國立政治大學語言學研究所碩士論文

National Chengchi University
Graduate Institute of Linguistics
Master Thesis

指導教授：萬依萍 博士
Advisor: Dr. I-Ping Wan

台灣華語聲調習得
Tone Acquisition in Taiwan Mandarin

研究生：楊涵絜 撰
Student: Han-chieh Yang
中華民國一○二年六月
June, 2013

Tone Acquisition in Taiwan Mandarin

By

Han-chieh Yang

**A Thesis Submitted to the
Graduate Institute of Linguistics
in Partial Fulfillment of the
Requirements for the Degree of
Master of Arts**

June 2013

致謝詞

我曾經試著想要算算看這本論文大概花了多少小時才完成，但是發現這個數字龐大到無法計算，光是每個禮拜舟車勞頓收集小朋友語料長達一年多，就大概超過 100 小時，加上之後譯寫錄音檔案和整理也大概花了 100 多個小時，研讀文獻少說 80 小時，分析和寫作的時間大概也有 200 小時，整本論文要完整呈現必須要花將近 500 小時呢！真的是一個非常大的工程！殊不知我第一次跟老師 meeting 時，老師才花 10 分鐘就把我腦中兩個很籠統的方向整理出一個完整的架構！讓我至今還是敬佩不已。

三年前剛進來語言所時，我就是跟著萬依萍老師做助理。從前我是一個行動力頗弱的人，做事慢、也容易猶豫不決，但是因為老師是 10-minutes person，解決問題是以 10 分鐘為單位，所以長久下來，我也學習到老師處理事情的超高效率。一天內 to-do list 有 10 件事情也能分批解決！這本論文可以順利催生也是在萬老師高效率的指導下才能完成。之前時常因為趕論文而焦慮，老師的鼓勵和安慰都讓我覺得倍感窩心。真的非常非常感謝老師一路上的提攜！

也非常感謝研究所每位老師的栽培和指導。謝謝開朗愛搞笑的蕭宇超老師、風度翩翩愛立領的何萬順老師、如仙女下凡的黃瓊之老師、上課嚴肅下課慈祥的徐嘉慧老師、總是以學生為重的賴惠玲老師、從大一看著我長大的張郇慧老師、笑容會煞到一推女學生的林祐瑜老師。特別感謝黃瓊之老師和林祐瑜老師來擔任我的口試委員，能得到你們的肯定真的非常非常開心，我之後一定會繼續加油的！

在這三年間，剛好經歷萬老師的可愛女兒– Vickie 的誕生，因此我才能有機會跟著老師一起開始兒語習得的研究。為了做長期觀察研究，我們常常必須揹著攝影機到處去受試兒童家裡錄音。錄音的過程雖然辛苦，但是能陪著小朋友一起成長，每次去拜訪都聽到小朋友有新的詞彙出現，真的很有成就感！其中最最特別的就是我第一個錄的小寶寶–甯甯，害羞怕生的甯甯竟然很喜歡我，每次去錄音都是瘋狂大笑、跑跳、還有講一堆小大人口吻的話。最好笑的是她都直接叫我「涵絜」。有一次她還問媽媽：「涵絜什麼時候回來？」是用「回來」喔！親愛的甯甯，你不知道涵絜姊姊有多感動～也謝謝甯甯媽媽在我最後一次去錄音時還頒發了獎狀和禮物給我！我會一輩子好好珍藏這些回憶的。

研究生涯的大部分時間都是待在萬老師的語音暨心理綜合實驗室裡工作，那裡就像我在政大的另一個家。學長姐、學弟妹都像家人一樣，幾乎天天見面、經常分享生活大小事～我在趕論文時，幸虧有溫柔體貼的心怡、超高效率的冠霆、語言學大師明哲、開心果彥棻、配合度超高的佳琳、認真負責的家正、可愛乖巧的翊倫，若不是你們絕佳的工作能力，我也無法專心寫作，真誠的感謝你們！也特別

要感謝在統計上給予我超即時幫助的曹維軒，若沒有你半夜幫我趕跑 SPSS，我的論文後半部一定會完蛋。祝福你們未來都能畢業順利、工作順利！

回想起剛上研究所時，雖然環境很熟悉，但是人事已非，很多以前大學要好的朋友都畢業離開學校了。幸虧我們班的同學都非常友善，也很好相處。特別是我們的四人團體–DHA 花。謝謝空中英語教室老師 David，我們從碩一就常常同一組但是彼此都不記得實在很好笑；謝謝美食玩樂總監兼姊妹淘艽艽 Angela，除了每次都要靠你安排行程，也只有你能夠分享所有心情（我也會永遠懷念我們的新加坡天堂之旅）；也要鄭重感謝對這篇論文貢獻良多的音韻小神童亨亨 Henry，每次論文只要卡卡，去找你講完之後都會像吃了Ｘ藥一樣通暢！哈哈哈！還有每次一聊起來總是沒完沒了，話題辛辣度都不輸康熙，好愛那段天天都演內心小劇場、天天鼓吹你吃宵夜的日子啊～和 DHA 花住在大公寓的日子大概是我這輩子最棒的外宿經驗了～

碩三這一年，我除了做助理工作、寫論文，還要準備聽語所的入學考試。三種壓力同時壓在肩上讓我常常焦慮不安，甚至想放棄考試。每當我撐不下去時，感謝你總是適時的發現我的不對勁，在我身邊聽我訴苦；謝謝你因為捨不得我太操勞，給了我很多退路讓我不至於把自己逼瘋；謝謝你總是認真幫我分析問題，又還是能支持我做的每一個決定；謝謝你懂得趕論文的難處，總是能紓解我容易緊張的心情。沒有你的陪伴和支持，就不會有這麼豐碩的成果。獻給何韋辰：Thanks for coming into my life, and thanks for loving me.

最後感謝爺爺、奶奶、爸爸、媽媽多年來的栽培，給我強大的後盾，讓我沒有後顧之憂可以全力在學業上打拚。尤其是奶奶不間斷的關心和期望，給了我很多信心。也很謝謝大姑姑的啟蒙，帶我進入語言學的領域，一路上給我建議，讓我能夠順利完成這個里程碑。感謝楊家大家族源源不絕的愛，培育我樂觀正向的力量。另外還要感謝所有透過 Facebook 陪伴我、幫我加油打氣的家人、朋友，感謝你們幫我在通過論文口試的 PO 文上按讚、留言，有你們真好！

我畢業了！耶～

2013/6/19

Table of content

Chapter 1 Introduction 1

1.1 Children's phonological acquisition 1

1.2 Mandarin tone 2

1.3 Research gap 3

1.4 Research Questions 5

1.5 The framework of the thesis 6

Chapter 2 Literature review 9

2.1 Language universals in first language acquisition 9

2.1.1 Syllable 10

2.1.2 Suprasegmentals 12

2.1.2.1 Intonation 12

2.1.2 2 Stress 12

2.1.3 Reduplication of motherese 13

2.2 Introduction to Mandarin tonal system 15

2.2.1 The tonal representation systems of Standard Mandarin 15

2.2.1.1 Chao (1968) 15

2.2.1.2 Yip (2002) 16

2.2.1.3 Lin (2007) 17

2.2.2 The tonal representation systems of Taiwan Mandarin 19

2.2.2.1 Shih (1988) 19

2.2.2.2 Fon (1997) 20

2.2.3 Summary 21

2.3 Theory of markedness of tone 22

2.4 Tonal acquisition studies in East Asia 23

2.4.1 Thai (Tuaycharoen 1977) 24

2.4.2 Cantonese (Tse 1978; So & Dodd 1995) 25

2.4.3 Taiwanese (Tsay 2000) 27

2.4.4 Overview of the cross-linguistic studies 29

2.5 Tone acquisition on Mandarin 31

2.5.1 Chao (1951) 31

2.5.2 Li & Thompson (1977)....31
2.5.3 Zhu (2002)....33
2.5.4 Summary 33
Chapter 3 Methodology 35
3.1 Data collection 35
3.1.1 Recruitment 36
3.1.2 Subject....36
3.1.3 Observational procedures 37
3.1.4 Recording equipments....38
3.2 Data analysis....39
3.2.1 Transcription and coding....39
3.2.2 Tone emergence ordering 43
3.2.3 Frequency and accuracy rate of tones 44
3.2.4 Substitution pattern in tonal errors 46
Chapter 4 Results and analysis 49
4.1 Overall analysis 50
4.1.1 Tone emergence ordering 52
4.1.2 Frequency and Accuracy rate 55
4.2 Subgroup analyses in monosyllabic and disyllabic tokens 59
4.2.1 Monosyllabic tokens 60
4.2.2 Disyllabic tokens 63
4.2.3 The tone combination [21-35] in the reduplication of motherese 66
4.2.4 Reanalysis in disyllabic tokens 70
4.2.5 Reanalysis in overall data 73
4.3 Substitution pattern in tonal errors 77
4.4 The age of acquisition of tones 79
Chapter 5 Discussion....81
5.1 Summary of the findings 81
5.2 Comparison with tonal acquisition studies in Mandarin 84
5.2.1 Age of acquisition 84
5.2.2 Order of tonal acquisition 85
5.2.3 The tonal combination of reduplication in Taiwan Mandarin 86

5.3 Cross-linguistic comparison 87
5.4 Concluding remarks 90
References 91

List of Tables

Table 2.1 The examples of children's simplified form in English and French 11

Table 2.2 The two intonation patterns produced by an English-speaking children 13

Table 2.3 Chao's (1968) tonal representation system of SC 16

Table 2.4 Lin's (2007) tonal feature and pitch value 16

Table 2.5 Lin's (2007) tonal representation system of SC 17

Table 2.6 Pitch values of the neutral tone 17

Table 2.7 Yip's (2001) tonal representation system of SC 18

Table 2.8 Fon's (1997) tonal representation system of TM 19

Table 2.9 Shih's (1997) tonal representation system of TM 20

Table 2.10 The tonal representation adopted in this thesis 22

Table 2.11 Thai tonal system 24

Table 2.12 Cantonese tonal system 26

Table 2.13 Taiwanese tonal system 28

Table 2.14 The average error rate of each tone in juncture position 28

Table 2.15 The acquisition ordering in Thai, Cantonese, and Taiwanese 29

Table 3.1 The data collecting information on subjects and recordings 37

Table 3.2 The sample of coding 42

Table 3.3 A sample matrix of substitution pattern in tonal errors 47

Table 4.1 Subject information 50

Table 4.2 Number of tokens in subcategories 51

Table 4.3 Age of emergence of tone 53

Table 4.4 Number of tokens and frequencies of tones in all syllabic tokens 56

Table 4.5 Number of tokens and accuracy rates of tones in all syllabic tokens 58

Table 4.6 Number of tokens of each syllable in monosyllabic and disyllabic tokens 60

Table 4.7 Number of tokens and frequencies of tones in monosyllabic tokens 61

Table 4.8 Number of tokens and accuracy rates of tones in monosyllabic tokens 62

Table 4.9 Number of tokens and frequencies of tones in disyllabic words at separate syllable positions 64

Table 4.10 Number of tokens and accuracy rates of tones in disyllabic words at separate syllable positions 65

Table 4.11 Tone combinations of motherese in common nouns and kinship terms 67

Table 4.12 Modified number of tokens and frequencies of tones in disyllabic words at separate syllable positions 71

Table 4.13 Modified number of tokens and accuracy rates of tones in disyllabic words at separate syllable positions 72

Table 4.14 Modified number of tokens and frequencies of tones in all syllabic tokens 74

Table 4.15 Modified number of tokens and accuracy rates of tones in all syllabic tokens ...75

Table 4.16 The matrix of substitution patterns of tonal errors in disyllabic words 77

Table 4.17 Age of emergence and stabilization of each tone 79

Table 5.1 Tone acquisition orderings in different measures 81

Table 5.2 Tone acquisition orderings of cross-linguistic studies 87

List of Figures

Figure 4.1 Frequencies of tones in all syllabic tokens 57

Figure 4.2 Accuracy rates of tones in all syllabic tokens 58

Figure 4.3 Frequencies of tones in monosyllabic tokens 61

Figure 4.4 Accuracy rates of tones in monosyllabic tokens 62

Figure 4.5 Frequencies of tones in disyllabic words at separate syllable positions 64

Figure 4.6 Accuracy rates of disyllabic words in separate syllable positions 65

Figure 4.7 Number of tokens in all combinations in disyllabic tokens 69

Figure 4.8 Modified tone frequency of disyllabic words in separate syllable positions 71

Figure 4.9 Modified accuracy rates of disyllabic words in separate syllable positions 72

Figure 4.10 Modified frequencies in overall data 74

Figure 4.11 Modified accuracy rates in overall data 75

Figure 4.12 Percentage of the realized tone in tonal errors 78

國立政治大學研究所碩士論文提要

研究所別：語言學研究所

論文名稱：台灣華語聲調習得

指導教授：萬依萍　博士

研究生：楊涵絜

論文提要內容：（共一冊，19,799 字，分五章）

本篇論文是針對六位以台灣華語為母語的嬰幼兒，採長期觀察的方式，研究華語聲調的習得，並詳細描述單音節詞和雙音節詞之中聲調出現順序、頻率、正確率、以及代換模式。本研究同時要用 Yip (2002)的標記理論來檢驗各種不同聲調語言中的共通性。

本研究一共觀察了有六位年齡在十個月至一歲一個月的嬰幼兒長達八個月。以兩個禮拜一次的頻率收集嬰幼兒和母親或照顧者之間的自然對話。並利用錄製回來的高規格影音檔做譯寫和分析。

結果顯示[55]最早出現，也是頻率最高、正確率最高的聲調。而[51]在聲調出現順序、頻率、及正確率都是排在第二。[35]和[21]就比較晚出現，和前面兩個聲調相比，頻率及正確率都較低。輕聲不管是出現順序、頻率、或正確率都排在最後。

本研究結果還發現台灣華語中有一特殊的聲調組合[21-35]。在台灣華語中，媽媽對幼兒說話時所使用的「媽媽語」很常把這個聲調組合套用在重疊詞中。而這個聲調組合也被幼兒高度模仿使用。因此本研究認為幼兒發出的這個高頻的聲調組合[21-35]有可能是受到照顧者的影響，並且也認為幼兒並非分開習得此兩個聲調，而是把此兩聲調當作一個整體來習得。

最後，將所有跨語言的分析結果拿來檢驗 Yip (2002)的標記理論後，我們發現幾乎所有語言都支持平調比曲折調早習得、降調比升調早習得。但是除了泰文之外，沒有語言能支持 Yip (2002)提出的低調比高調早習得。因此，語言習得的證據能證明平調、降調、高調比曲折調、升調、低調還普遍。

關鍵詞：兒童語言發展、聲調習得、出現順序、頻率、正確率、代換模式、台灣華語

Abstract

The purpose of this study is to describe children's tonal development by analyzing the tone emergence, frequency, accuracy rate, and substitution pattern, based on observing monosyllabic and disyllabic utterances in six Mandarin-speaking children in Taipei, Taiwan. This study also aims to examine several cross-linguistic data the theory of markedness presented by Yip (2002).

Six subjects are investigated with the age range from 0;10 to 1;6. The data collection is conducted fortnightly by the author and the research team. Based on video and sound files, a set of coding are employed for data analysis.

The results showed that the high-level tone [55] emerged the first, and it also ranked as the most frequent and stable tone. Falling tones [51] were consistently ranked in the second place within tone emergence, frequency, and accuracy rate. Rising tones [35] and low-level tones [21] appeared late, and were also less frequent and stabilized later than [55] and [51]. The neutral tone was emerged and stabilized the last appeared and last acquired tone.

This study also found the dominated tone combination [21-35] applied particularly in the reduplications of motherese in Taiwan Mandarin. The tone combination [21-35] was proposed to be influenced by motherese, and was acquired as a prosodic whole.

The results of this study and all the cross-linguistic data are examined in Yip's theory of markedness. The first two constraints obtained more evidence that the features of level and falling in tones were suggested to be the unmarked features in tonal languages. Regarding the third constraint, because most of the tone acquisition data indicated that high tones were acquired earlier than low tones, the more unmarked tone feature should be level, falling, and high.

Keywords: phonological development, tone acquisition, tone emergence, frequency and accuracy rate, substitution pattern, Taiwan Mandarin

Chapter 1

Introduction

1.1 Children's phonological acquisition

The purpose of this thesis is to discuss issues involving in tone acquisition from Mandarin-speaking children in Taiwan. By observing the tone acquisition for children in early developmental stages, this research aims to add literature on cross-linguistic studies and to test a number of questions raised in the former phonological acquisition theories.

Children's phonological development proceeds rapidly. In their early developmental stage, infants are sensitive to prosodic cues. Lenneberg (1967) pointed out that at the age of 0;8, children could distinguish rising or falling pitch contours. Many researchers also proved that the suprasegmental features, such as intonation and stress, were acquired earlier and better than segments (consonants and vowels) (Lenneberg, 1967; Kaplan, 1970; Demuth, 1996). The earliest age for the mastery of a tonal language system was at 1;4 in Thai-speaking children (Tuaycharoen, 1977). Other studies also presented that children have mastered their tonal systems with little errors after the age of 2;0 (Li & Thompson, 1977; Tse, 1978; So & Dodd, 1995). It is reported that by the age of five, children could fully acquire all the phonemes in their native languages (Cantwell & Baker, 1987). Linguists are highly interested in children's acquisition on suprasegmental features. To investigate the area of tone acquisition, there are three essential issues needed to be discussed in children's development of tone productions. They are: (1) the tone emergence ordering, (2) the frequency and accuracy rate, and (3) the substitution pattern in tonal errors.

1.2 Mandarin tone

Mandarin is the most widely spoken tone language, and it is the official language in both China and Taiwan. Tone is defined to be one kind of the suprasegmentals. It is a physical signal involving frequency (Hz). The change of frequency could result in pitch change, and tones would be produced by certain pitch register or contour. People speaking tonal languages use tones to distinguish lexical meanings. Standard Mandarin has four lexical tones, high, rising, low-falling-rising, falling, and one neutral tone. Chao (1930) have provided a 5-point scale to specify the tone values in Standard Mandarin where 1 represented low, 3 was mid, and 5 was high. According to this tonal representation, the tone letters of the four tones were [55], [35], [214], and [51]. Besides the four lexical tones, Lin (2007) mentioned that the neutral tone which only appears in the non-initial position was a low tone underlyingly, but its tone value would change depending on the preceding citation tone.

1.3 Research gap

By age-tracking children's speech tokens in early stages, researchers could provide evidence to see whether tones are acquired in a universal ordering. However, studies focusing on tonal acquisition were still few, and there were questions have not been answered. First, the orderings of tone acquisition in cross-linguistic tonal studies were inconsistent. Several researchers found that high tones and falling tones were acquired earlier than low tones and rising tones regardless of Mandarin, Cantonese, or Taiwanese (Li & Thompson, 1977 and Zhu, 2002 for Mandarin; Tse, 1978 and So & Dodd, 1995 for Cantonese; Tsay, 2001 for Taiwanese). However, evidence from the Thai data showed that the Thai-speaking children acquired the rising tone [224] earlier than the falling tone [51] (Tuaycharoen, 1977). Due to the inconsistent results, the conclusion of tone acquisition ordering could still not be made.

Besides, there was no sufficient research specifying the frequency and accuracy rate on immature tones. Before children fully master the tonal system, there would be an unstable period that a number of tonal errors would occur. The frequency and accuracy rate provide a good way to measure the preference and the degree of stabilization of tones. From the studies concerning tone acquisition in Mandarin, though Li and Thompson (1977) indicated that [35] and [21] were making tonal errors during two- and three-word stage, they did not provide specific accuracy rate in these two tones. Without the accuracy rate, it would be dangerous to determine a tone is “acquired” or not. The degree of development of tones was more specific in Zhu’s (2002) study. She applied a criterion (66.7% of accuracy rate) for determining the stabilization of tones, and found that [55] and [51] were stabilized earlier than [35] and [21], while the neutral tone was stabilized the last. Nevertheless, Zhu did not calculate the frequency in each tone, and the subjects recruited in her study were children who grew up in Beijing. It was reported that the Mandarin in Taiwan had many phonological differences from that in Beijing (Kubler, 1985; Duanmu, 2000). Due to different language environment, it is necessary for the current study to conduct the tonal acquisition study again in Taiwan Mandarin.

There were also few researchers analyzing monosyllabic and disyllabic tokens separately, and few studies compared the differences between these two types of speech tokens. When talking to young children, adults would shift the way of normal speech to a high frequency, reduplicated, and simplified way called "motherese" (Carroll, 2008). Regarding the motherese in Mandarin, care-takers also like to use reduplicated form in most lexicon, and the most preferred tone combination was a low-falling tone [21] followed by a rising tone [35] (Yang, 2012). For instance, standard form of the word *'shoes'* is pronounced as [ɕje35], but it is commonly used as [ɕje21-ɕje35] in the reduplicative form of motherese; some kinship terms such as [pa51-pa] *‘dad’* and [je35-je] *‘grandfather’* would also transform into [21-35] tone combination and become [pa21-pa35] *'daddy'* and [je21-ye35] *'grandpa'* in motherese. Thus, there would be a tremendous amount of tokens produced in the [21-35] combination. This would result in a bias in the whole number of occurrences. To deal with this problem, this study would manage to separate the disyllabic words from monosyllables, analyze the two types of tokens separately, and cope with the particular tone combination [21-35] separately.

Based on the research gap mentioned above, including the inconsistency on chronological orderings in different languages, the unspecific measurement on immature tones, and the unnoticed tone combination in motherese, it is worthwhile to discuss more on these issues and provide answers to the following research questions.

1.4 Research Questions

The research questions would be analyzed and discussed by following the sequence below:

(1) Regarding the issue of tone emergence ordering:

Which tone appears first? What is the chronological ordering of the emergence among all tones? At what age does the first tone appear? At what age would all the Mandarin tones be applied to children's meaningful utterances? Do children show similar tone emergence ordering cross-linguistically? Would the ordering reflect the degree of ease of articulation?

(2) Regarding the issues of frequency and accuracy rate (stabilization):

What is the most frequent tone produced by subjects? What is the ordering of frequency in the four tones, including the neutral tone? What is the ordering of accuracy rate in all the Mandarin tones? Do easier produced tones present higher accuracy rate? Do higher frequency tones also have higher accuracy rate? Do tones that appeared earlier have a higher frequency and accuracy rate in the occurrences? Would there be universal ordering of stabilization among different Asian tone acquisition studies?

(3) Regarding the issue of substitution pattern in tonal errors:

What kind of strategy would children use to replace the un-mastered tones? Would they replace the un-mastered tone with more frequent tones or more accurate tones? Which tone is more unstable and is replaced by other tones more frequently? Which tone is more likely to replace the unstable tones? Are there obvious individual differences among children's substitution strategies?

1.5 The framework of the thesis

In chapter 1, I provided a brief introduction to indicate what kind of issues are going to be discussed in the whole thesis, and what is the motivation for dealing with these issues; the four main research questions were also addressed. In Chapter 2, firstly I will mention the language universals in first-language acquisition in 2.1, and the Mandarin tonal system representations will be introduced in 2.2. Section 2.3 and 2.4 include studies specifically on tone acquisition. The comparison among tone acquisition studies of Asian languages is in 2.3; the foundation of tone acquisition studies in Mandarin is in 2.4. The last section 2.5 will introduce the tone acquisition theories. Chapter 3 contains the methodology which can be viewed in two separate parts. Section 3.1 is the data collection method describing how I obtained the speech tokens, and section 3.2 is the data analysis method explaining how the data were arranged. Chapter 4 will display the results and analysis in tables and graphs. Section 4.1 shows the overall results in all data. Section 4.2 applies different measure to analyze the monosyllabic and disyllabic tokens separately. In section 4.3, the substitution patterns in tonal errors are offered, and in section 4.4, the ages of emergence and stabilization in all subjects are summarized. The discussion and explanation are provided in chapter 5. Section 5.1 summarizes the findings in chapter 4. Section 5.2 and 5.3 are the comparison among different studies in Mandarin and different cross-linguistic studies. Section 5.4 is the concluding remarks and the suggestion for further research.

Chapter 2

Literature review

This chapter consists of three sections. Firstly, I will review studies with universals in children's phonological acquisition in section 2.1. Secondly, the various representations of Mandarin tonal system will be introduced in 2.2. Section 2.3 consists of a theory of tonal markedness. Later on, the important studies related to children's tone acquisition in Asian tone languages will be presented in 2.4. Lastly, the most related references focusing on Mandarin tonal acquisition will be shown in 2.5.

2.1 Language universals in first language acquisition

Many researchers proposed that there were chronological stages in children's early vocalization, which means that children across the world acquired languages by following similar steps. (Lenneberg 1967; Kaplan & Kaplan, 1971). The four stages of children's early vocalization divided by Kaplan & Kaplan (1971) are presented below:

Stage 1: Crying (0;0-0;6)-The cries are identified by parents as angry or physical pain.

Stage 2: Pseudocry and Noncry vocalization (0;4-0;5)- Begin to use the articulatory organs, and the utterances become different from crying.

Stage 3: Babbling and Intonated vocalization (0;5-0;8)- The vocalization becomes speech-like. Vowel-like and consonant-like sounds are combined into reduplicated syllable structure. Infants begin to imitate adults' intonation pattern.

Stage 4: Patterned Speech (0;9-0;12)- A large number of sounds appear, and child's first word begins to appear.

Although there were some overlapping among each stages and not all researchers agreed the division of the stages, the chronology of these development still gained consensus.

The reason why the chronology is proved universally may be attributed to the biological maturation of brain and motor control. Neurologists and psychologists proposed that babbling occurs automatically when the relevant structures in the brain have reached a critical level of maturation (Preyer, 1882). Before the maturation of the brain and motor control, Oller (1974) found that children would simplify the speech sound they perceived and construct a simplified sound system to ease the burden on their short-term memories and motor coordination.

2.1.1 Syllables

It was observed that children's early syllable structures in the babbling stage were simplified and would go through a regular progression: V, CV, and CVCV (Kaplan & Kaplan, 1971; Kies 1995).

1. V (The vocalic sounds are like *aaa*, *eee*, and *uuu*.)

2. CV (Children combine consonants and vowels into open syllables, such as *ta* and *ma*.)

3. CVCV (There are reduplications repeating frequently. The consonants and vowels are in ABAB form, such as *papa*, *nene*, and *tata*.)

The syllable structures began from simply vowels without consonants. Vowels are easier to produce than consonants for there is no need to exercise the articulators to impede the air flow in the vocal tract. To pronounce a vowel, the only articulator involved is the tongue, but consonants with different manners and places need more articulators working together such as lips, jaw, and velum. The progress from V to CV surely implied the consistent growth of children's muscular control.

Lightfoot (1982) also believed that language development were biologically determined, and children all over the world acquired languages with uniformity. There were several examples extracted from Demuth (2010) for how children in different languages used similar rules to truncate adults' target forms.

Table 2.1 The examples of children's simplified form in English and French

	Adults' form	**Children's form**	**Meaning**	**Age**
English	[spəˈk ɛ ti]	[ˈk ɛ ti]	'spaghetti'	1;2
French	[paˈtat]	[pəˈtæː]	'potato'	1;4

In Table 2.1, an English-speaking child imitated adult's [spəˈ kɛ ti] as [ˈ kɛ ti], and a French-speaking child produced [paˈtat] as [pəˈtæː] . Both children used the same phonological rule to simplify adults' trisyllabic words into disyllables, which were the CVCV structure in the study of Kaplan and Kaplan (1971) mentioned above. Children's immature muscular control unables them to produce consonant clusters like [sp] or close syllable like [tat]. Thus, children naturally simplified or truncated consonants when producing syllables more complicated than open syllables.

2.1.2 Suprasegmentals

Regarding the stages Kaplan & Kaplan (1971) presented above, they mentioned that in the third stage, infants began to imitate adults' intonation pattern. This finding suggested that children's suprasegmental acquisition, such as stress, intonation, pitch, and tone, started very early in the developmental process. Many studies proved that these suprasegmental characteristics were acquired very early and even earlier than segments (Lenneberg, 1967; Kaplan, 1970; Demuth, 1996). Crystal (1970) also cited that children at ages as young as 0;7 to 0;10 could exercise suprasegmental characters much stably and readily than segmental characters.

2.1.2.1 Intonation

For English-speaking children, it was found that infants could distinguish and produce falling and rising intonation contours at 0;8 (Kaplan 1970). When infants were in their late babbling stage, they could well imitate adults' intonation patterns despite with immature segments (Kies 1995). Kies provided an example to prove that an

English-speaking child could uttered two different intonation patterns at the age of 0;8.

Table 2.2 The two intonation patterns produced by an English-speaking child

(1)	o i tu ba ba ma	(Declaratives)
(2)	m-mu ga dow ba	(Interrogatives)

In these two sentences, although most of the syllables were simplified to V or CV forms, he had already developed the rising and falling in declarative and interrogative intonation patterns which refer to different meanings in English.

2.1.2.2 Stress

What's more, linguists also noticed that English-speaking children used stress to distinguish meanings as early as they uttered their first words (Kies 1995). Engel (1973) reported that one child used ['mama] to call his mother but [ma'ma] to his father at 0;10. The only difference of these two utterances laid on the stress position that 'mother' was stressed on the first syllable and 'father' was stressed on the second syllable. These examples could clarify that children do acquire prosodic features in their early phonological development.

2.1.3 Reduplication of motherese

Motherese, also named child-directed speech (CDS), refers to the speech form used by adults in talking to young children. Adults would naturally shift the way of normal speech to a high pitched, slow, reduplicated, and simplified way when talking to children (Carroll, 2008). There was a study pointing out that infants paid more attention to repetitive form than regular conversation, and the use of child-directed speech facilitated children to in language learning (Matychuk, 2005). The use of reduplication in motherese facilitated children's language development in several aspects: Phonologically, children could perceive the same syllable twice because the consonants and vowels in reduplications are identical in the first and second syllables. Morphologically, the syllables are simplified to CVCV forms, and the consonants and vowels are replaced to earlier acquired phonemes. Therefore, children better repeated syllables that were reduplicated.

Reduplication was the most common form applied in children's early production (Grunwell, 1982). As mentioned above in the study of Kaplan and Kaplan (1971), the CVCV structure were repeated frequently by children, and the consonants and vowels were in ABAB form, such as *papa*, *mama*, and *tata*. In the reduplication of motherese in Mandarin, Yang (2012) found that in Taiwan, most care-takers would produce reduplications in a particular tone combination which is a low-falling followed by a rising tone, such as [njow21-njow35] 'a cow' and [y21-y35] 'a fish.' Children also uttered most reduplications in [21-35] combination. It was obvious that children's preference of tone combination was identical to that of adults' reduplication in motherese. Whether children's form was affected by adults' input needs more specific methods to examine, but the examination would not be conducted in this current study.

2.2 Introduction to Mandarin tonal system

Mandarin tones has been classified and transcribed into several different systems. In this section, I will introduce tonal representations both in Standard Mandarin and Taiwan Mandarin in 2.2.1 and 2.2.2. A summary and the adopted tonal representation will be provided in 2.2.3.

2.2.1 The tonal representation systems of Standard Mandarin

The Standard Mandarin (henceforth SC) is the official language in China. The standard pronunciation of SC is based on Beijing dialect, but it absorbed many other Chinese dialects in different regions (Duanmu 2000). Despite the variance, the tonal representation systems of SC reviewed here are the descriptive view of Mandarin.

2.2.1.1 Chao (1968)

Chao (1968), also known as a musician, established the five-point scale to transcribe the pitch of tones in Standard Mandarin (henceforth SC). He proposed that the tone classification in Mandarin could be realized as pitch differences, and the pitch register was shown in a range numbered 1 to 5. The bigger the number was, the higher the pitch register would be.

Table 2.3 Chao's (1968) tonal representation system of SC

Tone #	Chinese Name	Tonal feature	Pitch value
1	*Yinping*	High Level	55
2	*Yangping*	High Rising	35
3	*Shangsheng*	Low Falling-Rising	214
4	*Qusheng*	High Falling	51

In Table 2.3, the four tones named *yinping*, *yangping*, *shangsheng*, *qusheng* were described as High Level [55], High Rising [35], Low Falling-Rising [214], and High Falling [51]. The five-point scale which was invented by Chao (1930) was widely used and cited in most research on tone languages for its convenience to depict a tone clearly and specifically. The tonal acquisition studies reviewed in the following sections also adopted Chao's five-point scale for describing pitch values.

2.2.1.2 Yip (2001)

Yip (2001) provided the link between the phonetic targets and their corresponding representations in Mandarin tonal system. Her tonal representation was shown below.

Table 2.4 Yip's (2001) tonal representation system of SC

Tone #	Pitch value	Tonal feature
1	55	H
2	35	MH
3	21	L
4	53	HM

In this tonal system, tones were considered to have a head and a contour in the following. The capital H and L in Tone 1 and Tone 3 represented level tones, and MH and HM in Tone 2 and Tone 4 were contour tones. With regard to the pitch values, the representation of Tone 1 and Tone 2 agreed with those in Chao (1968), but Tone 3 was presented as [21] without the final rising and Tone 4 was presented as [53] with partial falling from high to mid.

2.2.1.2 Lin (2007)

The system Lin (2007) adopted in her studies concerning SC was presented both with tone features and pitch values. She mentioned that most analyses of SC used only high and low distinction, but it was not specific enough. There were three distinctions of register in Lin's system, which were high, mid, and low, and they would be presented in capital letters (H, M, and L).

Table 2.5 Lin's (2007) tonal feature and pitch value

Pitch value	Tonal feature
4 or 5	H (high)
3	M (mid)
1 or 2	L (low)

In Table 2.5, it defined the corresponding tonal features of the pitch values respectively. The highest two numbers, 4 and 5, were categorized in H; number 3 which referred to a middle pitch sound fitted in M; the lowest two digits, 1 and 2, had the low pitch feature, so were sorted in L. The tonal features could be used to transcribe into the digital system.

Table 2.6 Lin's (2007) tonal representation system of SC

Tone #	Tonal feature	Pitch value
1	HH	55
2	MH	35
3	LH (in phrase final syllable)	214
	LL (before another tone)	21
4	HL (in phrase final syllable)	51
	HM (before another tone)	53

The four tones in Table 2.6 were transcribed into pitch register features. For Tone 1, the high level [55] pitch value mapped with HH in this system; Tone 2 which with [35] pitch value was presented by MH. There were respectively two representations in Tone 3 and Tone 4 for a more detail distinction regarding different syllable positions. The basic forms of these two tones were shown in LH for Tone 3 and HL for Tone 4. Yet, it was believed that final part of Tone 3 and Tone 4 would be reduced when followed by another tone. When these two tones were not in syllable-final position, the rising part of Tone 3 would be missing and presented as LL; the falling part would become [53] and be presented as HM.

The neutral tones which usually appear in grammatical particles or the second syllable of reduplicated kinship terms were precisely depicted in Lin's (2007) book. Phonologically, a neutral tone was unstressed and was a low tone underlyingly. However, the pitch values were found to be different when following different stressed tones.

Table 2.7 Pitch values of the neutral tone

Tone #	Tone feature	Example	
T1 + T0	55 + 2	[ma55 ma2]	'mother'
T2 + T0	35 + 3	[lai35 lə3]	'came'
T3 + T0	21 + 4	[tɕ je21 tɕ je4]	'older sister'
T4 + T0	53 + 1	[k^{h}an53 lə1]	'saw'

The Table 2.7 showed the pitch values after each tone. The tone values remained low when following T1 and T4, but when following T2 and T3, the values would rise to [3] and [4]. The examples in the right column showed where the neutral tone could appear and which pitch value would be surfaced respectively.

2.2.2 The tonal representation systems of Taiwan Mandarin (TM)

Many linguists observed that the Mandarin in Taiwan (henceforth TM) was distinct from SC which spoken in Mainland China (Kubler, 1985; Fon, 1997; Fon & Chiang 1999; Duanmu, 2000; Lin, 2007). The tone change in TM may result from the influence of Taiwanese, the dialect widely spoken in Taiwan. The specific tone values of TM have been reanalyzed acoustically and the studies and their tonal representations would be presented below.

2.2.2.1 Shih (1988)

An acoustic study conducted by Shih (1988) pointed out that the pronunciation of Tone 3 was particularly different from people between Mainland China and Taiwan. In TM, Tone 3 was often produced without the rising part no matter in word-final position or not. The same description was also stated by Kubler (1985) that in Taiwan Mandarin, Tone 3 tended to be pronounced as [21] instead of [214]. The representation Shih adopted was presented below.

Table 2.8 Shih's (1988) tonal representation system of TM

Tone #	**Tonal feature**
1	HH
2	MH
3	LL
4	HL

In Table 2.8, Shih used HH to represent the high-level tone, and MH to represent the rising tone. Tone 4 was presented in HL which was considered a complete falling. Due to the dialectical variance mentioned by Kubler (1985) and Shih (1988), Tone 3 which often produced without the rising part was showed to be LL in Shih's Taiwan Mandarin tonal system.

2.2.2.2 Fon (1997)

Fon (1997) measured the pitch height, duration, and slope steepness of the four tones by the Computerized Speech Lab, KAY Elemetrics. Analyzing the tones produced by the subject who was a 22-year-old female college student in National Taiwan University, Fon (1997) has characterized the tonal representation system for TM in Chao's five-point scale.

Table 2.9 Fon's (1997) tonal representation system of TM

Tone #	**Pitch value**
1	44
2	323
3	312
4	42

The pitch values of the four tones presented in Table 2.9 were totally different from the system invented by Chao (1968) which has been shown in Table 2.3. Fon (1997) argued that the system she proposed has acoustic fact, and the subject enrolled was born in Taiwan. Thus, the result of the tonal representation would be so distinct from that in Chao's (1968) study. In fact, the two main distinctions between SM and TM lied in the pitch range and the contour changes of Tone 2 and Tone 3. The pitch range narrowed to a four-point scale. The height in Tone 1 lowered from [55] to [44], and the pitch range for Tone 4 also narrowed from [51] to [42]. On the other hand, Fon presented Tone 2 as a dipping tone whose contour was [323], and the contour was similar to the pitch value in Tone 3 presented as [312].

A further perceptual study conducted by Fon and Chiang (2004) illustrated that Tone 2 and Tone 3 had crucial differences between duration, steepness, and height. The duration in Tone 2 was longer than Tone 3, and the slope of Tone 2 was not steeper than Tone 3. The higher pitch of Tone 2 in the ending point was a cue for distinguishing Tone 2 and Tone 3.

2.2.3 Summary

While there were several different tonal representation provided by phonologists and phoneticians in Standard Mandarin and Taiwan Mandarin, how to determine which tonal system is the most suitable one? Duanmu (2000) suggested that it was allowed to slightly modify the transcription of the tonal representation if only if the modified form does not cause meaning contrast in Mandarin. For example, the [21] and [11] in Mandarin do not distinguish meanings, so it does not matter whether we call Tone 3 as a low falling or a low level.

In this thesis, I will mainly adopt the five-point scale invented by Chao (1968) because it is the basic system adopted by most of the phonetic studies. To accommodate the change of tone in Taiwan Mandarin which proposed by Shih (1988) and Kubler (1985), Tone 3 would be modified to [21] for the dialect variation in TM.

Table 2.10 The tonal representation adopted in this thesis

Tone	Pitch value	Tonal feature
Tone 1	55	high-level
Tone 2	35	rising
Tone 3	21	low-level
Tone 4	51	falling
Neutral	neut	neutral

Table 2.10 showed the pitch value and tonal features which will be adopted in this study. The transcription used for the four tones respectively would be [55], [35], [21], and [51]. The corresponding tonal features of the tones would be high-level, rising, low-level, and falling. The neutral tone would be transcribed in the abbreviation 'neut.'

2.3 Theory of markedness of tone

Yip (2001) presented a theory of tonal markedness to distinguished marked features and unmarked feature in tone. The tonal markedness theory was derived from three types of data. The first type of data was from Hashimoto's (1987) survey on tone sandhi in 83 dialects of Chinese. Second, she provided data from Cheng's (1973) quantitative study of Chinese tones in which the tonal inventories of 73 dialects were studied. The third type of evidence was from the acquisition studies conducted by Clumeck (1980) and Li and Thompson (1978). The evidences could be generalized into three rules, and Yip stated that the markedness rules was used "to minimize articulatory effort."

Minimize articulatory effort

a. contour tones are more marked than level tones

b. rising tones are more marked than falling tones

c. high tones are more marked than low tones

Yip summarized Hashimoto's (1987) study that contour tones were more likely to be leveled in tone sandhi, and similar results were also found in children's acquisition data (Li & Thompson,1978; Clumeck, 1980). In numerous tonal inventories which presented by Cheng (1973), it was more likely to have falling tones in the tonal system than rising tones. The third constraint focused on comparing level tones. She found that high level tones were more marked because it needed more strength and could not "minimize articulatory effort." From this point of view and the evidences from numerous dialects, she concluded that low level tones were more unmarked. The three tonal markedness rules could be examined by the acquisition data from cross-linguistic studies to determine the universal features of tones.

2.4 Tone acquisition studies in East Asia

In this section, I will review several tone acquisition studies cross-linguistically, including Thai, Cantonese, and Taiwanese. Linguists are interested in whether there are language universals in first language acquisition. The tone acquisition studies reviewed here will be compared to the results of the current study in the discussion section. In this section, a Thai study focusing on phonetic and phonological development in early speech will be presented in 2.4.1. Two longitudinally conducted researches on Cantonese tone acquisition will be described in 2.4.2. A Taiwanese study, which was also conducted longitudinally, would be reviewed in 2.4.3. Then, there will be an overview of the cross-linguistic tonal acquisition studies in section 2.4.4.

2.4.1 Thai

Thai is the official language used in a southern Asian nation, Thailand. There are five tones in Thai's tonal system. There are three level tones, including High, Mid, and Low; and the other two tones are contour tones, which are the Rising tone and the Falling tone. Examples of contrastive meanings for the same syllable are presented in Table 2.11.

Table 2.11 Thai tonal system

Tone	Tone feature	Pitch value	Examples	
1	Mid	33	*khaa33*	'a grass'
2	Low	11	*khaa11*	'galangal'
3	Falling	51	*khaa51*	'to kill'
4	High	45	*khaa45*	'to engage in trade
5	Rising	224	*khaa224*	'leg'

Tuaycharoen (1977) observed the tone acquiring order and age from her son between the age of 0;3 to 1;6. The data collection was done by the author at home about twice a week. The recording collected the natural interaction between the child and his parents and grandparents whose mainly used language was Bangkok Thai. The finding revealed that [33] and [11] were first acquired at the first-word stage aged 0;11, and the next acquired tone was the rising tone [224] which was learned by 1;2. The high tone [45] and falling tone [51] appeared unstably at 1;3 to 1;6. To sum up, the order of the tone acquisition on Thai in this particular study was presented to be: [33], [11]> [224]> [45], [51].

2.4.2 Cantonese

Cantonese is a Chinese dialect spoken by people in Hong Kong and Macau in southern China. The tonal system is categorized into six contrastive tones, including four level tones and two rising tones. There are three extra-short tones (entering) which are allotones of three level tones and occur only on syllables which are closed by plosives. The four level tones are [55], [33], [22] and [11] and Tse (1978) categorized them into Upper Even, Upper Going, Lower Going, and Lower Even. The two contour tones are different from pitch height that the Upper Rising [25] starts from mid to high, and the Lower Rising [13] starts from low to mid. The three entering tones are not included in the tone acquisition studies below.

Table 2.12 Cantonese tonal system

Tone	Tone feature	Pitch value	Examples	
1	Upper Even	55	*ji55*	'a cloth'
2	Upper Going	33	*ji33*	'meaning'
3	Lower Going	22	*ji22*	'two'
4	Lower Even	11	*ji11*	'to doubt
5	Upper Rising	25	*ji25*	'a chair'
6	Lower Rising	13	*ji13*	'an ear'
7	High Entering	5	*sik5*	'to know'
8	Mid Entering	3	*sek3*	'tin'
9	Low Entering	1	*sik1*	'to eat'

Tse (1978) had conducted a longitudinal case study and collected the tone acquisition data from his son between the age of 0;10 to 2;8. He found that [55] and [11] were firstly mastered at the beginning of one-word stage at 1;4, and then [33] appeared by 1;8. By this age, the child was still in the one-word stage. [22], [25], and [13] were mastered in the two-word stage which was from 1;9. Tse pointed out that the child was confused about the two rising tones [25] and [13] during 1;9 to 1;10. At the age of 1;10, the child had completely mastered the tonal system. To briefly review the process of tone acquisition in this study, the order was [55], [11]> [33]> [22], [25], [13].

So and Dodd (1995) also presented a longitudinal study in Cantonese tone acquisition by observing four children from 1;2 to 2;0. The results showed some discrepancies comparing to Tse's (1978) study. The data showed that the first mastered tones were [55] and [33] at 1;4, and the [25] appeared second at 1;6. The four children's orderings remained the same before this age. After 1;8, individual differences were found. Two children acquired the Lower Going tone [22] first, one acquired [11] and [13] first, and the other one had not acquired any additional tones by the age of 1;10. At the age of 2;0, it is reported that all the four children had mastered the six tones in Cantonese. Because there were individual differences in this study, the conclusion could only be made in the following order: [55], [33]> [25]> [11], [13], [22].

2.4.3 Taiwanese

Taiwanese is a dialect of Southern Min Chinese spoken in Taiwan. There are seven lexical tones in this language, but each of it has two tone values. When syllables are in the "juncture positions," they are pronounced in the original tones; when they appear in the "context positions," the tone sandhi rule would be applied and they would be pronounced in different tones. There are three level tones with different pitch height, which are High, Mid, and Low tones. The contour tones are the High Falling tone and the Low Rising tone. The last two short tones are 'Rusheng' tones and were not discussed in the following tone acquisition study. The tone values in both juncture and context positions are shown below in Table 2.13.

Table 2.13 Taiwanese tonal system

	Juncture				Context	
Tone	**Tone feature**	**Pitch value**	**Examples**		**Tone feature**	**Pitch value**
1	High Level	55	*si55*	'a poem'	Mid level	33
2	Low Rising	13	*si13*	'time'	Mid level	33
3	High Falling	53	*si53*	'death'	High level	55
4	Low Level	11	*si11*	'four'	High falling	53
5	Mid Level	33	*si33*	'a temple'	Low level	11
6	Mid Ru	3	*sik3*	'color'	High ru/ High falling	5/53
7	High Ru	5	*sik5*	ripe	Low ru/ Low level	1/11

Tsay (2001) had conducted a longitudinal and big-scale research to observe a total of fourteen Taiwanese-speaking children aged 1;2 to 3;11 in southern Taiwan. The big-scale observation persisted for three years and recorded a total of 330 hours data. In the tonal acquisition analysis, Tsay focused on seven children (3 girls and 4 boys) aged 2;1 to 2;3 whose MLU were longer than two words and could produce utterances longer enough to apply tone sandhi in context position. The seven children's average error rates in each tone were calculated in the following.

Table 2.14 The average error rate of each tone in juncture position

Tone	[55]	[53]	[33]	[13]	[11]
Error rate	4%	6%	6%	8%	18%

In Table 2.14, it indicated that the errors on [55] tone were the fewest. The errors on [55] accounted for 4% of the total error tokens. The error rates on [53] and [33] were also few, accounting for 6 % respectively. There were 8% of errors on [13], and the most unstable tone was [11]. From the tone error rates in juncture position, Tsay pointed out that tones in high pitch such as [55] and [53] acquired more stable than those in low pitch, including [13] and [11]. In addition, the results also showed that the falling tone [53] had fewer errors than the rising tone [13]. To sum up, the tone acquisition order in Taiwanese children was [55]> [53], [33]> [13]> [11].

2.4.4 Overview of the cross-linguistic studies

After reviewing the three tonal systems in East Asia and the relevant tone acquisition studies above, we would like to compare the acquisition ordering cross-linguistically.

Table 2.15 The acquisition ordering in Thai, Cantonese, and Taiwanese

	Thai (Tuaycharoen 1977)	Cantonese (Tse 1978)	Cantonese (So & Dodd 1995)	Taiwanese (Tsay 2001)
Early	[33] [11]	[55] [11]	[55] [33]	[55]
	[224]	[33]	[25]	[53] [33]
↓	[45] [51]	[22] [25] [13]	[11] [13] [22]	[13]
Late				[11]

Although the tonal inventories in every language are different, we still can categorize them by tone features. The tones could first be divided to level tones and contour tones. Regarding the level tones, the different registers of tones could roughly be divided to high, mid, and low tones. Lin (2007) defined that the digit number of 5 and 4 in pitch value were high tones, 3 was mid, and 2 and 1 were low tones. With regard to contour tones, there are mainly two types which are falling and rising tone. Tones that end at a higher pitch than the starting pitch are considered rising tones; tones which end at a lower pitch than the starting pitch are viewed as falling tones.

From the cross-linguistic studies above, the common ground was that the level tones were acquired earlier than contour tones among all studies. The earliest acquired tones in these three languages, including the [33] and [11] in Thai, the [55], [33] and [11] in Cantonese, and the [55] in Taiwanese, were all level tones. There was also language-specific phenomenon. It seemed that falling tones were mastered earlier than rising tones in Taiwanese, but Thai had the opposite ordering. In Taiwanese, the falling tone [53] was produced more stable than the rising tone [13], but in Thai, the rising tone [224] seemed to be acquired earlier than the falling tone [51]. Regarding the ordering of level tones in different registers, the high-level tones tended to be acquired earlier than low-level tones. In Taiwanese, the orders showed sequential ranking from high-level tones to low-level tones that the [55] tone was acquired earlier than [33], and [33] was followed by [11]. Also, the Cantonese study conducted by So and Dodd (1995) found that [55] and [33] were acquired earlier than [22] and [11]. However, the Cantonese study presented by Tse (1978) indicated that [55] and [11] were acquired at the time same, [33] was acquired later, and [22] the last. The sequence of high tones and low tones also showed variations in Thai. Tuaycharoen (1977) suggested that [33] and [11] were acquired earlier than [45], which was the only HH tone in Thai. To sum up, level tones were acquired earlier than contour tones universally; falling tones were acquired more stable than rising tones cross-linguistically except for Thai; high-level tones seemed to acquired the earliest except for Thai, but it was uncertain whether the mid tones were acquired earlier than low tones.

2.5 Tone acquisition in Mandarin

2.5.1 Chao (1951)

When observing the process of tone acquisition, early studies were focusing on the age and ordering of the development of tones. A pioneering observation on Mandarin tone acquisition was resented by Chao (1951) who collected speech data from his 28-month-old grand-daughter. Chao found that his grand-daughter could have already distinguished the rising tone [35] and the low-falling tone [214], and could produce tones correctly in isolated words. Though she could distinguish the rising tone [35] and the low-falling tone [214], Chao noticed that she tended to replace the low-falling tone [214] with the rising tone [35]. However, Chao's description was based on only one child, and he did not exactly present the whole picture of tone acquisition in this study, the value of the results may be limit to a first glance in the tone acquisition field in Mandarin Chinese.

2.5.2 Li & Thompson (1977)

After Chao's (1951) contribution in Mandarin tone acquisition, Li and Thompson (1977) conducted a larger and systematic research focusing on Mandarin speaking children in Taiwan. They use four stages to sketch the tone acquisition process:

Stage I: The child's vocabulary is small. High and falling tones predominate irrespective of the tone of the adult form.

Stage II: The child is still at the one-word stage, but he has a larger vocabulary. The correct 4-way adult tone contrast has appeared, but sometimes there is confusion between rising and dipping tone words.

Stage III: The child is at the 2/3-word stage. Some rising and dipping tone errors remain. TS is beginning to be acquired.

Stage IV: Longer sentences are being produced. Rising and dipping tone errors are practically non-existent.

The four divided stages on tonal development presented the chronological ordering and the corresponding word length children uttered. The important findings in Li and Thompson (1977) were that the high-level tone [55] was the first acquired tone, and then the high-falling tone [51] was the second. The rising tone [35] and the low-falling-rising tone [214] were acquired the last. By the stages when children have not mastered all tones yet, the switch between [35] and [214] persisted throughout stage II and stage III. This report provided a more complete understanding of the age and ordering in Mandarin tone acquisition. They also provided two children’s substitution strategies in tonal errors. One child replaced all [35] and [214] with [55] and [51], and the other child had constant substitution between [35] and [214]. Although Li and Thompson did the first systematic tone acquisition study in Taiwan Mandarin and sketched the stages of development, the number of utterances or error rates of tones were not documented specifically, so the degree of development in the process of tone acquisition was not documented.

2.5.3 Zhu (2002)

A more recent work related to the tone acquisition in Mandarin set more specific criterion on stabilization of the acquired tones. Zhu (2002) conducted a longitudinal study in Beijing on four Mandarin-speaking children aged 0;10 to 1;2 in the beginning and 1;8 to 2;0 at the end. She provided the age of tone emergence and age of stabilization in each subject. The order of tone emergence was similar to that in tone stabilization. The criterion for deciding the tone emergence and stabilization was clearly cited in this study, and the tonal error patterns were presented in specific number of frequencies. The results showed that the high-level tone [55] was firstly emerged and stabilized. The second one was the falling tone [51], and rising [35] and falling-rising tones [214] were the last. When tonal errors occurred, the most frequent tone that realized to replace the error was the high-level tone, and high-level tone seemed to be replaced by the falling tone when produced wrongly. However, the subjects Zhu studied were from Beijing, and the Mandarin was different from that in Taiwan. It is valuable to see whether the development of tones would be different in children exposed to dialects in Taiwan.

2.5.4 Summary

Based on the three studies reviewed above, researchers agreed that the high-level [55] and falling [51] acquired earlier than the falling-rising [214] and the rising tone [35]. But the substitution pattern for whether [35] was easier to replace [214] or vice versa did not gain consensus. There was also no precise document describing the tonal acquisition process in developmental stages. Thus, the topic is worth for further research.

Chapter 3

Methodology

The methodology would include two parts: one is the data collection, and the other is the data analysis. The data were collected by the author and the research team in the Phonetics and Psycholinguistics Lab at National Chengchi University under the NSC project, “Consonant Acquisition in Taiwan Mandarin,” investigated by Professor Wan I-Ping (NSC 100-2410-H-004-187-).

For data collection in section 3.1, I will introduce how I recruited the participated families in 3.1.1, what my subjects’ backgrounds were in 3.1.2, how the observation proceeded in 3.1.3, and what recording equipments were used in 3.1.4. For the data analysis in section 3.2, I will illustrate how I transcribed the data in 3.2.1. From 3.2.2 to 3.2.4, I will show how I arranged the data in order to obtain the result of the tone emergence ordering, frequency, accuracy rate, and the substitution pattern in tonal errors.

3.1 Data collection

This section contains the process of recruitment in 3.1.1, the background information of the informants in 3.1.2, the observational procedures in 3.1.3, and the recording equipments in 3.1.4.

3.1.1 Recruitment

The participated families were recruited through an advertisement on a popular parent forum called Babyhome (http://www.babyhome.com.tw/), on the behalf of the NSC project investigated by Professor Wan I-Ping (NSC 100-2410-H-004-187-). In the non-profit advertisement forum, an article was pasted to declare the academic research purpose, and to ask for recruiting children aged from 0;8 to 1;0 who was at the beginning stage of their language development. Parents who wanted to join the research could sign up by filling out the online registration form designed by the "Google doc spread sheet," which could be customized by users. There were totally 16 families enrolled in the NSC project, but only 6 children fit in this study.

3.1.2 Subject

The six children were all from middle class families in Taipei City or New Taipei City. These families were all core families that the children only lived with their parents, and the informants were all taken cared by their mothers for the whole day. All Mothers used Mandarin Chinese to communicate with their children, so these children's first language was determined to be Mandarin.

Among the 6 subjects, three of them were males and three were females. From the beginning of the observation, their ages were between 0;10 to 1;1 (mean age= 0;11.67, SD= 0.8 months). The observation continued for eight months. At the end of the recording, the subjects grew up to the ages between 1;5 to 1;6 (mean age= 1;5.67, SD= 0.05 months). The six subjects were all healthy and had not detected with any hearing or intellectual impairment.

The subjects' language development was around the one-word stage that some of them had not produced any meaningful words yet, but some of them had already produced some meaningful words with clear lexical tones. Among these six children, three of them were the only child in their family, including subject #1, #3, and #4. The other three were the second child, including subject #2, #5, and #6 whose older siblings were all brothers, and the age gaps between the first and second children were smaller than four years old. The subjects' background information is presented below.

Table 3.1 The data collecting information on subjects and recordings

Subject	**Gender**	**Age range**	**Duration**
#1	M	1;1-1;6	6 months
#2	F	1;0-1;6	7 months
#3	F	1;0-1;6	7 months
#4	M	0;11-1;6	8 months
#5	F	0;10-1;5	8 months
#6	M	0;10-1;5	8 months

3.1.3 Observational procedures

The data collection started from January 2012 to the present. There were over six research assistants in the research team, and the team sent two assistants to an informant's house to record the spontaneous speech between the child and the mother every two weeks. The recording was about sixty minutes long for one time, but it might be shorter if the children felt tired and started to cry. During the recording, one of the assistants was in charge of the video-taping who had to step aside, stabilize the camera, and make sure to film the child's face and the object he/she was playing; the other assistant had to hold the digital sound recorder, stay near the child, and interact with the child. The research assistants would not bring any toy or reading-material to the family, but might use their own toys to ask questions or attract the child's attention. Because three of my informants had older siblings, sometimes the older child would join the free-play while recording. Although the older sibling usually performed better language competence than the younger one and would interfere the recording, some Mothers told us that the younger child uttered more words when playing with their older siblings.

The participated families were paid NT$80 per visit, and they could receive an album of their own video recordings as a souvenir at the end of the term in the research project. The rewards, equipments and cost were all supported by the NCS project (NSC 100-2410-H-004-187-).

3.1.4 Recording equipments

We used both video-recording and sound-recording equipments, which were Sony DCR-SR40 Handycam digital video camera recorder and the Sony ICD- UX513F digital voice recorder. The sizes of these equipments were both very small and functioned well. Both equipment provided high-quality digital files. The video files helped us decode the utterance meaning by children's gestures and eye movements, and the sounds file provided us high quality audio signals.

3.2 Data analysis

The subjects in the observation period were around one-word stage, and most of their utterances were short within two syllables, so the study plans to analyze only monosyllabic and disyllabic words. Although children in this stage could hardly produce perfect consonants and vowels, their tones developed better and earlier (Lenneberg, 1967; Kaplan, 1970; Demuth, 1996). To analyze the tones, the utterances with clear tones were all included. However, the referential meaning of children's utterances would sometimes be unclear. No matter the utterances had clear meanings or not, once the tones were recognizable, we would include them. In 3.2.1, I will introduce how I transcribed the utterances into speech tokens. The method used to track the tone emergence ordering will be explained in 3.2.2, the formulas to calculate frequency and accuracy rate will be shown in 3.2.3, and how to arrange the substitution pattern in tonal error will be presented in 3.2.4.

3.2.1 Transcription and coding

Each recording file was transcribed by two assistants at the same time. If there were disagreements, the token would be discussed or checked by another research assistant in the team. Though young children's utterances were sometimes fuzzy and hard to categorize, Yang (2010) has tested Mandarin native speakers' perception of tones and found that native speakers had the ability to perceive four tones by pitch contour and register. In order to show the reliability of our perception, we extracted some examples from the sound files that represented the recognizable and unrecognizable tokens respectively. The pitch contours of the exemplified tokens were presented by the computer phonetic software, Pratt, in fundamental frequency (F0). The recognizable tones were classified into different tone groups in Figure 3.1, and there are 4 examples with unrecognizable tones presented together without classification in Figure 3.2.

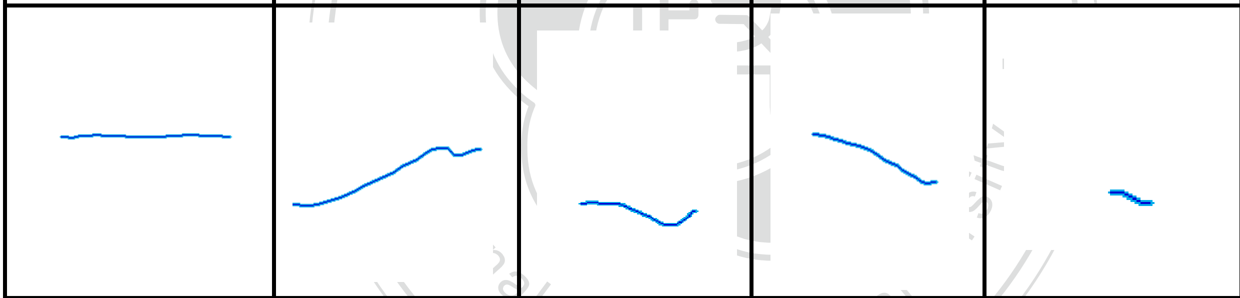
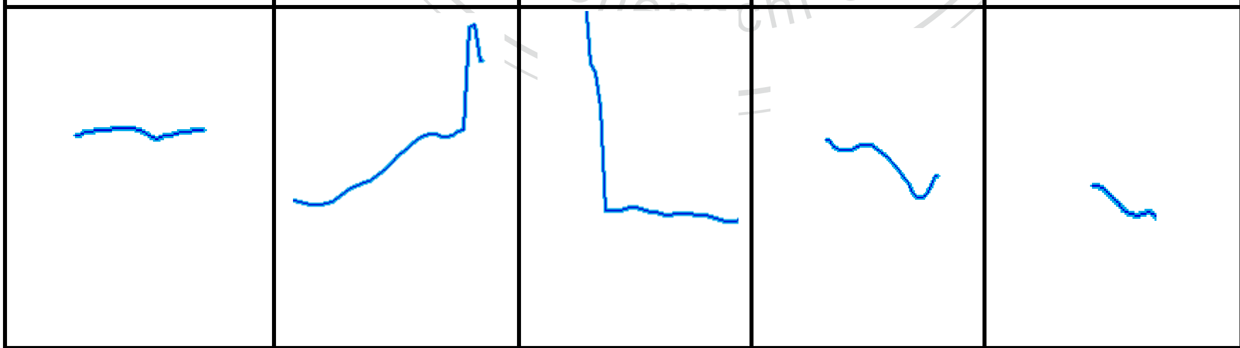
Recognizable tones					
T1 [55]	T2 [35]	T3 [21]	T4 [51]	T0	
					700Hz 75Hz
[ja55]	[ma35]	[ma21]	[tu51]	[mə]	
					700Hz 75Hz
[tɕɤ 55]	[jɛ 35]	[tɕ jo21]	[pa51]	[tə]	

Figure 3.1 The pitch contour of recognizable tones

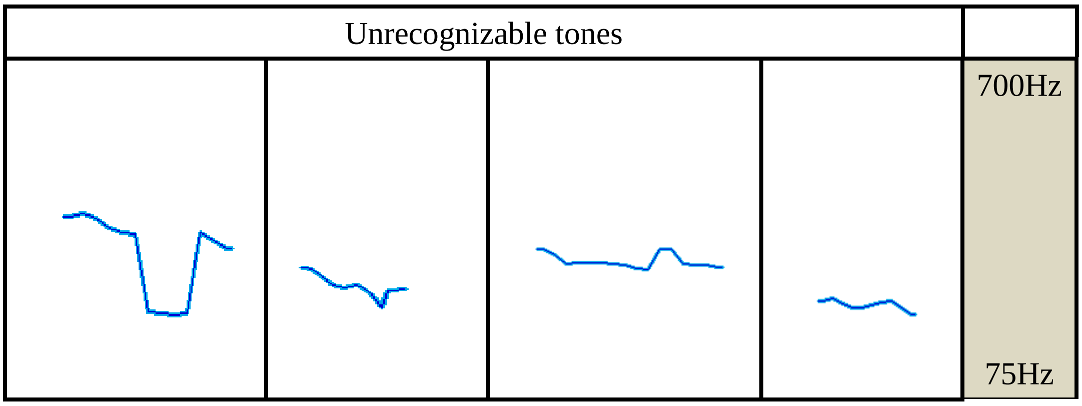


Figure 3.2 The pitch contour of unrecognizable tones

The pitch contours in the two rows of Figure 3.1 were recognizable tokens and were categorized into each tone group. The high-level tones [55] in the two examples were both flat and high in frequency. The contours of the rising tones [35] were climbing and the durations were long. The low-level tone [21] in the first row exactly presented a low-level contour. The second example in [21] showed a low-level contour preceded by a high-level neighboring tone in the context, so it started from a vertical line and then leveled off. The pitch contours in falling tones [51] went downward clearly, and the neutral tones were rather short than other tones. The neutral tones were examined to be significantly shorter than other lexical tones in Mandarin (Chen & Xu 2006). Thus, the short contours suggested that our ability of detecting the neutral tones was precise. The examples in Figure 3.2 were unrecognizable tones produced rapidly or sloppily. The movements of these pitch contours seemed more uncertain, so they were excluded. All in all, the pitch contours in Figure 3.1 and Figure 3.2 have testified our perception that our coding of the Mandarin tones was reliable.

All speech tokens were classified into two groups. The tokens without clear semantic meanings were grouped together and were transcribed into segments and tones. The other group contained the tokens with clear meanings and they were transcribed into four parts: segments, produced tones, word meanings, and target tones. The transcribed examples are presented below in Table 3.2.

Table 3.2 The sample of coding

	Segment	**Produced Tone**	**Meaning**	**Target tone**
without meaning	na	[1]		
	pipi	[51-21]		
with meaning	ja	[35]	a lamb	[35]
	tjotjo	[55-N]	a ball	[21-35]
	wawa	[35-55]	a doll	[35-55]

Firstly, the consonants and vowels were taken down by International Phonetic Alphabet (IPA). Secondly, the produced tones were coded with [55], [35], [21], [51], and the neutral tones were coded with a capital N. Thirdly, the intended meanings of speech tokens were also transcribed. Although children's articulation was underdeveloped, their meanings of utterancea were sometimes recognizable. The referential meaning could be identified by context or children's gestures. For example, if a child pointed at a ball and uttered [tjo21 tjo35], we would take down its intended meaning as '*a ball*' instead of '*to throw*.' Although the segments in this utterance sounded more like the verb *'to throw'* in Mandarin, it was more reliable to determine the target tones from context. If the utterance yields meaninglessness, or we could not recognize it by context, we would leave it blank. The meaningless tokens could only apply frequency analysis but could not apply other measures because it does not have target tones. Fourthly, in order to determine the correctness of tones, we had to decide what the target tones were in every meaningful token. Basically, the target tones were determined by their care-takers' tone model in motherese. However, every mother's motherese would be slightly different (Demuth 1993). For example, some care-takers may use [wa21-wa35] to refer to *'a doll'* but others would use [wa35-wa55] to refer to the same thing. Therefore, we had to make sure which tone was applied by each care-taker when they used motherese to talk to their children, and then we could determine what the target tones were and whether their children's tones were correct.

3.2.2 Tone emergence ordering

The tone emergence ordering is a common issue for studies concerning first language acquisition, and it is also the first step researchers could investigate in children's tonal development. In this study, I also applied the age-tracked method to take down the occurrences of every tone by ages. The criterion for tone emergence was defined by Vihman (1996) that a tone which was produced more than once in meaningful words would be considered an emerged tone. To compare whether there were individual differences between the subjects, I took down their individual ordering separately. For example, if a child first uttered [ma55] *'mother'* twice at 0;11, had [ta21] *'to hit'* and [tu51] *'rabbit'* at 1;0, and finally produced [je35] *'grandpa'* at 1;3, then, his tone emergence ordering would be [55]>[21],[51]>[35]. With this ordering, researchers could examine the universal rules cross-linguistically.

3.2.3 Frequency and accuracy rate of tones

The tone emergence ordering could only help us observe which tone appeared first, but it could not present the mastery of each tone. Analyzing the number of tokens and checking their correctness are the better ways to picture children's developmental process. The frequency could reveal children's preference of tones, and the accuracy rate could show the stabilization of tones. Li and Thompson (1977) revealed that children would avoid producing words that contain tones that they have not mastered yet. If the frequency of a certain tone is low, it may be explained that the tone is more problematic to children and has not yet been acquired. For instance, if a child's frequency of the rising tone [35] was observed to be much lower than other tones and was often replaced with other tones, it would indicate that the child had not yet acquired [35], and tended not to use this [35] when reproducing adults' speech. The tone frequency could be used to examine whether the more frequent tone would be acquired earlier, and whether the least frequent tone would be more problematic and would be acquired last.

When calculating the frequency of tones, because it is to calculate the number of occurrences, the utterance meaning does not matter. Thus, we included both tokens that with and without clear meanings in frequency analysis. The frequency of every tone would be computed by the formula presented below.

$$\text{Frequency} = \frac{\text{the number of tokens of a tone}}{\text{the number of tokens of all tones}} \times 100\%$$

The frequency would be applied five times for calculating four lexical tones and a neutral tone. The denominator should be the total number of syllables regardless of meaningful or meaningless tokens. The numerators would be the number of tokens of a particular tone. The fraction then should be presented as a percentage. The frequencies of all tones could be ranked into an ordering, too. The ordering of tone frequency could be used to compare to those in tone emergence and in accuracy rate.

Regarding the accuracy rate, it could provide a percentage to show the degree of stabilization. As mentioned above, the coding of each utterance included the word meaning and target tones. The purpose of transcribing the target tones was to determine the correctness of tones. Therefore, in the measurement of accuracy rate, we only included tokens with clear meanings. To judge the correctness of tones, we tended to use care-takers' target tones as the criterion. If children produce identical tones with their care-takers', we would count it as correct tones. For example, a mother used [pej55 pej55] to indicate *'a cup,'* and if the child imitates the tones [55-55] correctly, it would be counted as a correct token; if the child used [pej51 pej51] or other tone patterns which are not identical to the mom, then the tones would be classified into incorrect tokens.

The tone accuracy rate will be calculated by the formula which had been applied by many researchers (see Shriberg & Kwiatkowski, 1982; Sheiberg, et al., 1997)

$$\text{Accuracy rate} = \frac{\text{the number of correct tokens of a tone}}{\text{the number of targeted tokens of a tone}} \times 100\%$$

When calculating the accuracy rate, the denominator should be the number of targeted tone that the children intended to produce. And the numerator would be the number of correct tones which are determined both by children's referential meanings and care-takers' target tones. For example, given that the reduplication form of *'a pen'* the children learned is [pi35 pi21], when the child uttered [pi55 pi21], the tone of the first syllable would be considered a tone error. Hua & Dodd (1995) provided a criterion that a tone was viewed stable when two-third (66.7%) of the tones were produced correctly. With this criterion, we could precisely examine whether a tone is acquired or not by checking the accuracy rate. Zhu (2002) applied the 66.7% criterion of stabilization from Hua & Dodd (1995), and also applied a 90% criterion of stabilization to measure the advanced level of tone stabilization.

3.2.4 Substitution pattern in tonal errors

Meaningful tokens which were determined to be a tonal error would be further analyzed in this measure. Unstable and immature tones would sometimes be changed into other tones, and would be determined as a tonal error. Although the accuracy rate could pretty much depict the states for children's tonal development, it could not tell us how children deal with immature tones. Therefore, the substitution pattern was used to demonstrate which tones were more likely to be realized in replacing the error tones.

The substitution pattern in tonal errors would be presented in a matrix that the row represents the target tones, and the column represents the actual realized tones. The example of the matrix is presented in Table 3.3.

Table 3.3 A sample matrix of substitution pattern in tonal errors

Target tone / **Realized tone**	**[55]**	**[35]**	**[21]**	**[51]**	**Total**	
[55]		24	19	17	60	50.8%
[35]	3		13	10	26	22%
[21]	4	7		11	22	18.6%
[51]	0	5	5		10	8%
Total	7	36	37	38	118	
	5.9%	30.5%	31.3%	32.2%		

The matrix could present the number of tokens and the frequency of substitution in each tone. If a tone should be produced in [55] but is realized as [35], it will be put into the first column and the second row. The percentages on the right column would tell us which tone is more frequently used in replacing others in tonal errors, and the percentages on the bottom row tell us which tone tends to make more tonal errors. In this sample in Table 3.3, there are 3 tonal error tokens that replace [55] to [35], and there are totally 26 tonal errors that are realized into [35], accounting for 22 % of the total errors. There are only 7 tonal errors whose target tones are [55], accounting for 5.9% of the total errors. With this matrix, the substitution pattern in tonal errors could be displayed clearly.

Chapter 4

Results and Analysis

This section will present the results and general analysis of the data, and the analysis will follow the method introduced in chapter 3. Section 4.1 will be the overview of the overall data including monosyllabic and disyllabic tokens. The results regarding the age of tone emergence will be presented in 4.1.1, and the tone frequency and accuracy rate will be shown in graphs in 4.1.2. In section 4.2, to obtain more specific results, the monosyllabic and disyllabic tokens will be analyzed separately. The monosyllabic tokens will be analyzed in 4.2.1, and the disyllabic tokens will be analyzed in 4.2.2.

We found that the unusual high frequencies of [21] in the first syllable and [35] in the second syllable in disyllabic tokens might result from the tone combination [21-35] related to the reduplication of motherese, so the results of reduplication in motherese will be displayed particularly in 4.2.3. After we determined to exclude tokens produced in [21-35], the reanalysis of the modified disyllabic data will be shown in 4.2.4, and the reanalysis of the modified data combining both monosyllabic and disyllabic tokens will be put in 4.2.5. Then in section 4.3, the substitution patterns of tonal errors will be illustrated. Last but not least, the age of tone emergence and stabilization will be summarized in section 4.4.

4.1 Overall analysis

As mentioned in chapter 3, there were 16 children enrolled under Professor Wan's NSC project, but only 6 children fitted in this study. Among the 97 hours of observational data collected from the 6 subjects, only 44 hours of the recordings were adopted. Table 4.1 summarized the subject information including the subject number, gender, the age range during the observation, and the total number of tokens uttered by each subject.

Table 4.1 Subject information

Subject	Gender	Age range	Duration	Total number of tokens
#1	M	1;1-1;6	6 months	486
#2	F	1;0-1;6	7 months	181
#3	F	1;0-1;6	7 months	637
#4	M	0;11-1;6	8 months	585
#5	F	0;10-1;5	8 months	86
#6	M	0;10-1;5	8 months	87
			Total	**2062**

Three male and three female children were adopted in this study. The observation started from the age of 0;10 to 1;1 in the beginning and ended at the age of 1;5 to 1;6. The children enrolled were all between the babbling stage to one-word stage and their conceptual and lexical abilities were still under construction, so it was common for children to produce vocalization without recognizable meanings. However, the development of prosody has already started at the babbling stage, so it was applicable to collect tonal acquisition data from this stage. The observation lasted for 8 months, and there were originally 2286 tokens. Yet, the speech tokens which were spoken in Taiwanese or English, such as [ne55-ne1] *'milk'* and [ˈpaj-paj] *'bye-bye'* were not supposed to include in this Mandarin tonal study. Plus, the utterances which were more than two syllables or utterances whose tones were not clear would also be inappropriate to this topic. After excluding these tokens, there were a total number of 2062 tokens which could be analyzed in this study. Children's production had individual differences. As Table 4.1 shows, the most productive children were subject #1, #3, and #4 who uttered more than 400 tokens, while subject #5 and #6 produced the least number of tokens, which were under 100 tokens.

There were monosyllabic and disyllabic tokens in the data. Under these two big categories, there were two subgroups classified the tokens with meaning and tokens without meaning. The following table summarized the total number of tokens in different subcategories.

Table 4.2 Number of tokens in subcategories

Subject	**Monosyllabic tokens**		**Disyllabic tokens**		**Total**
	without meaning	**with meaning**	**without meaning**	**with meaning**	
#1	20	66	70	330	**486**
#2	96	40	27	18	**181**
#3	48	157	87	345	**637**
#4	65	25	204	291	**585**
#5	11	17	19	39	**86**
#6	42	1	43	1	**87**
Total	**282**	**306**	**450**	**1024**	**2062**

There were several findings derived from Table 4.2. First, it shows that most of the subjects, including #1, #3, #4, and #5, preferred producing disyllabic tokens than monosyllabic tokens. The disyllabic utterances outnumbered monosyllabic ones no matter in tokens with meanings or without meanings. There were 1474 disyllabic tokens, but only 588 in monosyllabic tokens. Second, tokens with clear meanings were more than those without clear meanings. The number of meaningful tokens was 1330, but that of the meaningless tokens was only 732. Third, the subjects who were more productive tended to produce more meaningful words. Subject #1, #3 and #4 were productive children that their numbers of tokens were much bigger than those of the other three subjects, and they all produced higher numbers of meaningful tokens than meaningless tokens. For example, subject #1 produced totally 396 meaningful tokens, but there were only 90 meaningless tokens among all his utterances. Yet, there were exceptions, such as subject #5 who produced the least number of tokens but had more meaningful tokens. The exception would be accounted for individual differences or sampling fluctuations.

4.1.1 Tone emergence ordering

Many researchers were curious about the ordering of tone emergence. The best timing for collecting this ordering was from the very beginning of children's one-word stage. In children's one-word stage, the referential meaning of the utterances became more clear, and children have learned to refer to a particular object by using a particular form. By tracing the meaningful words that children uttered, we were able to know which tone appeared first, and what was the ordering of tone emergence in early lexical development.

Table 4.3 Age of tone emergence

	[55]	**[35]**	**[21]**	**[51]**	**Neut**
#1	1;1	1;1	1;1	1;1	1;1
#2	1;0	1;3	1;3	1;1	1;6
#3	1;3	1;2	1;2	1;0	1;1
#4	1;0	1;0	1;0	1;1	1;1
#5	0;11	1;2	1;2	1;1	1;2
#6	1;1	NA	NA	NA	NA

Table 4.3 showed the age-tracked data of tone emergence in six individual subjects. The criterion for determining whether a tone had emerged was to see if the tone was produced more than once by children (Vihman, 1996). Noted that the data collection of subject #1 started from the age of 1;1. By that time, all his tones had already emerged, so the exact age for each tone's emergence was hard to detect from his data. The only thing we could make sure was that all his tones have emerged by the age of 1;1. Moreover, subject #6's lexical development was still before one-word stage that most his utterances were meaningless by the age of 1;6, so the analysis of tone emergence ordering would not be applicable in his utterances.

Most children uttered the high-level tone [55] in their first words, including subject #2, #4, #5, and #6 between the age from 0;11 to 1;0. Only subject #3's [55] appeared the last and later than other tones. Later on, at the age of 1;1, falling tone [51] had emerged in subject #2, #3, #4, and #5. Regarding the emergence of [35] and [21], these two tones seemed to emerged together after other lexical tones. Among all the subjects, except for subject #4 whose rising tone [35] and low-level [21] appeared together with [55] at 1;0, other subjects' [35] and [21] developed later than [55] and [51] at the age between 1;2 to 1;3. As for the neutral tone, it appeared as early as other citation tones between 1;1 to 1;2 in most subjects, but subject #2 did not develop the tone until 1;6. The differences of the ordering might mostly accounted for individual differences. However, it should be noted that the data collection method had flaws that from children's spontaneous speech, the opportunity for children to utter each tone was not equal, so there might be sampling fluctuation. For instance, a child may have already learned the rising tone [35] in 1;1, but we missed collecting her [35] at 1;1. Thus, the age of emergence of [35] would then be less accurate due to the sampling fluctuation.

Although the individual differences and the sampling fluctuation caused it hard to make conclusion, the overview of tone emergence still showed some similarities among subjects. First, [55] and [51] seemed to emerge earlier than [35] and [21] at their first-word stage. Table 4.3 indicated that both subject #2 and #5 uttered high-level tone [55] first and falling tone [51] the second by 1;1, and their [35] and [21] were appeared later by 1;2 and 1;3. Subject #3's [35] and [21] also appeared later than [51] by 1;2, but her [55] appeared the last at 1;3. It may result from sampling fluctuation again because her [55] also reached 90% of stabilization at 1;3. The data of stabilization will be presented later in the last section in this chapter. In subject #4's data, the four contrastive tones appeared roughly at the same time, and the data in subject #6 could only show that [55] was the first appeared tone in his tonal development.

Second, once children have reached the one-word stage, the four contrastive tones would emerge one after another in a short period of time. Subjects #2, #3, #4, and #5 spent about one to three months to contrast all the citation tones, but it took more than five months for subject #6 to apply all tones in his lexicon. Third, the neutral tone tended to appeared last comparing to the four citation tones. Data showed that the neutral tone was the last emerged tone in subject #2, #4, and #5. Only subject #3 was the exception that her neutral tone was the second appeared tone which showed up later than [51].

4.1.2 Frequency and Accuracy rate

In the previous section, I presented the ordering and ages of tone emergence, and found that the high-level tone [55] and falling tone [51] appeared earlier than [35] and [21], and the neutral tone tended to appear last. Thus, in order to see whether the frequency and accuracy rate of tones would rank in the same sequence as the tone emergence, the number of occurrences and the correctness of each tone would be computed in the section. The tone frequency would inform which tone is more frequently used and which tone is less used. In addition to tone frequency, the accuracy rate is also crucial for determining the degree of development. A tone with higher accuracy rate reflects a more mature and more stable degree of acquisition. By analyzing the tone frequency and accuracy rate, the detail of tone acquisition could be better described.

Among the 2062 tokens, there were 588 monosyllabic tokens and 1474 disyllabic tokens, and each of the disyllabic tokens had 2 syllables, so the 1474 disyllabic tokens would have 2948 syllables. Thus, there were totally 3536 syllables (588+1474x2=3536) in the data. As mentioned in chapter 3, the tokens without meaning could also apply frequency analysis. To see which tone was used more frequent, the 3536 syllables would all be included. The four tones were written in the tone values modified from Chao's (1968) tone number, and the neutral tone was also included in the tone frequency chart with the label of 'Neut.'

From the total of 3536 syllable tokens, the frequencies of all tones were calculated below and the number of tokens and frequencies were presented in Table 4.4 and Figure 4.1.

Table 4.4 Number of tokens and frequencies of tones in all syllabic tokens

	[55]	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	960	954	894	565	163	3536
Frequencies of tones	27.1%	27.0 %	25.3%	16.0%	4.6%	

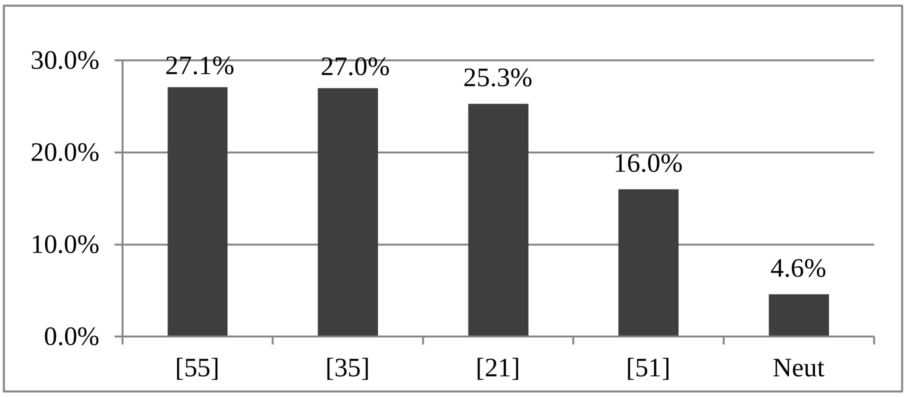


Figure 4.1 Frequencies of tones in all syllabic tokens

In Figure 4.1, the bar graph presented the frequency of each tone. The difference of the five percentages was statistically significant (χ^2=18.2, p<.001). The highest three bars, [55], [35], and [21] were all higher than 25%. Falling tone [51] ranked as the fourth place with 16% of appearance, and neutral tone appeared last frequently, accounting for 4.6%. Actually, it was unfair to compare the neutral tone with other tones in frequency, because in Mandarin, neutral tones only appear in weak stress that are usually in utterance-final position. If children acquire this phonological rule early in this stage, the neutral tone would scarcely be found in utterance-initial positions in monosyllabic and disyllabic tokens. In fact, Figure 4.1 showed that children preferred using [55], [35], and [21] the most, and the neutral tone appeared last frequently. The hypothesis that children learned the phonological rule of neutral tone in this stage might be approved. Then, to see whether the acquisition of tones could demonstrate the rule that ‘practice makes perfect,’ we need to compare the ranking with accuracy rate below.

The accuracy rate was calculated by taking meaningful tokens. Meaningful tokens had clear targets which encoded with target tones, so the correctness could then be decided. Among the 3536 syllables, only 2354 tones were syllables that had target tones, so the 2354 syllables were included in the accuracy rate analysis. The accuracy rates were calculated by dividing the number of correct tokens by the number of targeted tones. Table 4.5 presented the fraction of the accuracy rates in each tone, and the percentage of accuracy rates were shown in Figure 4.2.

Table 4.5 Number of tokens and accuracy rates of tones in all syllabic tokens

	[55]	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	468/529	586/642	598/703	298/356	91/111
Accuracy rates of tones	88.5%	90.3 %	85.1%	83.7%	82.0%

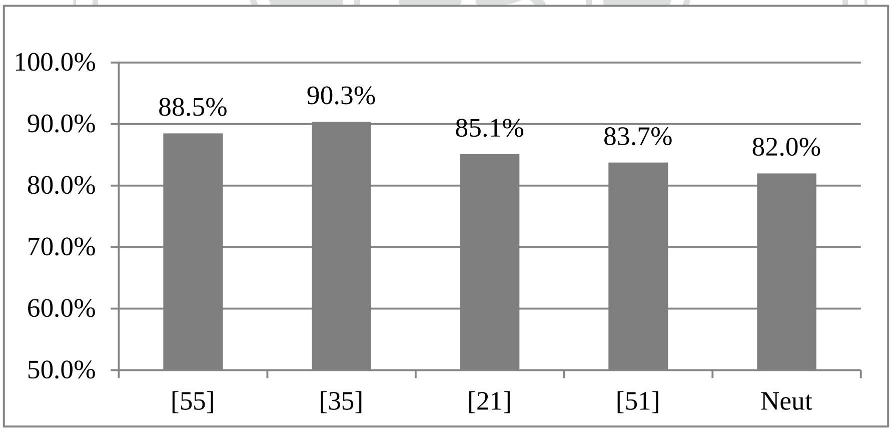


Figure 4.2 Accuracy rates of tones in all syllabic tokens

The result suggested that all the lexical tones including neutral tone showed high percentages in accuracy rate. The most stable tone was [35] that reached 90% of accuracy rate. The high-level tone [55] showed lower rate than [35], accounting for 88.5%. The third and fourth places were [21] and [51], and the neutral tone was the last. However, the numbers of accuracy rates in all tones were all high, and there was no significant differences found among the five percentages (χ^2=5.229, p=0.265). If the accuracy rates were not different statistically, we could only say that the degree of tonal development was similar during the age between 0;10 to 1;6.

Now that the accuracy rate in the overall data could not indicate which tone was more mature and which was not, the following analyses would separate the data into monosyllabic and disyllabic tokens and see whether there would be more specific results in the ordering of tonal acquisition.

4.2 Subgroup analyses in monosyllabic and disyllabic tokens

It is mentioned that the tokens collected in this study included monosyllables and disyllables. To see whether different types of tokens would affect the development in tonal acquisition, the two types of subgroups would be analyzed separately. During the observation, we found that younger children could only produce short utterances. There were few tokens that were longer than two syllables, so the tokens collected here were all one-to-two syllabic utterances. As mentioned in section 4.1, the number of tokens in disyllables was much higher than that in monosyllables. The table below provided more information about the number of tokens in different syllable positions in disyllabic tokens.

Table 4.6 Number of tokens of each syllable in monosyllabic and disyllabic tokens

Tone	Monosyllable	Disyllable 1	Disyllable 2	Total
[55]	162	445	353	960
[35]	127	238	589	954
[21]	127	614	153	894
[51]	171	176	218	565
Neut	1[1]	1	161	164
Total	588	1474	1474	3536

Table 4.6 listed the number of tokens in each tone that showed up in three different positions, including monosyllable, and the first and second syllable in disyllabic tokens. There was a significant position effect found in number of tokens among tones (χ^2=735, p<.001), which means that children tended to produced different number of tokens in different positions. Take the rising tone [35] for example, it appeared 127 times in monosyllabic tokens, but appeared 238 times in the first syllable position in disyllabic tokens, and the number increased dramatically in the second syllable of disyllabic tokens. The circumstances were not significantly found in [35] and [21], and it would be discussed later in the further exploration.

4.2.1 Monosyllabic tokens

The number of tokens and frequencies of monosyllabic tokens were presented in Table 4.7. The bar graph in Figure 4.3 makes it easier to compare the differences among frequencies of tones.

[1] The neutral tone in Mandarin is not supposed to be in the utterance-initial position. The neutral tone here was produced by children imitating the utterance-final affix.

Table 4.7 Number of tokens and frequencies of tones in monosyllabic tokens

	[55]	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	162	127	127	171	1	588
Frequencies of tones	27.6%	21.6 %	21.6%	29.1%	0.2%	

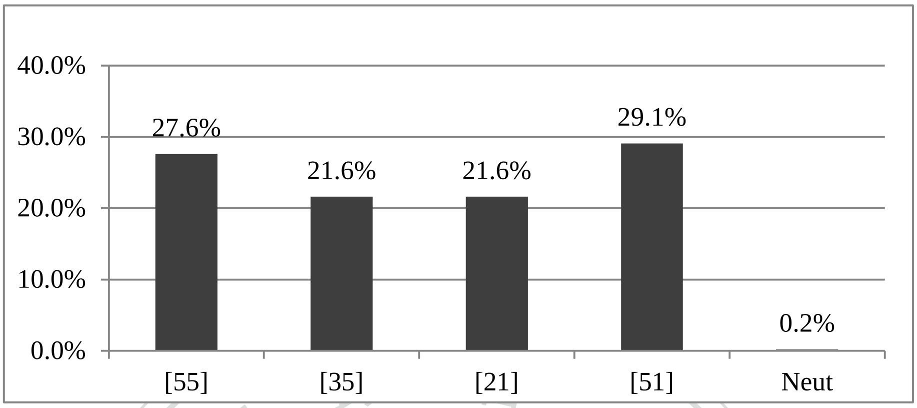


Figure 4.3 Frequencies of tones in monosyllabic tokens

Different from the frequencies presented in the overall data in Figure 4.1, the ranking of tone frequencies in monosyllabic tokens was [51]> [55]> [35], [21]> N, and their percentages were proved to have significant difference (χ^2=26.9, p<.001). The falling tone [51] accounted for 29.1% of the total number of monosyllabic use, and high-level tone [55] accounted for 27.6%. The other two lexical tones [35] and [21] reached the same percentage, 21.6%, which appeared less frequent than [55] and [51]. The neutral tone had only 0.2% of occurrences, because the phonological constraint in Mandarin does not allow it to appear in the utterance-initial position. Compare to Figure 4.1, the ordering of frequencies in Figure 4.3 could better correspond to the sequence of tone emergence that the [55] and [51] appeared earlier and more frequently than [35] and [21]. It seemed that the earlier appeared tones were produced more frequently in monosyllabic tokens.

With regard to the accuracy rates in monosyllabic tokens, the results here were also different from that of all syllabic data.

Table 4.8 Number of tokens and accuracy rates of tones in monosyllabic tokens

	[55]	**[35]**	**[21]**	**[51]**
Number of correct tokens / number of targeted tones	61/72	42/44	84/119	61/71
Accuracy rates of tones	84.7%	95.5 %	70.6%	85.9%

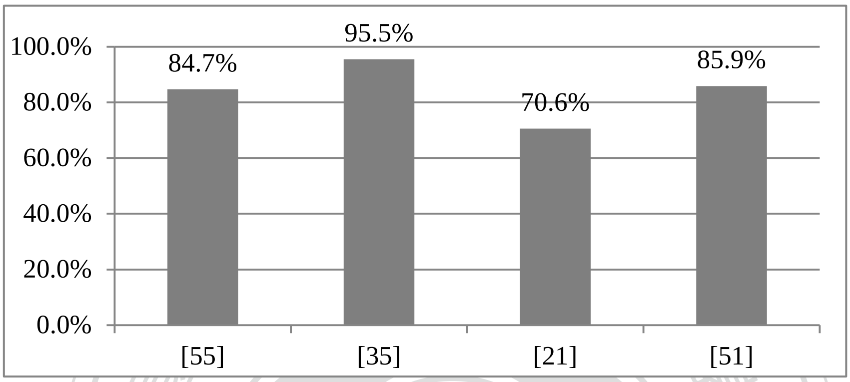


Figure 4.4 Accuracy rates of tones in monosyllabic tokens

The tone which obtained the highest accuracy rate was [35] that accounted for 95.5%, followed by [51] and [55] whose accuracy rates were 85.9% and 84.7%. The most immature tone in monosyllabic tokens was [21] whose degree of stabilization was at 70.6%. The accuracy rate in neutral tone was not applicable here in the monosyllabic tokens, because in Mandarin the neutral tone is only allowed to appear in the utterance-final position. The monosyllables were all in the utterance-initial positions, so it would be impossible for a neutral tone to produce correctly in the monosyllabic speech tokens. Therefore, the ranking of the accuracy rate of monosyllabic tokens was [35]> [51]> [55]> [21]. The accuracy rates showed similar ranking with tone emergence that [55] and [51] was more stable and appeared earlier than [21] except for [35]. It was more questionable why the later appeared tone [35] gained the higher rate in accuracy. According to the raw data, it was showed that children uttered only few [35] in monosyllables, but once they uttered [35], they produced it with high correctness. More interestingly, the monosyllables children uttered were sometimes the second syllable of the disyllabic tokens, such as [je35] in [je21-je35] 'grandpa' and [ma35] in [ma21-ma35] 'mother.' These examples though were part of the reduplications we still viewed them as correct tokens. Thus, the high accuracy rate of [35] might result from this situation.

4.2.2 Disyllabic tokens

When the monosyllabic tokens were analyzed separately in tone frequency and accuracy rate, we found that the ranking was different from that in the overall data. As for disyllabic tokens, would the results of frequencies and accuracy rates be similar to the results presented in overall data in section 4.1.2 or to the results in monosyllable data in section 4.2.1? In the following, the frequencies and accuracy rates of disyllabic tokens would be shown in separate syllable positions.

In table and graph below, disyllable 1 stood for the first syllable in disyllabic tokens, and disyllable 2 stood for the second syllable. The total number of tokens in disyllabic tokens, which was shown in Table 4.3, was 1474, so there were 1474 syllables in the first syllable position and also 1474 syllables in the second syllable position. The tone frequencies were computed and presented separately in two different positions.

Table 4.9 Number of tokens and frequencies of tones in disyllabic words at separate syllable positions

Disyllable 1	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	445	238	614	176	1	1474
Frequencies of tones	30.2%	16.1 %	41.7%	11.9%	0.1%	
Disyllable 2	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	353	589	153	218	161	1474
Frequencies of tones	23.9%	40.0 %	10.4%	14.8%	10.9%	

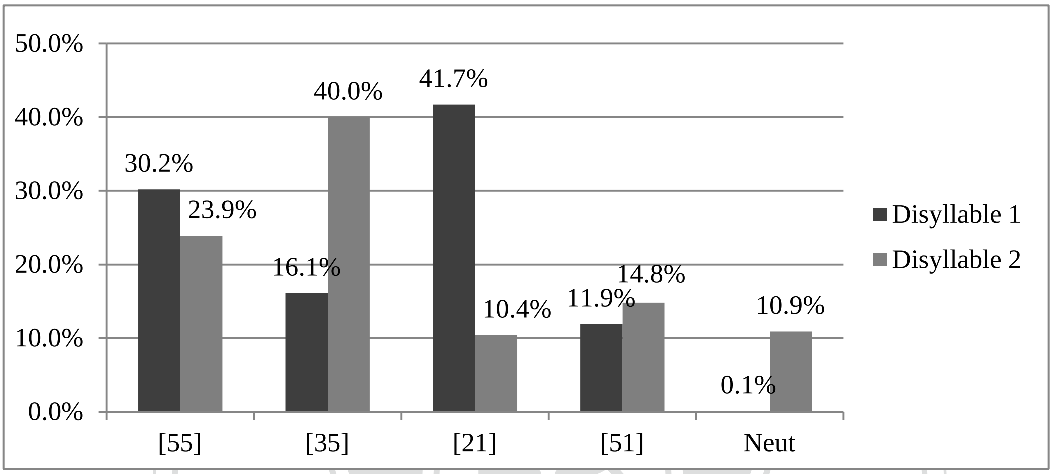


Figure 4.5 Frequencies of tones in disyllabic words at separate syllable positions

The gaps between first syllable and second syllable in [35] and [21] were huge that the frequency of [35] in the first syllable was 16.1% but it reached 40% in the second syllable; the similar situation also happened in [21], but the higher frequency was at the first syllable. That is to say, the first syllable in [21] and the second syllable in [35] were used a lot in children's disyllabic speech tokens. Other than [35] and [21], the percentages of accuracy rates were closer between the two syllable positions in high-level tones [55] and falling tone [51]. With regard to the neutral tone, it only occurred once in disyllabic 1 and accounted for only 0.1% of frequency, but it appeared 161 times in disyllable 2 which reached 10.9%. The distinction might indicate that children may have acquired the utterance-final constraint already.

The next table and graph showed the accuracy rates of the first and second syllable positions in disyllabic tokens. The results were computed individually in each tone in two syllable positions. The number of correct tokens and number of targeted tones were presented in fractions and the percentages of accuracy were also shown in the following.

Table 4.10 Number of tokens and accuracy rates of tones in disyllabic words at separate syllable positions

Disyllable 1	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	207/239	114/144	434/488	111/140	
Accuracy rates of tones	86.6%	79.2 %	88.9%	79.3%	
Disyllable 2	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	200/218	424/454	80/96	126/145	91/111
Accuracy rates of tones	91.7%	93.4 %	88.9%	79.3%	73.3%

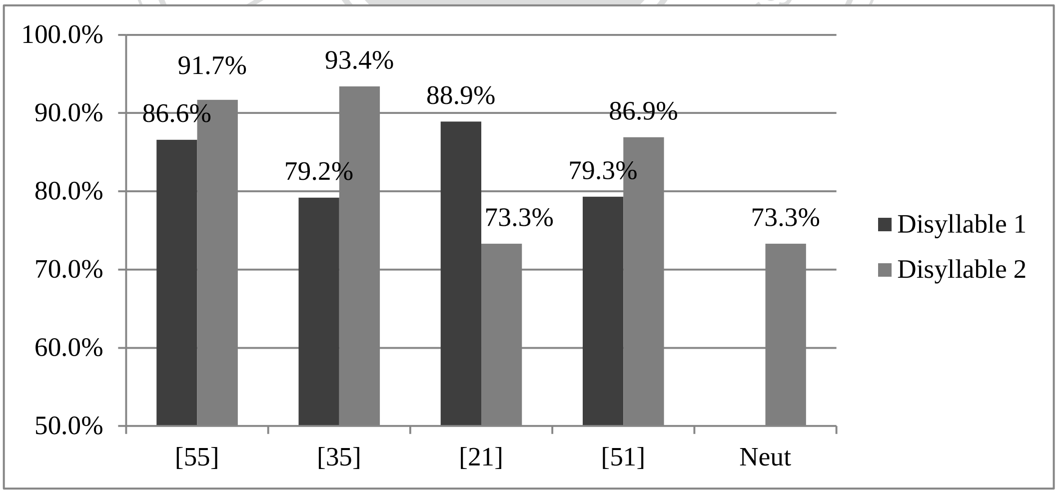


Figure 4.6 Accuracy rates of disyllabic words in separate syllable positions

First of all, in Figure 4.6, the accuracy rate in neutral tone in disyllable 1 was not applicable because of the phonological constraint, so only utterance-final position could be applied in accuracy analysis. Then, we could find that the accuracy rates in disyllable 2 were all higher than disyllable 1 except for [21] whose disyllable 1 was higher than disyllable 2. Moreover, the gaps between the two syllables in [35] and [21] were bigger than the gaps in [55] and [51]. The gaps between disyllable 1 and disyllable 2 in [35] and [21] were 14.2% and 15.6%, but the gaps in [55] and [51] were only 5.1% and 7.6%. The [21] in disyllable 1 and the [35] in disyllable 2 were also found extremely high percentages in frequency analysis in Figure 4.5. The reason for these two findings may both be caused by the combination [21-35] in reduplications; however, we could not make conclusion from these evidences, and there should be more examination in exploring this particular tone combination [21-35] in disyllabic tokens.

4.2.3 The tone combination [21-35] in the reduplication of motherese

Actually, the two tones [21] and [35] were not appeared separately in different tokens, but were emerged together in the same disyllabic tokens and formed a particular tonal combination [21-35]. In our observation, it was found that when care-takers were talking to children, they tended to transfer monosyllabic words into reduplicative forms which were in disyllables, and the most preferred tone combination in motherese was [21-35], such as [njow21-njow35] *'a cow'* and [y21-y35] *'a fish.'* Meanwhile, we found that children also produced more reduplications than non-reduplicatoins in disyllabic tokens. Among all disyllabic tokens, the reduplicative tokens (710 tokens) were twice more than non-reduplicative forms (314 tokens), and there were 513 reduplicative tokens which were produced in [21-35] combination. It seemed that children's production was influenced by the input from adults to some extent.

There were some examples that showed the tone combination applied in most adults' reduplications in motherese.

Table 4.11 Tone combinations of motherese in common nouns and kinship terms

Common nouns	Formal form	Reduplicated form	Gloss
[55] → [55-55]	tʂ ʰɤ 55	tʂ ʰɤ 55 tʂ ʰɤ 55	'a car'
[35] → [21-35]	y35	y21 y35	'a fish'
[21] → [21-35]	ʂ wej21	ʂ wej21 ʂ wej35	'water'
[51] → [51-51]	mɑ w51	mɑ w51 mɑ w51	'a hat'
Kinship terms	**Formal form**	**Reduplicated form**	**Gloss**
[55] → [21-35]	ku55 ku	ku21 ku35	'aunty'
[35] → [21-35]	je35 je	je21 je35	'grandpa'
[21] → [21-35]	naj21 naj	naj21 naj35	'grandma'
[51] → [21-35]	pa51 pa	pa21 pa35	'daddy'

In the reduplicated forms in common nouns, the words that were originally pronounced in [55] and [51] would not be changed into other tones. For example, the first common noun [tʂ ʰɤ 55] in the formal form was produced in [55], and it is also produced in [55] in the reduplicated form [tʂ ʰɤ 55-tʂ ʰɤ 55]. While monosyllabic words in [35] and [21] such as [yu35] *'a fish'* and [ʂ wej21] *'water'* would not produced in [yu35-yu35] or [ʂ wej35-ʂ wej21] [2] in reduplications in motherese, they would apply the tone combination [21-35] instead. When it comes to kinship terms, the combination [21-35] would be applied overwhelmingly in all circumstances. The lower rows in Table 4.5

[2] In Mandarin tone sandhi, if the low-level tone is followed by another low-level tone, the first low-level tone would become a rising tone.

provided four examples of kinship terms in their formal forms and reduplicated forms in motherese. It indicated that no matter what the original tones were in formal forms, the tone would be changed into [21-35]. For example, [ku55-ku] ‘aunty’ would be changed to [ku21-ku35] and [naj21-naj] ‘grandma’ would also be changed to [naj21-naj35]. Although there were variations that some care-takers also applied [21-55] in motherese, the most frequent combination was still [21-35].

After examining how adults used reduplications in motherese, we have to analyze the reduplicative forms in children’s productions. From children’s utterances of monosyllabic and disyllabic words, it was interesting to find that children preferred using reduplicative forms than monosyllabic words and non-reduplicated disyllables. In addition, children also like to adopt the tone combination [21-35] in most of their reduplications. For example, subject #4 was a boy who loved to play plastic balls very much. His total productions of *‘a ball’* were 95 times, and 94 times of the words were produced in reduplicative forms. Among the 94 reduplications of *‘a ball,’* there were 80 tokens realized in [21-35] tone combination and 3 tokens in [21-55] combination. We noted that subject #4’s mom also used [tɕjow21-tɕjow55] to refer to balls whose tone combination was [21-55], so balls which were produced in [21-55] were also counted as the correct tones. The example suggested that Mandarin-speaking children preferred reduplicating monosyllabic words and non-reduplicated disyllables, and [21-35] was the most frequently used tone combination in reduplications. The [21-35] tone combination was applied frequently by both adults and children. It seemed that children's [21-35] tokens were influenced by motherese. However, we could not determine whether children's production of [21-35] was directly from imitating adults' reduplicative forms in motherese. The effect between children's input and output of language needs further research in the future.

To see how greatly children applied [21-35] combination, we categorized all disyllabic word into 20 tonal combinations and found that the combination [21-35] mentioned previously showed extremely high number of tokens in Figure 4.7.

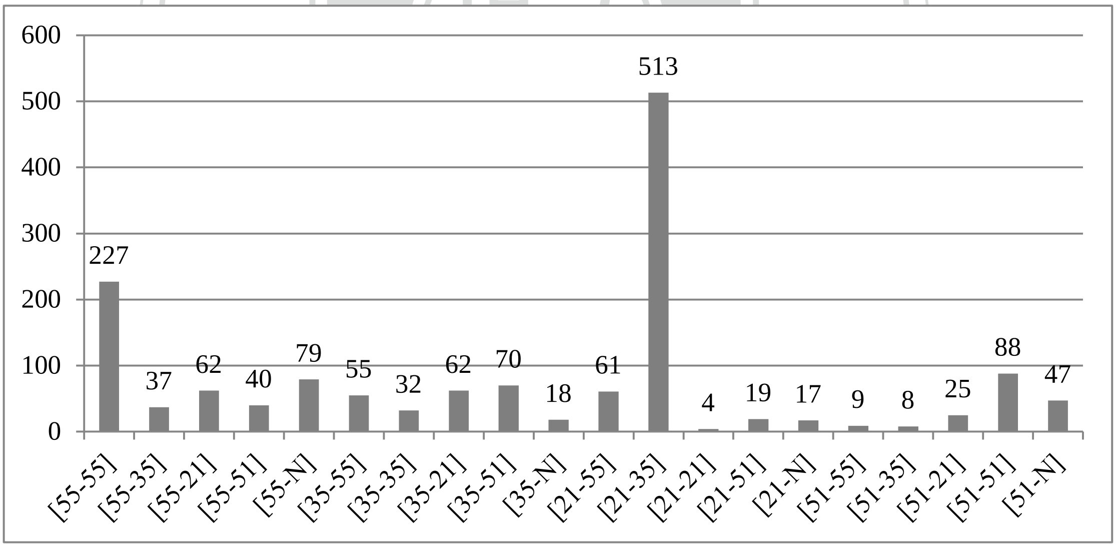


Figure 4.7 Number of tokens in all combinations in disyllabic tokens

Apparently, [21-35] had been produced 513 times and stood for the most frequent combination in disyllabic words, and its accuracy rate reached as high as 98.4%. If the overall speech tokens were not analyzed in separate syllable positions, the high frequent use of [21-35] would not be found and the results of children's tonal acquisition ordering would concluded inaccurately especially in [21] and [35]. Now that we have noticed that the high frequency and accuracy rate of [21] and [35] were caused by the particular combination [21-35] in motherese, we should do some adjustment in the data.

Statistically, the combination [21-35] was three times more than the standard deviation and was considered to be an outlier (Average=73.65, SD=114.13, Outlier>416.04.) This particular combination was incompatible with other tone combinations. Other than [21-35], the low-level tone [21] and the rising tone [35] showed up much less frequently in other combinations individually. If the [21-35] combination was included, the result may be greatly biased that the tone frequency and accuracy rate of [35] and [21] would be too high. Therefore, all disyllabic tokens which were produced in [21-35] combination should be excluded and the results including frequency and accuracy rate would be computed again in the following sections.

4.2.4 Reanalysis in disyllabic tokens

In this section, the problem of [21-35] combination mentioned in the previous paragraph had already been adjusted. That is, the outlier [21-35] was excluded, and the tone frequencies and accuracy rates were calculated again.

Table 4.12 Modified number of tokens and frequencies of tones in disyllabic words at separate syllable positions

Disyllable 1	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	445	238	101	176	1	961
Frequencies of tones	46.3%	24.8 %	10.5%	18.3%	0.1%	
Disyllable 2	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**	**Total**
Number of tokens	353	76	153	218	161	961
Frequencies of tones	36.7%	7.9 %	15.9%	22.7%	16.8%	

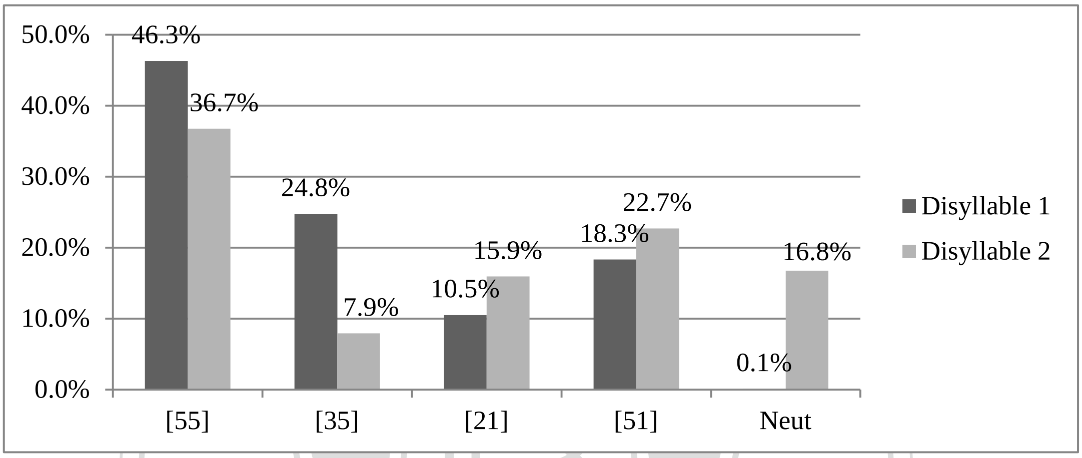


Figure 4.8 Modified tone frequency of disyllabic words in separate syllable positions

The Figure 4.8 excluded 513 tokens that were produced in [21-35] combination. Comparing the frequencies between Figure 4.5 and Figure 4.8, the new results in Figure 4.8 revealed that the percentage of [35] in the disyllable 2 have decreased a lot from 40% to 8%, and that of [21] in the disyllable 1 have also dropped from 41.7% to 10.7%. It seemed that there were few occurrences of [35] and [21] in combinations other than [21-35], and the previous high frequencies of [21] and [35] may be resulted from the mass application of [21-35] tone combination in motherese.

The accuracy rates were calculated again after excluding the tokens produced in [21-35] combination.

Table 4.13 Modified number of tokens and accuracy rates of tones in disyllabic words at separate syllable positions

Disyllable 1	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	207/227	113/141	49/99	111/149	
Accuracy rates of tones	91.2%	80.1 %	49.5%	74.5%	
Disyllable 2	**[55]**	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	196/205	38/65	79/89	126/144	91/111
Accuracy rates of tones	95.6%	58.5 %	68.3%	87.5%	73.3%

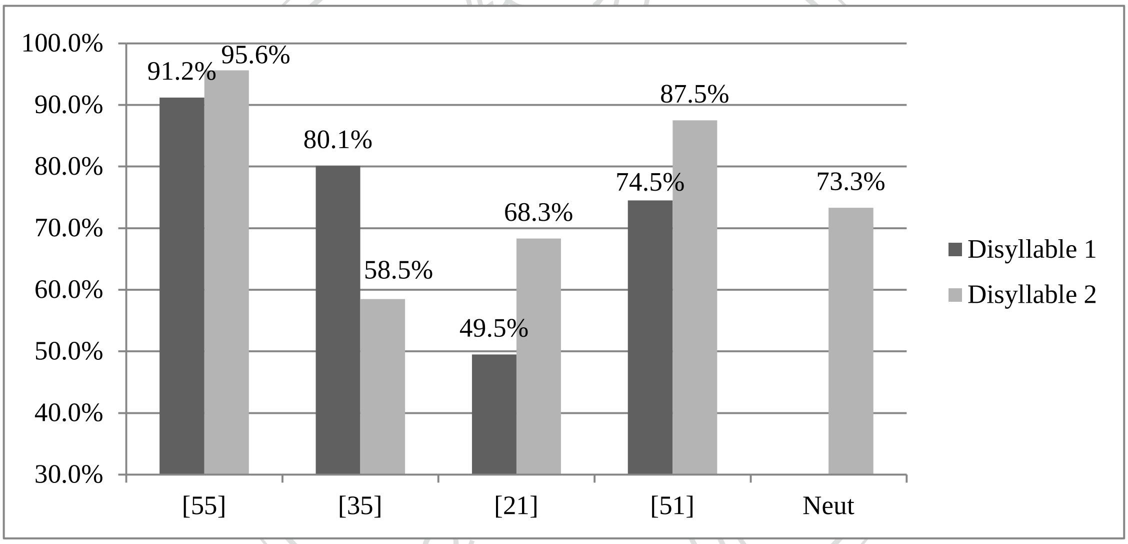


Figure 4.9 Modified accuracy rates of disyllabic words in separate syllable positions

The modified accuracy rates in Figure 4.9 showed obvious reduction of [35] in disyllable 2 and [21] in disyllable 1. Comparing the percentage between Figure 4.6 and Figure 4.9, the accuracy rate of [35] in disyllable 2 dropped from 93.4% to 58.5% and that of [21] also fell from 88.9% to 49.5%. It suggested that the stabilization of [35] and [21] in other combinations, such as [55-35] or [21-51] were not as mature as that in [21-35] combination, and the previous high accuracy rates in [21] and [35] would be resulted from the mass application of [21-35] tone combination in motherese. Other than [21] and [35], the accuracy rates between different syllable positions in high-level tones [55] and falling tones [51] were similar. The neutral tone, as mentioned before, was limited to the phonological constraints that it is illegal to appear in the utterance-initial position, so its accuracy rate in disyllable 1 was inapplicable.

To sum up, the results between section 4.2.2 and 4.2.3 demonstrated that the frequencies and accuracy rates dropped a lot in the first syllable position of low-level tones [21] and the second syllable position of rising tones [35] after excluding the tone combination [21-35]. It indicated that the occurrences and stabilization of [21] and [35] separately in other syllable positions were not as high as those in [21-35] combination. Therefore, we could conclude that [21] and [35] were acquired as a phonetic whole in [21-35] combination but were not acquired individually in monosyllables or other tone combinations in disyllables.

4.2.5 Reanalysis in overall data

After excluding the [21-35] combination in disyllabic tokens, the results of frequencies and stabilization in [21] and [35] became more reasonable. In the precious section, the results were presented separately in different syllable positions. In order to compare the ranking of frequency and accuracy rate among all tones, the syllables in monosyllabic and disyllabic tokens should be combined together. The modified overall data here contained all monosyllable and disyllabic tokens except for [21-35] combination.

With regard to the modified data in frequency analysis, the results were different from the original overall data in Figure 4.1.

Table 4.14 Modified number of tokens and frequencies of tones in all syllabic tokens

	[55]	[35]	[21]	[51]	Neut	Total
Number of tokens	960	441	381	565	163	2510
Frequencies of tones	38.2%	17.6 %	15.2%	22.5%	6.5%	

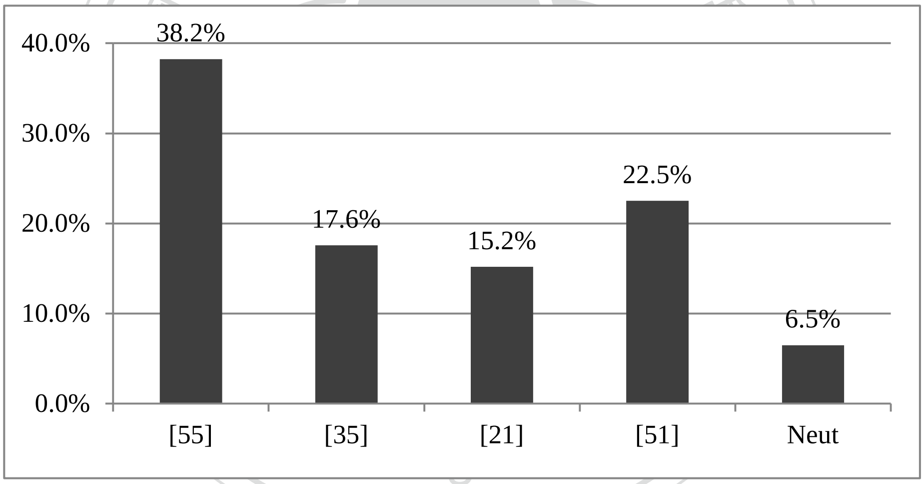


Figure 4.10 Modified frequencies in overall data

Now that we excluded the tone [21-35] combination from the overall data, we could finally rank the frequencies among all Mandarin tones. Based on the significantly different results presented in Figure 4.3 (χ^2=27.9, p<.001), firstly, the high-level tone [55] was the most frequently used tone, accounting for 38.2% of the total 2510 syllables (3536-513x2=2510). Secondly, [51] were used less frequently, with the percentage of 22.5%. The third and fourth places were [35] and [21], which obtained 17.6% and 15.2%

respectively. The least used tone was neutral tone, accounting for 6.5%. The frequency ranking according to Figure 4.10 would be [55]> [51]> [35]> [21]> N.

The same modified data were applied to the calculation in accuracy rate. The accuracy rates below were arranged without separating different syllable positions after excluding [21-35] combination.

Table 4.15 Modified number of tokens and accuracy rates of tones in all syllabic tokens

	[55]	**[35]**	**[21]**	**[51]**	**Neut**
Number of correct tokens / number of targeted tones	464/504	193/250	212/307	298/364	91/111
Accuracy rates of tones	92.1%	77.2 %	69.1%	81.9%	73.3%

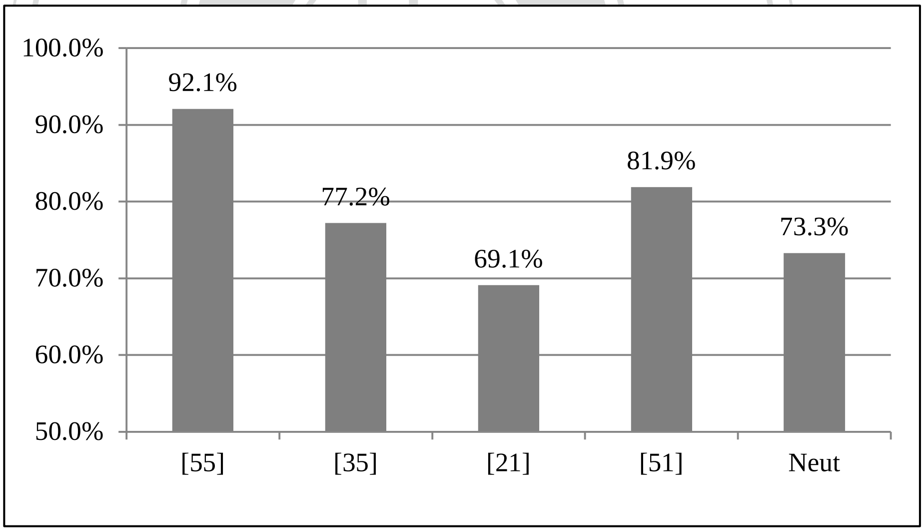


Figure 4.11 Modified accuracy rates in overall data

In the tonal accuracy rates in Figure 4.11, the differences of the five percentages was significantly different (χ^2=18.157, p<.001). The most frequent tone [55] also ranked as the most accurate tone, and it reached as high as 92.1% of correctness. The falling tone [51] showed lower accuracy rate than [55], which accounted for 81.9% of accuracy. The rising tone [35] and neutral tone were rather unstable, and their accuracy rates reached 77.2% and 73.3%. The low-level tone [21] had the lowest accuracy rate which was 69.1%. As a consequence, the Figure 4.11 showed that children produced [55] more stable than [51], [35] and neutral tone were less stable than [51], and [21] was the most immature tone among all tones. The ranking of the accuracy rates of the tones would be [55]> [51]> [35]> Neut> [21].

All in all, the ranking of the frequencies before excluding [21-35] was [55]> [35]> [21]> [51]> Neut, and became [55]> [51]> [35]> [21]> Neut after modification. After the adjustment, the ranking of tone emergence coincides with that in tone frequency. It revealed that the tone that appeared earlier was used more frequently by children. As for the ranking of accuracy rate, it was originally [35]> [55]> [21]> [51]> Neut, but it changed to [55]> [51]> [35]> Neut> [21] after excluding [21-35] combination. The unreasonable high accuracy of [35] in the first result switched to the third place at the reanalysis, and [21] became the least stable tone. The ranking of accuracy rate in reanalysis was similar to that in tone emergence and tone frequency except for the sequence of [21] and the neutral tone. However, the neutral tone was not compatible in frequency ranking, because there is positional constraint for it to appear only in utterance-final. Thus, we could hardly compare the development of neutral tone with other lexical tones in frequency analysis.

4.3 Substitution pattern in tonal errors

When a tone has not fully acquired yet, children would make consistent tone errors. If one tone is often replaced with another tone, the low accuracy rate would suggest that the tone has not reached stabilization (Dodd 1995). Tones that were different from the target tone would be counted as a tonal error. The tonal error pattern was also studied by many linguists, because it could reveal whether children would use a more unmarked tone to replace a more marked tone.

The substitution pattern in the following figure showed the target tones and the realized tones in a matrix. For instance, if a child wrongly produced [njow21-njow35] *'a cow'* as [njo55-njo35], the target tone in the first syllable would be [21]. [55] would be the selected tone that realized in the tonal error.

Table 4.16 The matrix of substitution patterns of tonal errors in disyllabic words

Target tone / **Realized tone**	**[55]**	**[35]**	**[21]**	**[51]**	**Neutral**	**Total**	
[55]		32	32	49	7	**120**	37.6%
[35]	24		63	7	0	**94**	29.5%
[21]	24	11		14	4	**53**	16.6%
[51]	13	14	8		9	**44**	13.8%
Neutral	0	5	2	1		**8**	2.5%
Total	**61**	**62**	**105**	**71**	**20**	**319**	
	19.1%	19.4%	32.9%	22.3%	6.3%		

There were totally 319 tonal errors in the overall data. Table 4.16 showed the matrix in specific numbers of tonal errors that the upper row represented the target tones and the left column represented the tones that realized in substituting the error tones. The percentages on the right-hand side indicated which tone was more likely to replace the inaccurate tones. The percentages in the lower row represented which tone was more frequent to make errors. The most frequent tone which children chose in replacing an error was [55] in 37.6%, and the most frequent tone in making tonal errors was [21] in 32.9%. The percentages on the right column were put into the bar graph in Figure 4.12.

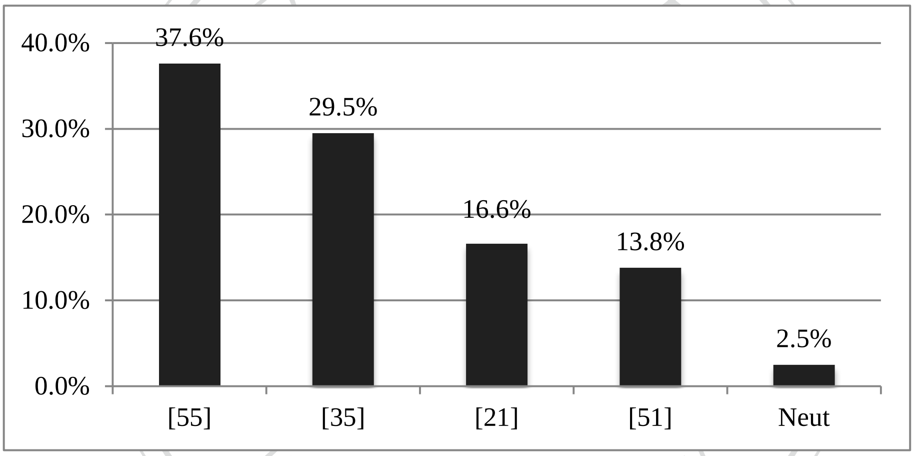


Figure 4.12 Percentage of the realized tone in tonal errors

In Figure 4.12, [55] was the first choice to replace other tones in errors which accounted for 37.6%. The [35] was also frequently used in replacing immature tones, accounting for 29.6%. The low-level tone [21] was realized less frequently in tonal errors, accounting for 16.6%. The falling tone [51] was a highly frequent and mature tone, but it did not replace tonal errors frequently, accounting for 13.8%. The neutral tone was the last choice children choose in substituting for tonal errors (2.5%). To sum up, when a tonal error occurred, children preferred to replace the immature tone with [55] the most, with [35] the second, with [21] and [51] the third, and with neutral tones last. The ranking of the realized tones in substitution would be [55]> [35]> [21]> [51]> Neut.

4.4 The age of tone emergence and stabilization

The table in the following summarized the age when tones were emerged and stabilized in every subject. The table also recorded the age when children stabilized a tone in 66.7% and 90% of correctness respectively. The 66.7% of the stabilization represented the fact that two third of the tone were produced correctly, and 90% of the stabilization represented higher degree of correctness.

Table 4.17 Age of emergence and stabilization of each tone

Tone		**#1**	**#2**	**#3**	**#4**	**#5**	**#6**
[55]	Emergence	1;1	1;0	1;3	1;0	0;11	1;1
	Stabilization 66.7%	1;1	1;3	1;3	1;0	1;2	NA
	Stabilization 90%	1;3	1;5	1;3	1;0	1;2	NA
[35]	Emergence	1;1	1;3	1;2	1;0	1;2	NA
	Stabilization 66.7%	1;1	1;3	1;4	1;0	1;5	NA
	Stabilization 90%	1;1	1;3	1;4	1;0	NA	NA
[21]	Emergence	1;1	1;3	1;2	1;0	1;2	NA
	Stabilization 66.7%	1;1	1;3	1;2	1;0	1;2	NA
	Stabilization 90%	1;1	1;3	1;3	1;3	1;2	NA
[51]	Emergence	1;1	1;1	1;0	1;1	1;1	NA
	Stabilization 66.7%	1;4	1;1	1;1	1;4	1;4	NA
	Stabilization 90%	1;6	1;6	1;2	1;4	1;4	NA
Neut	Emergence	1;1	1;6	1;1	1;1	1;2	NA
	Stabilization 66.7%	1;1	1;6	1;2	1;4	1;2	NA
	Stabilization 90%	1;1	NA	1;2	1;4	1;5	NA

Note: the capital letters NA which stood for “not applicable” indicates that the tone has not emerged or stabilized yet during the observation.

The age of emergence had been presented in the beginning of this chapter, the age of stabilization in two different degrees were added in this table. There were several things which could be presented in Table 4.17. We could compare the age of stabilization of the same tone in different subjects or compare the same subject in different tones. The individual differences could be seen in the first kind of comparison, and the chronological ordering could be seen in the second kind of comparison. From the comparison, we could figure out several questions, including which tone was emerged first in each subject, which subject stabilized all tones first, at what age did he/she stabilized all the tones, how long did subjects take to stabilize a tone after the tone was emerged, and roughly at what age have all subjects stabilized all tones.

Overall, the earlier appeared tone was [55] that emerged at 0;11 produced by subject #5. Subject #3 and #4 were the earliest children to have stabilized all tones in the degree of 90% stabilization at the age of 1;4, and subject #1 has reached the same goal at 1;6. Subject #2 and #5 have not stabilized neutral tone and [35] yet respectively. Subject #6 has not stabilized any tone yet, and there was only [55] emerged during the observation age from 0;10 to 1;5.

Chapter 5

Discussion

5.1 Summary of the findings

In this study, the acquisition of Mandarin tones in six children aged 0;10 to 1;6 is observed. To describe children's tonal development specifically, we age-tracked the age of emergence, calculated the frequencies and accuracy rates, and presented the substitution pattern in tonal errors in the overall data. Moreover, we classified the total of 2062 tokens into monosyllabic and disyllabic tokens, and have noticed that there were significant differences in frequencies or accuracy rates among different syllable positions due to a mass usage of the tone combination [21-35] in reduplications. After all tokens of the combination [21-35] were excluded, we reanalyzed the data again and obtained the results below:

Table 5.1 Tone acquisition orderings in different measures

(1) Tone emergence	[55] > [51] > [35],[21] > N
(2) Tone frequency	[55] > [51] >[35] > [21] > N
(3) Accuracy rate	[55] > [51] > [35] > N > [21].
(4) Substitution pattern	[55]> [35]> [21]> [51]> N

(1) Tone emergence: the high-level tone [55] tended to emerged the first, followed by the falling tone [51]. Rising tones [35] and low-level tones [21] were appeared later. Neutral tones appeared last or together with the last appeared citation tone.

(2) Tone frequency: high-level tones [55] had the highest frequency of occurrences, followed by falling tones [51]. The frequencies of [35] and [21] were originally high, but after excluding the outlier combination [21-35], their frequencies became lower than [51]. The occurrence of neutral tones is limited to utterance-final position, so its frequency revealed to be the last.

(3) Accuracy rate: results showed that high-level tones [55] were the most stabilized tone. The rates on falling tones [51] and neutral tones were similar and lower than [55]. Rising tones [35] were acquired less stable than [51], and the most unstable tones was the low-level tone [21].

(4) Substitution pattern: high-level tones [55] realized the most in tonal errors. Rising tones [35] seemed to replace immature tones more often than [21]. The tone that least occurred in tonal errors were falling tones [51]. Neutral tones would scarcely replace other tones.

(5) The age of tone emergence of stabilization in each subjects were summarized in Table 4.17 and it presented the individual differences and the chronological ordering in each tone.

Coincidentally, the high-level tone [55] emerged the first, and it also ranked as the most frequent and stable tone. Falling tones [51] were consistently ranked in the second place within tone emergence, frequency, and accuracy rate. Rising tones [35] and low-level tones [21] appeared late, and were also less frequent and stabilized later than [55] and [51]. For the time being, the four lexical tones, [55], [51], [35] and [21], ranked in the same order among tone emergence, tone frequency, and accuracy rate. From the similarities mentioned above, it seemed that the tone children acquired earlier was also the tone children preferred to use. The fact that the more frequently used tone obtained higher accuracy showed that children would avoid producing immature tones in the early developmental stages.

However, there were inconsistent rankings of the neutral tone between tone accuracy rate and other measures that it occurred last but stabilized earlier than [21]. For the tone emergence, the neutral tone could be seen to emerge at the same time with the last appeared lexical tone. Although the neutral tone seemed to appear last in most subjects, it still showed up in a short time after the lexical tones or even at the same time with the last appeared lexical tone. For tone frequency, due to the phonological constraint that neutral tone could only appear in the utterance-final position, it was unfair to rank neutral tones with other lexical tones in frequency ranking. Yet, the neutral tone could still be ranked in accuracy rate, and its ranking was higher than [21]. In short, the neutral tone was emerged and stabilized roughly at the same time with the last appeared and last acquired lexical tone.

The other inconsistency was found in the ranking of substitution pattern. The substitution pattern had different order from tone emergence, tone frequency, and accuracy rate that the second highest tone [51] in frequency turned out to be the last appeared tone in replacing tonal errors. It means that children were less preferred to replace a tonal error to [51]. Nevertheless, from Table 4.16, we found that [51] was the most preferred tone in replacing tonal errors in neutral tone. The reason could explain why [51] was the less preferred tone in tonal errors. Because the neutral tone has the least occurrences in lexical use that it only occurred in the particular position and in specific types of lexicon, the chance for [51] to replace others in tonal errors were also few.

5.2 Comparison with tonal acquisition studies in Mandarin

5.2.1 Age of acquisition

Several studies have pointed out the age children acquired the tonal system. Chao (1951) stated the age of distinguishing [35] and [21] was at 28 months old, and Jeng (1979) showed that by the age of 1;7, his two subjects were able to produce all tones accurately. Zhu (2002) presented more specifically that [55] was first acquired at the age between 1;2 to 1;5, and [21] was acquired last at the age between 1;4 to 1;9. By the age of 1;9, all lexical tones were acquired by all of the four subjects. First of all, the age that Chao (1951) presented was too old that in the following literature and the results in this study, children were able to distinguish the four contrastive tone at as early as 1;2 in Zhu's study and 1;1 in this study. Secondly, the age of acquiring the four tones that Jeng

(1979) pointed out was 1;7, and that in Zhu's study was 1;9. In the current study, five of my subjects have reached 66.7% of stabilization by the age of 1;6. My subjects seemed to acquire the four lexical tones a little bit earlier than those in Jeng and Zhu's studies, but the gaps between these ages were small. In substance, though children had individual differences, the earliest age that could master the four contrastive tone was at about 1;6.

5.2.2 Order of tonal acquisition

With regard to the ordering of tonal acquisition, Li and Thompson (1977) reported that [55] was the earliest acquired tone, and [51] was the second. The dipping tone [214] and the rising tone [35] were acquired last. Zhu (2002) described the same ordering that the high-level tones [55] and falling tones [51] stabilized earlier than rising tones [35], and falling-rising tones [214] were the last to stabilized. The order of stabilization in this current study was similar to these two previous studies that the first stabilized tone was [55], followed by [51], and the rising tone [35] and low-level tone [21] were stabilized last. It was surprising that the sequences of tonal acquisition were similar from children who lived in Taiwan thirty-five years ago and children who acquired Mandarin in Beijing. All of these studies showed that high-level [55] was the most unmarked tone, because children tended to acquire this tone first consistently. The more marked tones were [35] and [21] for the reason that they were acquired late by children. Ohala (1978) proposed the articulatory effort theory that the rising tones needed more energy and were more difficult to produce. This theory could be verified from all the results in the three studies.

5.2.3 The tonal combination of reduplication in Taiwan Mandarin

In chapter 4.2, the issue of the tonal combination [21-35] in reduplication in motherese was presented. However, there was almost no study noticing the particular use of tonal combination [21-35] of disyllabic reduplications in the Mandarin tonal acquisition in Taiwan. In many cross-linguistic studies, it was reported that reduplication was the major form in children's early production (Grunwell, 1982). The combination [21-35] is widely used in motherese by care-takers who spoke Mandarin in Taiwan. The tonal acquisition would definitely be influenced by the tonal combination in reduplications.

There was also no research proposed that the tones in reduplicative forms may not be acquired individually, but may be learned together as a prosodic whole. The evidence presented in this study showed that in [21-35] combination, the accuracy rates of [21] and [35] individually were low, but they become much higher in this mass used combination [21-35]. It may be caused by the influences in adults' input that care-takers tended to transfer words into reduplicative forms that were mostly in [21-35] tone combination. Researchers believe that children were sensitive to the repeated and regular form of speech sound and could imitate the same pattern more easily (McClelland et al., 1986). The repeated tone pattern [21-35] might be learned more quickly in children's mind than the individual [21] and [35] due to the more occurrences of the form [21-35]. But more examination should be conducted to actually analyze the frequency of input in these two tones; otherwise, the prediction would only be an assumption.

5.3 Cross-linguistic comparison

In chapter 2, several cross-linguistic studies in tonal acquisition were reviewed. These studies provided the developmental sequences of tones in several languages or dialects. In the following, I would like to compare the results of the current study to the previous tonal studies, and see whether language acquisition is universal or language-specific. I will also apply the cross-linguistic evidences to examine the theory of markedness presented by Yip (2002).

The result of the current study showed that the tone acquisition ordering in Mandarin was [55]> [51]> [35]> [21]. To compare the orderings cross-linguistically, I gathered the tone acquisition orderings below.

Table 5.2 Tone acquisition orderings of cross-linguistic studies

Cross-linguistic studies	**Tone acquisition orderings**
Thai (Tuaycharoen 1977)	[33],[11] > [224] > [45],[51]
Cantonese (Tse 1978)	[55],[11] > [33] > [13],[22],[25]
Cantonese (So & Dodd 1995)	[55],[33] > [25]> [11],[13],[22]
Taiwanese (Tsay 2001)	[55] > [53],[33] > [13] > [11]
Mandarin (Li & Thompson 1977)	[55] > [51] > [214],[35]
Mandarin (Zhu 2002)	[55] > [51] > [214],[35]

In Table 5.2, though the tone values in their inventories were different from each other, we could still compare and contrast the chronological orderings by viewing the tones in features. The level tones could be characterized by high, mid, and low tones, and contour tones could be categorized into falling and rising tones.

These cross-linguistic data could be used to examine the three constraints in the tonal markedness theory presented by Yip (2002).

a. contour tones are more marked than level tones

b. rising tones are more marked than falling tones

c. high tones are more marked than low tones

All studies including the results in this study agreed the first constraint, contour tones were more marked than level tones. The first acquired tones were all level tones and all the contour tones were acquired later. The high-level tones [55] in Taiwanese, Mandarin, and Cantonese found to be the first acquired tones, and the mid-level [33] and low-level tones [11] in Thai and Cantonese also developed earlier than contour tones. There were no studies showed that contour tones were acquired earlier than level tones, so it could be generalized that level tones were more unmarked universally.

If we considered the second constraint, we could find that the sequences were different between the study in Thai and those in other language. The falling tones [53] in Taiwanese and [51] and Mandarin were acquired earlier than the rising tones, [13] and [35], but the sequence was opposite in Thai that Thai-speaking children in Tuaycharoen's study learned the rising tone [224] earlier than the falling tone [51]. Cantonese has no falling tones, so it is not compatible with this rule. As a matter of fact, the result from Thai was derived from only one subject. It was doubtful whether the contrast could be attributed to individual differences.

The last markedness rule could only apply to level tones, so contour tones would not be included in this comparison. Comparing the acquisition ordering among high-level, mid-level and low-level tones, only Thai fulfilled this constraint that its low level tone acquired earlier than the high-level tone. This constraint could hardly find evidences from other cross-linguistic data except for Thai. The first acquired level tones in Thai were mid-level [33] and low-level [11] tones, and the high-level tone [45][3] in Thai was acquired the last. However, in Taiwanese and Mandarin, the high-level tones [55] were acquired the earliest. Especially in Taiwanese, the three level tones were acquired in the sequence of [55]> [33]> [11] which revealed the opposite ordering to the third constraint Yip proposed. The results from this current study also disapproved the third constraint that the high-level tone [55] was acquired earlier than the low-level tone [21] in Taiwan Mandarin. In Cantonese, the evidence from Tse's (1978) study was not comparable because it showed the spontaneous acquisition for high-level tone [55] and low-level tone [11]. The other Cantonese study done by So and Dodd (1995) suggested that high-level

---

[3] Duanmu (2000) pointed out the of five-point scale of tonal representation had problems. For example, the differences between [21] and [11] could not be detected by Mandarin native speakers. That is, if the slight falling or slight rising would not contrast meanings, they could be seen as level tones. In the Thai tonal system, there is a [45] tone. Because there is no high level tones similar to this tone, we could treat this tone as a high-level tone.

tones were more unmarked than low-level tones that the high-level [55] and mid-level [33] were acquired earlier than the mid-low-level [22] and the low-level [11].

On the whole, the first two constraints obtained more evidence that the features of level and falling in tones were suggested to be the unmarked features in tonal languages. Regarding the third constraint, because the data from four languages indicated that high tones were acquired earlier than low tones, the more unmarked tone feature should be level, falling, and high.

5.4 Concluding remarks

This research described how children learned to produce tones in Mandarin Chinese. From the analysis and discussion, the high-level tone [55] emerged the first, and it also ranked as the most frequent and stable tone. Falling tones [51] were consistently ranked in the second place within tone emergence, frequency, and accuracy rate. Rising tones [35] and low-level tones [21] appeared late, and were also less frequent and stabilized later than [55] and [51]. The neutral tone was emerged and stabilized roughly at the same time with the last appeared and last acquired lexical tone. The rankings of tone emergence, frequency, and accuracy rate were identical, which indicated that the tone that emerged earlier tended to be produced more frequently and would be stabilized earlier, and the tone that was used more frequently would have higher accuracy rate.

Concerning children’s reduplicative form, results showed that the tone combination

[21-35] dominated the tokens in disyllabic words. The tone combination [21-35] was proposed to be influenced by motherese, and was acquired as a prosodic whole by children. Further studies could be done in analyzing the correlation between the frequencies of [21-35] in care-takers and children's production.

The results in the current study were applied in cross-linguistic comparison including Thai, Cantonese, and Taiwanese. The tonal markedness presented by Yip (2002) was examined though the acquisition data in these languages. The comparison showed that level tones and falling tones which were the more unmarked features were acquired earlier in most of the tone acquisition data. While low tones which were shown to be the more unmarked feature tended to be acquired later cross-linguistically. Therefore, the unmarked features should be level, falling and high among all tonal languages in Asia.

This study provided another tone acquisition data to examine the universal rules in tonal features. Although the results were consistent with most of the previous studies, this study still found the unnoticed tone combination [21-35] applied particularly in the motherese in Taiwan Mandarin. The finding related to the tone combination was still preliminary. In the future study, researcher could collect care-takers' production and examine the relation between children's input and output and manage to resolve more questions in the language acquisition field.

References

Cantwell, D. & Baker, L. (1987). *Developmental speech and language disorders.* New York: The Guilford Press.

Carroll, D. W. (2008). *Psychology of Language* (Fifth Edition). Thomson Wadsworth.

Chao, Y. R. (1930). A system of tone letters. *Le Mâitre Phonetique* 45. 24-47

Chao, Y. R. (1951). The Cantian idiolect: an analysis of the Chinese spoken by a twenty-eight-months-old child. In W. J. Fischel (ed.), *Semantic and oriental studies.* Berkeley and Los Angeles: University of California Press.

Chao, Y. R. (1968). *A grammar of spoken Chinese.* Berkeley and Los Angeles: University of California Press.

Cheng, C. C. (1973). A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1. 93-110.

Clumeck, H. (1980). The acquisition of tone. In G. D. Allen and S. Hawkins (eds.), *Child Phonology*. *Vol: Production*. 257-275.

Crystal, D. (1970). Prosodic systems and language acquisition. In P. Leon (ed.), *Prosodic feature analysis.* Montreal: Didier.

Danmu, S. (2000). *The phonology of Standard Chinese.* London: Oxford university press.

Demuth, K. (1996). The prosodic structure of early words. In J. Morgan & K. Demuth (eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition.* Mahwah, N.J.: Lawrence Erlbaum Associates. 171-184.

Engel, W. R. (1973). An example of linguistic consciousness In childhood. In Ferguson, C. A. and Slobin, D. I. (eds), *Studies of child language development.* New York: Holt, Rinehart, Winston.

Fon, J. & Chiang, W.-Y. (1999). What does Chao have to say about tones?---a case study of Taiwan Mandarin. *Journal of Chinese Linguistic* 27.1. 15-37.

Fon, J., Chiang, W.-Y., & Cheung, H. (2004). Production and perception of two dipping tones (T2 and T3) in Taiwan Mandarin. *Journal of Chinese Linguistics* 32.2. 249-280.

Fon, Y.-J. J. (1997). *What are tones really like?* Thesis, National Taiwan University.

Grunwell, P. (1982). *Clinical Phonology*. London: Croom Helm.

Hashimoto, Anne Oi-Kan Yue. (1987). Tone sandhi across Chinese dialects. In Chinese Language Society of Hong Kong (ed.), *Wang Li Memorial Volumes: English Volume*. 445-474.

Jakobson, R. (1941). *Child language, aphasia, and phonological universal*. The

Hague: Mouton.

Kaplan, E. & Kaplan, G. (1971). The Prelinguistic Child. In J. Eliot (ed.), *Human Development and Cognitive Process.* New York: Holt, Rinehart, Winston. 359-381.

Kaplan, E. (1970). Intonation and language acquisition. *PRCLD* I. 1-21.

Kies, D. (November 2011). Language development in children. *Modern English Grammar*. Retrieved from http://papyr.com/hypertextbooks/grammar/lgdev.htm

Kubler, C. (1985). *The development of Mandarin in Taiwan: a case study of language contact.* Taipei: Student Books.

Lenneberg, H. (1967). *Biological Foundations of Language*. New York: John Wiley and Sons.

Li, C. N. & Thompson, S. A. (1977). Acquisition of tone in Mandarin-speaking children. *Journal of Child Language* 4,2. 185-199.

Li, C. N. & Thompson, S. A. (1978). The acquisition of tone. In V. Fromkin (ed.), The production of tone. In V Fromkin (ed.), *Tone: a linguistic survey.* New York: Academic Press. 271-284.

Lightfoot, D. (1982). *The Language Lottery: Toward a Biology of Grammars*. Cambridge, Mass.: The M.I.T. Press.

Lin, Y.-H. (2007). *The sounds of Chinese.* London: Cambridge University press.

Matychuk, P. (2005). The role of child-directed speech in language acquisition: a case study. *Language Sciences* 27. 3. 301-379.

Ohala, J. J. (1978). The production of tone. In V. Fromkin (ed.), *Tone: a linguistic survey.* New York: Academic Press.

Oller, D.K. (1974). Simplification as the goal of phonological process in child speech. *Language Learning* 24. 299-303.

Preyer, W. (1882). Dies seeles des kindes, leizig. *The mind of the child v2*. New York: Appleton. 1888-1890.

Shih, C.-L. (1998). *Tone and intonation in mandarin.* Working paper, Cornell phonetics laboratory.

So, K. H. & Dodd, B. J. (1995). Acquisition of phonology by Cantonese-speaking children. *Journal of Child Language* 22. 3. 473-495.

Tsay, J. (2001). Phonetic parameters of tone acquisition in Taiwanese. In M. Nakayama (ed.), *Issues in East Asian Language Acquisition*. Tokyo: Kuroshio Publishers. 205-226.

Tse, K-P. (1978). Tone acquisition in Cantonese: a longitudinal case study. *Journal of Child Language* 5, 2: 191-204.

Tuaycharoen, P. (1979). An account of speech development of a Thai child: from babbling to speech. In T.L. Thongkum, V. Panupong, P. Kullavanijava and M.R. Kalaya Tingsabadh (eds.), *Studies in Tai and Mon-Khmer Phonetics and Phonology*. In Honour of Eugénie J.A. Henderson, Chulalongkorn University Press. 261-277.

Vihman, M. M. (1996). *Phonological Development*. Oxford: Blackwell.

Yang, B. (2010). *A model of Mandarin tone categories--a study of perception and production.* PhD. Dissertation, University of Iowa.

Yang, H. C. (2012). Tone acquisition in Taiwan Mandarin- the frequency and accuracy of tone in 11- to 15-month-old children. Paper presented at *45th Conference on Sino-Tibetan Languages & Linguistics*.

Yip, M. (2002). *Tone*. Cambridge, UK: Cambridge University Press.

Zhu, H. (2002). *Phonological development in specific contexts*. Multilingual Matters Ltd.