutch children acquiring 12 vowels and compared it to the Russian results in order to interpret the influence of the number of vowels per language. The targeted words were six Russian vowels: /a/, /o/, /u/, /i/, /e/, and /y/. Results indicated that the vowel space becomes narrower as age increases. **(Figure 6-7)** van der Stelt et al. suggested that the vowel space for adult is smaller than for the infants due to the infant's smaller oral cavity which consequently produces higher frequencies.

Interestingly, Buhr (1980) examined the vocal production of an infant (aged 16-64 weeks) for longitudinal perceptual and acoustic analysis. First three formants for English vowels: /i/, /I/, /e/, /ɛ/, /æ/, /a/, /o/, /ʌ/, /ɔ/, /u/, and /U/ were measured. Results indicated that the front vowels seem to be acquired before the back vowels. In terms of vowel space, by 16 weeks, vowel space was found considerably overlapped. By 24 weeks, individual vowels become wider. By 41 weeks, vowel space became wider. By 62 weeks, the spread of individual vowels began to narrow, but the whole vowel space of all vowels is the same as those in week 41. However, by 64 week, the vowel space become narrower and more stable. It seems that vowel space would eventually become narrower as age increases.

In terms of vowel space, previous studies show three different patterns: a) the vowel space of children is *higher* than that of adult, b) the vowel space of children gets *narrower* as age increases, c) the vowel space of children *wider* at first but is *narrower* as age increases.

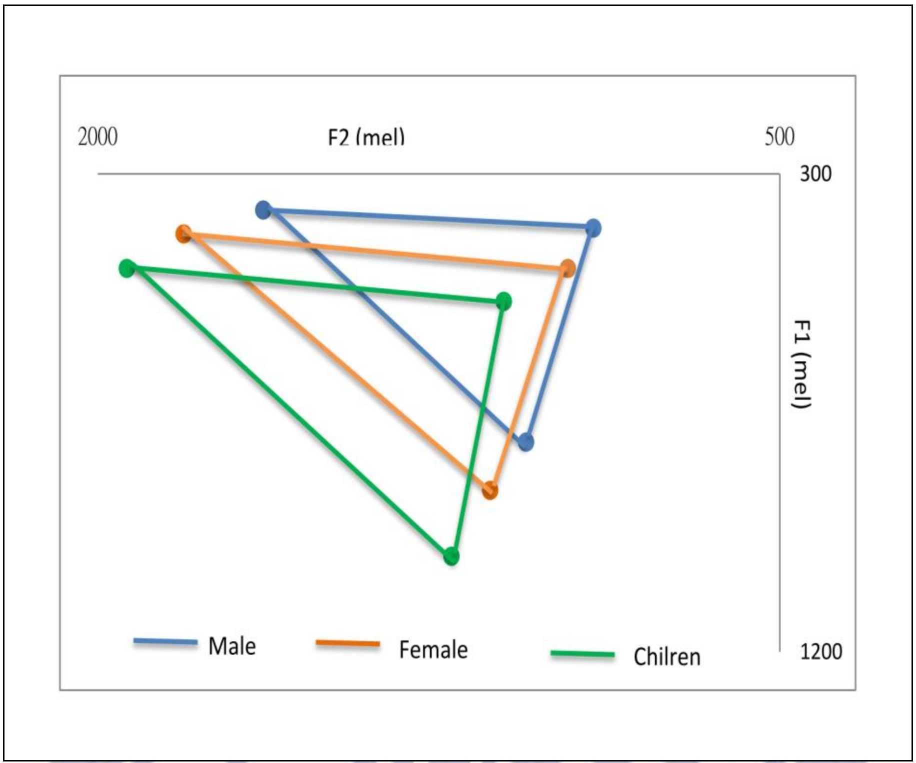


Figure 3 Vowel spaces of children and adults. (Data from Peterson & Barney, 1952)

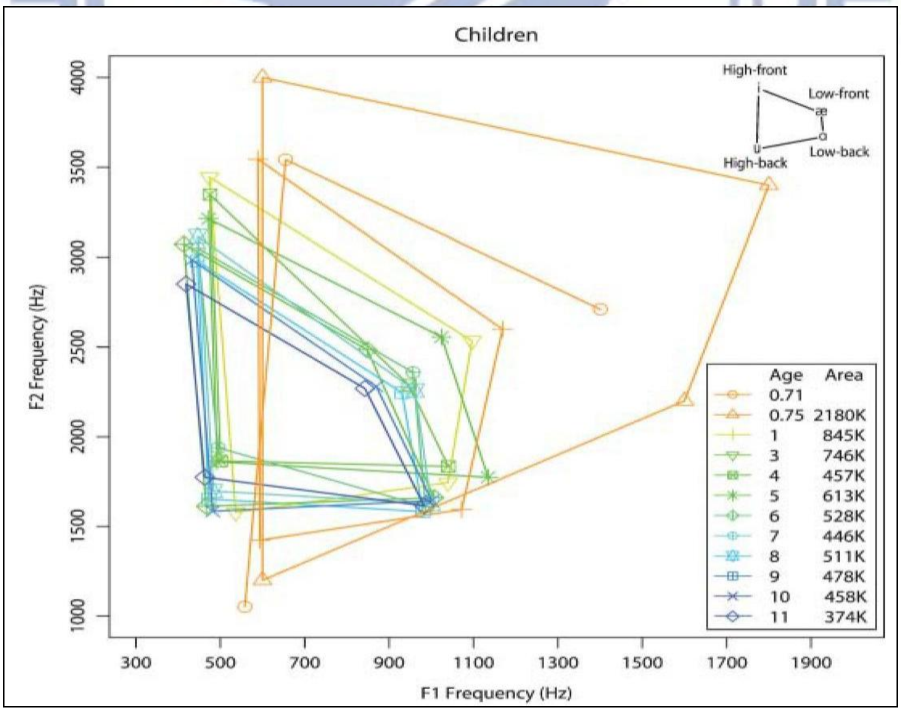


Figure 4 Vowel spaces of infants and children. (Figure of Vorperian and Kent, 2007)

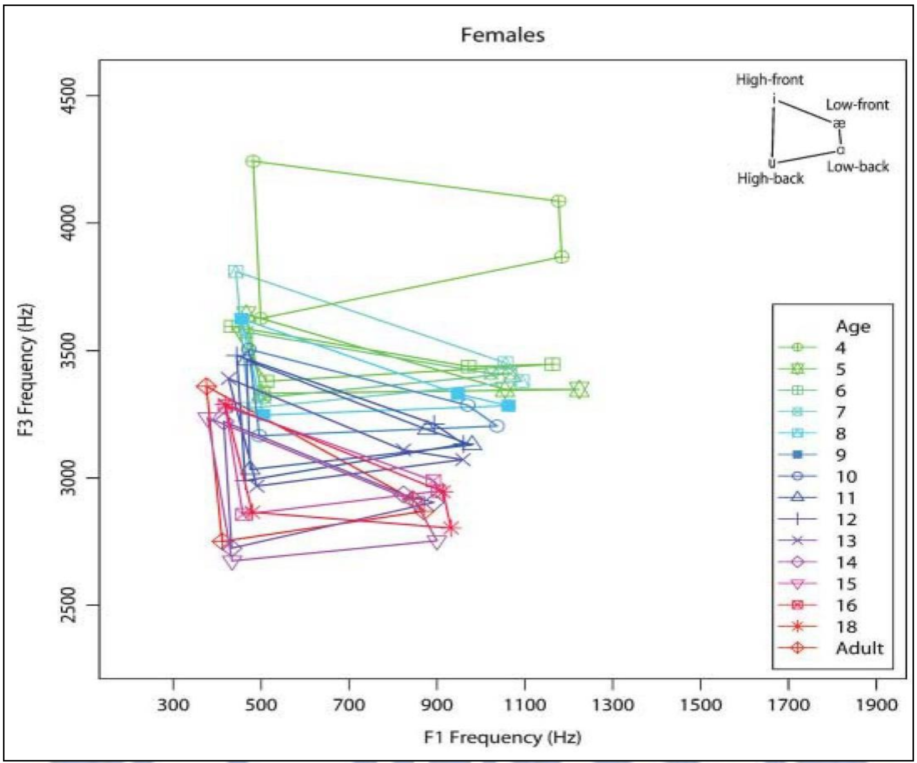


Figure 5 Vowel spaces of children and adults. (Figure from Vorperian and Kent, 2007)

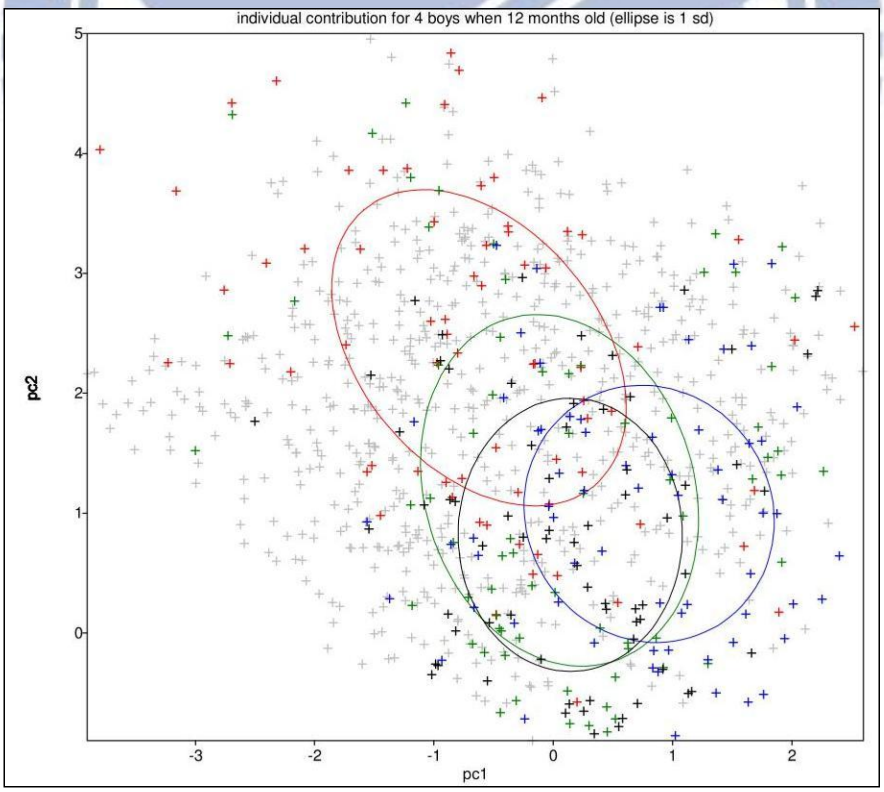


Figure 6 Vowel space of 12 months children. (Figure from van der Stelt et al., 2006)

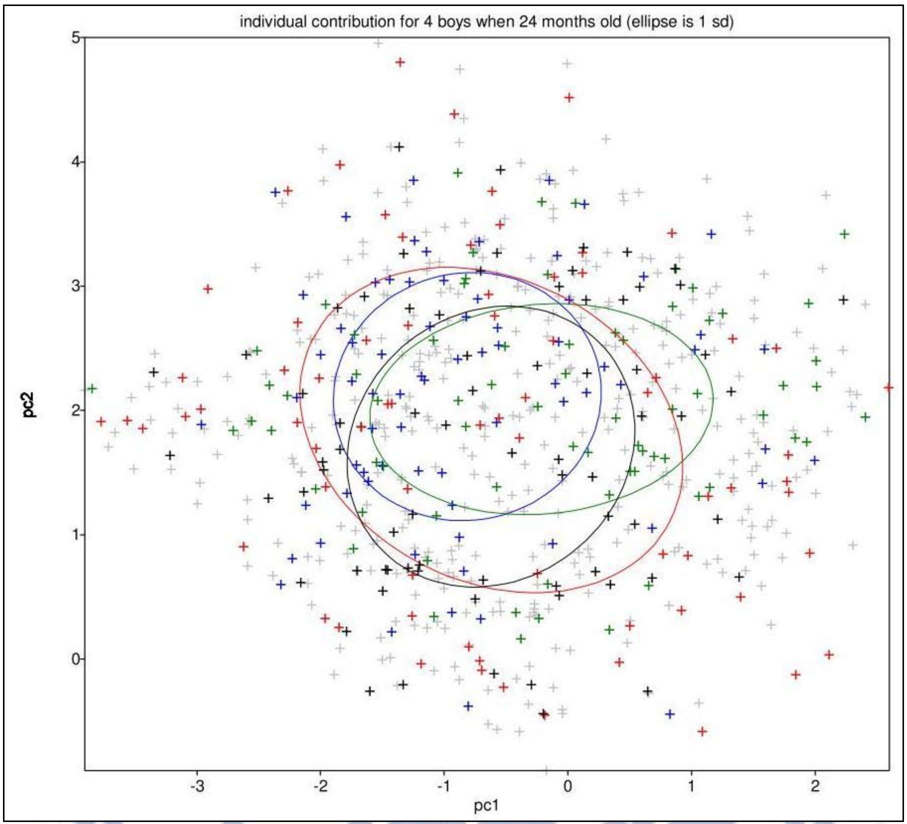


Figure 7 Vowel space of 24 months children. (Figure from van der Stelt et al., 2006)

Vowel acquisition of Chinese language

Similar studies have been conducted in Chinese language. Lee (2007) investigated vowel development of 1-to-2-year-old Southern Min-speaking children in Taiwan. The F1 and F2 of six vowels /i/, /e/, /a/, /u/, /ɔ/ and /ə/ were measured and compared with adults. Her results indicated that these vowels overlapped in their vowel spaces at an early stage, however it decreased as age increased **(Figure 8-9)**. However, the vowel space of children are narrower than that of adults **(Figure 10).**

Zhu (2002) investigated a longitudinal study for the acquisition of consonants, vowels and tones produced by four Chinese-speaking children (10 months to 2 years). The speech samples were transcribed, and imitated productions were marked in the data analysis. Here the only focus on the results of vowel. Results indicated that for the monophthong, central low vowel /a/ and back high vowel /u/ were the earliest to emerge, but the retroflex vowel /ɚ/ and the back vowel /o/ seemed to be the last vowels to emerge. For the diphthongs, /ei/ emerged first, and /yɛ/ emerge last. For the triphthongs, /ioʊ/ emerged first, and /uae/ was last.

Strokes and Wong (2002) studied the vowel production of 40 Cantonese-speaking infants (10-27 months). There were 100 transcriptions which are meaningful speech produced by each children. Only data from subjects aged 15-27 months were included because most of 10-13 months children did not produce any real words. Results indicated that front vowels [ɛ], [i] [œ] and one back vowel [a] appear at the age of 10-13 months. Back vowels [ɔ], [u] appear at age of 15-18 months. [y] was showed to appear at 20-27 months. Results also indicated that mid vowel [ɔ] and low vowel [a] were mastered at 15-18 months. Front mid vowel [ɛ] and high front vowel [i], mid back [ɔ] and low central vowel [a] were mastered at 24-27 months. In this study, most of the rounded vowels had not yet been mastered.

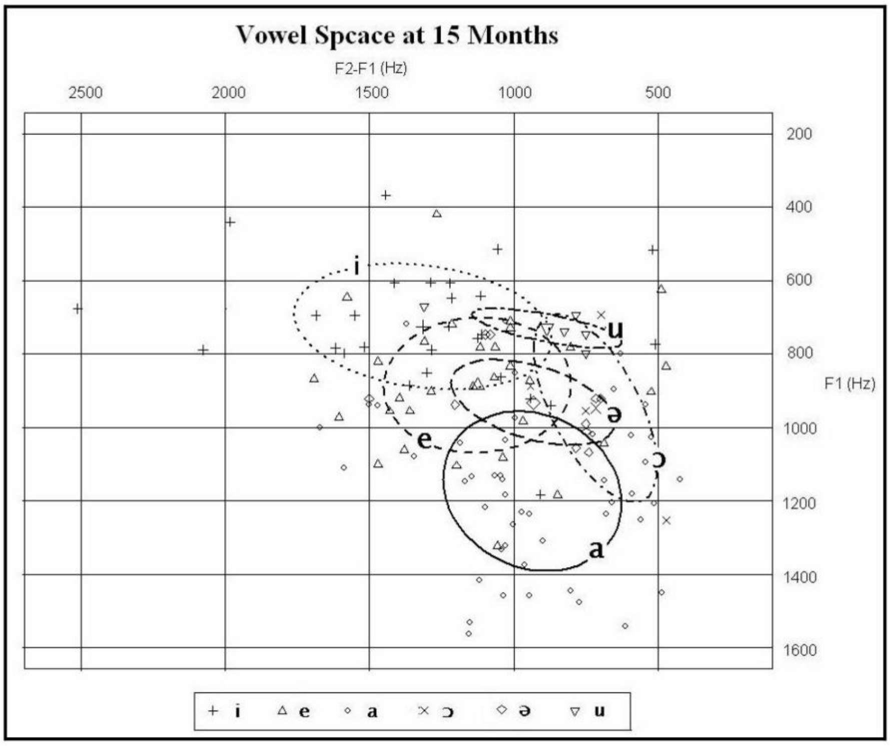


Figure 8 Vowel space of 15 months children. (figure from Lee, 2007)

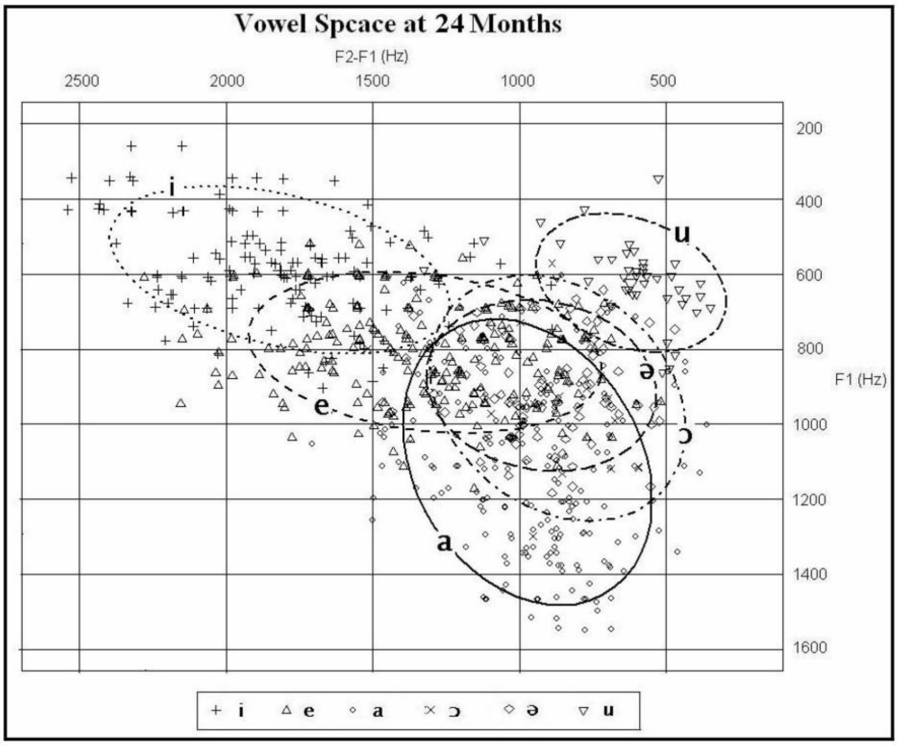


Figure 9 Vowel space of 24 months children. (figure from Lee, 2007)

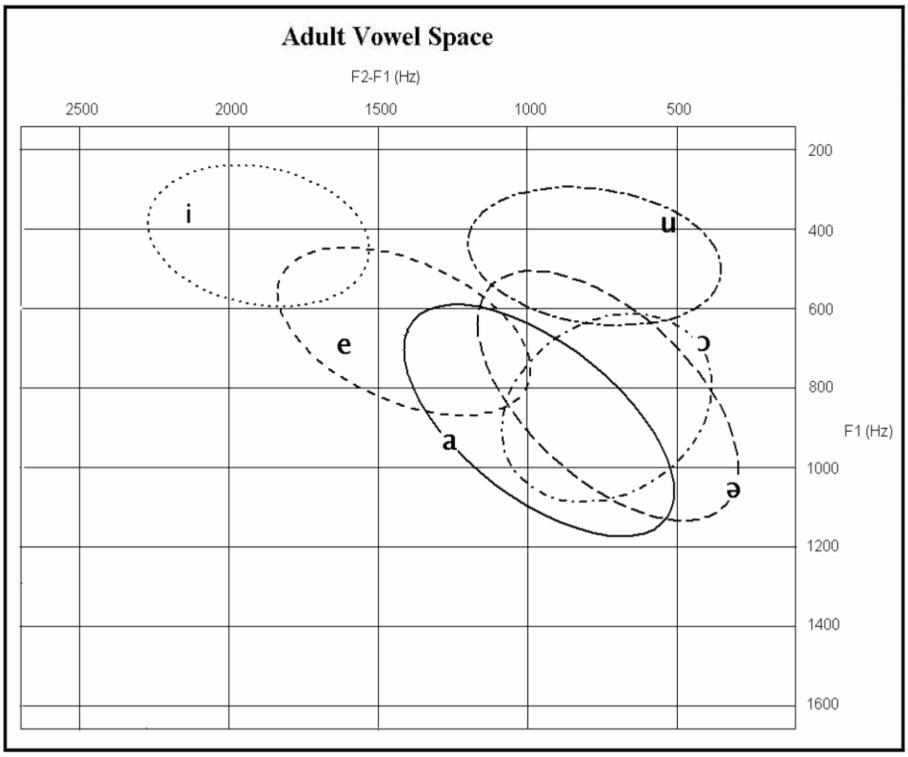


Figure 10 The vowel space of Adult (figure from Lee, 2007)

1.2.2 Suprasegmentals

Stress

Ingram (1989) proposed that there are five stages which children learn stress. *Stage 1*: Only monosyllables are produced. *Stage 2:* Disyllabic forms appear, invariably with initial stress. Initial unstressed syllables are deleted. *Stage 3:* Trisyllabic words are realized as disyllabic words. *Stage 4:* Initial unstressed syllables are attempted in disyllabic words, but not so much as in trisyllabic words. *Stage 5:* The child achieves an adult-like pronunciation. These results were also found in Fikkert's (1994) study. Fikkert investigated stress of Dutch-speaking children and compared it to that of adults. Results indicate that the children's form at an early stage contain a single quantity-insensitive trochaic foot. After this, children still contain exactly one foot, but the monosyllabic forms are now disyllabic, and none of the stress parameters are changed. At a later stage, children produced both feet with the same degree of stress. While entering the last stage, children produced an adult-like stress pattern.

Tone

Yip (1989) bisects the tonal space into two discretes: registers and contour tones. Register tone refers to the upper and the lower portions of the vocal range, and contour tones refers to a high-rising and a low-falling tone. In Mandarin, there are four tones: T1 (High level), T2 (Rising), T3 (Falling-rising) and T4 (High-falling). Suprasegments were generally found to be acquired earlier than segments (Chang, 1991).

Chang (1991) reviewed tonal studies (MaCarthy, 1930; Jakobson, 1968) and concluded that the suprasegmental nature of tone makes it the earliest systems of Mandarin phonology mastered by children (Li & Thompson, 1977; Su, 1985). Tonal errors were rare even at an early stage in children. Li and Thompson (1977) suggested that lexical tones were for differentiating the diverse lexical meanings, so when children begin to use words, tones also acquired. Physiological factors also might influence tonal acquisition. To produce segmental contrasts requires different kinds of coordination between different mechanisms, this might cause the suprasements to be acquired earlier than segments.

In terms of specific order of tone acquisition, studies show different results. For example, Vihman (1996) suggested that falling pitch movement was a natural gesture of speech production and required less physiological effort than rising tones, therefore, falling tones would be acquired earlier than rising tone.

Another complexity in Mandarin tone acquisition is tone sandhi. The most common is that when two tone3's are adjacent, the first tone3 will change into tone2. The other sandhi tone is when a tone3 is preceded by a tone 1,2, or 4. This tone 3 will be pronounced as a half third tone, which means that it is not a fully complete tone three. If a tone 4 occurs before another tone 4, it will become a tone 2. To examine the tone pattern, people studied the fundamental frequency contours of Mandarin Chinese. In Zhu's (2002) study, subjects replace tone 2 into tone 4 and tone 3 was replaced by tone 2. Zhu suggested that children occasionally used citation tones when tones should be changed according to tone sandhi rules.

Intonation

Cruttenden (1997) suggested that there are four periods which could identify the early vocal development of infant: (1) crying period, (2) babbling, (3) one-word period, and (4) two-word period. During the crying period, children cry, suck and strain. During the babbling stage, children often show mimicry of adult pitch patterns. Children might use one identical pattern which has been acquired from an adult's speech to form different words. In the late babbling period, children may actually produce whole sentence intonations called "jargon intonation". After it, children appear to contrast a fall and rise of intonation in one or two syllables. During one-word, children begin to distinguish between high and low varieties which are learned from adult's speech. For example, they use high-rise for echoing and attention-getting, while a low-rise is used for listing or questioning. A high-fall from a low-fall is used as a more emphatic tone to indicate surprise and insistence. The fall-rise may appear until two-word periods are used for warning or contrast.

Syllable structure

In terms of syllable structure, Jeng (2009) pointed out that children's consonant error production does not only include misarticulated consonants, but also that errors result from the coarticulation of consonants and vowels. Subjects included 306 children, age ranging from 2.5 to 6 years. All subjects divided into six groups according to their age. Subjects were asked to name all the pictures, a total of 32 mono syllables and disyllables, including 28 disyllables and 4 mono syllables. Tokens chosen from Mandarin 21 consonants (also is the target consonants). These consonants were shown in both a “simple phonetic environment” and “complex phonetic environment” once. The target phonemes are /p^h/, /f/, /t/, /t^h/, /l/, /ʂ/, /ts/, /ʐ/, /ts^h/, /s/.

The easiest syllable structure is a CV structure, while the hardest syllable structures include CVN, CVVN, and CVVV. The result shows that the consonant errors decrease as the age of the children increase. Therefore, pronunciation accuracy of consonants in an easy co-articulate phonetic environment is higher than that in harder co-articulate phonetic environment; consonant accuracy in easy syllable structures is also higher than that of complex syllable structure. This shows that the accuracy of consonant production of pre-school children will be influenced by where it is placed.

It is well known that easy structures are acquired faster for children. It is well known that children acquire easy structures faster. So single structure emerges first, second is the disyllabic structure, and last is multisyllabic. An isolated consonant or vowel sound is easiest to learn, children then acquire CV structure, while VC and CVC structures are acquire later, and finally a complete word (Chen, 2004).

1.3 Mandarin acquisition

Consonants

Cheung (2000) examines Taiwan children’s perception and production of 16 syllable initial consonants. This study aims to test whether children's perception is developed before their production. 20 Subjects aged between 3 to 4 first named pictures and then imitated the names of novel cartoon figures. The transcription was performed by two adult listeners. The production is coded “correct”, if the pronunciation was identical to an adult’s. After the production tests, subjects participated in a perception test. The experimenter prepared four pictures (A and B groups) and put two pictures of A group into player, then put other two pictures of B group. The subjects were asked to answer the question: "Whose name is this?" Cheung concluded that /s-ʂ/ and /ts-tʂ/ are not well controlled both in production and perception tests. While the production of /ts, ts^h/ were the most difficult ones to master. On the other hand, /p-m/, /n-l/, /t-k/, /p-p$^{h/}$, /p-t/ have a high accuracy rate in production and perception tests. Cheung suggested that perception and production develop gradually, but production might proceed slightly faster.

With respect to manner of articulation, affricates and fricatives are the most difficult to acquire (Chang, 1986). Wan (2002) suggested that coronals are common in the early acquisition in Mandarin. Retroflexes are coronal with an extra production effort–the curling of the tongue. However the four Mandarin retroflexes, /ʈʂ, ʈʂʰ, ʂ, ʐ/, are mostly affricates which may cause retroflexes to be even more difficult to produce. Jeng studied 301 Taiwanese children with ages ranging from 2.5 - 6. She suggested that the most difficult consonants were coronals. She also discovered that bilabials were acquired when the child reaches an age of 2.5, and that retroflexes were acquired after the age of six (Table. 1 in 1.2.1).

In acquisition of Mandarin consonants, the results of these two production tests from Cheung (2000) could be summarized as 1) voicing, 2) manner and 3) place. In terms of voicing, voiced sounds are more controlled than voiceless sounds. In terms of manner, stops, nasals and approximants were best controlled, followed by fricatives and affricates. In terms of place, labials, alveolar and velar are best controlled, and the most difficult for children were coronals. Jeng (2004) and Wan (1984) both suggested that retroflexes are the most difficult segment and were acquired late for children.

Vowel

Chen (2004) studied the production of Mandarin-speaking children from birth to 36 months using perceptual transcription and acoustic analysis. Results indicated that the formant frequencies of vowels get lower as the child grows. The acquired sequence of vowels is monophthong followed by diphthong and then triphthong.

Zee and Lee (2001) investigated the 7 vowels in Beijing Mandarin in order to present the spectral characteristics of these vowels. Their subjects included 20 native speakers of BM, 10 male and 10 female, aged between 18 and 22. A total of 1600 tokens were analyzed. The F1-F3 average values were presented **(Appendix III. Table 2-3).** Results showed that the vowel type and nasal type affected the F1-F2 of the average vowel positions for vowels which were followed by a nasal ending. In addition, there is gender difference in vowel ellipses. The vowel ellipses are more left and lower for female compared to male speakers **(Appendix IV).**

The acquisition order of Mandarin vowels is monophthong, diphthong, and finally triphthong (Chen, 2004). The vowel space of Mandarin vowels show different patterns in different studies. The study of Peterson and Barney (1952) indicated that the difference of vowel space between children and adult is the F1 range. Children have a higher F1 and F2 than adults. Children show a higher F1 frequency of vowel space. Different vowels overlapped in their vowel spaces in early stage, however it decreased as they grow (Lee, 2007).

Tone

Li and Thompson (1977) studied the acquisition of lexical tone in 17 Mandarin-speaking children (1:6-3:0). Experimenters collected eight months longitudinal data, and visited each child every 3 weeks. A total of about 80 hours of data were collected. Experimenters uses naming task to elicitate children’s pronunciation. The following results were found: (1) tone acquisition is accomplished within a short period of time; (2) tone is mastered before segmentals; (3) Mandarin high-level and falling tones are acquired before the rising and dipping tones; (4) the rising and dipping tones are substituted for each other; (5) unstressed syllables are treated as if they were stressed, the tone assigned to them being an approximation of the phonetically conditioned pitch which they carry; (6) the tone sandhi associated with the dipping tone in Mandarin is acquired with very little error as soon as propositional utterances begin to be created. In general, high level tones are unmarked and are most frequently used to replace other tones. This might cause the high level tones to be acquired earlier than other tones. Li and Thompson (1977) reviewed several studies and argued that the ease of articulation and perception in rising tones were more difficult than high falling tones, and that high falling tones occur more often than rising tones in the world's language.

1.4 Phonological acquisition in cross-language context

Most of studies have examined the developmental patterns of children from different language backgrounds. Similarities and differences exist in the acquisition of phonological units. In terms of error patterns, Mandarin-speaking children showed a similar pattern to English-speaking children for simplifications in production (Grunwell, 1981). However, Mandarin-speaking children showed typical error patterns like syllable-initial consonant deletion and backing, which are not found in English-speaking children. Furthermore, in terms of the features of aspiration, affrication and retroflex were acquired in English, Cantonese, Russian. However, Japanese shows the opposite pattern (Olmsted, 1971; So & Dodd, 1995; Timm, 1977; Yasuda, 1970, cited in Locke, 1983).

Many error patterns or acquisition order of phonological units may vary from one language to another. It is even more interesting to investigate phonological acquisition of bilingual children to see whether the languages around them will influence their acquisition, and whether or not they are affected by the universal constraint or the language-specific constraint.

Grosjean (1982) suggested that most people grow up speaking two or more languages. Many bilingual speakers tend to use one language in some settings and the other language in others. In general, infants appear to discriminate the rhythmical patterns that differentiate some languages, Bertoncini, & Mehler, (1998) suggested that stress-timing in English and French could be discriminated by infants when they are at a very young age.

Maneva and Genesee (2002) examined an English-French bilingual infant (recorded at 10-13 and 15 months) with his mother (English) and father (French). Results indicated that the child produced more CV syllables in English than French, which is identical to the results of monolinguals in English and in French. This infant's CV sequence in English contained more stops, but contained more approximants and liquids in French. The infant also produced more monosyllabic and disyllabic sounds in English, but more polysyllabic sounds in French, resulting in longer babbled utterances in French. In their study, the child showed a language-specific constraint for acquiring both English and French.

Previous researches suggested that when children are exposed to two languages, there might be a mixing of languages during the earlier stages of acquisition (Vihman, 1996). Children produced words from both languages in the same utterance, or they attached morphological endings from one language to word-stems from the other.

The present study focuses on children with Shanghainese L1 background learning Mandarin as an L2, to see whether the languages around them will influence their acquisition, and whether they are affected by the universal constraint or the language-specific constraint. Below is a brief introduction of Mandarin and Shanghainese.

1.5 Introduction of Mandarin and Shanghainese

There are 6 monophthongal vowels (/i/, /y/, /ɛ/, /u/, /o/, /a/), 21 consonants (/p/, /p^h/, /m/, /f/, /t/, /t^h/, /n/, /l/, /k/, /k^h/, /h/, /tɕ/, /tɕ^h/, /ɕ/, /ʈʂ/, /ʈʂ^h/, /ʂ/, /ʐ/, /ts/, /ts^h/, /s/) and 4 lexical tones and a neutral tone in Mandarin. The Mandarin vowel system is rather complex with a structure of V, VV, and triphthong, VVV. (e.g., iao, iou). The syllable structure is simple; there are no word-initials or final clusters and only nasals can be in the coda position. The four lexical tones are high level (T1), rising (T2), falling rising (T3) and falling tone (T4) (Chang, 1991). T2 and T3 tend to be the longest in tone duration, while the T4 is the shortest (Howie, 1976).

There are 9 monophthong (/i/, /y/, /ø/, /ɛ/, /u/, /o/, /ɔ/, /a/, /ɤ/), 33 consonants (/p/, /p^h/, /b/, /m/, /f/, /v/, /ʔm/, /t/, /t^h/, /d/, /n/, /ʔn/, /l/, /l̪/, /k/, /k^h/, /g/, /h/, /ɦ/, /ʔ/, /ts/, /ts^h/, /s/, /z/, /ʥ/, /ʑ/, /ɲ/, /ʔɲ/, /ŋ/, /ʔŋ/) and 5 citation tones in Shanghainese. Shanghainese tones are merged from eight tones to five tones among which T1 (high falling), T2 (high rising) and T3 (low rising) are long, while T4 (high) and T5 (low rising) are short (Zhu, 1999). The sounds of Shanghainese are categorized in initials and rimes (rhyme). Initial is the first part of syllable, usually a consonant, and rime is the part that follows. Phil (1997) studied the five citation tones of Shanghai **(Figure 12).** The first three tones are long and last two tones are short which shows an identical result to Zhu (1999).

When comparing the tones in Mandarin and in Shanghainese, the most obvious difference is that there is no real falling tone (like Mandarin) in Shanghainese. Although T1 in Shanghainese is a high falling tone, the pitch range is not as wide as that of Mandarin T4. /ø/, /ɔ/, /ɤ/ are not in Mandarin vowel inventory, and the 16 consonants /b/, /v/, /ʔm/, /d/, /ʔn/, /l̪/, /g/, /ɦ/, /ʔ/, /z/, /ʥ/, /ʑ/, /ɲ/, /ʔɲ/, /ŋ/, /ʔŋ/ are not in Mandarin consonant inventory. On a side note, there are no retroflexes in Shanghainese.

Wang et al (2003) reviewed previous studies about perception and production of Mandarin tones. For native speakers the tonal pattern is an integral part of each word, but this pattern did not exist in non-tonal speakers. Jongman et al (1997) indicated that tone 1 starts at a high F0 onset and maintains a level pattern. Tone 2 and tone 3 can be characterized in terms of two measures: Turning Point and △ F0. Turning Point is the point in time at which the contour changes from falling to rising. △ F0 is the change in F0 from the onset of the tone to the Turning Point. Typically, Tone 2 has an earlier Turning Point, and smaller △ F0 than tone 3. Tone 4 starts at a peak and gradually falls to a low point. Tone 3 has a lower onset than tone 4. We could use these acoustic properties to examine the tone sandhi realization **(Figure 11).** Tone system is acquired quickly and mastered before segmental system. The high and falling tones are acquired earlier than the rising and dipping tones (Li & Thompson, 1977).

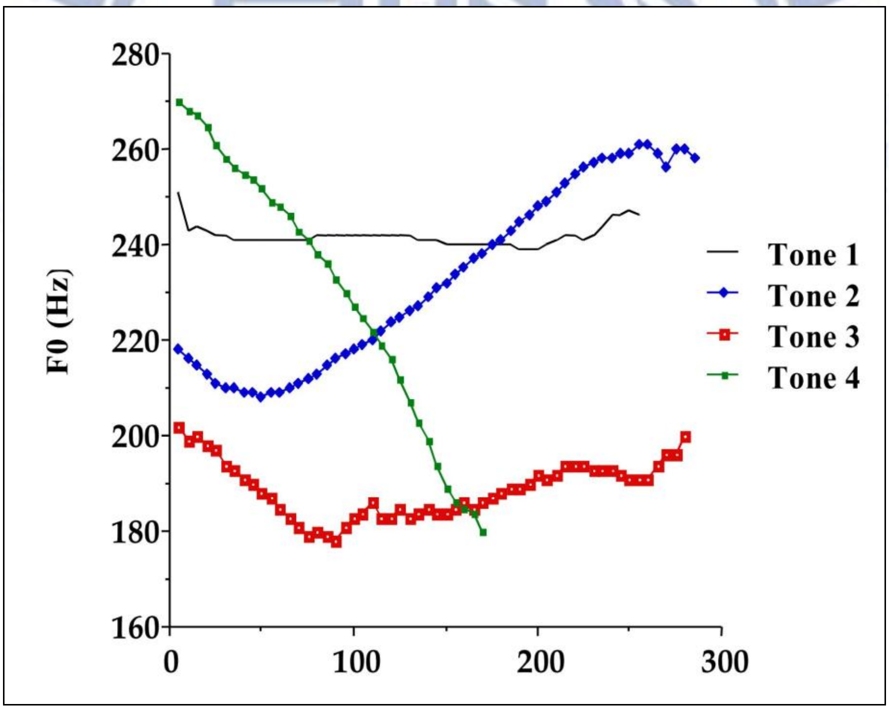


Figure 11 F0 contours of Mandarin adult (Figure from Jongman, 1997)

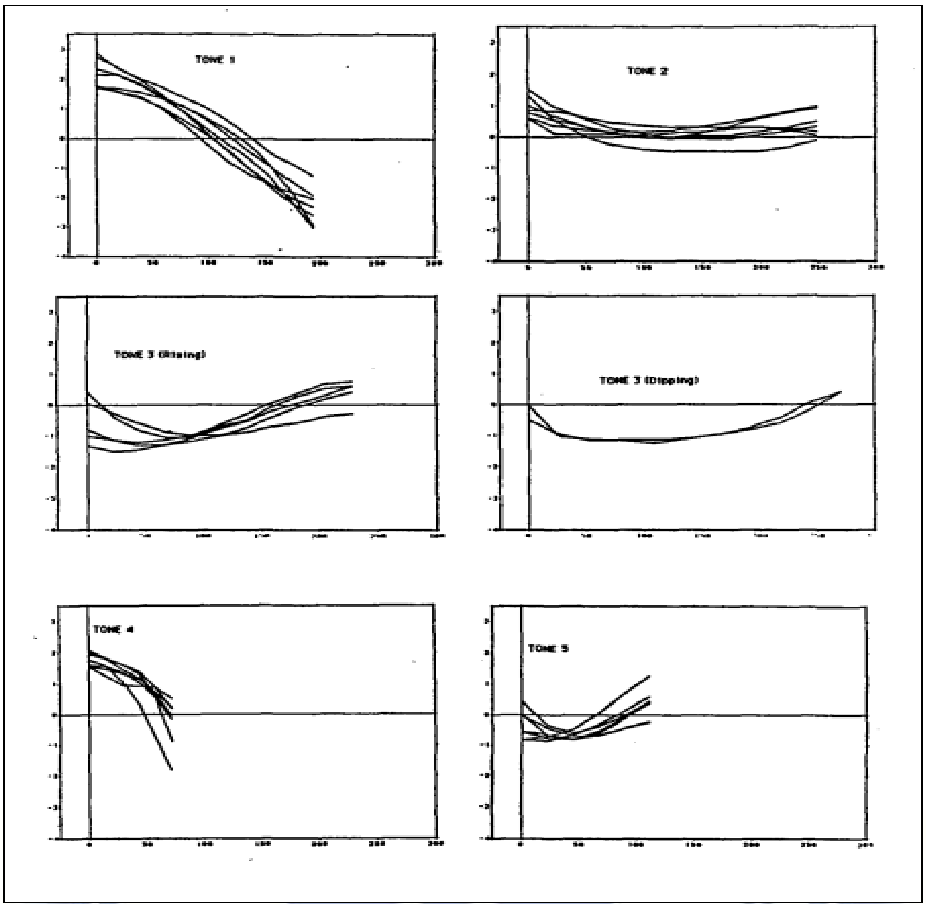


Figure 12 Vertical scale shows normalized F0. Horizontal scale shows mean absolute duration in centiseconds. (Figure from Phil Rose, 1993)

1.6 The present study

The present study is a subprogram of the cross-linguistic study of protracted phonological development (PPD), which included children from Shanghai (30 typically developing; 30 with PPD) and Canada (pilot) (led by Dr. Bernhardt. UBC Clinical speech-language pathology). The preliminary results from 30 typically developing 4-year-old children indicated that the most common error pattern was deretroflexion, followed by retroflexion of nonretroflexed sibilants. The children with PPD showed many of the same errors as the typically developing ones, but with a greater frequency.

The recording data from 30 typically developing children from Shanghai were transcribed in the present study, and the acoustic analysis further investigated the acoustic realization of consonants, vowels and tones produced by them. The transcription was collected in a excel sheet which will be presented in the methodology section. The present study hypothesizes that because of universal developmental constraints; children show some similar phonological patterns when compared to children with different L1 background. Nonetheless there exists some diverse patterns which reflect structures, segments and phonotactics of their first language (Shanghainese).

For example, although Shanghainese and Mandarin are both tone languages, different tone systems might influence Shanghai children in producing Mandarin tones. The production of vowels might go through the universal acquisition sequences, and might also be influenced by L1 background. Furthermore, the complexities of consonants in both L1 and L2 may cause Shanghai children to have difficulties when speaking Mandarin.

In the present study, the examiner makes the following assumptions: (1) Tone will be the earliest acquired, so the error pronunciation in tones might be less than other segment units. (2) Vowels are acquired and produced earlier, with fewer speech errors than consonants. (3) The vowel space of children will be wider than an adult's. (4) Among typically developing children, they have more difficulties in producing retroflexes.

Chapter 2. Methodology

In this section, subjects' background, materials, procedure, transcription and acoustic analysis are introduced. The acoustic analysis included three parts: retroflex, vowel and tone analyses.

2.1 Subjects

Thirty (male: 16, female: 14) typically developing children. Age range of subjects is four years and one month to four years eleven months; mean age is four years and eight months. All participants reside in Shanghai and attend kindergarten for five days a week. Based on the teachers' reports, all participants were exposed to the local language (Shanghainese) at home. However, only Mandarin is used in the kindergartens, and more than 90% of the children communicate with their parents in Mandarin.

2.2 Material

A list of words known by preschoolers was used to elicit the consonant, vowels, and syllable structures of Mandarin. Item selection for the list took into account phonology, imageability, and other factors. The word list targets three major aspects: segment, major syllable structure and place sequence. Each target appears at least twice in different words and tone sequences at least once. The list contains monosyllables (48%), disyllables (50%) and multisyllabic compounds (2%) **(Appendix II).**

2.3 Procedure

The stimuli are depicted pictorially in color photographs. The examiner showed the children one picture at a time and asked them to name the item. If the child did not know the name, the examiner would give them a hint using sentences as: Is this an X or a Y. If the hint failed, the examiner would tell the child the name and had him or her repeat it.

2.4 Transcription

In transcription, the examiner listened to and transcribed the recorded data of 30

Shanghainese children by using Praat. The subjects' recorded data was collected in different files, (one word for 30 subjects in a file);there is a total of 80 sound files for different words. The transcription results were collected in an excel sheet **(Fig. 13).** The 30 subjects were coded from one to thirty in the first column. Because our subject data was collected from Dr. Barnhart's study, the subject number they coded was put in the second column. In order to be more specific, the examiner divided each character and tone into a different column (one syllable/tone in one column). The syllable structure was coded in the sheet. The target words were listed on the left and the actual production of children was listed on the right side of the sheet. After the actual production, in the far right column, the occurring error was coded from one to nine (there were nine errors found in subjects).

code	subject	words				tone				structure		words				tone				mark
01	s27	bao				1				CVV		bao				1				
01	s27	qun	zi			2	0			CVC.CV(V)		qun	zi			2	0			
01	s27	huai	le			4	0			CVVC.CV		huai	le			4	0			
01	s27	shui				3				CVVV		shui				3				
01	s27	wa	zi			4	0			CV.CV(V)		wa	zi			4	0			
01	s27	ü				2				V		ü				2				
01	s27	niu	nai			2	3			CVVV.CVV		niu	nai			2	3			
01	s27	bei	zi			1	0			CVV.CV(V)		bei	zi			1	0			
01	s27	tang				2				CVC.CVC		tang				2				

Figure 13 A part of excel sheet of transcription.

2.5 Acoustic analysis

2.5.1 Retroflexes

The F3 values at the onset, mid-point, and offset of the vowels following retroflexes were measured. A very low F3 formant frequency is the most clearly defining acoustic property (MacGowan, 2004). Through measuring F3, the acoustic properties of retroflex of Shanghain children can be clearly seen. TF32 (Jeng, 2004) was used to measure the M1-M4 of retroflexes.

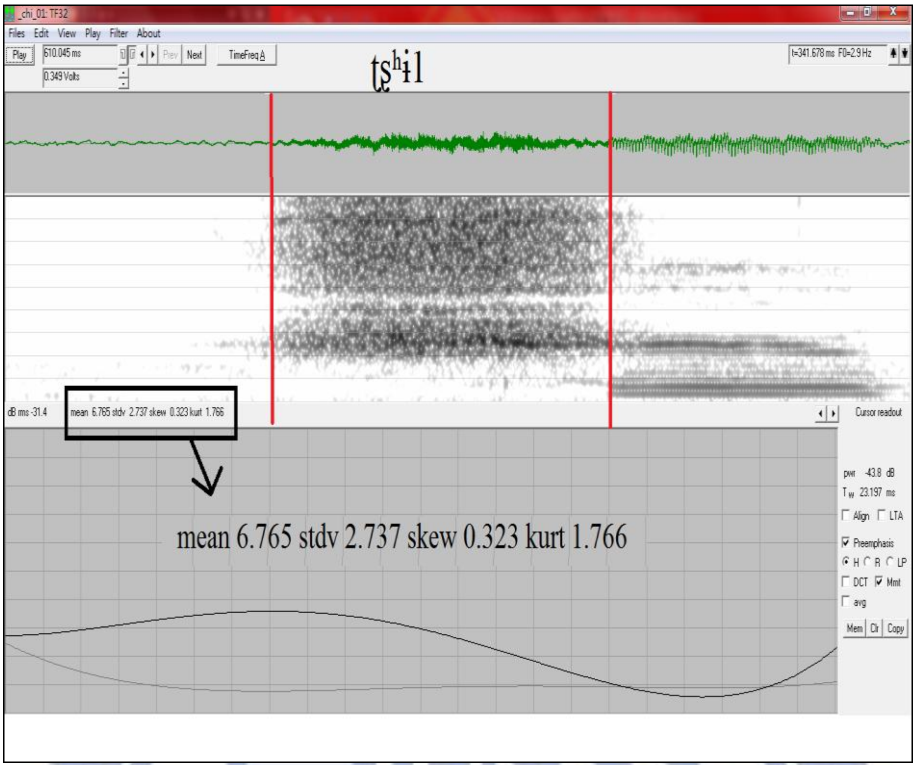


Figure 14 Mandarin tʂʰɨ1 produced by Shanghai children.

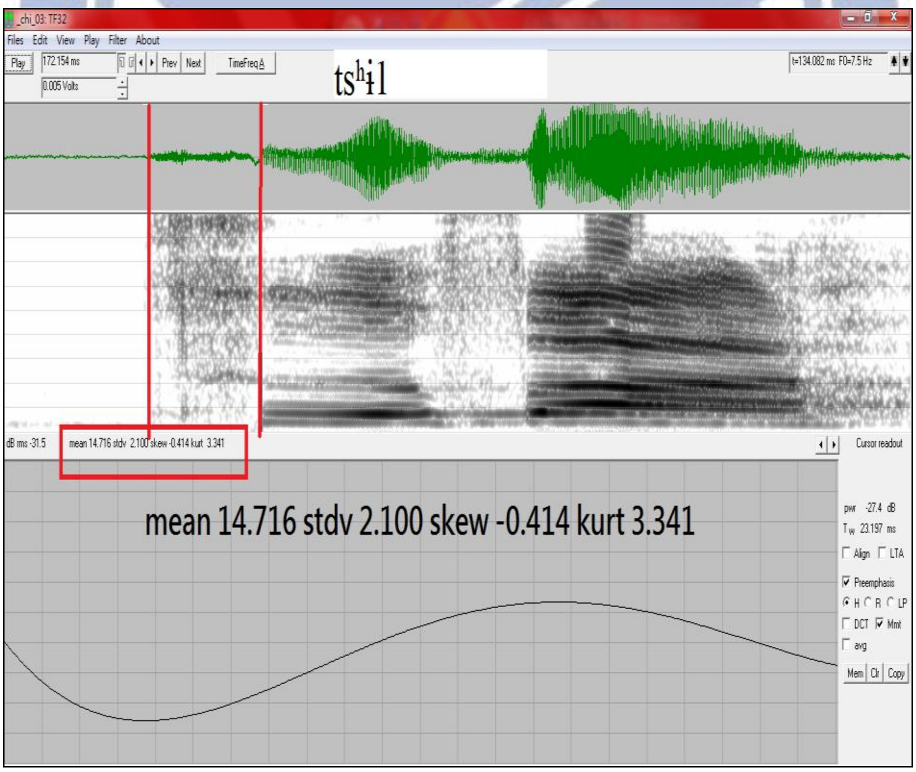


Figure 15 Mandarin tsʰɨ1 produced by Shanghai children.

2.5.2 Vowels

For vowels, the first two formants were measured in order to draw the vowel space of Shanghain children. The vowel space was compared with that of Shanghainese (data from Chen, 2008) and Mandarin adults (data from Zee & Lee, 2001) to see whether there is L1-L2 influence.

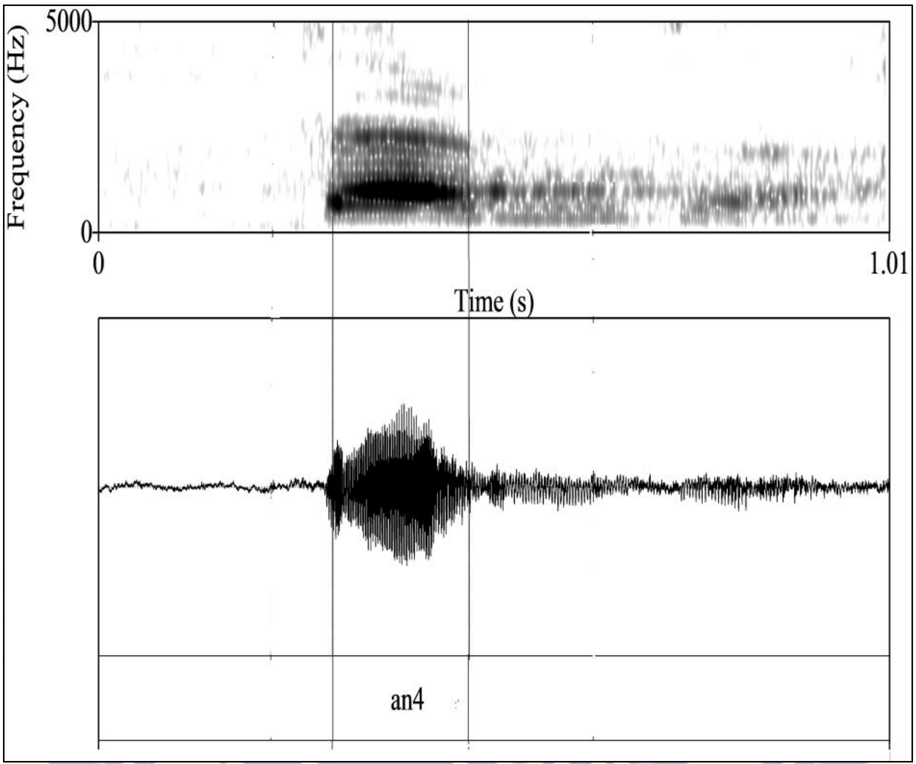


Figure 16 F1- 3 measurement. Tier one shows boundary of ʦan4 "stand" produced by subject 8. Two boundary lines show the nuclei portion. The first boundary line is the onset, the second line is offset, mid-point is in the middle of the nuclei portion.

2.5.3 Tones

F0 of all target words were measured using an F0 normalization Praat script by Xu. Each tone was measure for ten points (every 10%), and averaged values were used to draw the tone picture. In order to get the result files, boundaries need to be put on the TextGrid which could mark out the nuclei portion. This script will get the F0 information from this marked portion. Five Shanghainese citation tones were collected from personal communication with Dr. Jie Zhang.

Chapter 3. Results

3.1 Transcription results

Transcription data yielded the following speech errors: (1) Non-retroflex was produced as retroflex, (2) Deretroflexion: produced a retroflex sound as non-retroflex, (3) Substitution: Mandarin /ʐ/ was misarticulated as /l/, /n/, or deretroflexed, (4) Nasal confusion: velar nasal was produced as alveolar nasal. (5) /h/ substitution: /h/ was produced as /f/ or was deleted. (6) /tɕ^h, tɕ, ɕ/ confusion: these three consonants are randomly substituted by one another, (7) Vowel confusion: diphthong was displaced by monophthong, (8) Tone confusion: the most common error was T3 produced as T2, and then T4 was produced as T1, (9) Other errors which were less found in subjects, for example, /k/ replaced by /d/. In terms of this error result, errors for suprasegmental (29%) are fewer than that of segment (71%). Subjects made the most errors in tones (29%), vowels (15%) and retroflex (7% + 32% + 6% = 45%) (Fig. 15). Thus, retroflexes, vowels, and tones are further examined in the acoustic analysis.

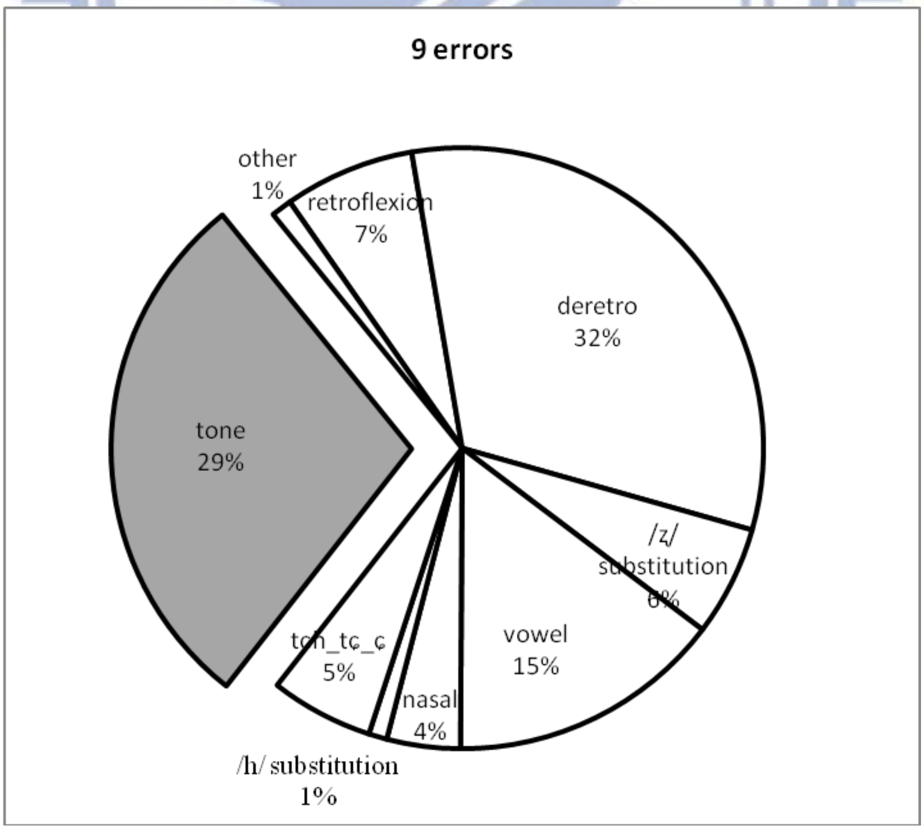


Figure 17 Pie chart of 9 errors.

Voicing

In Mandarin, there are nine voiceless sibilants / tʂ, tʂ^h, ʂ, tɕ, tɕ^h, ɕ, ts, ts^h,s/ and one voiced sibilant /ʐ/. In the nine voiceless sibilants, / tʂ, ʂ ,tɕ, ɕ, ts, s/ are unaspirated sounds, and /tʂ^h, tɕ^h, ts^h/ are aspirated sounds. In all the error patterns, errors for voiceless unaspirated sounds (65%) are the most frequent, followed by voiceless aspirated (21%), and voiced sibilant (14%) last **(Figure 18).**

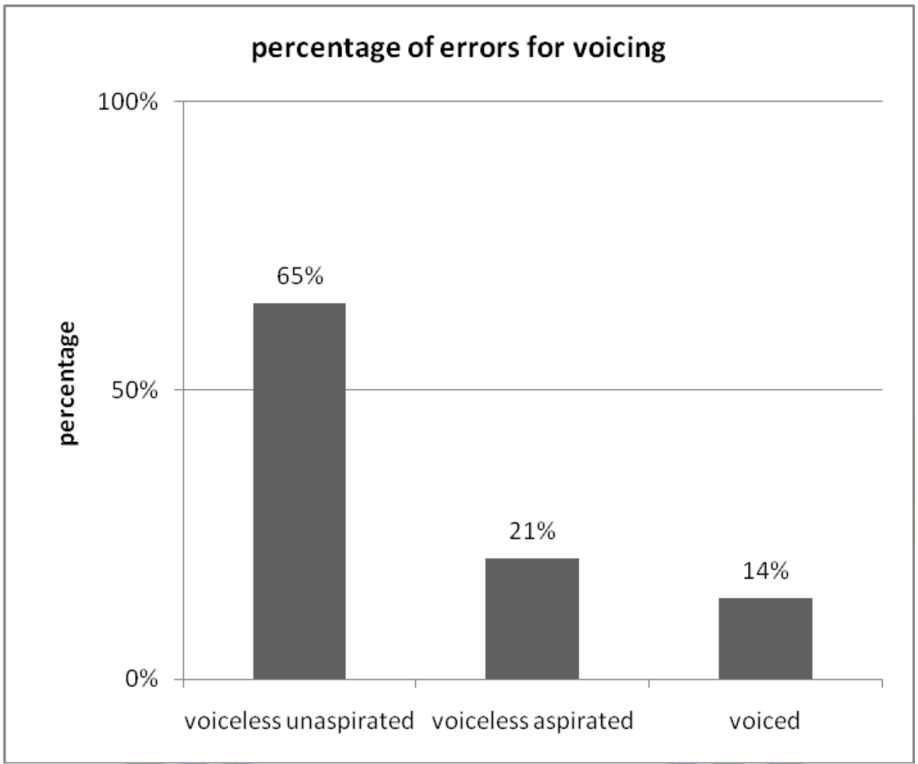


Figure 18 Error percentage for different voicing patterns.

Place of articulation (POA)

In all the error patterns, the errors occurring in the **coronal** are the most frequent (94%), which include retroflex confusion (70%), /ʐ/ substitution (15%), /tɕ^h/, /tɕ/, /ɕ/ confusion (15%), and alveolar nasal confusion or deletion (9%). The errors for each manner of articulation misarticulated in coronals **(Fig 19)** indicated that the affricates are the easiest misarticulated MOA in coronals (47%), followed by fricative (42%), nasal (6%), stops (5%). Other errors in other places of articulation are fewer. For example, the errors in **bilabial** is about 4 %, **dorsal** is about 3 % (velar nasal confusion, 2.3 % and /h/ substitution, 0.2%) **(Fig 20).** It is worthy to note that the most frequently error place of articulation in vowel production are when vowels are next to **coronals** (87.9%), **bilabials** (6.6%), **dorsals** (5.4%).

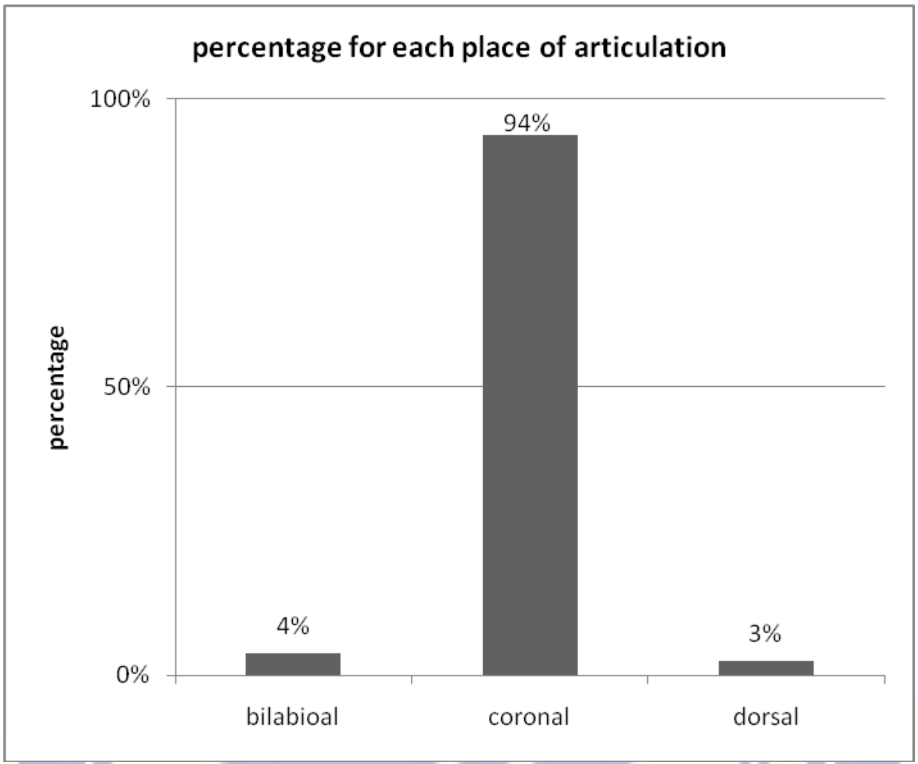


Figure 19 Error percentage for each POA

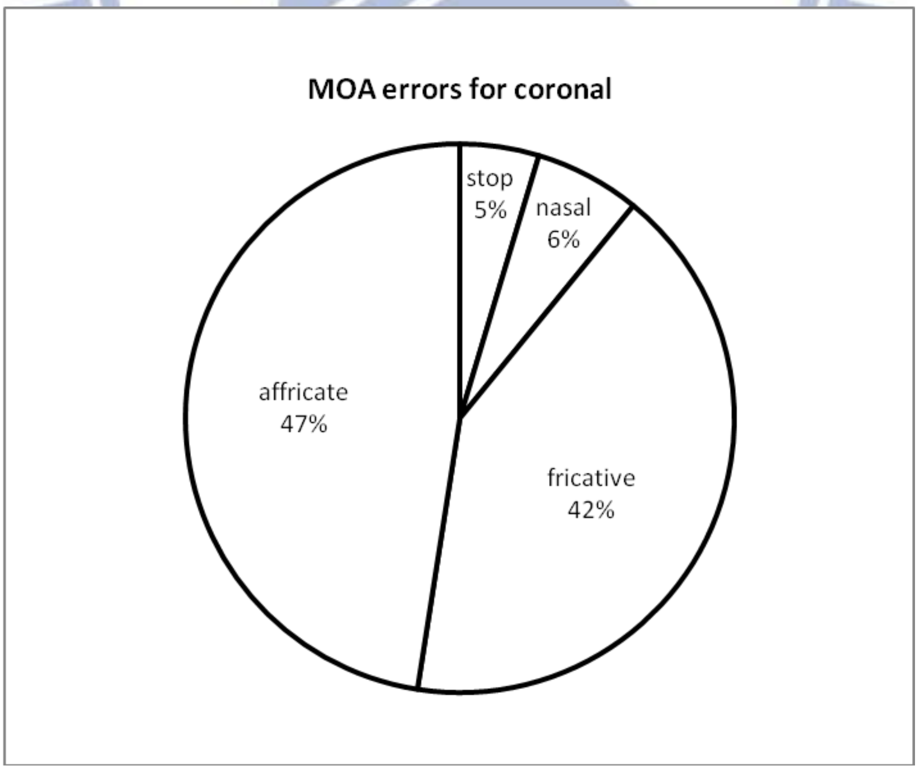


Figure 20 Each MOA errors which occurs in coronals.

Manner of articulation (MOA)

Among the nine errors, retroflex confusion is the most frequently error. In terms of manner of articulation, the errors of affricates (66.31%) are the most frequent (/tʂ, tʂ^{h}, ʂ, tɕ, tɕ$^{h/}$) followed by fricative (18.03%) *(/*ʐ and ɕ*/)*. Next were the errors for nasals (8.75%) (alveolar and velar nasal confusion). Lastly are stops (6.89%) (/b, p, d, t, g and k/). Approximants (/l/) are mastered well. There is no liquid in Mandarin **(Fig. 21).**

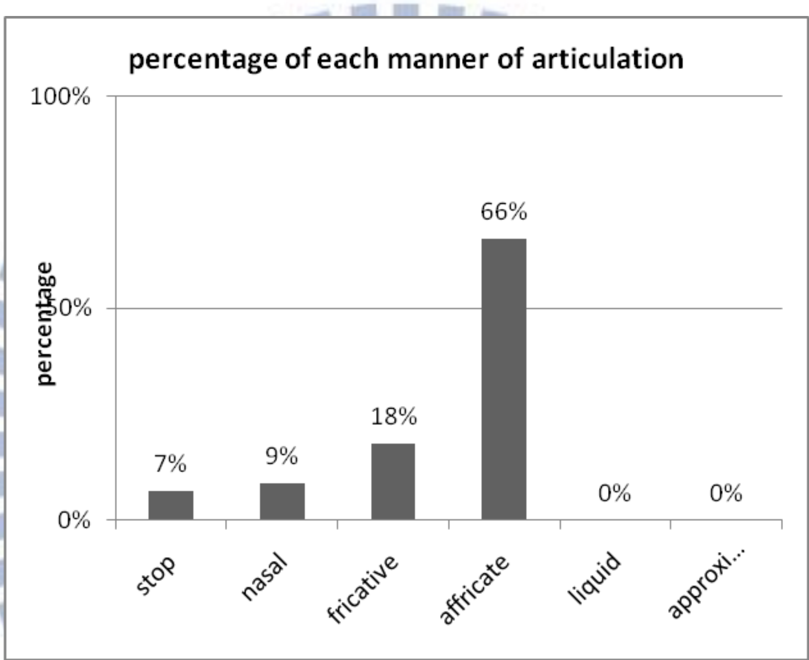


Figure 21 Percentage of errors in each MOA

Vowels

The vowel /u/ was the most frequently error (27%), followed by /a/ (26%), /e/ (19%), /o/ (15%), the last is /i/ (12%) **(Fig 22)**. The most common vowel which children used to substituted other vowels is /a/ (38.04%), followed by /e/ (21.73%), /u/ (16.3%), /i/ (10.86%), /o/ (9.78%), /ü/ is the last (3.26%) **(Fig 23).**

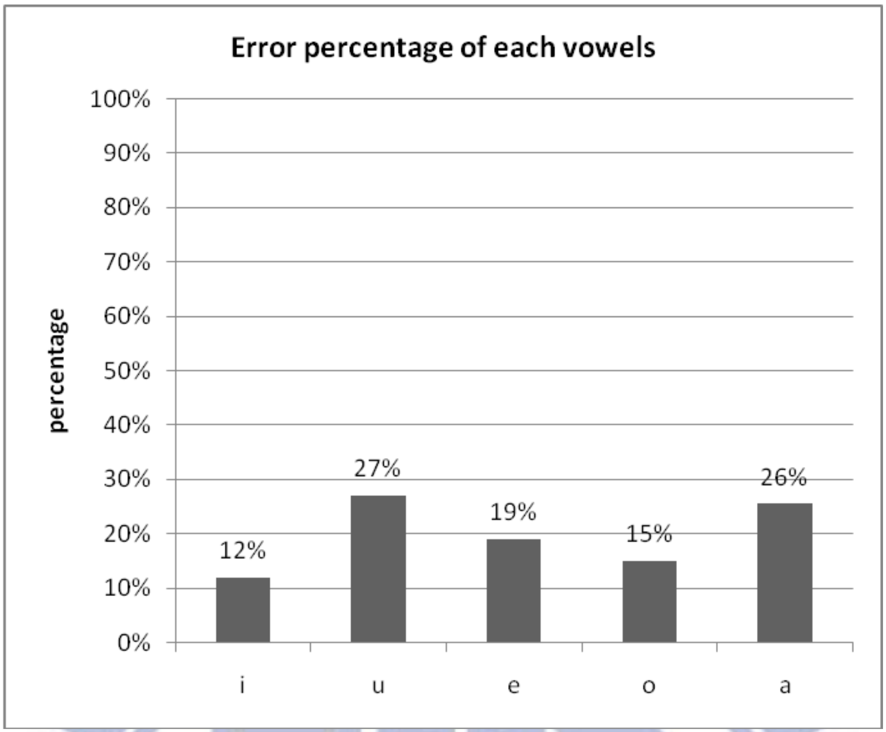


Figure 22 Error percentage of each vowels

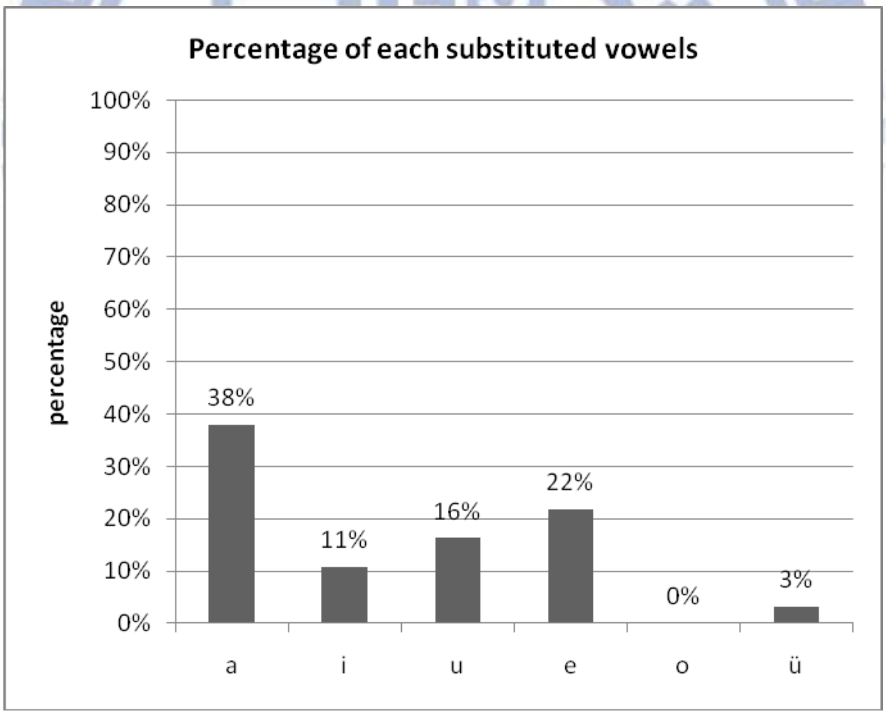


Figure 23 Occurrence percentage of each substituted vowels.

Errors in vowel production, the most frequent is VV structure (30%), followed by VVN (28 %), then VN (22%), the last is V (20%), VVV (1%) **(Fig 24).**

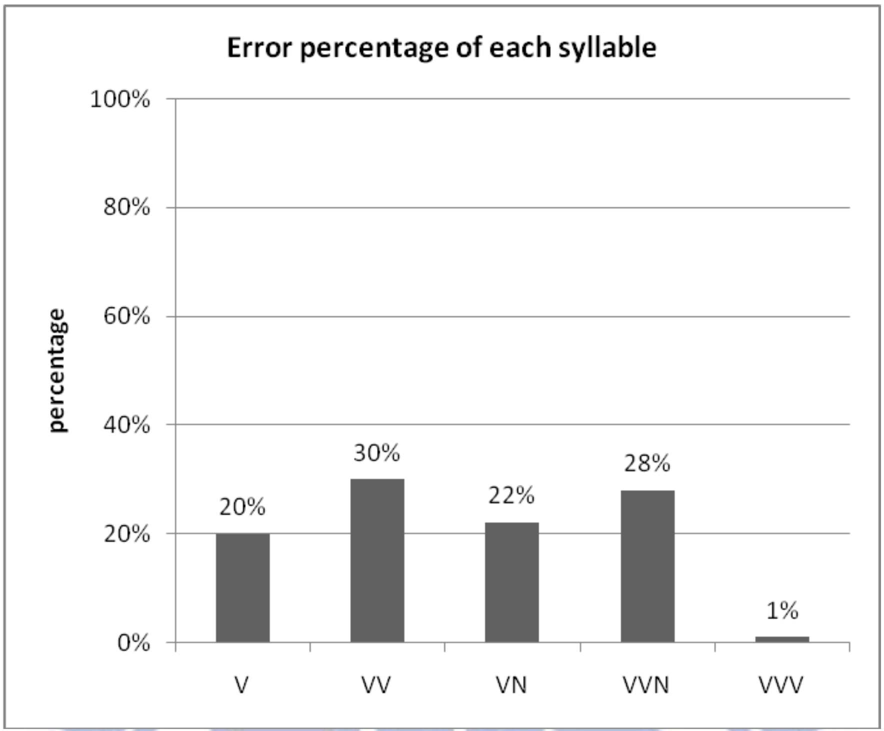


Figure 24 Percentage of vowel errors in each syllable.

Tone

There is a total of 209 tone errors found. Tone 3 was misarticulated the most (105 errors, 50%), followed by Tone 1 and 4 (40 errors, 19%), Tone 2 was the last (24 errors, 12%) **(Fig 25).** For Tone 1 errors, the most frequently error was Tone 1 as Tone 2 (75%), Tone 1 as Tone 4 is next (20%), Tone 1 as Tone 3 is the last (5%) **(Fig 27)**. For Tone 2 errors, the most frequent error was Tone 2 as Tone 4 (42 %), followed by Tone 2 as Tone 1 (37%), and then Tone 2 as Tone 3 (21%) **(Fig 28)**. For the Tone 3 errors, Tone 3 as Tone 2 is the most frequently error (48%), followed by Tone 3 as Tone 4 (32%), and Tone 3 as Tone 1 (20%) **(Fig 29).** For Tone 4 errors, subjects misarticulated Tone 4 as Tone 1 most frequently (52.5%), Tone 4 as Tone 2 next (27.5%), and finally Tone 4 as Tone 3 (20%) **(Fig 30).** The easiest tone which used to substitute other tones is Tone 2 **(Fig 26).**

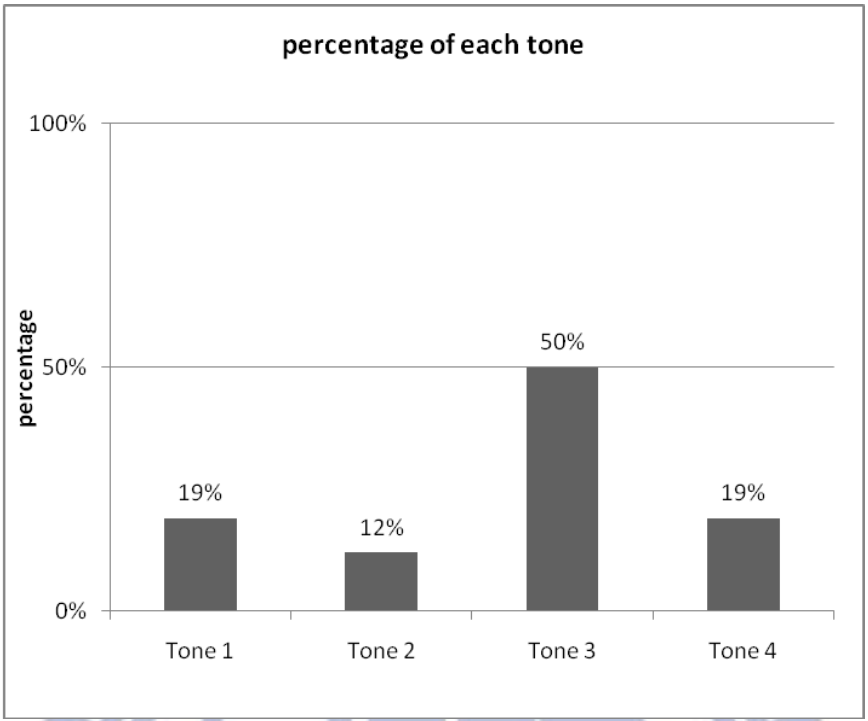


Figure 25 Error percentage of each tone.

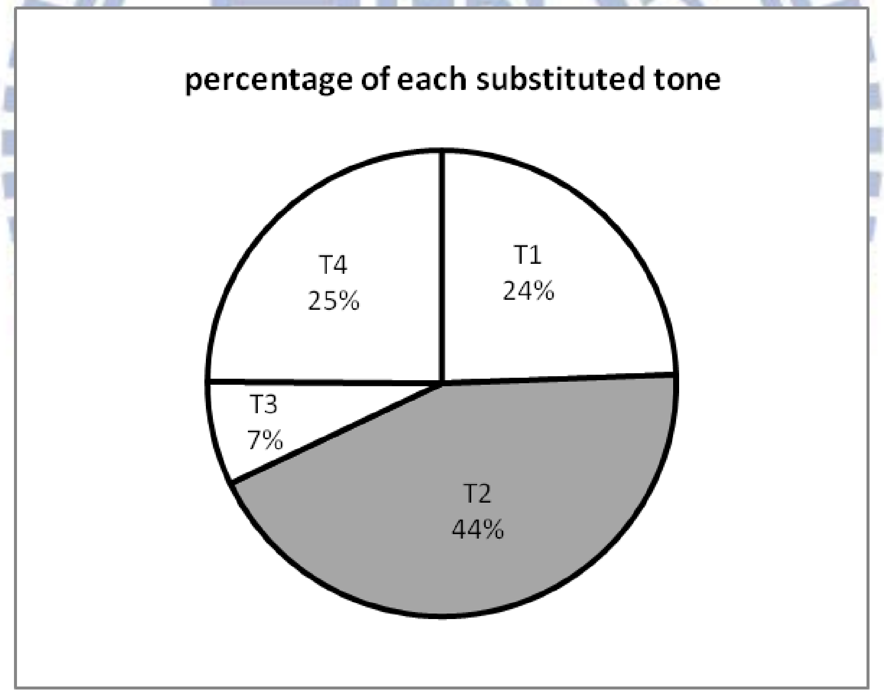


Figure 26 Percentage of each substituted tone.

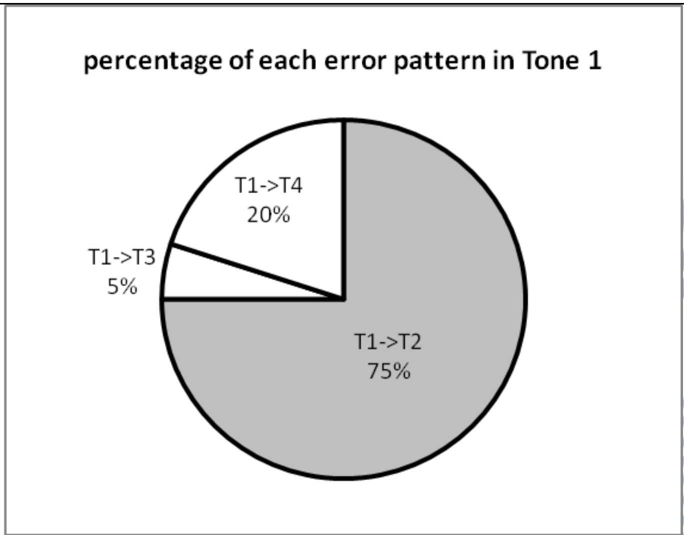


Figure 27 Percentage of each error patterns in Tone 1

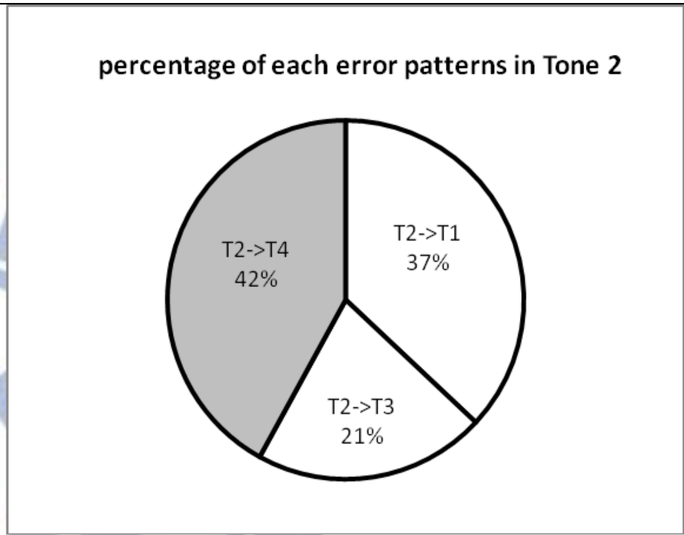


Figure 28 Percentage of each error patterns in Tone 2.

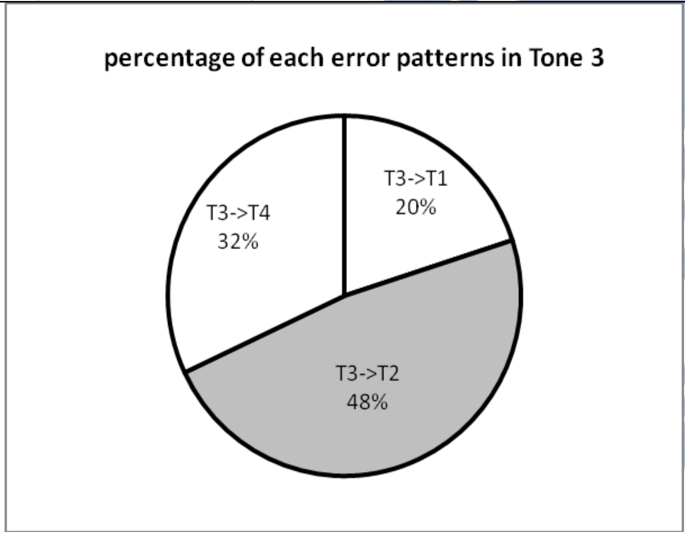


Figure 29 Percentage of error patterns in Tone 3

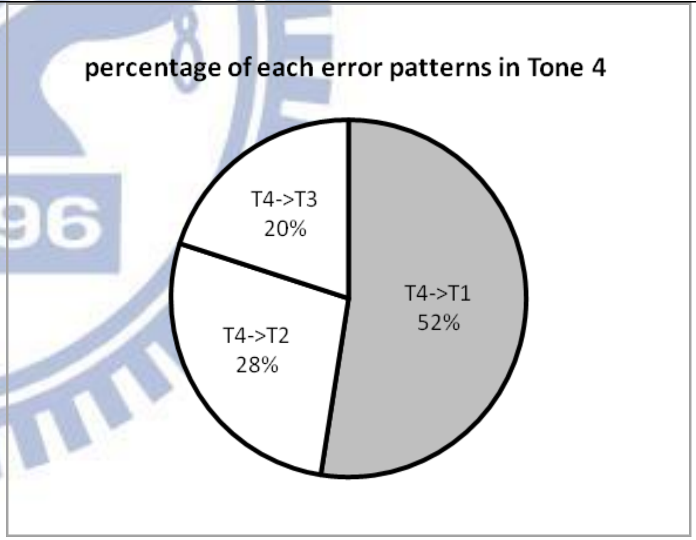


Figure 30 Percentage of error patterns in Tone 4.

3.2 Acoustic analysis

3.2.1 Retroflex

F3 values in voiced retroflex consonants

Result indicated that the F3 average values for voiced retroflex sibilants (ʐɤŋ1, ʐou4 and ʐe4) in correct pronunciation showed F3 dipping, but in incorrect pronunciation showed no F3 dipping **(Fig 31).**

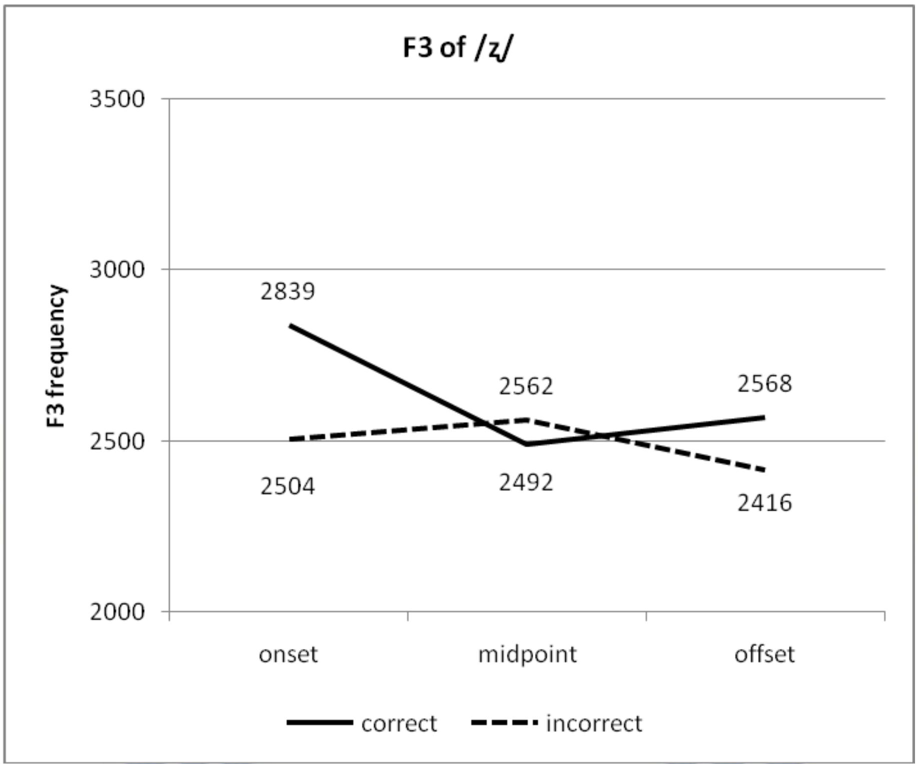


Figure 31 F3 values of /ʐ/ in correct and incorrect pronunciation.

F3 values in vowels (different vowels comparison)

In this section, different vowels in different rhymes were compared in order to make sure that whether different rhymes structures affect F3 values. There are four parts in this section: 1) retroflex + V, 2) retroflex +VV, 3) retroflex +VN, 4) retroflex +VVN.

monophthong /i/, /e/, /u/

As expected, the dipping F3 value of correct pronunciation was found in all three monophthongs (retroflex +V), and the F3 did not dip in incorrect pronunciation for all monophthongs **(Fig 32-34).**

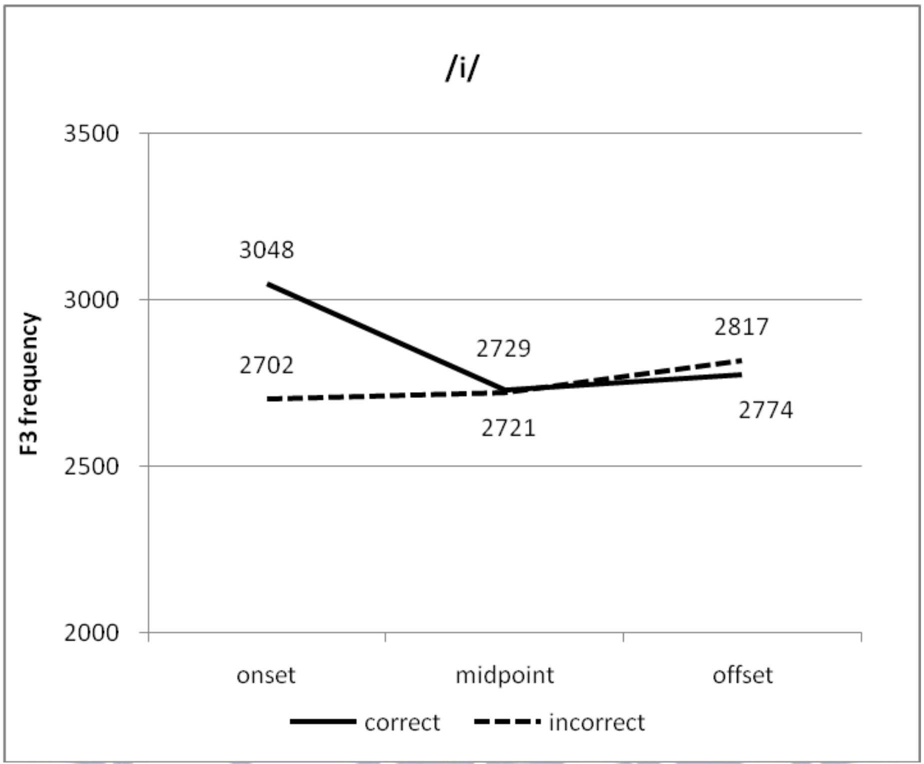


Figure 32 Correct and incorrect pronunciation for retroflex + /i/.

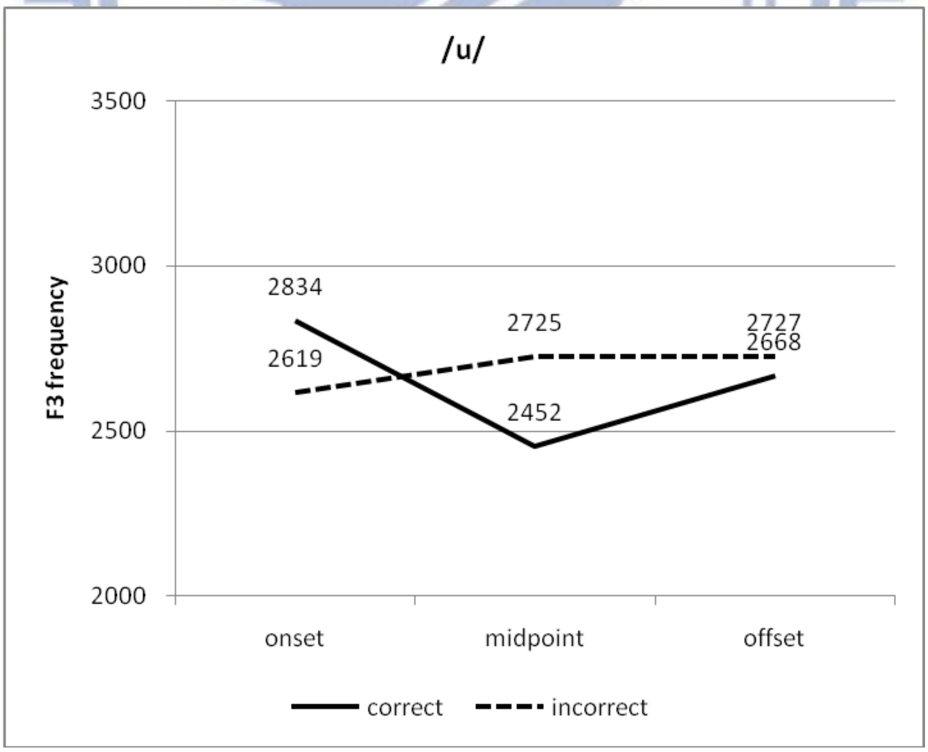


Figure 33 Correct and incorrect pronunciation for retroflex + /u/

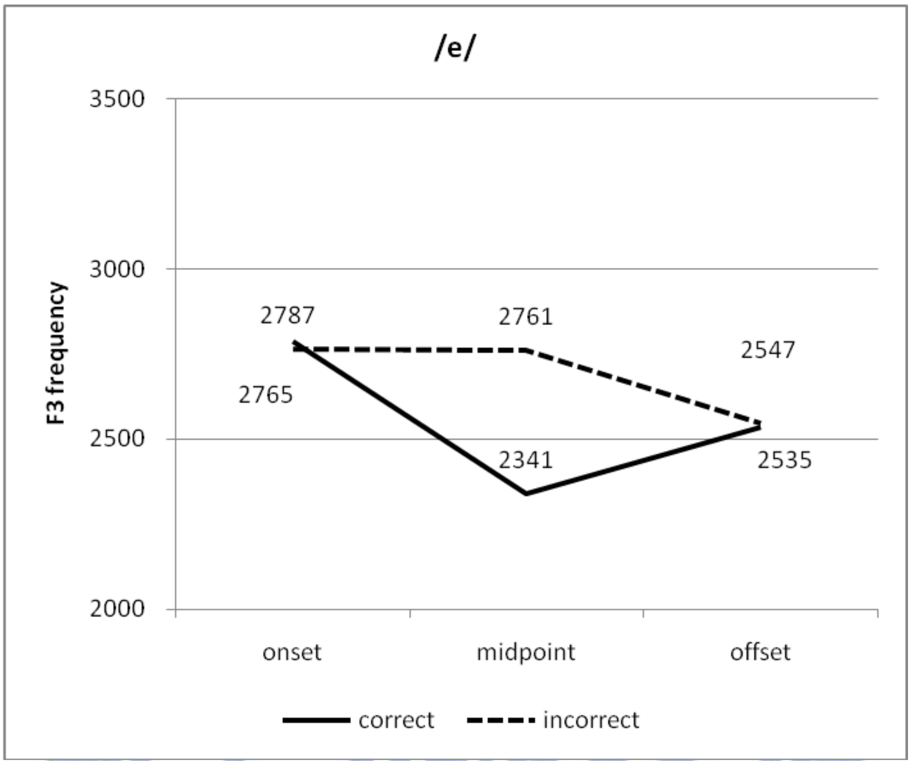


Figure 34 Correct and incorrect pronunciation for retroflex + /e/

diphthong /ua/, /ou/

The F3 values dipped in both /ua/ and /ou/ rhyms were found **(Fig 35-36).**

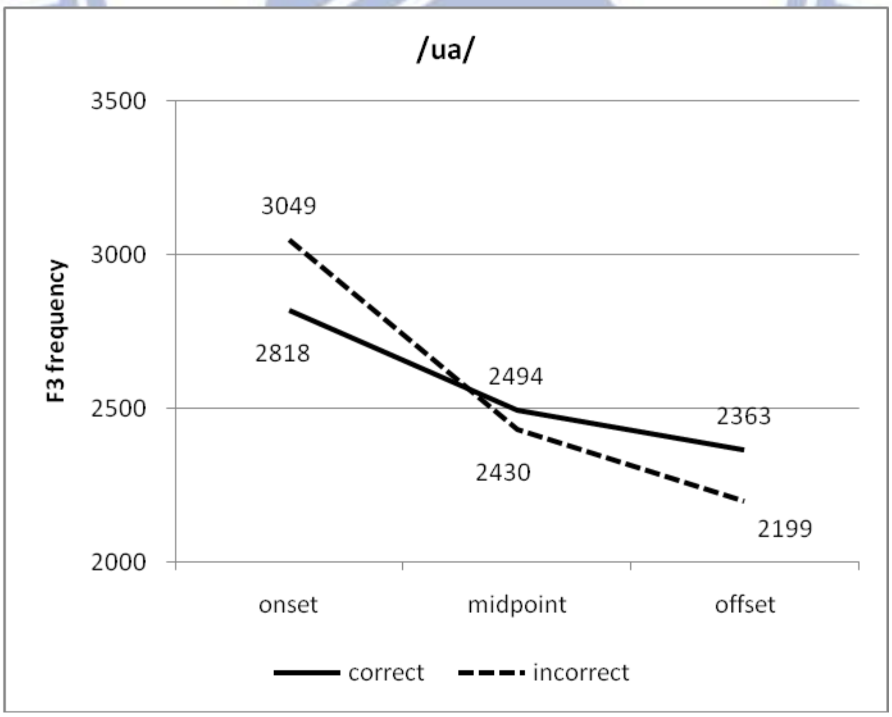


Figure 35 Correct and incorrect pronunciation for retroflex + /ua/.

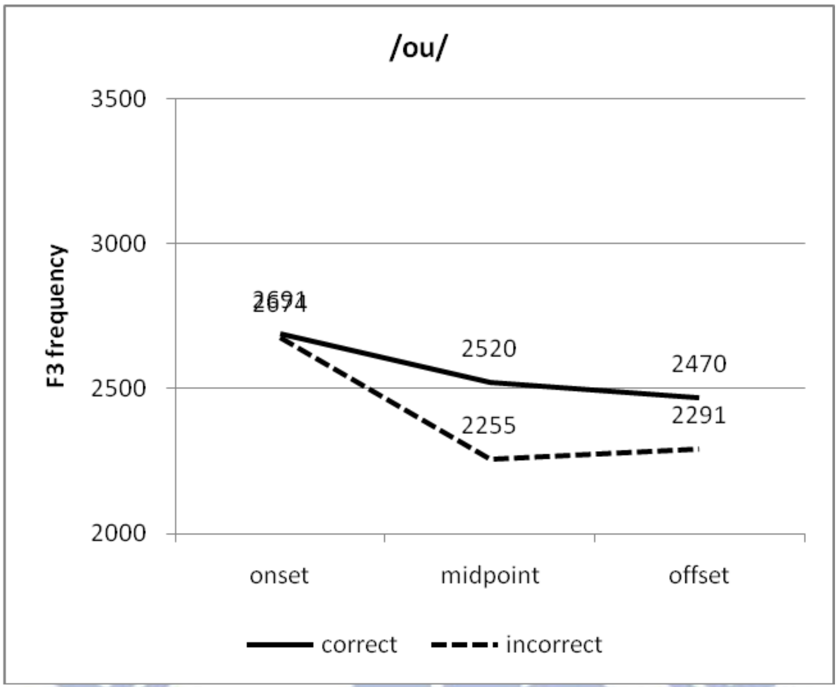


Figure 36 Correct and incorrect pronunciation for retroflex + /ou/

VN /en/, /an/

For VN /en/, the F3 dipping was found in correct pronunciation, but not in incorrect pronunciation. However, for /an/, the F3 dipping was found in both correct and incorrect pronunciation **(Fig 37-38).**

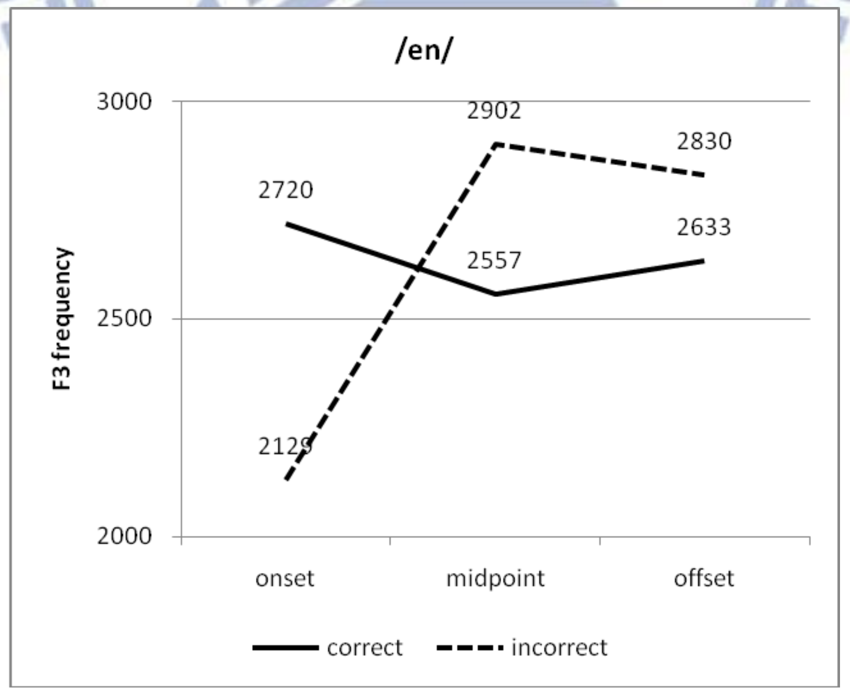


Figure 37 Correct and incorrect pronunciation for retroflex + /en/

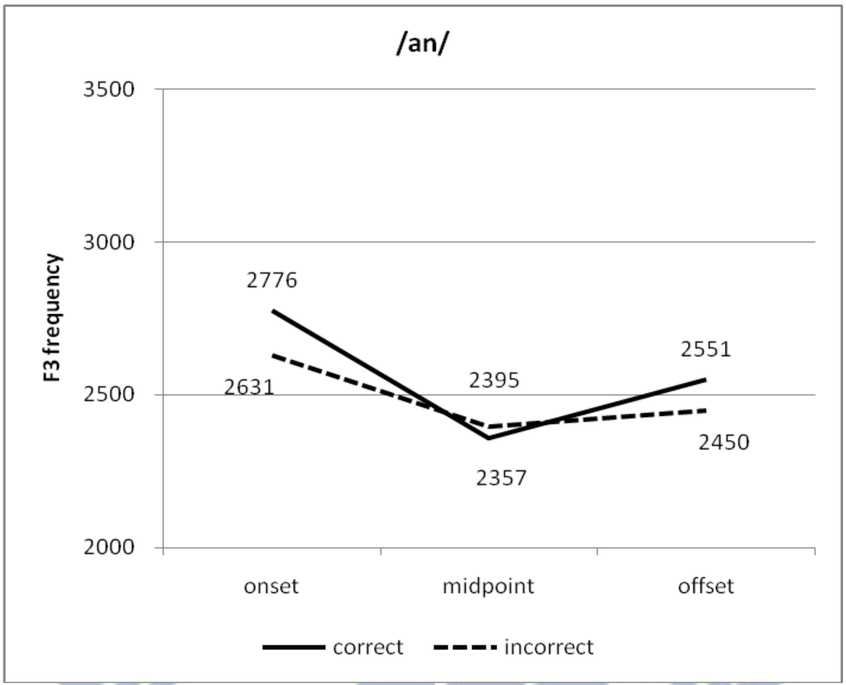


Figure 38 Correct and incorrect pronunciation for retroflex + /an/

VVN /uaŋ/

For /uaŋ/, F3 dipping was found in both correct and incorrect pronunciation **(Fig 39)**.

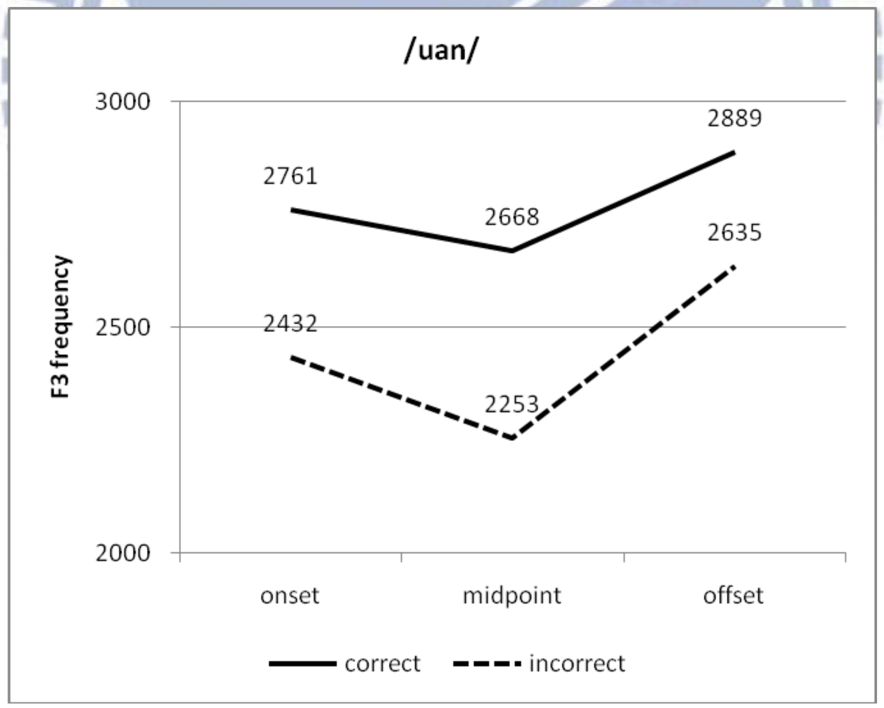


Figure 39 Correct and incorrect pronunciation for retroflex + /uaŋ/

In the next section, the examiner further measured the spectral moment of different syllable structures.

M1 - M4 of voiceless retroflex consonants in four different rhymes

M1 (central gravity of spectrum)

In this section, M1, the central gravity of retroflex, was measured (Jeng, 2006). As in the previous section, four rhymes were measured: retroflex +V, retroflex +VV, retroflex + VN and retroflex +VVN. Results indicated that the M1 value of retroflex in correct pronunciation was significantly lower than that of incorrect pronunciation for all the four syllable structures. **(Fig. 40-44)** According to Jeng (2006), M1 value is the most robust for distinguishing retroflexion.

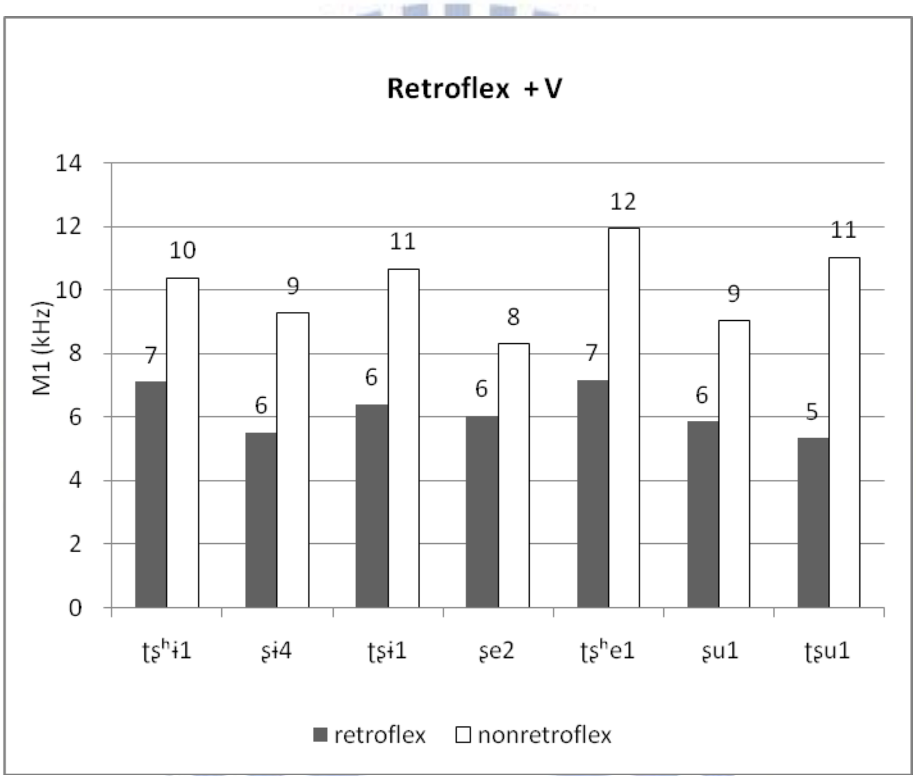


Figure 40 M1 values in correct and incorrect pronunciation for retroflex +V

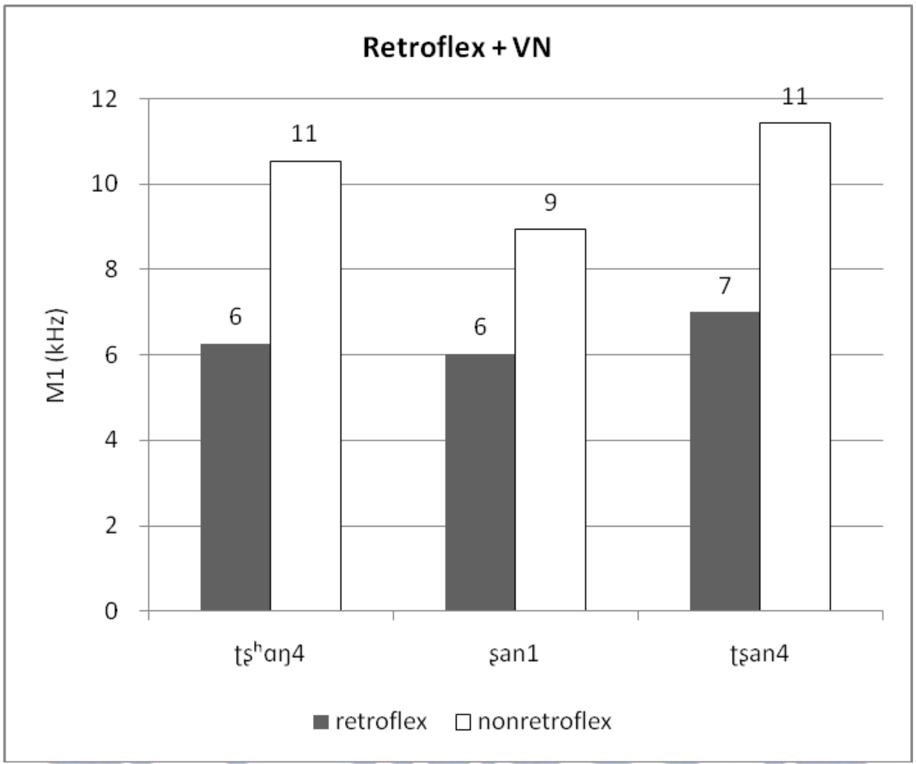


Figure 41 M1 values in correct and incorrect pronunciation for retroflex +VN

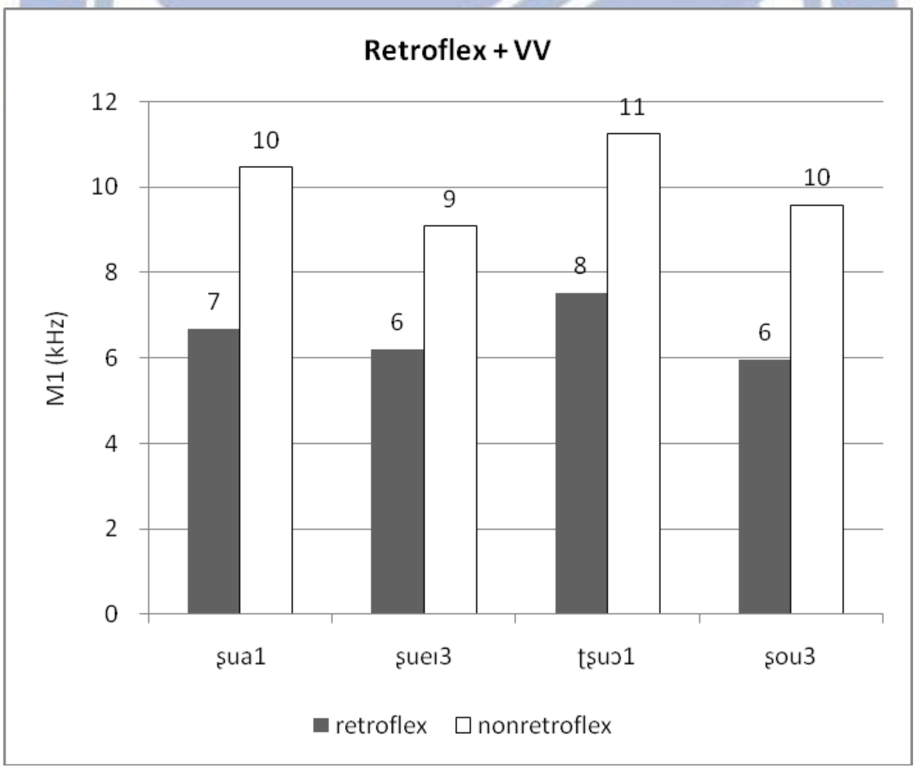


Figure 42 M1 values in correct and incorrect pronunciation for retroflex +VV

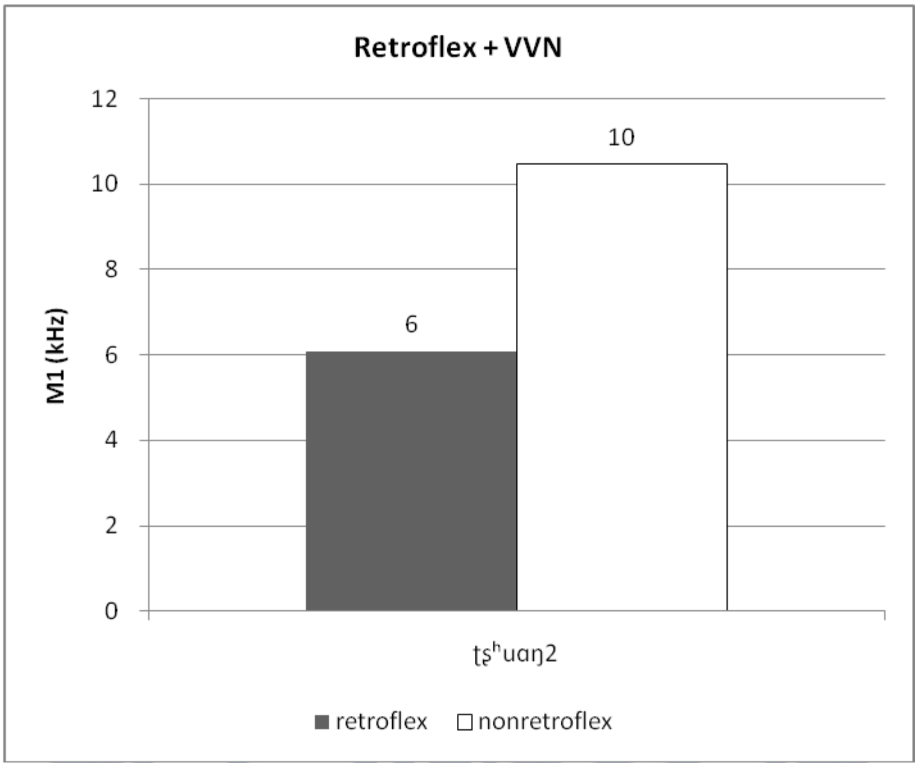


Figure 43 M1 values in correct and incorrect pronunciation for retroflex +VVN

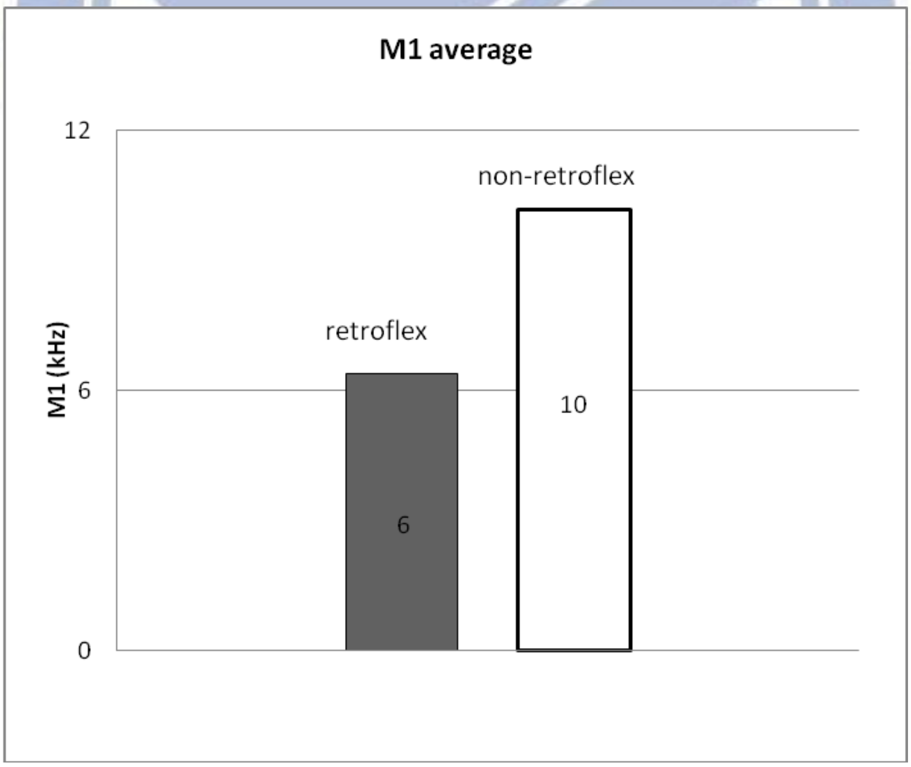


Figure 44 M1 average for four syllables structures.

M2 (diversity)

Four rhymes were measured: retroflex +V, retroflex +VV, retroflex + VN and retroflex +VVN. Results indicated that the M2 value of retroflex in correct pronunciation was lower than that of incorrect pronunciation for all the four syllable structures **(Fig. 45-49)**, which were different to the results of Jeng (2009).

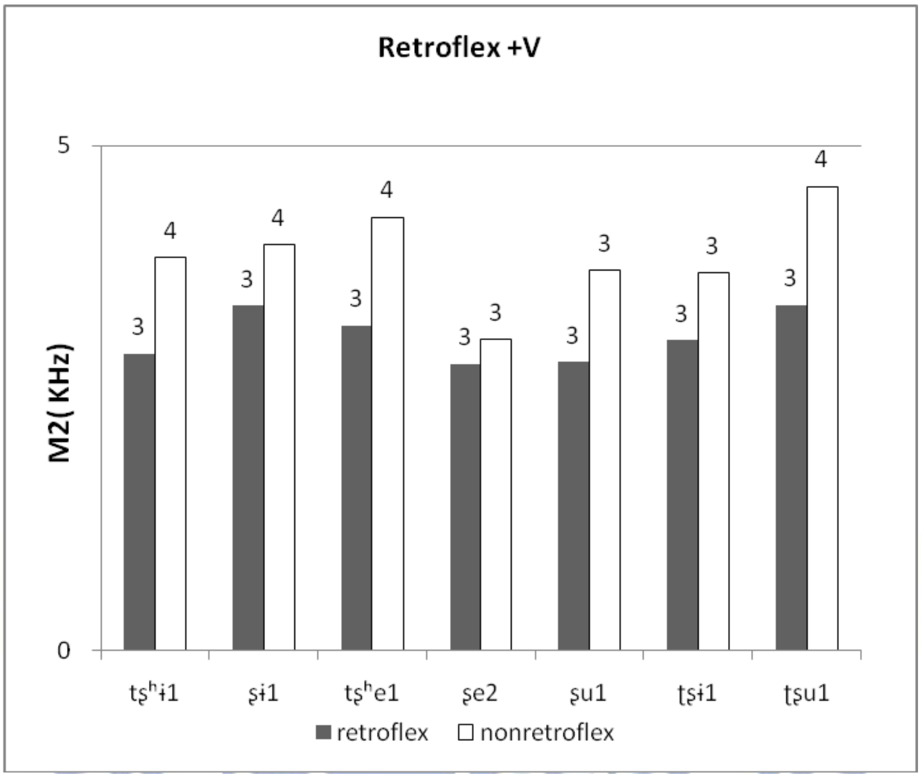


Figure 45 M2 values in correct and incorrect pronunciation for retroflex +V

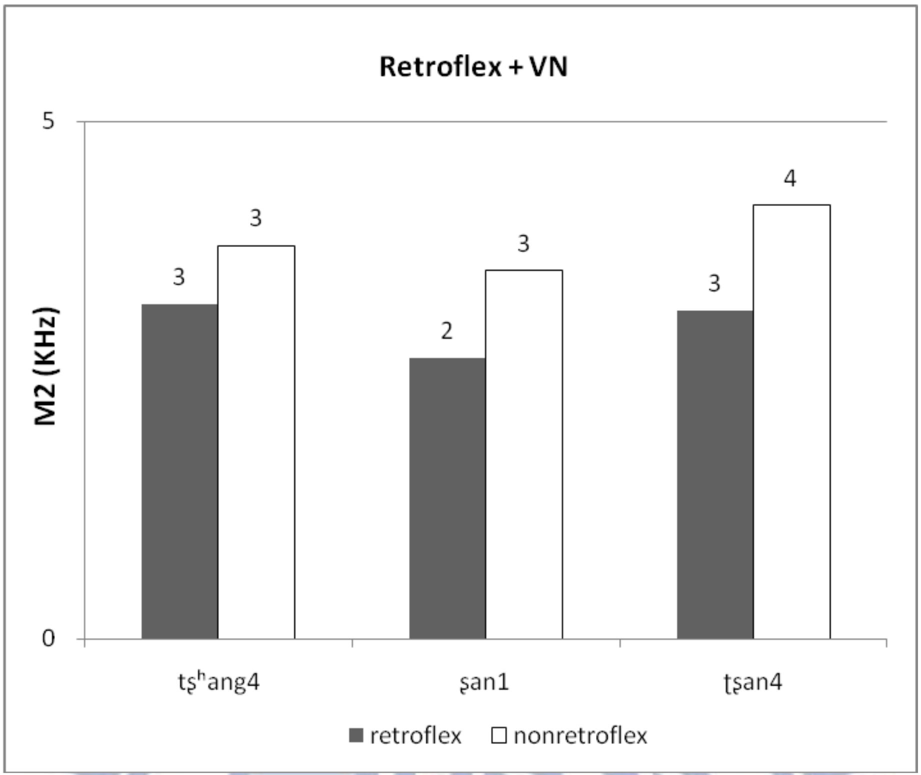


Figure 46 M2 values in correct and incorrect pronunciation for retroflex +VN

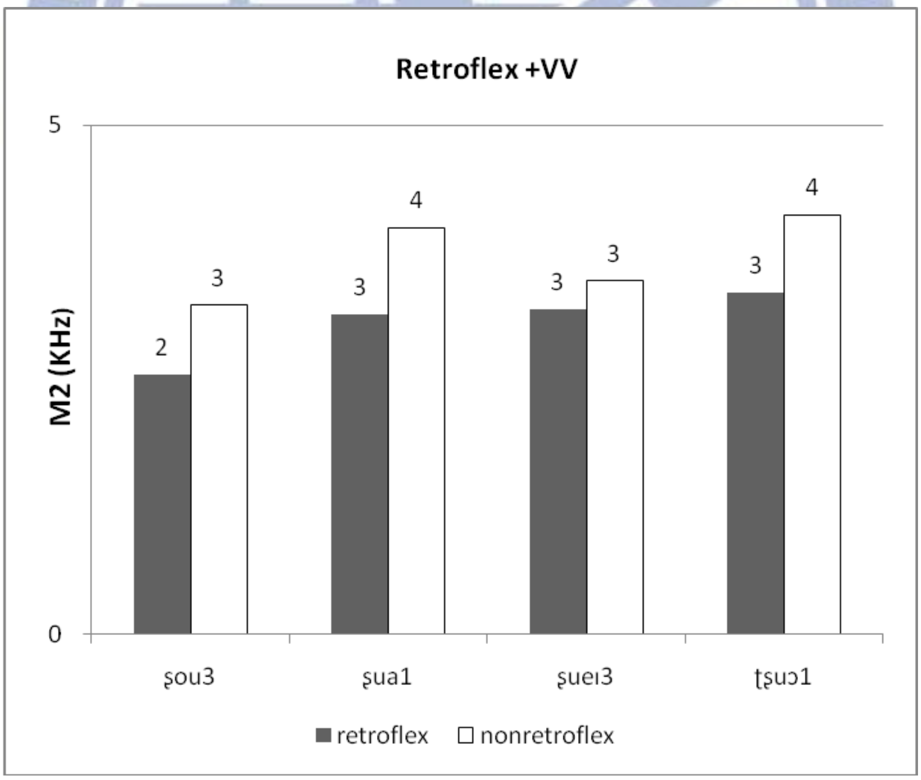


Figure 47 M2 values in correct and incorrect pronunciation for retroflex +VV

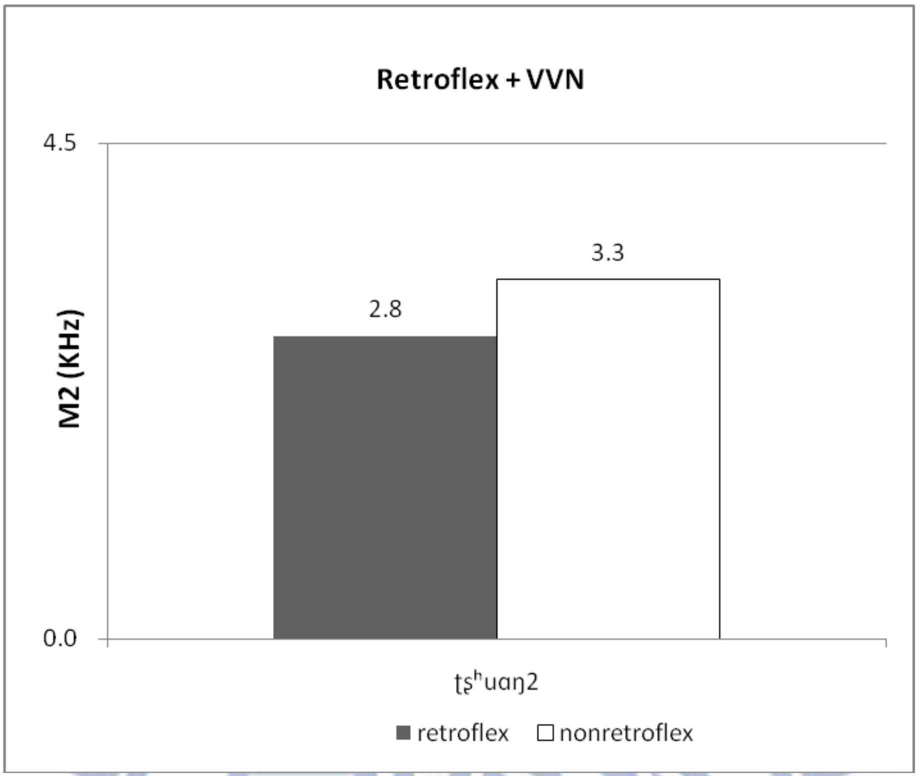


Figure 48 M2 values in correct and incorrect pronunciation for retroflex +VVN

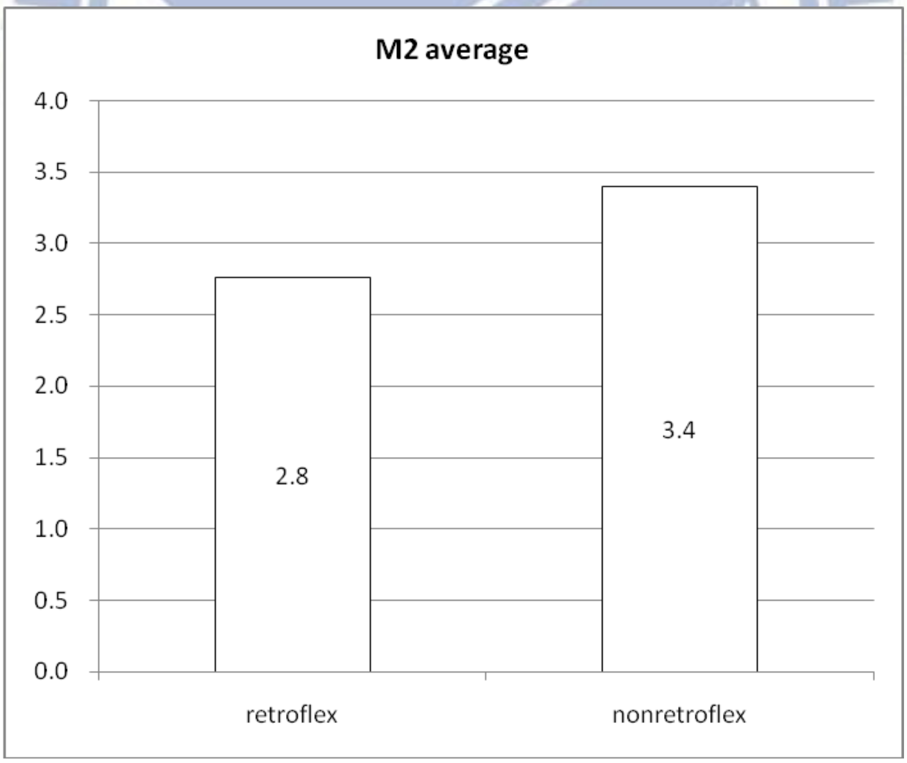


Figure 49 M2 average for four syllables structures.

M3 (spectral tilt)

In this section, M3, the spectral tilt of retroflex, was measured (Jeng, 2006). Results indicated that the M3 value of retroflex in correct pronunciation was significantly higher than that of incorrect pronunciation for all the four syllable structures **(Fig. 50-54).**

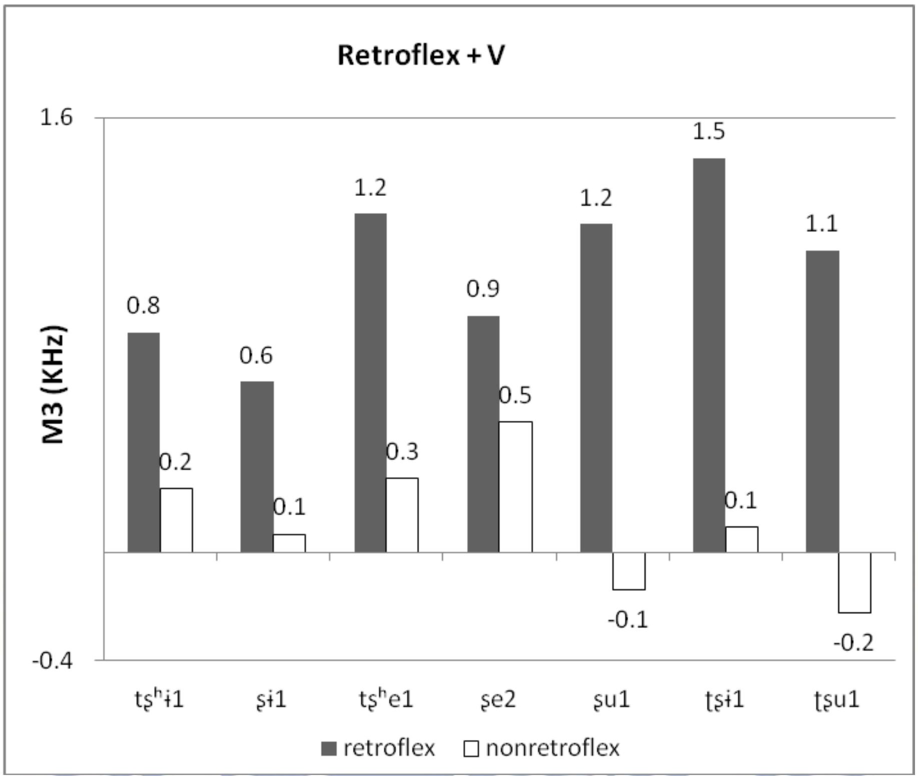


Figure 50 M3 values in correct and incorrect pronunciation for retroflex + V

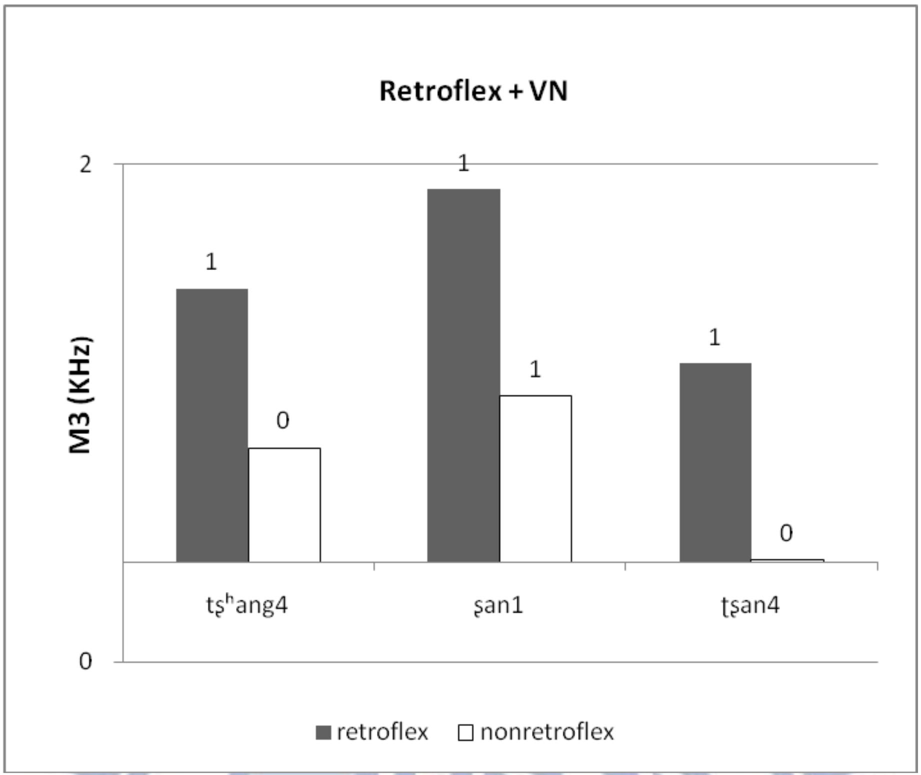


Figure 51 M3 values in correct and incorrect pronunciation for retroflex + VN

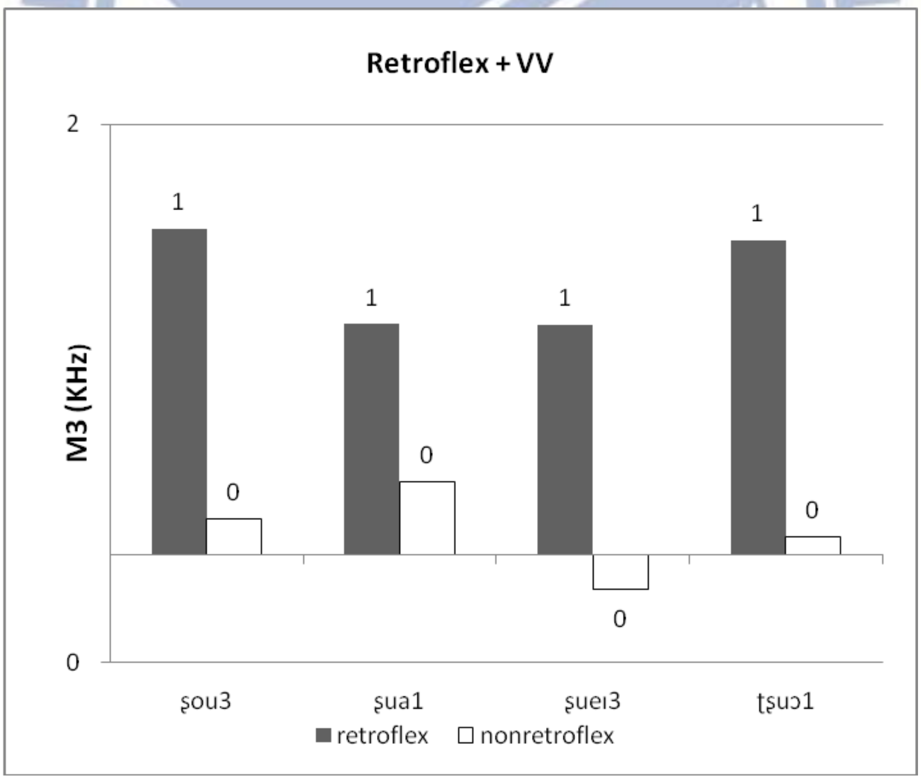


Figure 52 M3 values in correct and incorrect pronunciation for retroflex + VV

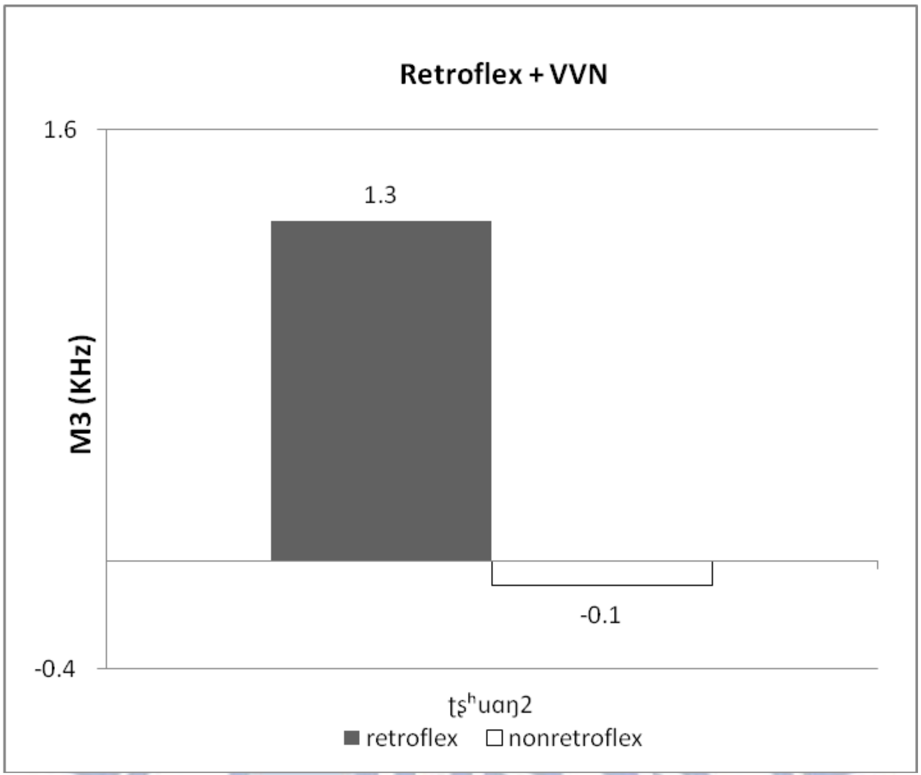


Figure 53 M3 values in correct and incorrect pronunciation for retroflex + VVN

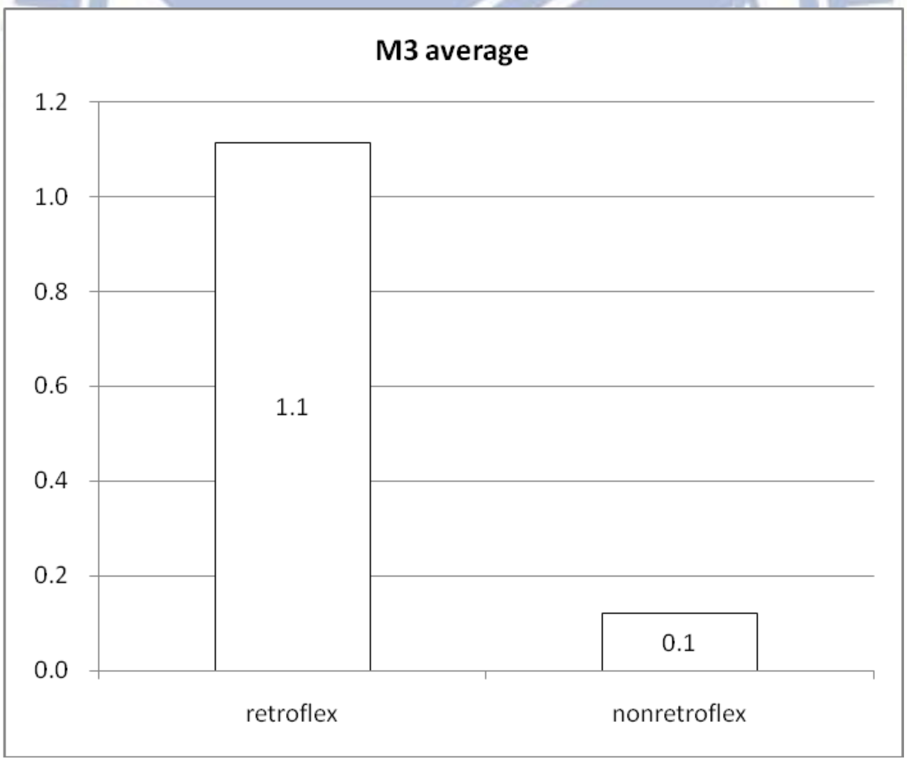


Figure 54 M3 average for four syllables structures.

M4 (kurtosis)

M4, the central gravity of retroflex, was measured (Jeng, 2006). There are four rhymes measured. Results indicated that the M1 value of retroflex in correct pronunciation was significantly higher than that of incorrect pronunciation for all the four syllable structures **(Fig. 55-59)** which were different to the results of Jeng (2009).

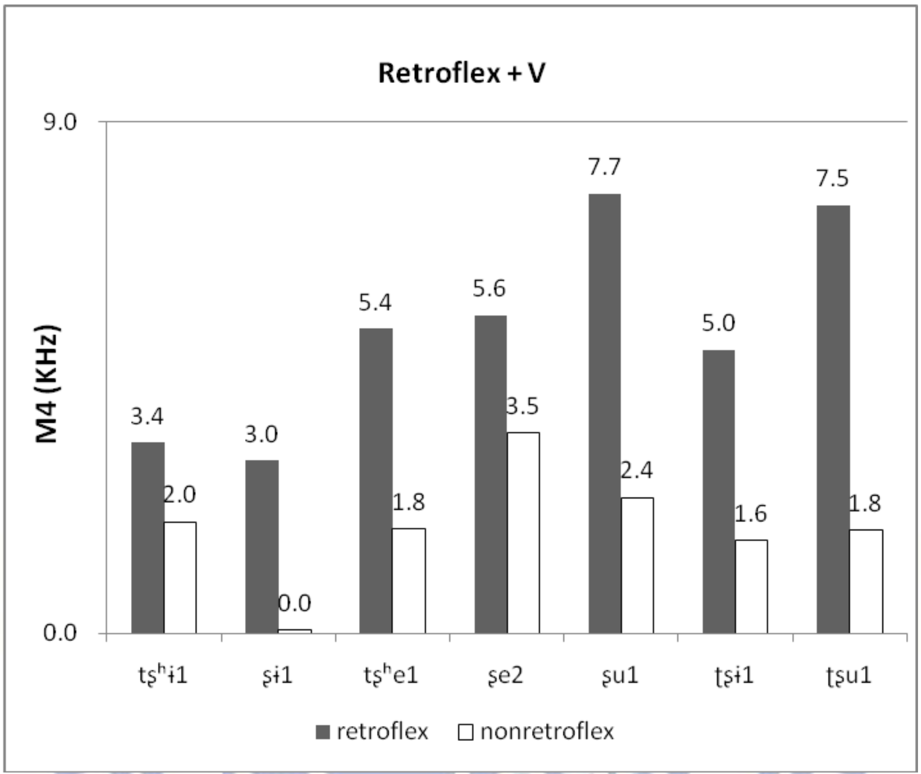


Figure 55 M4 values in correct and incorrect pronunciation for retroflex +V

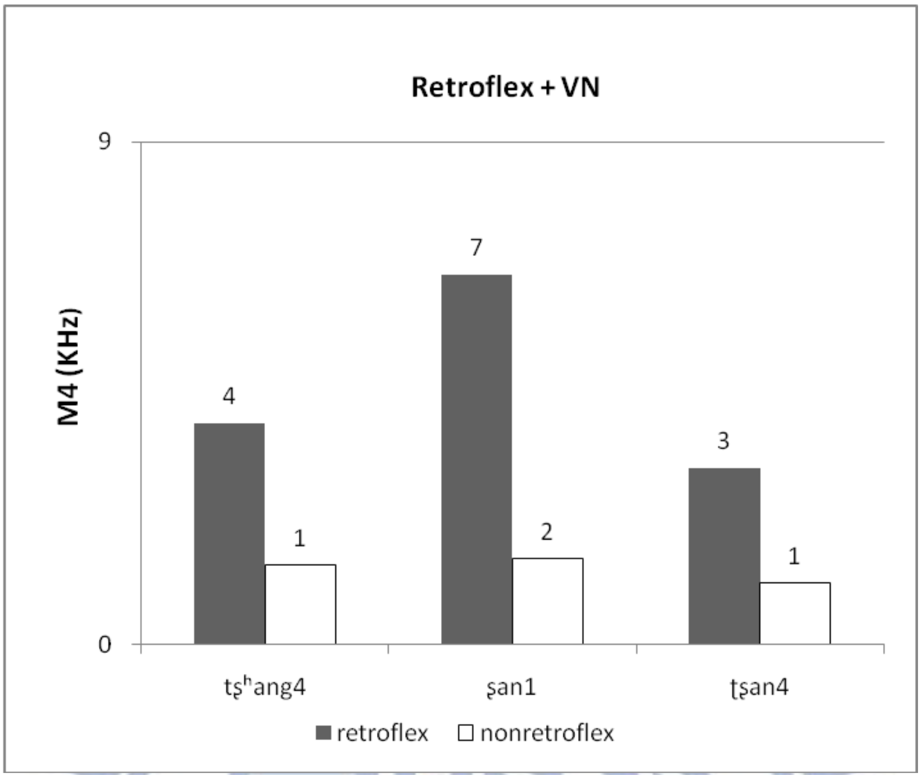


Figure 56 M4 values in correct and incorrect pronunciation for retroflex +VN

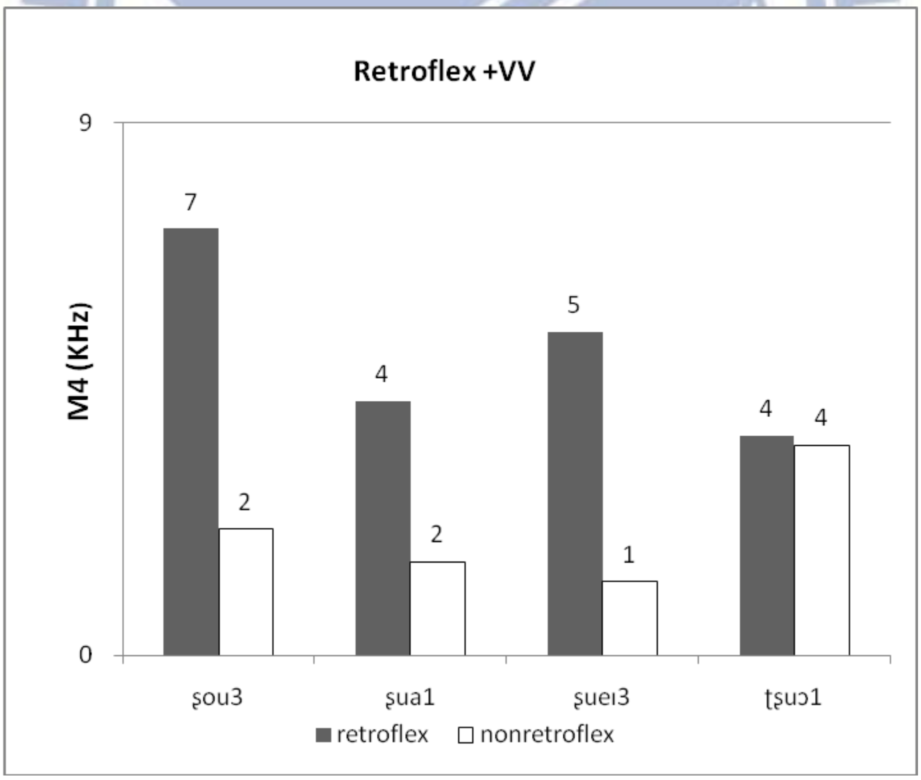


Figure 57 M4 values in correct and incorrect pronunciation for retroflex +VV

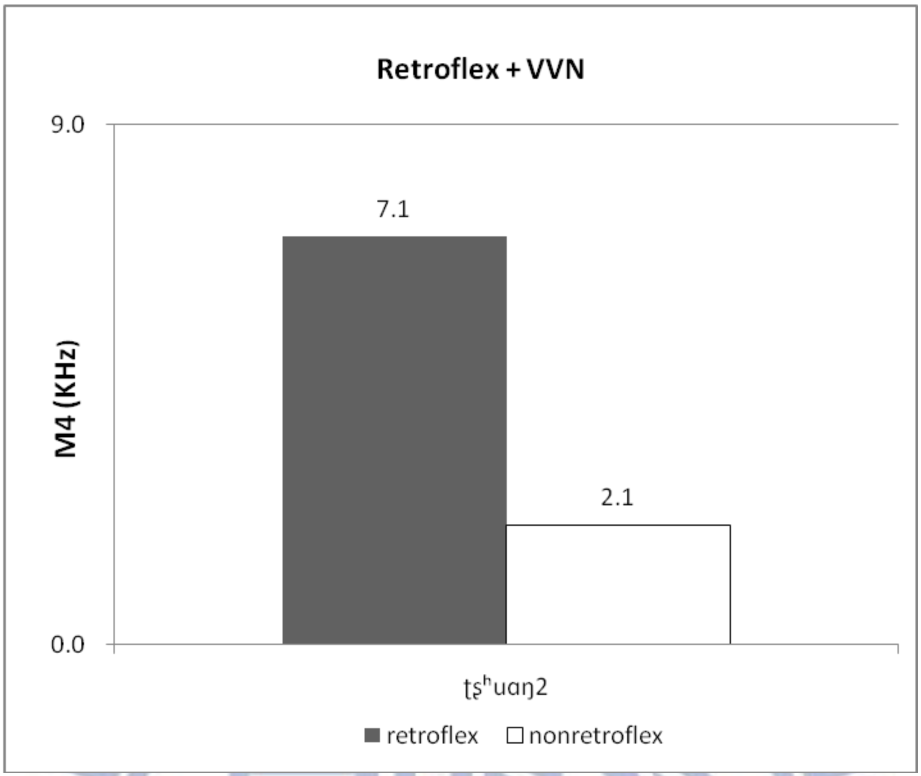


Figure 58 M4 values in correct and incorrect pronunciation for retroflex +VVN
)

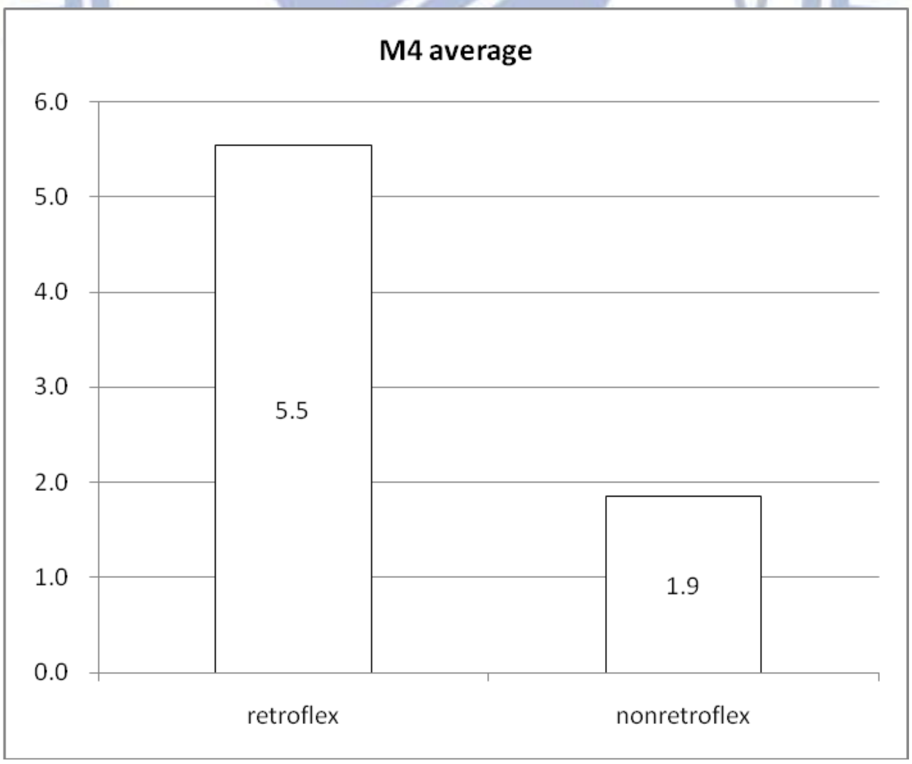


Figure 59 M4 average for four syllables structures.

3.2.2 Vowels

The mid-points of F1, F2 of monophthongs were measured. The peripheral vowel /i/, /a/, /u/ were used to draw the range of vowel space as a triangle. Results indicated that the vowel space is narrow in children's production **(Fig.60)**. Shanghainese (Chen, 2008) show a narrower vowel space than that of Mandarin (Zee & Lee, 2001) **(Fig. 61)**. The three vowel spaces comparison were showed in **Fig. 62.**

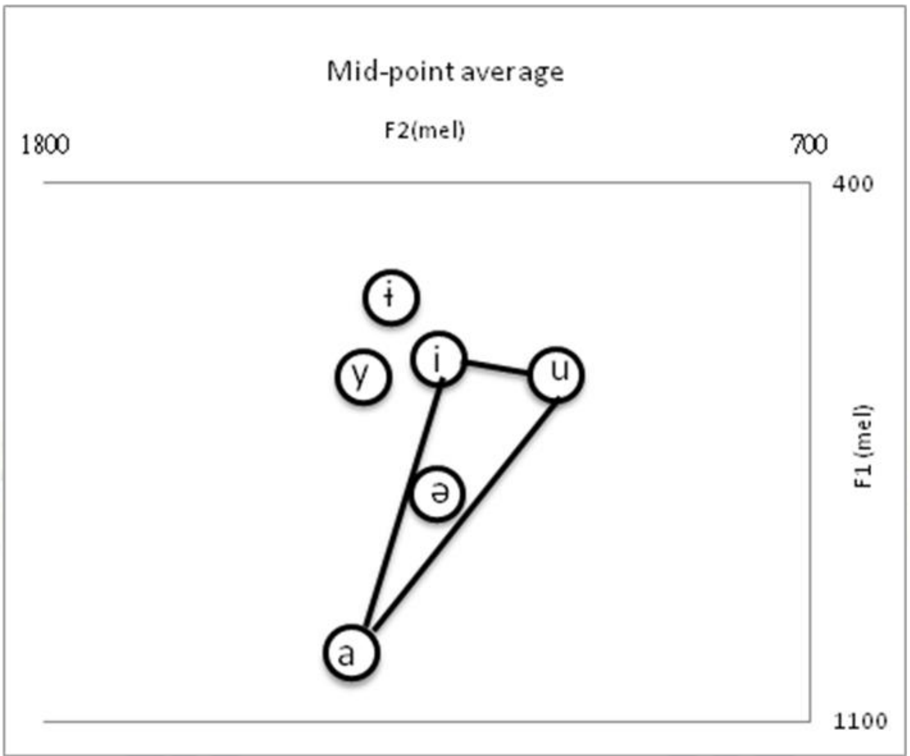


Figure 60 Averages of six monophthongs on the mid-point by Shanghainese children

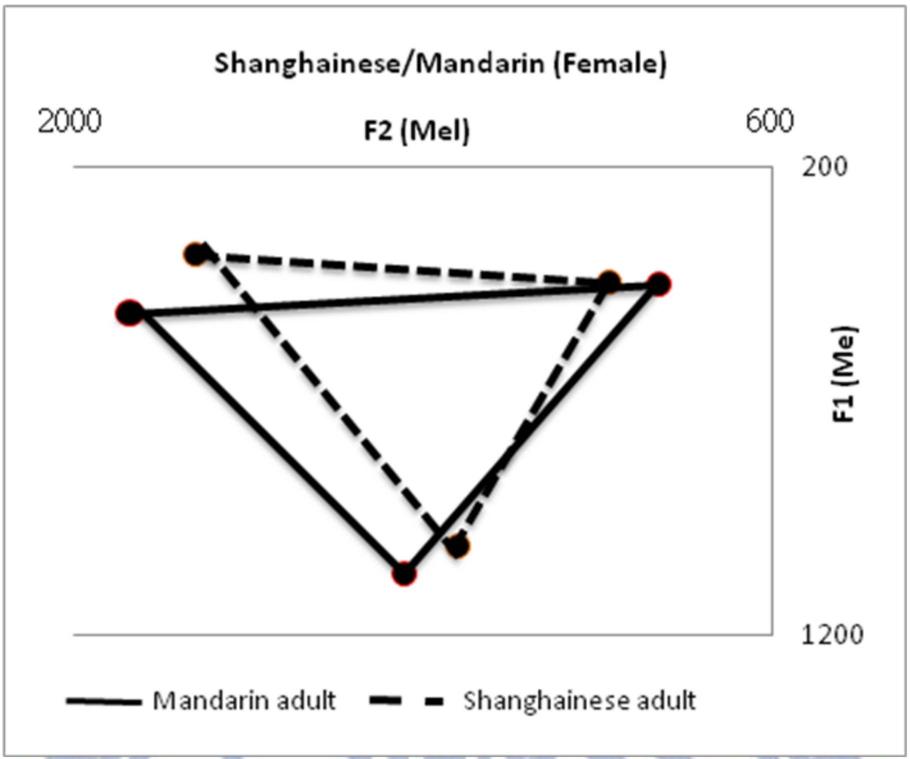


Figure 61 Shanghainese and Mandarin female vowel space comparison.

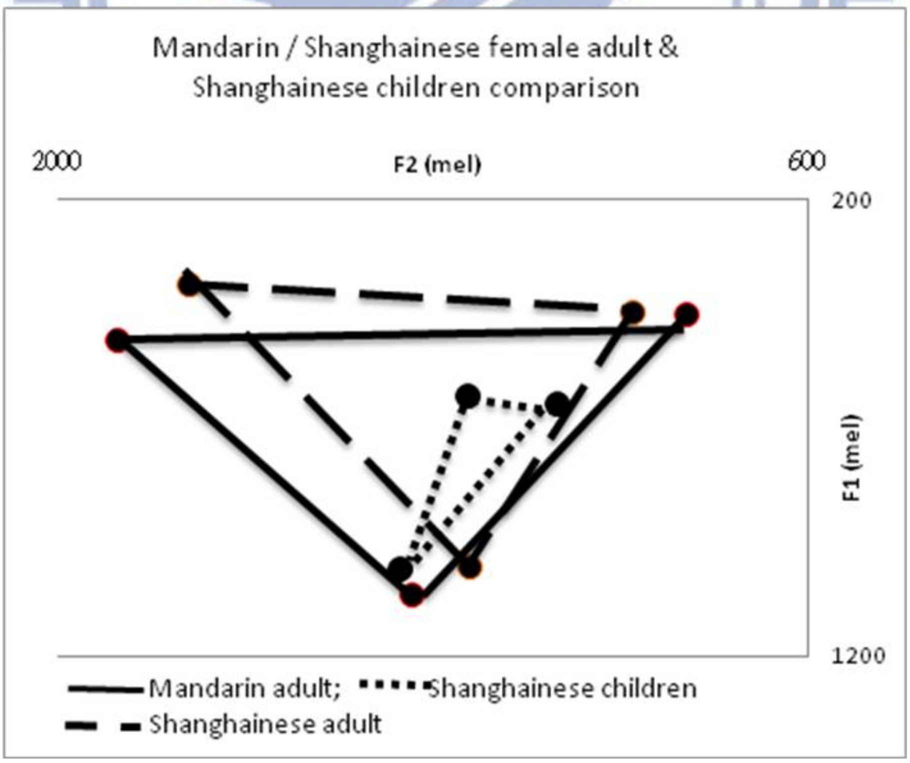


Figure 62 Shanghainese, Mandarin adult and Shanghainese children vowel space comparison.

3.2.3 Tones

The four tones of 30 subjects in final position were shown in Figure 63. Results indicated that the F0 contours of children are like that of adults **(Fig. 63-64)**. Results also indicated that the Turning Point of Tone 2 is about 30 percent of the tone, which is earlier than that of Tone 3 (about 60 percent). Tone 2△ F0 is 6.6259Hz (onset 249 – Turning Point 242=6) which is smaller than Tone 3△ F0 (onset 155 - Turning Point 115=40). The adult norm for tone duration is T3/T2 longer than T1, and T4 is shortest **(Fig. 64)**, the tone duration of children **(Fig. 65)** perform an adult-like pattern. The four tones are significantly different **(Fig. 67)**. These children managed to utilize F0 cue as well as duration cue. However, the pitch range of T4 is not as wide as that of adults, which might be the results of their L1 (Shanghainese) **(Fig.66)**, which does not have real falling tone as Mandarin T4.

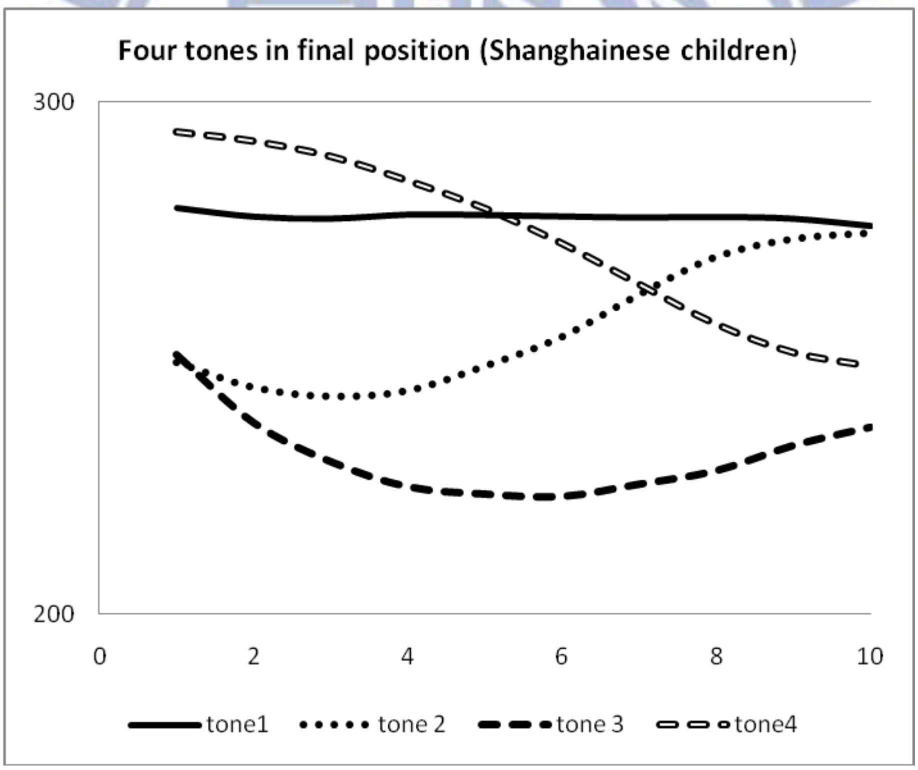


Figure 63 Four tones in final position of 4-year-old Shanghainese children

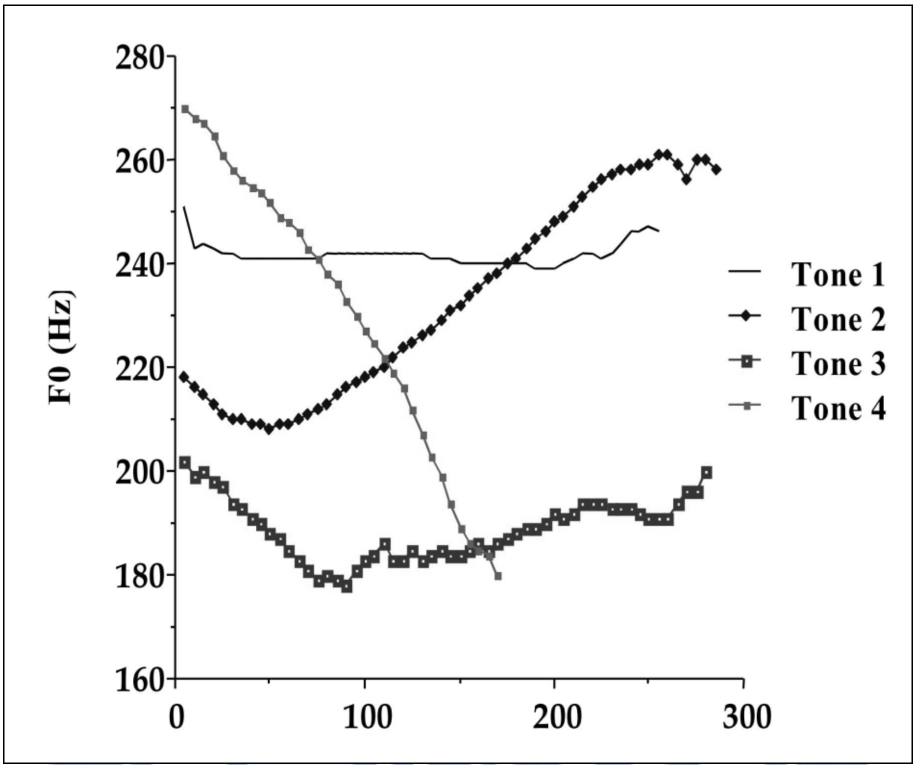


Figure 64 F0 contours for each of the four Mandarin Chinese tones by a female adult speaker (from Moore and Jongman, 1997)

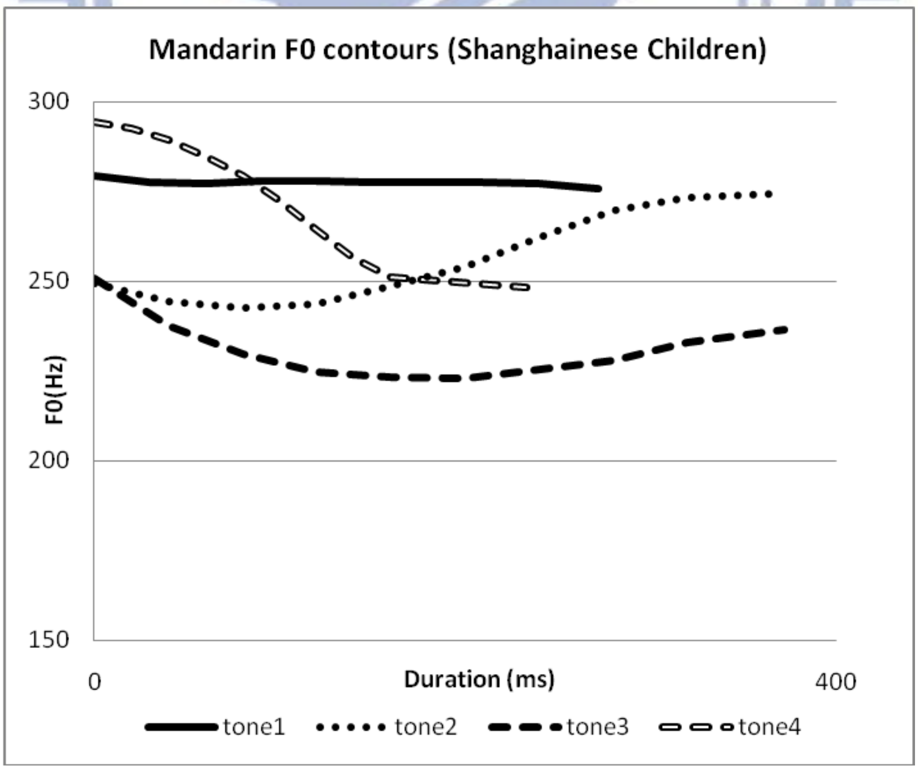


Figure 65 F0 contours for four Mandarin tones produced by Shanghainese children.

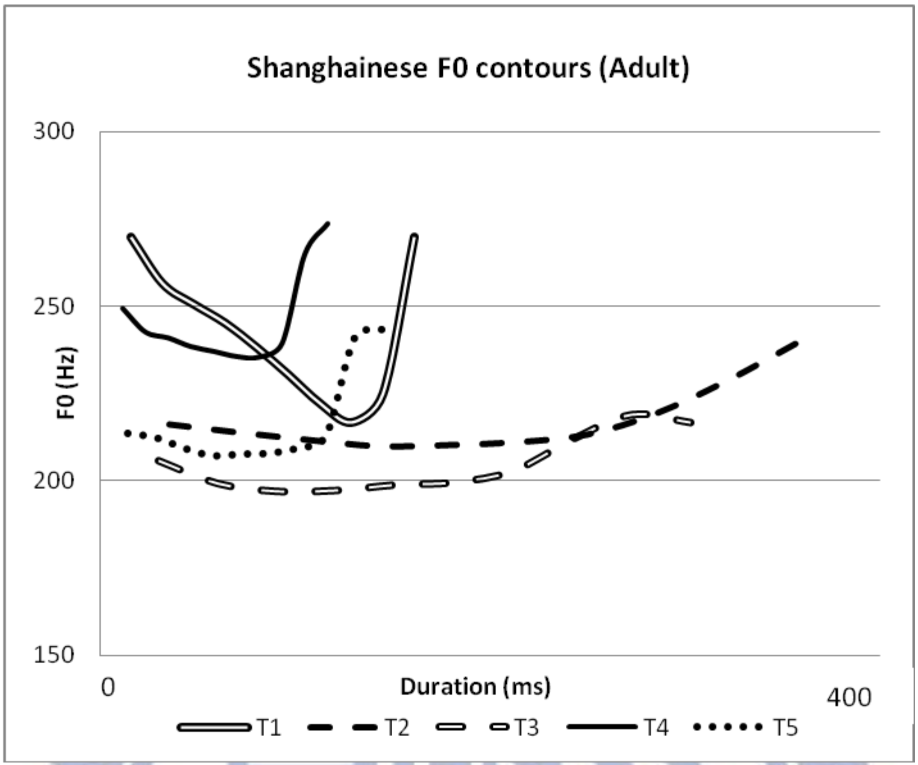


Figure 66 F0 contours for each of the five tones by Shanghainese adult speaker.

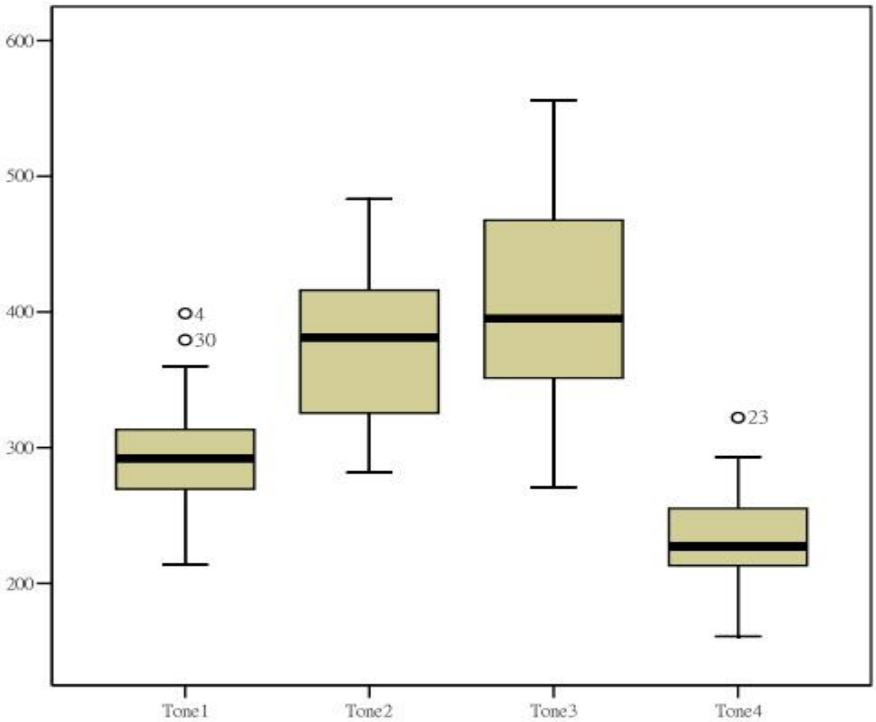


Figure 67 Box plot of the four tones which produced by subjects.

Chapter 4. Discussion and Conclusions

The current study examines the phonetic realizations of Mandarin by 4-year-old Shanghainese children. The possible effects of their first language (Shanghainese) on their second language (Mandarin)is also discussed. Eighty target words were recorded from 30 children. The production data was transcribed and an acoustic analysis was conducted on retroflexes, vowels, and tones. Transcription data yielded nine speech error patterns: (1) Non-retroflex produced as retroflex, (2) Deretroflexion, (3) Substitution, (4) Nasal confusion, (5) /h/ substitution, (6) /tɕ^h, tɕ, ɕ/ confusion, (7) Vowel confusion, (8) Tone confusion, and (9) Others. Details of these errors will be discussed below. Acoustic analysis indicated an inconsistency between F3 dipping and rhyme structures, a detailed discussion will follow. In line with Jeng's (2009) observation, the findings indicated the most robust cues for retroflexion is M1 (central gravity) in spectral moment. The vowel space of Shanghainese children was found narrower than that of Mandarin-speaking and Shanghainese-speaking adults. This result will be discussed in terms of universal development trends as well as with first language interference. Regarding tone acquisition, the children's F0 contours and duration patterns are similar to that of Mandarin adult speakers, which echoes Li and Thompson's (1977) findings, suprasegmental acquisition generally precedes segmental acquisition.

4.1 Transcription analysis

Nine speech errors were found from the transcription data: (1) Non-retroflex produced as retroflex, (2) Retroflex produced as non-retroflex (3) Substitution: Mandarin /ʐ/ misarticulated as /l/, /n/, or deretroflexed, (4) Nasal confusion: velar nasal produced as alveolar nasal or final nasal deletion. (5) /h/ substitution: /h/ produced as /f/ or deleted. (6) /tɕ^h, tɕ, ɕ/ confusion: these three consonants are randomly substituted by one another, (7) Vowel confusion: diphthong produced as monophthong, or certain vowels were substituted by one another (8) Tone confusion: the most common error: T3 produced as T2, T4 as T1, and (9) Others. These errors are grouped and discussed at the segmental level: consonant and vowel as well as suprasegmental levels below.

Consonant

Regarding voicing, Bernhardt and Stemberger (1998) indicated that voiced segments are acquired before voiceless ones. However, voicing is not contrastive in Mandarin, so the discussion focused on the other contrastive laryngeal feature- aspiration. In general, unaspirated consonants are usually acquired before aspirated ones (Bernhardt and Stemberger, 1998). More solid evidence supporting this claim can be found in prevailing cases of /tɕ^{h}/ realized as [tɕ]. This type of substitution was not observed in aspiration-contrasted stops which might be the result of stops being one of the earlier acquired MOA, hence facilitating the acquisition of aspiration unlike affricates which are acquired later and may hinder the acquisition process.

In terms of MOA, the agreement among scholars is that the stops are acquired first, followed by the nasals, fricatives, and finally the affricates. The liquids are constantly avoided in the early stages (Goodluck, 1991). The results of present study show a universal development in manner of articulation which indicate that affricates were the most frequent errors, followed by fricatives, then nasals, and stops were the last.

In terms of POA, Bernhardt and Stemberger (1998) suggested that labials are one of the earliest place of articulation to be acquired followed by coronals and then by dorsals. Surprisingly, the error rates these POA present in the present study are bilabials (4%), coronals (94%), and dorsals (3%). That is, these children have mastered bilabials and dorsals but have extreme difficulties with coronals. Furthermore, children produced the most errors for coronal affricates (47%), coronal fricatives (42%), coronal nasals (6%) and coronal stops (5%) (cf. Fig. 18). This result is postulated to be caused by the large inventory of coronal affricates in Mandarin. Interestingly, these Shanghainese children were found to master dorsals quite well. It postulated to be the result of the greater dorsal inventory in Shanghainese - /ʔm/, /ʔn/, /k/, /k^{h}/, /g/, /h/, /ɦ/, /ʔ/, /ʔɲ/, /ŋ/, /ʔŋ/.

Given that most retroflexes in Mandarin are affricates, which are one of the most difficult POA for children to acquire, many errors were found in our data. Examples such as /tʂ^{h}/ realized as unretroflexed /ts^{h}/ while /tʂ/ were realized as [ts]. The result concurs with the universal developmental that retroflex is acquired fairly late in children's learning process. It also echoes the finding of Jeng (2009) which indicated that children do not acquire retroflex until they are six years old.

Vowel

The vowel development of Shanghainese children in the present study is ranked by error percentage: /u/ (27%), /a/ (26%), /e/ (19%), /o/ (15%), and /i/ (12%). Although previous research showed different patterns in vowel development, some studies indicate that the low vowels [æ], [ɑ], [a], mid vowels [ʌ], [ɛ] and high vowel [i] are early mastered, while back rounded vowel [u] is a late mastery (Bernhardt & Stemberger, 1998; Ito, 1990; Fee, 1991) (see 1.2.1 for detailed review). The high frequcy of /u/, and /a/ misarticulation agree with the universal development; in general, the rounded vowel [u] is mastered late and it is assumed that [a] has more features to be acquired later (Bernhardt & Stemberger, 1998).

Interestingly, in most of the misarticulated cases, the above mentioned vowels were overwhelmly substituted by /a/ (38% of the cases), followed by /e/ (21.73%), /u/ (16.3%), /i/ (10.86%), /o/ (9.78%), and /y/ (3.26%). Although [a] is mastered late, it appeared frequently in early stages (Ito, 1990; Fee, 1991). Jakobson (1968) suggested that the openness of vowels is the earliest contrasted feature for children. Children in the present study might use a vowel which occurs most frequently for substituting other vowels. The vowels might be substituted by another easy distinguished vowel, i.e., openness: narrow[i] versus wide [a]. Thus, children choose [a] to substitute other vowels, due to the fact that [a] occurs most frequently and is the most mouth-opened vowel.

Syllable structure

In terms of the syllable structure, monophthong V was best controlled by children, and VV was the worst; which echoes the finding of Chen (2004). Chen suggested that children acquire monophthongs first, followed by diphthong, the triphthong is the last. As the examiner predicted, monophthongs are easier for children to produce, but the unexpected pattern, VVN, was mastered better than VV. It postulated that the nasal coda might not cause difficulties for children to produce, the most difficult for children is two vowels combined together. VVV (1%) shows a low error percentage, but there is only one VVV target word (jiao), so it might not be a reliable result.

Tones

There are total of 209 tone errors found. Tone 3 was misarticulated the most, followed by Tone 1 and 4, Tone 2 was the last. The delay T3 master is the same as the results of Li and Thompson (1977) and Su (1985). Generally T2 is acquired later than T1 and T4, however results in the present study show different patterns. According to Shen's (1989) results, American learners who learn Mandarin tones acquired T2 earliest, followed by T3, T1 then T4. It is thus assumed that the tone acquisition order in L1 is different from that of L2, and the L1 language background would influence L2 tone acquisition order. Furthermore, there is no actual T1 and T4 in Shanghainese so it might cause the difficulties in master Mandarin T1 and T4 for Shanghainese children.

For T1 errors, the most frequent error was T1 as T2, T1 as T4, and T1 as T3. Children produce T1 as T2 most, which might be due to the fact that children master T2 better than T1 as mentioned in the above. For T 2 errors, the most frequent error was T2 as T4, followed by T2 as T1, and T2 as T3 is the last. For the T3 errors, T3 as T2 is the most frequent error, followed by T3 as T4, and T3 as T1 is the last. These patterns are correlated to the tone acquisition order in that children use the best mastered tone to substitute a tone they cannot control well.

For T4 errors, subjects misarticulated T4 as T1 most frequently, T4 as T2 next, and T4 as T3 last. This pattern (children did not misarticulate T4 as T2 more than T1) is unexpected. It postulated that Mandarin T1-T4 contrast is more difficult than T2-T4 contrast for Shanghainese children. Children master T2 better than T1 and T3, that's why they could clearly identify the differences between T4 and T2. On the other hand, children could not yet master T1 and T4 well, so they might get confused between these two tones. When misarticulation occurs, children seem to choose the tone which is the easiest one for them to master.

The tone acquisition of Shanghainese subjects indicated that they master T2 best, followed by T1 and 4, and T3 is the most difficult for them. The results of Li and Thompson (1977) indicated T1 and T4 were acquired earlier than T2 and T3, and they concluded that rising tones (T2 and T3) are acquired later than the falling tone (T4). The result in the present study shows different patterns in the acquisition order, however T3 is the most difficult tone for children to master in both studies.

4.2 Acoustic analysis

Retroflex

One of the major contributions for the present study is an in depth acoustic investigation of retroflexes. Previous studies measured F3 in the consonant parts of retroflexes; however, the retroflexes in those studies are voiced retroflex consonants. In the present study, the retroflex consonants in Mandarin include voiced and voiceless retroflex consonants. The formants of voiceless retroflex consonants cannot be measured, which is the reason we measured the F3 of the following vowels. Although there is no powerful support research for this measurement, we aim to provide a reference value for future studies. In the present study, we separated retroflexes into voiced and voiceless. There are three measurements in the retroflex section: a) F3 of the following vowel of all retroflex consonants, b) F3 of voiced retroflex consonant itself, and c) M1-M4 values of voiceless retroflex consonants.

The result of F3 in voiced retroflex consonants indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for three different rhymes (VVN, VV, and V). In order to make sure that the following vowels will affect the F3 dipping patterns, the experimenter further compared F3 values in different following vowels (section 3.2.1). Results indicated that F3 dipping was found in correct pronunciation, but not in incorrect pronunciation for retroflexes preceding V (/i/, /e/, /u/) and VN (/en/) rhymes. However, F3 dipping was found in correct as well as in incorrect pronunciation of retroflex preceding VV (/ua/, /ou/) and VVN (/uaŋ/) rhymes. The result of retroflex in transcription and acoustic analysis show an asymmetry pattern. Even experimenter hear a deretroflex sound in transcription, the result of acoustic analysis show that the F3 dipped in the deretroflex sound. It is postulated that children are in the progress in acquiring retroflex, and produce an retroflex acoustic cue (F3 dipping) in both retroflex and deretroflex sounds. It is also postulated that the unexpected F3 dipping in both correct and incorrect pronunciation is due to the complexity of rhyme or the difficulties in producing an affricate. According to the results above, the monophthong (retroflex + V) rhyme for children is the best mastered at this age, followed by retroflex +VN rhyme which are not stable. In other hand, the diphthong (retroflex + VV and VVN) rhymes are the most difficult to master.

These results show a universal development in the syllable structure acquisition. Furthermore, there is no retroflex consonant in Shanghainese; therefore it might be the reason that Shanghainese children misarticulated the most error for retroflex.

For the result of M1-M4, M1 and M2 value of correct pronunciation (retroflex) is significantly lower than that of incorrect pronunciation (nonretroflex). M3 and M4 of correct pronunciation are higher than that of incorrect pronunciation. The results of the present study showed different patterns to the result of Jeng (2009), in which M1 and M4 of retroflexes are lower than that of deretroflexes, and M2 and M3 of retroflexes are higher than that of deretroflexes. In other words, Jeng (2006) suggested that retroflexes in Mandarin have lower central gravity frequency, low but weak energy and wide energy distribution (section 1.2.1). It is postulated that the different results between the two studies is due to the fact that the subjects in the present study were children from Shanghai, which means and the production of retroflex might be influenced by their L1. However the subjects from Jeng (2009) were 19-25 years old adults. The production of retroflex by children and adults might show different patterns. According to the literature review, different researches show different patterns. Forrest et al., (1988) investigated ten English adult speakers; results indicated that M3 is the most efficient parameter to distinguish /s/ and /ʃ/ which is different to Jeng's study. Flipen et al., (1999) investigated /s/ produced by 26 teenagers and found out that M1 and M3 are the most salient characteristics to represent /s/. It is postulated that the different results shown in the present study might be affected by the language background and age of the subjects.

Vowel

Comparisons among Shanghainese speaking children, Mandarin speaking and Shanghainese speaking female adults yield three sets of results of vowel production: (a) the vowel space of Shanghainese children was narrower than that of Mandarin adults, (b) the vowel space of Shanghainese adults (producing Shanghainese) was found narrower than that of Mandarin speaking adults. (c) The vowel space of Shanghainese children was found narrower than that of Shanghainese adults and Mandarin-speaking adults. Generally, children's vowel space shows three different patterns (section 1.2.1 *vowels*). Results from the present research show that the vowel space of children is narrower than that of adults. It is not sure whether this pattern correlates to the universal development. However, since Shanghainese-speaking adults have a narrower vowel space than Mandarin-speaking adults **(c.f. fig 62)**, it is thus postulated that the dense vowel space of children might be influenced by their L1, Shanghainese.

Tones

The F0 contours and duration pattern of tones produced by the children resemble adult norms. The four tones are significantly different: T3 > T2 > T1 > T4 **(Fig. 67)**. This result suggests that they have managed to utilize F0 and duration cue as adult speakers. This result agrees with the result of Li and Thompson (1977), in which they argued that tone acquisition is accomplished within a short period of time, and tones are mastered before segmentals. Moreover, the rising and dipping tones are substituted for each other. However, the acquisition order was different to the result of Li and Thompson (1977). Interestingly, in the tone production of children, the T4 pitch range is not as wide as the adult norm **(cf. Fig. 64-65)**, which is posited to be the effect of their L1 (Shanghainese) **(cf. Fig.66)** which does not have a high falling tone as Mandarin T4.

Conclusion

In conclusion, Transcription and acoustic analysis from normal speaking children provided evidence of the effect of universal development and that of L1 background. For consonants, bilabials acquired earlier than coronals were in line with the universal development, but dorsals mastered better than coronals might be due to the L1 background, in which there are eleven dorsals in Shanghainese but only three dorsals in Mandarin. Stops are acquired the earliest, followed by nasals, fricatives, and affricates are the last, which show a universal development trend. The misarticulation of retroflexes might be affected by the universal development trend--- retroflexes were generally acquired after the age of six and influenced by their L1---as well as the fact that there are no retroflexes in Shanghainese. The high frequcy of /u/, /a/ misarticulation agree with the universal development. The delayed master of T3 agrees with the universal development. However, the early mastery of T2 might be the L1 effect.

For acoustic analysis, the vowel space of Shanghainese children was found narrower than that of Shanghainese and Mandarin speaking adults. Since previous studies show different vowel space acquisition patterns, this finding is assumed to be the effect from their L1. Finally, the children perform similarly to adults in tone production, which corresponds to the general development trend---suprasegment acquired before segment. On the other hand, the narrower T4 pitch range may be caused by their L1.

4.3 Limitations of the Study

In this study, there are still many things which need to be improved. In order to develop this study better, recording material, recording procedure, and acoustic analysis need to be changed. In the study, the stimuli were an 80-word list which targets each segment, major word structure and place sequence. (B.M. Bernhardt). While measuring these data, the first limitation concern is the difficulty to organize the target words. In the present study, acoustic analysis are separated into three parts: consonant, vowel, and tone. In the consonant section, F3 value of the retroflex and non-retroflex sounds were measured. However, the retroflex and non-retroflex sounds did not have the minimal pair; their following vowels may cause effects to the analysis. For example, *shan1* "mountain" is a retroflex sound but there is no non-retroflex sound *san1* "three" in the word list. This asymmetry also happened in vowels and tones data. If comparing vowel /i/, there should be /i/ in four tones. (For example, *i1* "cloth", *i2* "move", *i3* "chair", *i4* "meaning"). As mentioned above, the vowel structures and tones are special in Mandarin Chinese. These complexities might induce children's production errors. So more tone sandhi phrases in the word list is needed as the present study has only two

The second limitation is that it is difficult to complete a word list which satisfied all the experimental needs. There are still some absent words which are not allowed in Mandarin. For direct comparison across retroflexes and non-retroflexes, vowels and four tones causes need to be taken, in order to balance the list even further. A proposed word list construction is as follows (Table.1) On the other hand, children are given a picture-naming task in which they will see many pictures and then speak the name out loud. For example, while children see a picture of a panda, he/she should say "panda", and then move on to the next picture. In my opinion, pictures like cartoons are more preferred because they could attract children's attention and are easier for them to understand and recognize **(Example in Figure. 68).**

The third limitation is that the recoding data were not recorded in a quiet room without any possible interference. While recording children's production, the experimenter's sounds are also recorded. Sometimes the experimenter's sounds covered the child's sound. The experimenter teaches children how to produce a word, but the child did not speak it out. It makes the acoustic analysis become difficult. However, to record children's production indeed has some difficulties. If we ask a child to stay in a quiet room by himself / herself, the child might be nervous and the experiment will not be finished or proved successful. The experimenter should be together with the child in a quiet room without any interference yet allowing the child to feel comfortable and safe. Also the experimenter needs to be careful not to record his/her voice in the recording data.

The fourth limitation is only retroflex/nonretroflex sounds, vowels and tones were measured. However, there are many other errors that occurred. For example, the replacement of /r/, (by /l/, or /n/). In this study the differences between /l/, /r/ and deretroflex sounds were not measured. To make children replace /r/ by /l/ might have some reasons, for example, the following vowel or the syllable structure. Children also have difficulties in producing diphthongs. They might change a diphthong into a monophthong, and different children might choose different vowels to replace the diphthong. It is difficult to put a boundary between two vowels (diphthong) specific; the onset, mid-point and offset of the vowel cannot be consistent. The other error is the final nasal deletion or confusion (alveolar vs. velar nasal). This error is only judged by perception. It is hard to do the acoustic analysis for this error because there is no reliable way to measure a final nasal and know whether it is an alveolar or velar nasal. If possible, in future study, all errors need to be measured by a reliable way.

The purpose of this study is to investigate the universal constraint and effect of L1 on children's production. However, while comparing the universal constraints to our results, many universal constraints are in the early stage like babbling. Our subjects are four years old; there might be some differences within these two stages. In the present study, examiner could only rely on the researches that have been conducted already. It is hoped that in the future study, more reliable and objective data could be found to support the results of the present study.

In the present study, subjects showed little or no effect of L1 transfer to L2.

Although L1 of subjects is Shanghainese, only Mandarin was used in the kindergartens, and more than 90% of the children also communicated with their parents in Mandarin. Most subjects use Mandarin more in their daily life. This may have subjects to speak Mandarin fluently as their L1, even more fluently than L1. In the future study, the subject group should be restricted to subjects who speak Shanghainese (L1) more than their L2. The effect of L1 transfer to L2 might be more significant. Also, in the future study, experimenter could study the younger children production while their language acquisition is still developing, or examined by a longitudinal study.

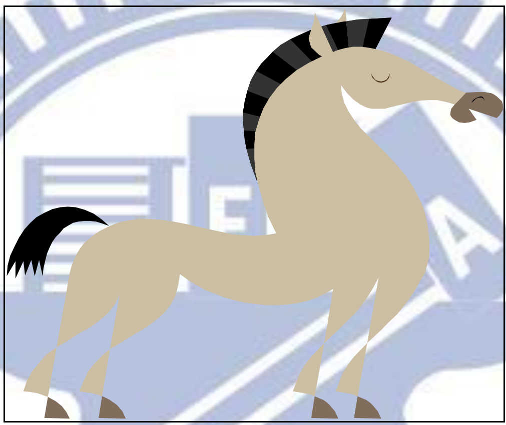


Figure 68 Media picture for target word ma3 "horse"

Table 2. New Target Words

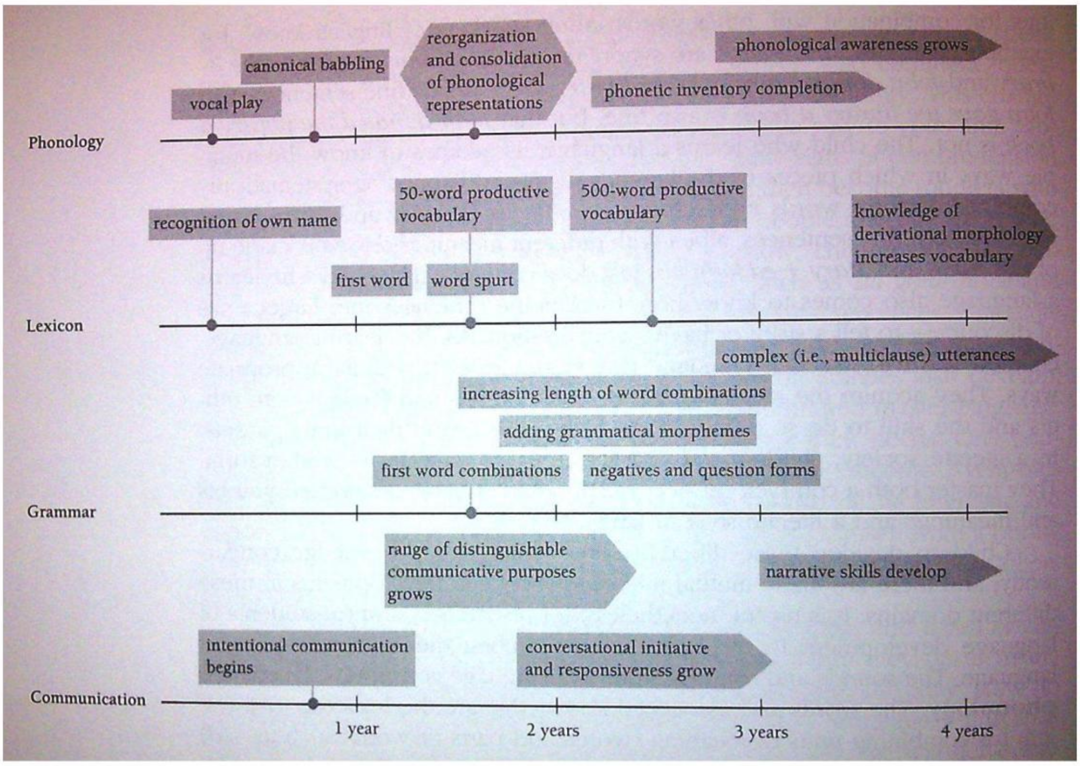
	labial	glossary	coronal	glossary	non-retro	glossary	retroflex	glossary	dorsal	glossary
CV										
tone1	ba1,ma1	cling,mother	xi1	west	ji1	chicken	zhi1,chi1	juice, eat	ke1	to carve
tone2	ba2,ma2	pull out, sesame	xi2	to practice	ji2	reach	zhi2,chi2	straight,late	ke2	to cough
tone3	ba3,ma3	grip,horse	xi3	wash	ji3	squeeze	zhi3,chi3	paper, ruler	ke3	thirsty
tone4	ba4,ma4	father, scold	xi4	fine	ji4	remember	zhi4,chi4	mole, red	ke4	quest
CVN										
tone1	peng1	cook	tan1	beach	can1	meal	chang1	prosperous	wan1	curve
tone2	peng2	friend	tan2	talk	can2	deficient	chang2	long	wan2	play
tone3	peng3	hold up	tan3	blanket	can3	tragic	chang3	factory	wan3	bowl
tone4	peng4	bump	tan4	sigh	can4	brilliant	chang4	sing	wan4	ten thousand
CVV										
tone1	bao1,pao1	bag, throw	xie1	scorpion	zao1	worse	zhao1	beckon	wei1	to threaten
tone2	bao2,pao2	thin, robe	xie2	shoe	zao2	drilling	zhao2	to be fascinated	wei2	to surround
tone3	bao3,pao3	fully, run	xie3	write	zao3	early	zhao3	to look for	wei3	tail
tone4	bao4,pao4	hug, firecrackers	xie4	thank	zao4	create	zhao4	to illuminate	wei4	feed
CVVN										
tone1			xian1	celestial	qian1	thousand	chuang1	window	huan1	cheerful
tone2			xian2	salty	qian2	money	chuang2	bed	huan2	to return
tone3			xian3	danger	qian3	shallow	chuang3	to rush	huan3	to slow
tone4			xian4	thread	qian4	owe	chuang4	to start	huan4	to exchange
CVVV										
tone1	miao1	meow	jiao1	teach	qiao1	knock				
tone2	miao2	seedling	jiao2	to chew	qiao2	bridge				
tone3	miao3	second	jiao3	foot	qiao3	skillful				
tone4	miao4	temple	jiao4	to shout	qiao4	raise				

Appendix

I. Word list from PPD program

SCREEN LIST				ADULT TARGET-STANDARD MANDARIN & TAIWAN MANDARIN									
#	Pinyin BM	Pinyin TM	English	Phonetics BM	Phonetics TM	Wd Length	Syllable structur	Tone BM			Tone TM		
2	er3duo0	er3duo1	ear	ɚ^{MLM}tuo^{0}	ɚ^{MLM}tuo^{1}	2	V.CVV	3	0		3	1	
3	zui3(ba)0	zui3(ba)1	mouth	tsuei^{MLM}(pa)^{o}	tsuei^{MLM}(pa)^{o}	1(2)	CVVV.(CV)	3	(0)		3	(1)	
4	shou3	shou3	hand	ʂou^{MLM}	ʂou^{MLM}	1	CVV	3			3		
5	tou2fa0	tou2fa3	hair	t^{h}ou^{MH}fa^{0}	t^{h}ou^{MH}fa^{3}	2	CVV.CV	2	0		2	3	
6	jiao3	jiao3	foot	tɕiao^{MLM}	tɕiao^{MLM}	1	CVVV	3			3		
7	xie2(zi0)	xie2(zi0)	shoes	ɕie^{MH}(tsi(ə)^{0})	ɕie^{MH}(tsi(ə)^{0})	1(2)	CVV.(CV(V))	2	(0)		2	(0)	
9	ping2guo3	ping2guo3	apple	p^{h}iŋ^{MH}kuo^{MLM}	p^{h}iŋ^{MH}kuo^{MLM}	2	CVC.CVC	2	3		2	3	
10	xi1gua0	xi1gua1	watermelon	ɕi^{H}kua^{0}	ɕi^{H}kua^{1}	2	CV.CVV	1	0		1	1	
11	xiang1jiao1	xian1jiao1	banana	ɕiaŋ^{H}tɕiao^{H}	ɕiaŋ^{H}tɕiao^{H}	2	CVVC.CVVV	1	1		1	1	
12	rou4	rou4	meat	ɹou^{HL}	ɹou^{HL}	1	CVV	4			4		
13	cai4	cai4	vegetable	ts^{h}ai^{HL}	ts^{h}ai^{HL}	1	CVV	4			4		
14	wan3	wan3	bowl	wan^{MLM}	wan^{MLM}	1	CVC	3			3		
15	kuai4zi0	kuai4zi0	chopstick	k^{h}uai^{HL}(tsi(ə)^{0})	k^{h}uai^{HL}(tsi(ə)^{0})	2	CVVV.CV(V)	4	0		4	0	
17	zhuo1zi0	zhuo1zi0	table	tʂuo^{H}(tsi(ə)^{0})	tʂuo^{H}(tsi(ə)^{0})	2	CVV.CV(V)	1	0		1	0	
19	xi2lian3	zi2lian3	wash face	ɕi^{MH}lian^{MLM}	ɕi^{MH}lian^{MLM}	2	CV.CVVC	2	3		2	3	
20	shua1ya2	shua1ya2	brush teeth	ʂua^{H}ja^{MH}	ʂua^{H}ja^{MH}	2	CVV.CV	1	2		1	2	
21	chuang2	chuang2	bed	tʂ^{h}uaŋ^{MH}	tʂ^{h}uaŋ^{MH}	1	CVVC	2			2		
22	men2	men2	gate	mɤn^{MH}	mɤn^{MH}	1	CVC	2			2		
23	deng1	deng1	light	tɤŋ^{H}	tɤŋ^{H}	1	CVC	1			1		
24	ü2san3	ü2san3	umbrella	y^{MH}san^{MLM}	y^{MH}san^{MLM}	2	V.CVC	2	3		2	3	
25	tai4yang0	tai4yang2	sun	t^{h}ai^{HL}jaŋ^{0}	t^{h}ai^{HL}jaŋ^{2}	2	CVV.CVC	4	0		4	2	
26	yue4liang0	yue4liang4	moon	jye^{HL}liaŋ^{0}	jye^{HL}liaŋ^{0}	2	CVV.CVVC	4	0		4	4	

II. Language acquisition aspects (Hoff, 2005)



III. Vowel space measurement

Table 3. Vowel values of Mandarin 5-year-old children. (Data from Lee et al, 1999)

	a	I	u
F1	1166Hz	467Hz	656Hz
F2	1750Hz	3071Hz	1434Hz

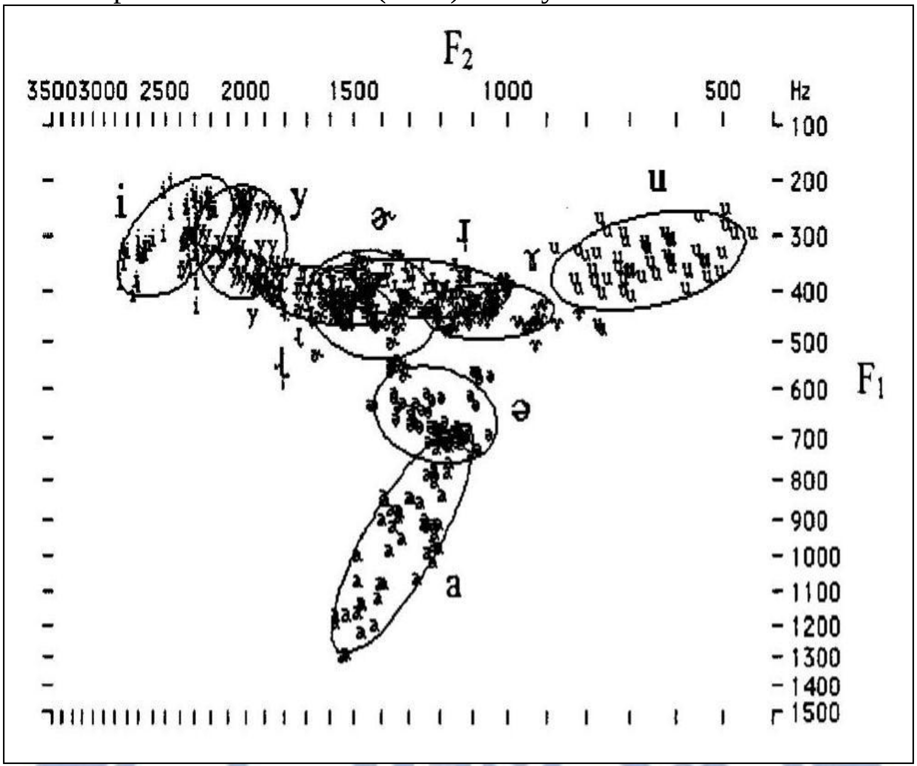
Table 4. Mean values of vowel by BM males. (Figure from Zee & Lee, 2001)

Vowel	F_1		F_2		F_3	
/s.a.	mean	s.d.	mean	s.d.	mean	s.d.
[i]	300.42	56.33	2443.08	206.60	3384.78	254.92
[y]	310.10	52.83	2030.64	127.64	2418.58	185.00
[ɤ]	441.40	27.15	1059.34	92.41	2861.22	194.70
[u]	345.16	46.86	661.32	108.11	2783.00	225.76
[a]	956.98	157.23	1328.26	125.47	2813.06	233.33
[ə]	654.10	53.19	1220.38	101.17	2258.26	210.00
[ɚ]	428.34	53.37	1434.62	125.93	1810.82	100.22
[ɹ]	396.50	27.98	1295.94	172.51	2922.58	183.27
[ɻ]	411.02	29.19	1632.48	148.13	1962.10	205.38

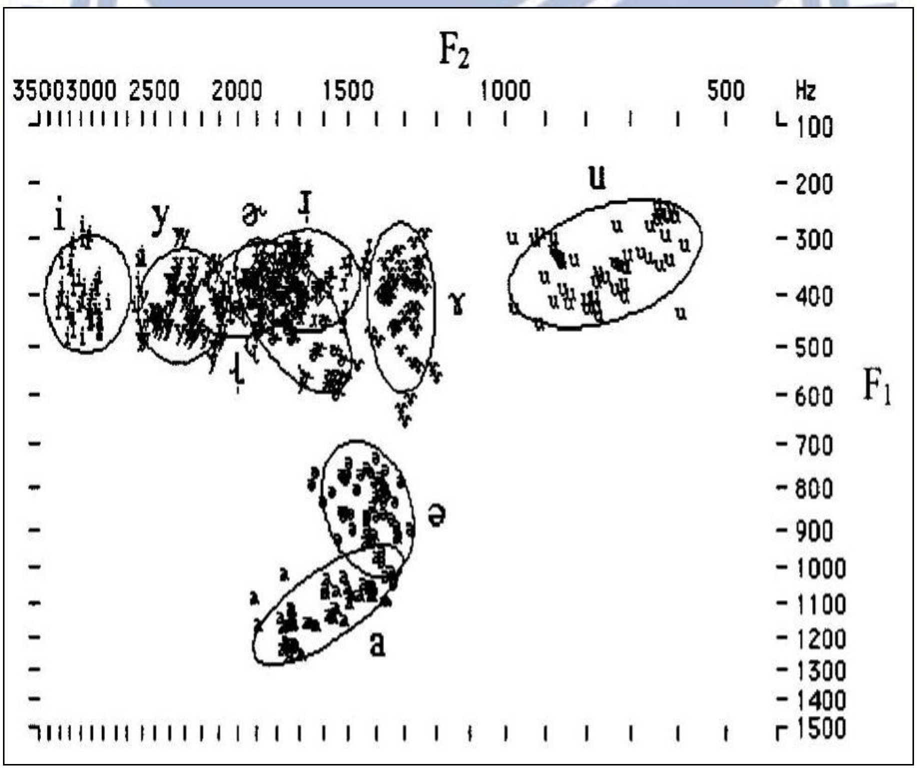
Table 5. Mean values of vowels by BM females. (Figure from Zee and Lee, 2001)

Vowel	F_1		F_2		F_3	
/s.a.	mean	s.d.	mean	s.d.	mean	s.d.
[i]	401.24	55.41	3036.76	185.03	3847.56	262.88
[y]	423.84	54.54	2327.36	141.18	2999.88	180.40
[ɤ]	426.52	83.18	1314.42	57.93	3460.70	268.04
[u]	345.02	59.72	758.68	111.73	3308.82	275.95
[a]	1104.04	85.90	1593.64	153.91	3188.38	239.25
[ə]	852.24	85.44	1432.84	85.31	2925.38	301.44
[ɚ]	443.18	74.94	1736.34	130.87	2196.22	163.28
[ɹ]	376.28	48.03	1680.32	115.28	3500.84	235.86
[ɻ]	392.80	45.16	1929.02	144.52	2465.92	350.25

IV. vowel ellipses in Zee and Lee (2001)'s study



vowel ellipses of male speakers. Figure from Zee and Lee (2001)



vowel ellipses of female speakers. Figure from Zee and Lee (2001)

References

Bernhardt B. H.; Stemberger J.P. (1998). *Handbook of phonological development.* Academic Press. ISBN: 9780120928309

Bertoncini, J. & Mehler, J. (1998). Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance.* Vol. 24. No.3, 756-766.

Buhr, R. (1980). The emergence of vowels in an infant. *Journal of Speech and Hearing Research* **23**(1): 73.

Carmichael, L. (2002). Second language acquisition of suprasegmental phonology. Unpublished Thesis. University of Washington Department of Linguistics.

Chang, H. (1991). Acquisition of Mandarin Chinese: A review of recent research in Taiwan. *Proceedings of National Science Council, ROC. Part C: Humanities and Social Sciences* **1**: 110-126.

Chen, L. (2004). Prelinguistic vocalic production in the first year of life: the acoustic - phonetic analysis. *Acoustical Society of R.O.C.* Vol. 17

Chen, Y. (2008). The acoustic realization of vowels of Shanghai Chinese. *Journal of Phonetics* **36**(4): 629-648.

Chen, Y. (2008). Revisiting the Phonetics and Phonology of Shanghai Tone Sandhi. *Speech of Prosody*, 253-256.

Cheung, Hintat. (2000). Three to Four-years Old Children's Perception and Production of Mandarin Consonants. *Language and Linguistics* 1.2: 19-38

Chomsky, N. & Halle, M. (1968). *The sound pattern of English.* New York: Harper and Row.

Clark, E. V. (2009). *First language acquisition*. Cambridge University Press. ISBN: 978-0-521-51413-2.

Clark, J., C. Yallop, et al. (2007). *An introduction to phonetics and phonology*, Wiley-Blackwell.

Clark, E., V. (2009). *First language acquisition.* Cambridge University Press. ISBN 978-0-521-51413-2.

Cruttenden, A. (1997). *Intonation.* Cambridge University Press.ISBN: 0-521-59825-7.

Delattre, P., and Freeman, D.C. (1968). A dialect study of American r's by X-ray motion picture. *Linguistics.* 44, 29-68.

Fee, E. J. (1991). Underspecification, parameters, and the acquisition of vowels. *Unpublished doctoral dissertation: University of British Columbia.*

Fikkert, P. (1994). *On the acquisition of prosodic structure.* Holland Academic Graphics, The hague Printed by ICG Printing, Dordrecht. ISBN: 90-74537-05-7

Forrest, K., Weismer, G., Milenkovic, P. & Dougall, R..(1988). Statistical analysis of word-initial voiceless obstruents : Preliminary data. *Journal of Acoustical Society of America*, 84, 115-123.

Goodluck, H. (1991). *Language Acquisition: A linguistic introduction*. BLACKWELL. Oxford UK & Cambridge MA.

Grunwell, P. (1981). *The nature of phonological disability in children.* New York: Academic Press.

Grosjean (1982). *Life with two languages: an introduction to bilingualism.* Library of Congress Cataloging in Publication Data. ISBN: 0-674-53092-6

Halle, M. & Stevens, K. (1997). The postalveolar fricatives of Polish. In: S. Kiritani, H. Hirose & H. Fujisaki (Eds.). *Speech Production and Language: In Honor of Osamu Fujimura.* Berlin. Mouton de Gruyter: 177-193.

Hamann, S. (2003/2005). The phonetic and phonology of retroflexs.

Hannahs, S. J., Yong-Scholten, Martha. (1997). *Focus on phonological acquisition.* Language acquisition and language disorders, Vol. 16. (eds)
Hoff, E. (2005). *Language development*, Wadsworth Pub Co. ISBN: 0-534-64170-9
Howie, J. (1976). *Acoustical studies of Mandarin vowels and tones*, Cambridge University Press.
Hu, F. (2008). The Three Sibilants in Standard Chinese. *International Seminar on Speech Production*. Vol. 8
Ingram, D. (1989). *First language acquisition: Method, description and explanation.* Cambridge University Press. ISBN: 0-521-34109-4.
Ito, K. (1990). *Kodoma no kotoba*. Tokyo: Keiso Shobo.
Jakobson, R. (1968). *Child language, aphasia and phonological universals.* The Hague: Mouton.
Jongman A. et al(1997). Perception and Production of Mandarin Chinese Tone. *Handbook of Chinese Psycholinguisitcs.* E. Bates, L.H. Tan, and O.J.L Tzeng (eds.). Cambride University Press.
Jeng, J. (2004). The acoustic spectral characteristics of retroflexed fricatives and affricates in Taiwwan Mandarin. *國立台南大學「人文研究學報」第 40 卷第 1 期*: 27-48.
Jeng. J. (2009). The Inconsistency of Mandarin Consonant Production in the Different Rime Contexts for the Preschool Children. *台灣聽力語言學會雜誌, 第 24 期.*
Kang, J. (1998). Comparison of speech intelligibility between English and Chinese. *The Journal of the Acoustical Society of America* **103**: 1213-1216.
Lee, S., Potamianos, A. & Narayanan, S. (1999). Acoustics of children's speech: Developmental changes of temporal and spectral parameters. *Journal of Acoustical Society of America.* 105 (3): 1455-1468.
Lee, Y., W. (2007) *An acoustic study of vowel development in a child acquiring Taiwan Southern Min.* M.A. Thesis.
Leonard, L. (1995). Phonological impairment. In P. Fletcher and B. MacWhinney (eds) *The Hangdbook of child language.* Page: 573-602. New York : Blackwell.
Levelt, C. (1994). *On the acquisition of Place.* Doctoral dissertation. HIL Dissertations in Linguistics. Dordrecht: ICP Printing.
Li, C. & S. Thompson (1977). "The acquisition of tone in Mandarin-speaking children." *Journal of Child Language* **4**(02): 185-199..
Lieberman, P. (1984). *The biology and evolution of language*, Harvard University Press. ISBN: 0-674-07412-2.
Locke, J. (1983). *Phonological acquisition and change.* New York: Academic Press.
McGowan, R., S. Nittrouer, et al. (2004). Development of [ɹ] in young, Midwestern, American children. *The Journal of the Acoustical Society of America* **115**: 871.
McGuire, G. (2007). English listeners' perception of Polish alveopalatal and retroflex voiceless sibilants: A pilot study. *UC Berkeley Phonology Lab Annual Report:* 391¡V416.
McLaughin, B. & Erlbaum L. (1984) *Second-language acquisition in childhood: Preschool children.* Vol. 1
Maneva, B. & Genesee, F. (2002) *Bilingual babbling: Evidence for language differentiation in dual language acquisition.* In B. Skarbela et al. (Eds.), Boston University Conference on Language Development 26 Proceedings (pp. 383-392). Somerville, MA: Cascadilla Press.
Milenkovic, P. (2004). TF32 Computer Program Madison, WI: University of Wisconsin-Madison Department of Electrical Engineering.
Miracle, W. C. (1989). Tone production of American students of Chinese: A

preliminary acoustic study. *Journal of the Chinese Language Teachers Association*, Vol. 24. 49-65.
O'Grady, D., Dobrovolsky, M., Katamba, F. (1997). *Contemporary Linguistics: An introduction*. Pearson Education. ISBN: 0582 24691 1 Ppr.
Paschall, L., (1983). Development at 18 months. In J. V. Irwin and S. P. Wong (Eds) *Phonological Development in Children 18 to 72 Months* (Carbondale, IL: Southern Illinois University Press), pp. 27-54.
Peterson & Barney. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, Vol. 24, No. 2, 175-184.
Phil R. (1993). A linguistic-phonetic acoustic analysis of Shanghai tones. *Australian Journal of Linguistic*. Vol. 13, 185-220.
Pollock, K. E. and Keiser, N. J., (1990). An examination of vowel errors in phonologically disordered children. *Clinical Linguistic and Phonetics,* Vol.4, 161-178.
Preston, M. S., Yeni-Komshian, G., & Stark, R. (1967). Voicing in initial stop consonants produced by children in the prelinguistic period from different language communities. *Annual Repor*t, 2, 305-323, Neurocommunications Laboratory, Baltimore, MD: Johns Hopkins University School of Medicine.
Reetz & Jongman. (2009). *Phonetics. Transcription, Production, Acoustics, and Perception.* Wiley-Blackwell.
Recasens, D. (1983). Place cues for nasal consonants with special reference to Catalan. *Journal of the acoustical society of America* **73**(4): 1346-1353.
Selby, J.C., Robb, M.P. and Gilbert, H. R., (2000). Normal vowel articulations between 15 and 36 months of age. *Clinical Linguistics and Phonetics*. Vol. 14, 255-265.
Slobin, D. (ed.) (1985) *The crosslinguistic study of language acquisition.* NJ: Erlbaum.
Stemberger, J. P (2008). Class Lecture.
Su, A.T. (1985). *The acquisition of Mandarin phonology by Taiwanese children.* Unpublished MA thesis, Fu Jen Catholic University, Taipei.
Stevens, K. (1998). *Acoustic Phonetics.* Cambridge MA: MIT Press.
Strokes, S.T. and Wong, I. M. (2002). Vowel and diphthong development in Cantonese-speaking children. *Clinical Linguistics & Phonetics*, 2002, Vol. 16, No.8, 597-671.
van der Stelt, J. M., Lyakso E., Pols, L.C.W. and Wempe, T.G. (2006). Developing a Russian vowel space in infancy: the first two years. *SPECOM' 2006*, St. Petersburg.
Vihman, M. (1996). *Phonological development :The origins of language in child.* Oxford: Blackwell.
Vorperian, H. and R. Kent (2007). Vowel acoustic space development in children: a synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research* **50**(6): 1510.
Wan, I. (1999). Mandarin Phonology: Evidence from speech errors, Ph.D. dissertation, University of New York at Buffalo.
Wan, I. (2002). Asymmetry in Mandarin consonant articulations: Evidence from slips of the tongue. *Concentric: Studies in English Literature and Linguistics* **28**(2).
Wan, I. (2007). On correlating aphasic errors with speech errors in Mandarin. *Senior Editors: Paul Robertson and John Adamson* **2**(2): 53.
Wang, Y., A. Jongman, et al. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the*

Acoustical Society of America **113**: 1033.
White, L. (1989) *Universal Grammar and second language acquisition.* ISBN 1-55619-092-1 (hb).
Wode, H. (1977). Four early stages in the development of L1 negation. *Journal of Child Language.* Vol 4. pp. 87-102.
Yip M. (1989). Contour tones. *Phonology* Vol. 6: 149-174
Zee, E. & Lee, W. S. (2001). An acoustical analysis of the vowels in Beijing Mandarin. *EUROSPEECH-2001*, 643-646.
Zhu H. (2002). *Phonological development in specific context: studies of Chinese-speaking children.* ISBN: 1-85359-588-8.
Zhu S. Xiaonong. (1999). *Shanghai Tonetics*. Munich: Lincom. ISBN 3-89586-584-2.
王南梅, 費珮妮, 黃珣, 陳靜文 (1984). <三歲至六歲學齡前兒童國語語音發展結構>, 《聽語會刊》1: 12-17