The Dissertation Committee for Jie Yang
Certifies that this is the approved version of the following dissertation:

SEGMENTAL/LEXICAL INFLUENCES ON TONE ACCURACY IN MANDARIN-SPEAKING CHILDREN

Committee:

Barbara L. Davis, Supervisor

Randy L. Diehl

Chang Liu

Peter F. MacNeilage

Li Sheng

SEGMENTAL/LEXICAL INFLUENCES ON TONE ACCURACY IN MANDARIN-SPEAKING CHILDREN

by

Jie Yang, B.A., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

To my loving parents

who witnessed my first words

Acknowledgements

I am in debt to my mentor, Barbara Davis, for her guidance throughout the years of my doctoral study. Her encouragement and emotional supports have helped me through the most difficult time of my dissertating experience. My gratitude to her love and caring to all aspects of my life beyond academics can hardly be conveyed by words. I was fortunate to have Peter MacNeilage on my committee. Together with Barbara Davis, his insightful feedback on my master thesis served as the starting point of this dissertation. My deep appreciation goes to Randy Diehl for his knowledge and all the inspiring conversations I had with him. I thank Chang Liu and Li Sheng for their expertise. Mostly importantly, I thank them for their friendship and showing me how to grow as a young scholar.

This dissertation would not have been completed without the help of many people. I am truly grateful to the participating families for allowing me into their homes and for their dedicated supports over the course of data collection. I thank the lovely children who shared far more words than the two little words "thank you" can possibly express. I thank Jordan Bubis, Sara Flynn Fletcher, Kelly Jerden, Avery Rains and Molly Reilly for assistance in data entry. I thank Hui Chen for her careful work in transcription reliability. I thank Craig Corcoran for his tremendous help in computation programming. Without him, I would not know when I could finish all the coding and calculations manually. I thank Cory Hubbard and Lourdes Jones for grant administration. I also thank the many others from CSD who have provided encouragement and supports along the way. Thank you, Craig, Chris, Carmen, Dana, Joan, Namhee, Rajani, Suzanne, Ying and all my fellow doctoral students.

I owe a lot to my parents, who have waited so long to see the completion of this dissertation. I can't thank them enough for their love and supports during the years of my education. I also thank my in-laws for their encouragement and prayers.

I want to thank my husband, Jie, who has travelled the entire journey with me as my best friend, supporter and provider. He has always been a faithful listener to my research projects and lent me his engineering wisdom on interpreting and presenting behavioral data. Finally, to my newest family member, my daughter Yaya: Thank you for your perfect timing. You are the joy of my life and my motivation for accomplishing this dissertation on time.

This research is supported by a National Science Foundation Doctoral Dissertation Improvement Grant #0844080.

SEGMENTAL/LEXICAL INFLUENCES ON TONE ACCURACY IN MANDARIN-SPEAKING CHILDREN

Jie Yang, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Barbara L. Davis

Emergence of accuracy is a first step toward acquiring adult-like phonological abilities. Mandarin Chinese is a tone language where speakers employ both tonal and segmental properties to code lexical meanings. Study of this dual-level complexity of tone and segment enables a broader view of how phonology is acquired than the view afforded by study of Indo-European languages. Understanding the interaction between phonatory properties for tone and articulatory adjustments for segments in emergence of early words helps to understand more generally the first steps toward the complex system embodied in phonology.

The present study investigated tone acquisition in relation to segmental and lexical development in Mandarin-speaking children in the earliest word stages. Spontaneous speech samples were collected longitudinally from 12 to 24 months from four Mandarin-speaking children. The relationships between tone accuracy, segmental accuracy, and word-level variables were examined quantitatively over time.

Results indicated that tone accuracy is not always higher than segmental accuracy. The relationship between these two seems to be influenced by the physiological complexity

of tonal shapes and children's developmental age. Autonomy of control over phonatory adjustments for tone and articulatory adjustments for segments was already apparent. Children were not sensitive to the contrastivity (characterized by Productive Tone Neighborhood Density) involved in tonal categories with a vocabulary of less than 50 words. Associations between production accuracy and word-level variables (articulatory complexity, neighborhood density and word frequency) established based on later developmental periods were not found in younger Mandarin-speaking children.

Findings suggest that tonal acquisition at the onset of speech development is not a passive process where innate phonological knowledge is revealed solely through children's maturation. Rather, phonological knowledge is established on the basis of children's pre-linguistic motor capacities in concert with cognitive learning occurring via the expansion of their lexicon. Tones and segments may be produced as holistic entities in early words. Tone acquisition at the onset of word learning is more child-centered in that availability of tonal forms to the child's production system underlies accuracy. Influences from lexical properties of word would only be apparent when phonological knowledge of tonal categories is established with vocabulary expansion.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Human speech embodies a complex system which integrates language-specific phonological knowledge with behavioral capacities for linguistic message generation. Studies of Indo-European languages have indicated that children's phonological knowledge of consonants and vowels emerges based on their pre-linguistic production and perception system capacities, sharpened through the experience of producing and listening to early words (e.g. Davis, MacNeilage, & Matyear, 2002; McCune & Vihman, 2001). Properties of words (word frequency and neighborhood density) have been shown to influence children's early output accuracy as well (Gierut & Morrisette, 1998; Sosa, 2008; Storkel, 2004).

Phonological acquisition of tone languages, which employ both tonal properties based on phonatory adjustments and vowel and consonant segments based on articulatory adjustments, to code lexical meanings has not been well studied. Mandarin employs all aspects of the production system found in modern languages. Understanding the interactions between phonatory and articulatory adjustments in the emergence of early word forms in Mandarin will help to understand the first steps toward the complex system embodied in phonology more broadly. The dissertation research will investigate **tone acquisition in relation to segmental and lexical acquisition in Mandarin-speaking children during the earliest period of lexical learning.** This process will be viewed from the perspective of emergence of a complex system.

Mandarin Chinese, a tone language, uses four pitch contours (high-level, high-rising, low fall-rise and high-falling) to contrast lexical meanings (Chao, 1968). A two-segment syllable [ma] produced with the four pitch contours means “mother”, “hemp”, “horse” and “scold”. Present research indicates that children are capable of producing these four pitch contours during canonical babbling (Lee, Lee, & Chen, 2005; Yang, 2006). The question remains as to how and when children can accurately associate capacities for producing pitch contours to establish phonological-level tonal contrast of tone in word forms.

Earlier studies proposed that children accurately attach pitch contours to target words before 24 months (Clumeck, 1980; Li & Thompson, 1977; Zhu & Dodd, 2000). This reported milestone is chronologically earlier than reports on accurate production of consonants and vowels. In contrast, recent findings suggest that tone acquisition is an item-by-item process, where children can produce tone accurately on only a subset of words with the same tone but not all of them (Yang & Lee, 2006). However, the relationship between tone and segmental acquisition has rarely been investigated on a quantitative basis where data has been evaluated statistically. Possible factors predicting word-level variation in emergence of tone accuracy have been based on small data corpora with individual children.

To study tone acquisition in relation to segments and words, we assume that the phonological knowledge of tone emerges as a dynamic system (Thelen & Smith, 1996) in an individual human child, based on early motor capabilities interactive with other components of the language system. Tones are intimately associated with word forms as they form the basis of meaning contrast signaled in word forms. Preliminary evidence suggests that the accurate production of tone initially appears to be achieved only on words with a smaller number of similar contrasting forms (Yang & Davis, 2008). For example, both "妈" [ma1] *mother* and "七" [tɕʰi1] *seven* are Tone 1 words. However, "妈" *mother* was produced with higher tone accuracy compared to "七" *seven*. The reason is that "妈" *mother* has one word that contrast with it in tone in the child's production output ("马" [ma3] *horse*). "七" seven has three contrasts ("旗" [tɕʰi2] *flag*, "起" [tɕʰi3] *rise*, "气" [tɕʰi 4] *air*).

To consider emergence of tonal accuracy in the early lexicon, sixty-minute spontaneous speech samples were recorded bi-weekly from 4 typical-developing Mandarin-speaking children (three boys and one girl) between 12 and 24 months of age. Words with recognizable lexical targets produced by the child were measured for tone accuracy, segment accuracy and word-level variables. We considered word frequency (in both ambient language and child production), neighborhood density in children's expressive lexicon and word complexity (in both adult target and child production) as word-level variables. The specific aims of this study are to:

Aim 1. Examine the relationship between tone accuracy and segmental accuracy over time. **Hypothesis:** Children will produce segments more accurately than word level tonal contrasts.

Aim 2. Examine the relationship between tone accuracy and word-level variables over time in each child. **Hypothesis:** Tone accuracy of a word will be negatively related to the number of tonally similar words in a child's expressive lexicon. Tones will be produced more accurately in words that have less similar words. This relationship will weaken as the child's vocabulary increases.

In Chapter 2, theoretical approaches to phonological acquisition and previous studies of tone acquisition relative to the two aims of the present study are reviewed. The phonological system of Mandarin Chinese, including segment inventory, tonal system and syllable structure, is also introduced. In Chapter 3, detailed study design and methods are described. Results of analysis of this study are summarized in Chapter 4. Chapter 5 discusses the implications of the findings.

This study enables consideration of the theoretical underpinnings of phonological acquisition by including the supra-segmental dimension of language complexity, not possible in study of acquisition in non-tone languages. According to study hypotheses, implementation of characteristics of tone found in the ambient phonological system is not a process where innate and highly specific phonological grammar is revealed in children through maturation. On the contrary, acquisition of tone accuracy in earliest lexical acquisition is an item-by-item process. The complexity of ambient tonal knowledge emerges through a dynamic process, where children's production and perceptual capabilities to perceive and produce pitch contours interact with the requirements for building sound - meaning correspondences within the lexical subcomponent.

Study of tone acquisition and its relationship with development of segments and the lexicon has implications for understanding the course of typical development in Mandarin. The study also contributes to systematic understanding of emergence of complexity in languages related to the phonatory dimension in tone languages, where the F_0 dimension is used much more broadly at a segment level than is typical of languages that use F_0 for voicing, aspiration, or utterance level intonation distinctions only.

Chapter 2: Background

2.1. Theories of Phonological Acquisition

In the process of phonological acquisition, children must acquire both the phonetic ability to produce sounds in the target language and the phonological knowledge to employ the sound patterns of their language contrastively. Phonetic ability characterizes children's behavioral capability to produce speech sounds. Phonological ability describes underlying knowledge that children have of their ambient language sound system, on the basis of which they can map phonetic ability onto lexical items in an adult-like way.

For phonological acquisition, Jakobson (1941/1968) postulated a discontinuity relationship between pre-linguistic babbling and early speech. Research carried out in the recent decades has effectively questioned Jakobson's claim by showing that babbling is related to and paves the way for early speech. (Boysson-Bardies & Vihman, 1991; Boysson-Bardies et al., 1992; Chen, 2005; Davis et al., 2002; Lee et al., 2005; Levitt & Utman, 1992). During babbling, children produce a basic set of the sound patterns within rhythmic syllable-like sequences of alternation between open and closed mouth configurations (Davis & MacNeilage, 1995). These patterns of sounds and serial properties are used to map pre-linguistic phonetic abilities onto early lexical items (Davis et al., 2002)

In considering phonological acquisition, two opposing theoretical perspectives characterize the underlying mechanisms of the process. **Formalist approaches** assume *a priori* or innate phonological representations which are said to be universal to all world languages (termed underlying grammar, UG, Chomsky & Halle, 1968). As children are exposed to the ambient target language, perceptual input provides a triggering function to unfold children's *a priori* representational capacities to equip them with language-specific phonological competence. Phonological acquisition is regarded as a passive maturational process that reveals underlying phonological competence. Children's own production, perceptual and cognitive capabilities are viewed as peripheral performance factors (Chomsky & Halle, 1968; Bernhardt & Stemberger, 1998; Prince & Smolensky, 1993). Variability during the process of maturation is considered to be caused by the differences in the ranking of constraints. **Functional approaches** address the issue of variability by considering phonological acquisition as an active process of pattern-finding and refinement. Phonological knowledge is seen as emerging, on the basis of children's biological capacities, through the local interaction of relevant component systems (Davis et al., 2002; Menn, 1976; Vihman, 1993). The acquisition of sound system capacities emerges over time and word level factors are considered as playing a critical role (Beckman & Edwards, 2000; Lindblom, 1992).

In addition to the segmental and syllabic levels of the output system, the phonologically contrastive use of pitch or fundamental frequency (Lehiste, 1970), needs to be acquired by children learning tone to enable successful mastery of their ambient phonological system. Languages, such as Mandarin, Cantonese and Thai are characterized by use of tonal level contrasts. Similar to the acquisition of segments, children's acquisition of a tonal system also involves acquiring the phonetic capability to produce pitch contours of the language they are acquiring, and the phonological knowledge to map pitch contours onto target words in their language.

Recent studies on the pitch characteristics of babbling in Mandarin- and Cantonese-speaking infants indicate that they are already capable of producing the pitch contours used in their native language in babbling utterances (Chen, 2005, Lee et al., 2005). The question remains as to how accurately they can map these pitch contour capacities onto lexical items. Compared to the large number of studies on the acquisition of segmental and syllable level patterns in early words, research in the acquisition of tone is relatively sparse. Furthermore, unlike the acquisition of English or other non-tone languages, children's acquisition of the phonological system of a tone language is complicated by the dual task of acquiring both the segmental and the suprasegmental system for conveying meanings in early salient lexical items. However, quantitative studies of the relationship between the acquisition of tone and segment are scarce.

Since tones are superimposed on the syllable level of the lexicon, the establishment of tonal knowledge is also hypothesized to be related to lexical acquisition. Previous studies have pointed out the existence of word-level variations in children's tone production (Yang, 2006; Yang & Lee, 2006). No attempt has been made to consider the source of these variations. Whether the relation between tone and segmental acquisition, or tone and lexical acquisition will stay constant during the entire acquisition process is also unknown. No investigations of the range of individual variations in tone acquisition have yet been made. To achieve better understanding of phonological acquisition of tone languages, examination of the acquisition of tone in relation with the acquisition of segments and the early lexicon can help to explore sources of variability in production accuracy. The proposed study will address these issues by gathering longitudinal data on tone acquisition in children learning Mandarin Chinese, a tone language with the largest number of native speakers world-wide (867.2 million, Gordon, 2005).

Our research assumes a continuity relationship between pre-linguistic babbling and early speech, where children are already equipped with the production capability to produce pitch contours in their language at the onset of word use. Thus, the task of the young child in achieving phonological accuracy is to manage the contrastive use of these tones and map them onto salient words. Therefore, we hypothesize that the degree of contrastiveness of these tones in the language and in each child's expressive lexicon will influence the child's output accuracy.

2.2. The phonological system of Mandarin

The phonological system of Mandarin Chinese provides a necessary background for evaluating acquisition of tonal accuracy.

2.2.1. Segments

Table 1 and Table 2 illustrate the phonemic system of consonants and vowels (Lee & Zee, 2003).

Table 1: Vowels of Mandarin.

Monophthongs Diphthongs Triphthongs

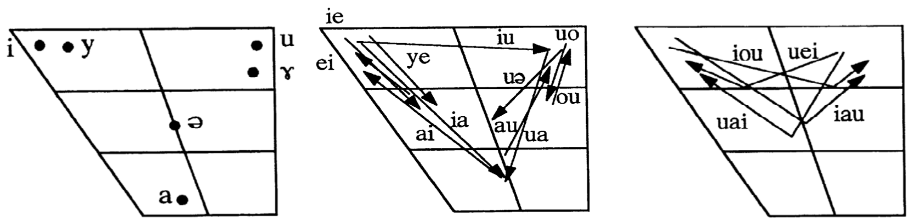


Table 2: Consonants of Mandarin.

	Bilabial	Labio-dental	Alveolar	Alveo-palatal	Palatal	Velar
Stop	p p^h		t t^h			k k^h
Fricative		f	s		ɕ	x
Affricate			ts ts^h	tʂ $tʂ^h$	tɕ $tɕ^h$	
Nasal	m		n			ŋ
Approximant (Glide)	w				j	
Approximant (Liquid)			l	r		

2.2.2. Tones

Mandarin has four tones: Tone 1 is high level (LH); Tone 2 is high rising (RH); Tone 3 is low fall-rise (FRL) and Tone 4 is high falling (FH) (Chao, 1968). With these four tones, the syllable [ma] is contrasted to mean “mother”, “hemp”, “horse”, and “scold”. Besides transcribing the four tone by a combination of pitch contour and height., these can also be described using Chao’s five-point scale tone letters as having tone values 55, 35, 214 and 51 respectively (Chao, 1930).

The pitch shape of the four tones in Mandarin can be represented in Figure 1 (From Xu, 1997).

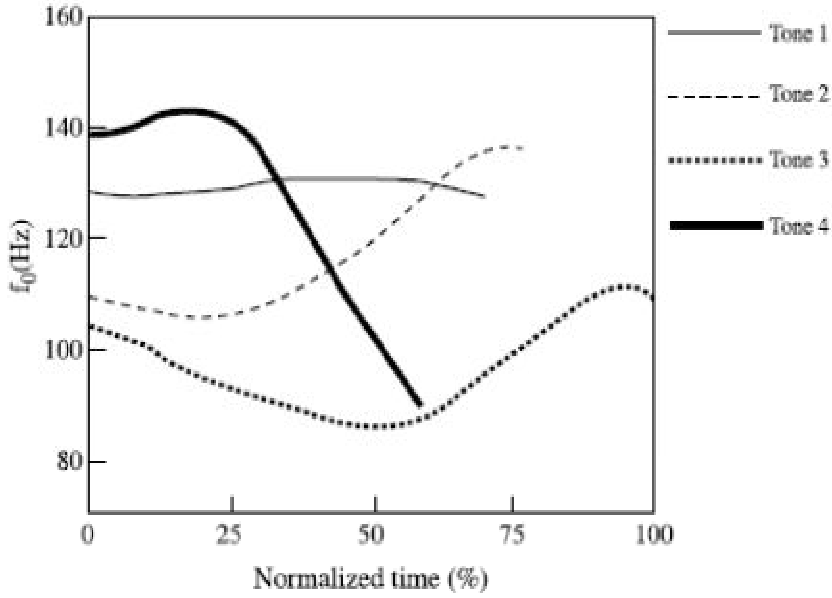


Figure 1: Pitch Shape of the Four Mandarin Tones.

There is a fifth tone in Mandarin, which occurs in unstressed syllable, referring to as Tone 0 or neutral tone. The pitch value of the neutral tone varies depending on the preceding full tone. Almost any morpheme in one of the four regular tones can be in the neutral tone under certain condition. There are also a small number of morphemes, such as suffix and particles, which are always in the neutral tone and do not belong to the four tones. The pitch of neutral tone is half low after Tone 1, middle low after Tone 2, half high after Tone 3 and low after Tone 4 (Chao, 1968).

As in other tone languages, the phenomenon of tone sandhi, which refers to changes in lexical tone when words occur in combination, also exists in Mandarin. Some tone sandhi rules are obligatory and some are optional. For bi-syllable tone sandhi, Tone 3 sandhi and Tone 4 sandhi which are presented as in (1) are obligatory, while tone 2 sandhi in tri-syllabic words is optional as presented in (2).

(1) Obligatory rules: Tone 3 sandhi $214 \rightarrow 35 / \_\_\_ \ 214$

$214 \rightarrow 21 / \_\_\_ \ \{55, 35, 51\}$

Tone 4 sandhi $51 \rightarrow 53 / \_\_\_ \ 51$

(2) Optional rules: Tone 2 sandhi $35 \rightarrow 55 / \{35, 55\} \_\_\_ \ \{55, 35, 214, 51\}$

There are also tone sandhi rules which are related to specific morphemes, such as *yi1* “one” *bu4* “not”, *qi1* “seven”, *ba1* “eight”. Take *bu4* “not” in (3) below for example, when it is pronounced before a Tone 4 word, such as *yao4* “want", the original falling tone changed into a rising tone.

(3) Morphologically-conditioned rule: *bu* 51 → *bu* 35/___ 51

2.2.3. Syllables

Mandarin syllables are composed of onsets and rimes. Except the velar nasal [ŋ], each of other 23 consonants can occur in syllable onset position. However, no cluster is allowed. Rimes may be a monophthong, a diphthong, a triphthong, a monothong or diphthong + nasal [n] or [ŋ] sequence. Triphthongs cannot be followed by nasals. The two types of syllable structure are illustrated in Figure 2.

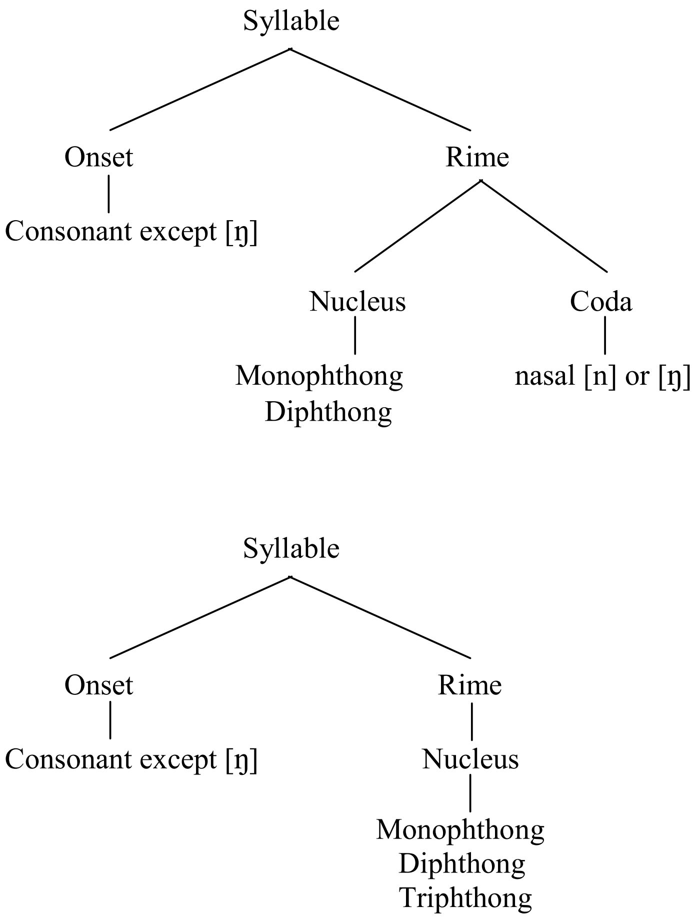


Figure 2: Two Types of Syllable Structure in Mandarin.

There are about 400 syllable types in Mandarin. One syllable corresponds to one Mandarin morpheme. The four Mandarin tones are almost equally distributed over those syllable types (Duanmu, 2006; Howie, 1976). Together with the monosyllabic nature of Mandarin words, each morpheme has 3 other morphemes that differ from it only in tone. In other words, with regard to ambient Mandarin language, the neighborhood density of tone is equal across morphemes.

The background for the study aims will be reviewed in the next section.

2.3. TONE ACCURACY AND SEGMENTAL ACCURACY

2.3.1. Tone and phonation

Tone, together with intonation, is a linguistic utilization of pitch (or fundamental frequency F_0, Lehiste, 1970). The difference between the two lies in the level at which pitch variation functions, i.e. tone is the contrastive use of pitch at word level, while intonation is linguistically significant functioning of pitch at phrasal or sentence level. Languages that use tones to distinguish words are termed tone-languages (e.g., Mandarin, Cantonese, Thai etc.), while those do not use tones to distinguish words are called non-tone languages (e.g., English and other Indo-European languages). Lexical tone is realized through the manipulation of phonatory adjustments. Tone-related phonation spreads over the domain of a lexical word. Phonation adjustments are also employed at segmental level (e.g. voicing distinction of segments [p] and [b] in English; aspiration distinction in Mandarin [p] vs. [p^h]). They are features of a particular segment and cannot exceed the domain of that segment (Lehiste, 1970).

2.3.2. Acquisition of phonatory contrasts

Development of phonatory distinctions in production either at the segmental (voicing and aspiration) or the lexical level (tone) has been documented in the early speech acquisition literature. For both acquisition of voicing/aspiration and acquisition of tone, young children need to refine production system-related phonatory control to achieve adult-like accuracy for target words. Although languages employ voicing or aspiration contrastively in their phonological system, different languages have diverse ways of realizing these contrasts. Similarly, in tone languages, the use of pitch for lexical tone might differ as well. Therefore, children's acquisition of these phonatory contrasts must be specifically attuned to their target language requirements. Understanding the acquisition of segmental-level phonatory contrasts (voicing and aspiration) will provide a more complete picture for understanding the acquisition of tone, a lexical level phonatory contrast system used in Mandarin.

2.3.2.1. Acquisition of voicing and aspiration

The acquisition of voicing and aspiration has been characterized by the realization of voice onset time (VOT) in children's words. VOT expresses the timing relations between the release of occlusion for a stop consonant and the onset of the vocal fold vibration.

Macken and Barton (1980) studied VOT of three pairs of voiceless-voiced stops, /p-b/, /t-d/ and /k, g/ produced by English-learning children between 18 and 30 months. The children's spontaneous speech was recorded biweekly for 20-30 minutes. The analysis was restricted to their production of the stops in the initial position of words. Based on study results, the authors proposed a three-stage model for voicing development. At the initial stage, children were able to produce adult-like voiced stops with VOT falling within the adult range at the start of the recording (<20 ms). However, although three of the children started to differentiate voiced and voiceless contrasts, the VOT value of their voiceless stops were still not long enough for adult transcribers to perceive the voiceless counterpart. After this intermediate stage, children were able to produce voiceless stops similar to adult VOTs.

A more recent study by Lowenstein and Nittrouer (2008) examined the acquisition of English VOT by audio-taping spontaneous speech of seven children at 2 month intervals from first words (around 15 months) to the appearance of three-word sentences (around 30 months). VOT of word-initial stops was analyzed. Counting the numbers of words produced with voiced and voiceless initial stops in their targets indicated that the children initially avoided producing words with voiceless initial stops. Mean VOT was 12 ms at the youngest age examined and increased to 40 ms at 23–24 months. The production of voiceless stops was more variable than for voiced stops, shown by the within-child standard deviations of VOT.

Unlike English, stop consonants in Mandarin Chinese are not contrasted using voicing but by aspiration, between the stop burst and the onset of voicing of the vowel following the stop. Aspiration is also measured by VOT. The development of phonemic use of aspiration in Mandarin has been studied (Chen, 2007). Chen examined the development of aspiration in a Mandarin-learning boy longitudinally from 17 to 35 months. He found a developmental pattern similar to the three-stage model proposed by Macken and Barton (1980). Prior to 18 months, aspiration contrasts were not evidenced in the child's production. Aspiration contrasts emerged first at the alveolar position when the boy was 18 months. The distinction was extended to bilabial and velar stops from 19 months onwards. But the VOT values of the contrastive pairs in child's production were not adult-like. From 22 months, average VOT of the child's voiceless unaspirated and voiceless aspirated stops fell within the adult-appropriate phonetic region: 0 to +30 msec.

2.3.2.2. Acquisition of tone

The present study focuses on the acquisition of lexical-level phonation contrast, i.e., tonal contrast. Compared to the large number of studies on the acquisition of segmental and syllable level patterns in early words, research in the acquisition of tone is relatively sparse. Existing studies of tone acquisition have been conducted in different tone languages, including Mandarin, Cantonese, and Thai, using both cross-sectional and longitudinal methods. These studies have focused on the sequence of acquisition of tone contours and the production error-patterns (see Clumeck, 1980; Vihman, 1996, for review).

Mandarin

Chao (1968) made the first attempt to tackle the problem of tonal acquisition in a study of a twenty-eight month old Mandarin-speaking child named Canta. Although the analysis of tone constituted only a small proportion of the data analyzed for this child, Chao made some pioneering claims. He asserted that (1) Tones are acquired early; (2) There exists confusion between rising and dipping tones, with rising tones largely replaced by dipping tones as compared to dipping-rising substitution; (3) Tone3 sandhi rule only begin to be learned.

Li and Thompson (1977) carried out a longitudinal/cross-sectional study dealing with 17 Mandarin-speaking children living in Taipei. They were aged 18 to 34 at the start of the study. During a period of 7 months, data were collected from spontaneous speech and picture identification tasks. Their research clarified Chao (1968)'s claim about early mastery of tone. Chao's second and third claims were also confirmed with the larger group of children. This larger scale systematic investigation specified the chronology for mastery of the four tones. The authors delineated four stages in the process of tonal acquisition which correspond somewhat to certain stages in syntactic development. They found that at the earliest one-word stage, high-level and high-falling tones predominate irrespective of the adult target characteristics. The second stage is characterized by a larger vocabulary, the appearance of four way contrast of tone and by occasional confusion of high-rising and low-dipping tones. When children move into the two-three-word stage, rising-dipping confusion still remains, but they begin to acquire tone sandhi rules. At the fourth stage, longer sentences are produced and rising-dipping confusion disappears.

The relative difficulty of the acquisition of dipping and rising tones was also found in Clumeck (1977). Two Mandarin-speaking children were followed longitudinally. However, the analysis included only words that were uttered either in isolation or in utterance final position. Calculating the percent accuracy for each tonal token, Clumeck found that both children showed almost complete mastery of the high-level and high-falling tones. Both had difficulties with the high-rising and low-dipping tones. But in Clumeck (1980), the participant acquired rising tone first, which he explained as influences by its high frequency in the reduplicated nouns of the child's input.

Results in Hsu (1996) were not consistent with previous studies. Twenty-eight Mandarin-speaking children from one year to two years six months were examined longitudinally. Although the children mastered phonatory control over tone production quite early as postulated in previous studies, they did not produce all words accurately. Of the children Hsu observed, level tones and falling tones appeared first at about one year of age. Rising tone began to appear shortly after one year, but not accurately on all the words marked with Tone 2. Tone 3 was the most difficult one for children to acquire and began to emerge at around 1 year 6 months. By 2 years 6 months, with some confusion between Tone 2 and Tone 3, all of his participants had acquired the four tones. However, it took the children longer to assign the tones accurately to all the lexical items they attempted. Apart from the Tone 2 and Tone 3 confusion, level and falling tones were substituted for each other at very early stages. Level and falling tones were also used to replace Tone 2 target words.

Zhu and Dodd (2000) and Zhu (2002) furnished parallel data from a cross-sectional study on the phonological acquisition of 129 monolingual Mandarin-speaking children, aged 18 months to 4 years 6 months. They used a picture identification task and a picture description task, with an imitation task as supplement in case that children failed to produce the target word or phrase spontaneously. They found that tone errors were rare, even in the youngest group of children. However, no chronology of the acquisition of tonal inventories was addressed.

Wong, Schwartz & Jenkins (2005) challenged the claim on early acquisition of tones as well as some of the error pattern found in previous studies. Both the perception and production of Mandarin lexical tones in monosyllabic words were investigated in thirteen 3-year-old children learning Mandarin as their first language in the United States. Picture-pointing and picture-naming tasks were used to elicit words. Unlike previous research on tonal acquisition, speech data of four of the children's Mandarin-speaking mothers was analyzed as a control. In order to avoid the bias in tone transcription usually carried out by a single transcriber, the authors recruited ten Mandarin-speaking judges to identify the production of tone in children's and mother's speech. The findings indicated that "3-year-old Mandarin-speaking children's tone production in monosyllabic words was not yet adult-like" (p.1076). With regard to the acquisition of tonal inventories, the authors found that Tone 2 was produced with comparable accuracy to Tone 1 and Tone 4. Tone 3 was significantly more difficult for the children. The study also reported that half of the Tone 1 target words were replaced by falling contours and half by rising contours.

Cantonese

Cantonese has a more complicated system of tonal distinctions than Mandarin, with level, rising, and falling tones on upper and lower registers. It also has three entering tones which differ in height: high-entering, mid-entering and low-entering tone. Altogether there are nine tones in Cantonese.

Tse (1978) studied the acquisition of Cantonese tone based on the longitudinal observation of his son Y.L. from 0; 6 to 2; 6. Similar to the acquisition of Mandarin tone, Cantonese tonal acquisition was completed earlier than segments. High level tones and falling tones were acquired in the first stage corresponding to the first three months of the single word stage. Mid-level tones, high-rising tones and the three entering tones were acquired in the remaining four months of the single word stage. The low-rising tones and low-level tones were acquired last, when Y.L. began to use two word constructions. Contour tones (such as rising and falling tones) are more difficult to master than level tones among suprasegmental patterns of Cantonese.

Tse (1993) reported a longitudinal study on the development of the phonological system of one Cantonese child, K.C., from 1; 2 to 3; 0, Tonal acquisition was one of the sub-aspects of the investigation. K.C.'s tonal acquisition process was divided into four stages, each of which agreed with Tse (1978), in that high and mid-level tones were acquired first and high-pitched tones were usually acquired before low-pitched tones.

So & Dodd (2000) confirmed data in Tse (1993) for four children observed longitudinally. Two of the three level tones (high-level and mid-level tones) were acquired first, followed by the high-rising tones. The three entering tones were acquired before the low-falling and low-rising tones. However, tonal acquisition in these four children was completed by age two, different from the report for K.C. in Tse (1993) whose low-rising tone was still not accurately produced at that age.

<u>Thai</u>

There are five tones in Thai. Three of them are level tones differing in pitch height: high, mid and low. The other two are falling and rising tones (Yip, 2002).

Tuaycharoen (1977) reported that her Thai participant used mid and low level tones from the age of 11 months, when his first words appeared. At this stage, the mid and low level tones substituted for the falling, high level and rising tones, with the substitution of low level tone for high level and falling tones occurring most frequently. A rising tone appeared on lexical items when the child was 14 months old. Falling and high tones did not appear until the end of the fifteenth month. The tones were all reported as being mastered by the child's twenty-third month.

Summary

Previous studies addressed two major issues in the process of tonal acquisition across languages: (1) Age of acquisition (2) Order of acquisition of tonal inventories. There are no congruent findings regarding age of acquisition. Some studies reported that children successfully acquire lexical tonal contrasts by around their third year, earlier if the language does not have too many alternations (e.g. Mandarin). Other studies, such as the recent study of Wong et al. (2005) suggested that children's production of tone is still not adult-like by age 3. With regard to the order of acquisition of tonal inventories, for Mandarin, studies consistently reported that Tone 3 (fall-rise) is the last tone to be acquire because of the complexity involved in the manipulation of pitch contours. Children need to lower their pitch and then rise again within the time domain of a syllable. For the other three tones, most studies suggested the following order: Tone 1>>Tone 4 >>Tone 2. However, Hsu (1996) found that Tone 1 and Tone 4 in Mandarin appeared at the same time. A recent acoustic studies of tone acquisition (Wong, 2012) suggested that 3-year-old Mandarin-speaking children are not capable of maintaining a level pitch over a syllable resulting in productions of Tone 1 that were still not adult-like.

Several limitations of the previous studies may contribute to the discrepant findings. One of them is the procedural differences in tonal identification and transcription. As Wong et al. (2005) and Yip (2002) pointed out, adult transcribers always find it difficult to transcribe child data which cannot be properly captioned by adult phonological system. Together with possible influence of tone expectation created by contextual cues (Oller & Eilers, 1975), transcription reliability is highly required. However, most of the previous studies rely solely on the perception of a single transcriber to determine tone contours without reporting transcription reliability. Secondly, criteria for determining the time of tone mastery were not specified in previous studies. For Mandarin tone, only Zhu (2002) explicitly stated the criteria for the emergence and stabilization of tone categories. For Cantonese, So & Dodd (1995) judged a tone as acquired when it was used contrastively on at least 50% of opportunities or correctly on 90% of opportunities. Thirdly, the distribution of tones in children's lexicon, which is relevant in determining the order of acquisition of tones within a certain language and across languages, was largely neglected in previous studies. Hsu (1996) observed the existence of lexical level variations where children are able to produce some but not all words in the same tone categories, but no further analysis was provided in this regard. Since tone is encoded on words in lexical tone languages, children's production of tone is revealed in the lexical items containing certain tones. Therefore, the investigation of tonal acquisition should relate itself to the acquisition of lexicon.

2.3.3. Tone acquisition in relation to segment acquisition

The studies reviewed previously also reported patterns with regard to the chronological relationship between the acquisition of tone and the acquisition of segments. Children have been described as producing tones in their language quite accurately at a stage when some segments were still not accurately produced. For example, Tuaycharoen (1977) pointed out that her Thai participant mastered the tone system by 23 months but had not mastered diphthongs, triphthongs, or initial consonant clusters. Li and Thompson (1977) reported that many of their participants' utterances were tonally accurate but incomprehensible due to segmental "errors". They gave an example of one of their participants, who, at 26 months, produced rare tone errors but showed very unstable capacities for producing affricates and exhibited a complete lack of [l] in the segmental inventory.

Explanations for the early acquisition of tone have been proposed from perceptual, functional and physiological perspectives. Tse (1978) suggested that the perceptual saliency of tone as compared to the segmental aspects of speech can be used to account for the rapidity and ease of tonal acquisition. In contrast, Li and Thompson (1977), Zhu & Dodd (2000) and Zhu (2002) regard the degree of complexity of contrasts and rules involved in phonological acquisition, i.e. phonological saliency, as a possible determinant of the chronological order of tonal and segmental acquisition. Zhu pointed out that the acquisition of tone may be earlier due to its greater capacity for differentiating lexical meaning and fulfilling children's communicative intentions. Allen and Hawkins (1980) suggest that physiological correlates could be used to account for the difference in developmental schedules for tone and segmental accuracy: tonal contrasts are signaled largely by changes in fundamental frequency originating in the larynx while segmental contrasts involved a wide array of articulatory processes using glottal, sub-glottal and durational mechanisms.

The literature in this area is typified by individual and small sample case studies, precluding generalizations about the course of acquisition. Other large scale studies (Zhu & Dodd, 2000; Zhu, 2002) are descriptive rather than presenting quantitative data that has been evaluated statistically. Claims about chronological order of tonal and segmental acquisition have been made based on qualitative observational data by making reference to the time when tones and segments were considered as "acquired" (where "acquired" was not precisely defined). The developmental relationships between tone and segmental production have been only generally characterized in this literature.

As most tone languages have more segments than tones, the number of lexical items bearing each of the tones is far greater than the number of lexical items containing each of the segments. The functional load of tone-word mapping is therefore much higher than that of segment-word mapping. Since the accuracy of either tones or segments is manifested by how well a child can attach a correct phonetic form to the corresponding lexical item, when children need to make decisions among more potential tone-word pairs than segment-word pairs, the higher functional load of the former might result in correspondingly lower tone accuracy as compared to segmental accuracy. A pilot case study (Yang & Davis, 2008) suggested that segmental accuracy is higher than tone accuracy between 12 and 27 months. The present study statistically evaluates the relationships between tone and segmental accuracy in a larger longitudinal database with more participants.

2.4. Tone accuracy and word-level variables

2.4.1. Acquisition of segment and word

In studies focusing on segmental acquisition in earliest periods of speech development, "word" has been regarded as playing a central role. Word learning provides children with the experience of producing or attempting to produce a variety of segments, syllables and word shapes, which potentially enhance children's speech output accuracy and facilitates the emergence of adult-like phonological knowledge. At the initial stages of acquisition, the whole word has been proposed as the basic production units for the child and production accuracy for segments is described as varying at a word level. (Beckman & Edwards, 2000; Ferguson & Farwell, 1975; Lindblom, 1992; Menn, 1976; Stoel-Gammon & Cooper, 1984). Two potential sources to account for word-level variation have been proposed: Lexical factors and phonetic factors.

2.4.1.1. Lexical factors

Consideration of lexical factors involved in phonological acquisition has been derived from findings on language processing in adults. This body of research emphasizes the role played by the lexical and sub-lexical patterns in the ambient language. Two lexical level factors have been considered: (1) **Word frequency**, which refers to the occurrence of a specific lexical item in the ambient language; (2) **Neighborhood density (ND)**, which characterizes the similarity across words. ND is calculated as the number of words that differ from the target by one phoneme only or the number of contrastive minimal pairs of the target (Luce & Pisoni, 1998). Previous findings suggested that these two factors affect word processing in both perception and production in adults. High-frequency words are associated with faster word recognition and are produced more quickly and accurately (e.g. Ellis, 2002). A segment is more likely to be produced accurately in higher frequency words than in lower frequency words (Gierut, Morrisette, & Champion, 1999). Regarding neighborhood density, dense neighborhood tends to inhibit word recognition and word production in adults because of the competition effects among phonologically similar forms (e.g. Luce & Pisoni, 1998). There are also studies suggesting a facilitative effect of neighborhood density on word production (e.g. Vitevitch & Sommers, 2003).

The role of word frequency and neighborhood density involved in phonological acquisition, especially in speech production of children is still being established. Most studies have considered the potential effects of word frequency and neighborhood density on segmental accuracy in children aged four and older. The majority of the studies have used non-word repetition tasks, where non-words are constructed according to controlled lexical factors and child productions are elicited to consider production accuracy. Non-words from dense neighborhoods that contain frequent sound sequences were reported to be produced more accurately than non-words from sparse neighborhoods (Beckman & Edwards, 2000; Edwards, Beckman, & Munson, 2004. Zamuner, 2009). This controlled experimental approach allows investigators to manipulate lexical variables and has yielded quite uniform results regarding the role of lexical factors in segmental accuracy. However, production of non-words reflects different processes than those involved in the production of real words. For example, production of real words requires a child to retrieve a stored representation from long-term memory based on their production system capacity without benefit from a repetition task (Hoff, Core, & Bridges, 2008).

Available studies of the effect of frequency and neighborhood density on the production of real words suggested that dense neighborhoods facilitate production. In a longitudinal study of fifteen children from 24 to 29 months, Sosa (2008) found that words from dense neighborhood were produced more accurately and with less variability than words from sparse neighborhood. On the other hand, the influence of word frequency on production accuracy didn't show a clear pattern. Two types of frequency measures were used in previous studies: one is frequency in children's own production and the other is frequency based on word counts from adult corpora. Tyler and Edwards (1993) calculated the frequency of words used in the speech samples of two children. They found that accurate productions of voiceless stops emerged first in words that were produced more frequently by children. In contrast, in a diary case study by Velten (1943), frequently produced words remained inaccurate longer than less frequent words when a new phonemic contrast entered the child's phonological system. In studies based on adult corpora, Sosa (2008) found that 24-to-29-months children's production is less variable in high frequency words than low frequency words. However, production accuracy did not seem to have clear relationship with word frequency.

In summary, findings from child-based studies of non-words and real-words indicate that children are sensitive to the lexical properties of their ambient language and this is reflected in their own productions. These properties influence their phonological production variability and accuracy. However, methodological issues make comparison across studies and generalization of results difficult (Stoel-Gammon, 2011).

First, the youngest age involved in these studies are 24 months, where children in an English language environment were reported to have an average productive vocabulary of 250-300 words. It is unknown how lexical properties influence production accuracy in children at the onset of word learning where they have an even smaller lexicon (e.g. less than 100 words). Although study of the influence of neighborhood density on the perception of young children (Hollich, Juscyk, & Luce, 2002) suggests that neighborhood effects could be found in individuals with a small vocabulary size, such as in children aged 24 months or under, few studies have examined the effect of neighborhood density the production accuracy in developmentally younger children. This emerging stage of acquisition poses potential questions for the investigation of word-sound relationships. Take neighborhood density for example. Children who just start to build a lexicon produce words that might not have any neighbors. The neighborhood density of these words increases as children expand their vocabulary. We would thus expect the influence on production accuracy to be different from that in children who already have a relatively larger lexicon.

Secondly, word properties determined by word counts from adult use in spoken language and/or text are not the same as a database of children's own productions, especially in younger children. For example, words frequently used by children might not show equivalent frequency in adults. Frequently used adult words might still be absent from children's vocabulary. Lexicon organization will also vary among individual children. Therefore, it would be more informative to include child production database or even words from a particular child as the basis to measure lexical properties.

Third, the majority of these studies are conducted with children learning English. Cross-linguistic studies, especially with languages that have distinct phonological and lexical patterns from English, will expand our understanding of universal and language-specific properties of the interactions between lexical and phonological acquisition. For example, Mandarin Chinese, which uses both segmental and tonal level contrasts in its phonological system, will shed light on these questions from a suprasegmental perspective.

2.4.1.2. Phonetic factors

Articulatory or phonetic complexity and phonetic context in words have been considered as possible phonetic factors to account for word-level variation in accuracy during segmental acquisition. **Articulatory complexity** relates to the relative production difficulty of sound in articulation. Stokes & Surrendering (2005) indicated that articulatory complexity is one of the factors predicting consonant accuracy across-languages. Articulatorily more complex consonants tend to be produced less accurately than articulatorily less complex ones. However, due to the lack of consistent measures, articulatory complexity at a whole-word level has rarely been considered to account for word-level variation. A segment contained in a target word with high Whole-word Articulatory Complexity would be predicted to have lower production accuracy compared with the same segment contained in a word that is articulatorily less difficult.

Phonetic complexity in the early word period has been approached from several vantage points. One recent attempt has been made to quantify phonetic complexity to enable researchers to examine articulatory factors involved in children's phonological acquisition independent of accuracy measures. As the complexity issue in phonetic sciences is still unsolved, it is also premature to consider the measure of phonetic complexity based solely on articulatory or phonatory efforts. The Word Complexity Measure (WCM) (Stoel-Gammon, 2010), implemented for study of English learning children, was designed based on a child's production. Complexity is determined by findings of investigations of phonological acquisition (e.g. Stoel-Gammon & Cooper, 1984; Robb & Bleile, 1994) and developmental norms (e.g. Prather, Hedrick, & Kern, 1975). Later acquired sounds, syllable shapes, and word patterns are counted as more complex compared to earlier phonetic units, although we acknowledge that no independent measure of production ease in acquisition is presently available. We will retain the terminology of "Measures of Phonetic Complexity", but we precisely refer to the timing of emergence of phonemes in the speech development of young children. Late emergent phonemes (i.e., “more difficult”) get more points with the WCM than earlier emergent phonemes, which are considered ‘easier’ by this index. WCM allows researchers and clinicians to compare the phonetic and phonological aspects of productions across children and over time. It can also be used to compare the complexity of a child’s productions with that of the target words.

Another phonetically motivated factor influencing accuracy might be the **phonetic context** in which the target sound occurs. Davis, MacNeilage and Matyear (2005) reported that the same vowel, occurring in the context of different neighboring consonants was related to differences in accuracy. This was based on the Frame/Content Theory (MacNeilage & Davis, 1990) which states that rhythmic mandibular oscillations reflect motoric constraints on young children related to rhythmic jaw movements without capacities for independent tongue control within and across syllables. Three intrasyllabic CV combinations, or “frames,” are said to occur with higher than chance frequencies in babbling (Davis & MacNeilage, 1995) and early words (Davis et al., 2002). Labial consonants occur with central vowels, alveolar consonants with front vowels, and velar consonants with back vowels. Accordingly, children are more likely to achieve higher vowel accuracy when the target vowel occurs in lexical items with a consonant-vowel sequence that fits the predicted CV combinations or ‘frames’ that have been found to be characteristic of babbled syllables. Front vowels are more likely to be produced more accurately with neighboring alveolar consonants. Mid vowels have higher accuracy when produced with neighboring labial consonants. Back vowels showed higher accuracy with neighboring dorsal consonants (Davis, MacNeilage & Matyear, 2005).

In sum, word-level production variations might be influenced by lexical and phonetic properties of the words that children attempt to produce in their earliest lexicon.

2.4.2. Acquisition of tone and word

Previous proposals about the early acquisition of tones has been challenged by recent studies suggesting that tone production still shows large variations after 24 months (Tse , 1993; Wong, Schwartz & Jenkins, 2005). Word-level variations have also been observed: children's mastery of tone in certain words does not predict their ability to produce all words in the same tone categories accurately. Yang (2006) and Yang and Lee (2006) proposed a lexical-dependent hypothesis for early tonal acquisition based on a longitudinal study of the acquisition of Mandarin tone. By examining two Mandarin-speaking children's spontaneous speech data from 12-27 months, the study found that for lexical items within the same tone category, tonal accuracy was not the same over the entire observation period. Tonal substitution patterns for lexical items bearing the same tone differed as well. These results suggested that lexical items bearing the same tone may follow idiosyncratic paths in approaching accuracy relative to lexical targets attempted. However, whether these word-level variations were related to the lexical properties or the phonetic properties of the word were unknown based on these earlier studies.

2.4.2.1. Lexical factors

For the factor of **word frequency**, a pilot study of one child by Yang & Davis (2008) indicated that word frequency in both ambient language and child own production are not reliable predictors for tone accuracy. As tones are almost equally distributed over syllables in adult Mandarin speakers, neighborhood density of tone is equal across words in ambient Mandarin (Duanmu, 2006; Howie, 1976). However, in children's productive lexicons, tones may not necessarily exhibit such an even distribution. Some words may have more contrastive tone minimal pairs than other words. Neighborhood density of tone within each child's own lexicon was termed **Productive Tone Neighborhood Density (PTND)**. According to preliminary pilot results (Yang & Davis, 2008) for a Mandarin child ages 12 to 27 months, within the same tone category, words that have smaller number of contrastive tone minimal pairs or with fewer number of neighbors in child's productive lexicon were produced more accurately with regard to tone than words with a larger number of contrastive minimal pairs. Words with low PTND may be perceptually more salient and cognitively less demanding for children to map their abilities for producing tones to meaning. Additional support for the effect of neighborhood density on tone accuracy comes from an artificial tone learning system. Tsay (2001) found that an artificial language with only one tone level was much more easily acquired than a language with more than one tone levels.

Developmentally, when children start to acquire tone in a systematic manner in the later acquisition stages, where tones are considered as phonological category, the influence of PTND on tone accuracy would become less strong. Therefore, children's accurate production of tone can be viewed as not necessarily determined by pre-existing phonological knowledge but through a constant dynamic process of refinement of production capacities with the expansion of an individualized lexicon.

2.4.2.2. Phonetic factors

Similar to phonetic factors in the acquisition of non-tone languages, **articulatory complexity** might be a possible aspect of importance in determining word-level variations in tone accuracy. For articulatory complexity, within the same tone category, words that are more complex in articulation requirements might have lower tone accuracy (e. g. [l] requires more complex tongue positioning than [d]).

As in non-tone languages, phonetic context might also potentially contribute to word-level variation in tone accuracy. Tone languages encode word meaning by both segmental level (articulatory) and tonal level (phonatory) contrasts. Thus we will consider **segmental accuracy** as the phonetic context for tone production. Specifically, different words within the same tone category produced more accurately in terms of segments will show higher accuracy in tone compared to those produced less accurately in terms of segments. Alternatively, a trade-off relationship between the two aspects of production output would predict that segmental level accuracy might be achieved at the cost of tone accuracy or vice versa.

However, tones are realized by pitch changes implemented through vocal fold vibration rate in the laryngeal system, while segments are produced by manipulation of relative position of articulators in supra-laryngeal system. In mature speech production by human adults, phonation and articulation processes are physiologically independent. If such independence exists in young children, we would not expect tone accuracy be influenced by segmental level factors on lexical items. On the contrary, it might also be possible that phonation and articulation have not yet achieved adult-like autonomy. Therefore, segmental level influences on tone would be in effect. Yang and Davis (2008) found that tone and segmental production were relatively independent in one Mandarin-learning child, as there was no correlation between tone and segmental level accuracy. Neither tradeoff nor facilitation occurred relative to segmental accuracy and tone accuracy.

Summary

Studies on the acquisition of tone languages are relatively sparse in comparison to studies of non-tone languages. How children manage available phonatory and articulatory capabilities to accurately produce tones and segments in Mandarin relative to the requirements of contrastiveness in early words have not been considered comprehensively. Considering tone acquisition in relation with segment acquisition and word learning, the present study may help to address basic questions on the nature of phonological acquisition of a language which requires implementation of both articulatory and phonatory adjustments for contrasting meaning in the lexicon. As cross-linguistic comparison of the mechanisms of phonological acquisition of segmental and suprasegmental properties requires understanding of diverse language trajectories in languages with and without tonal dimensions, the research can help to understand the underlying components driving the acquisition of phonological complexity in the earliest stages of development by young humans.

Chapter 3: Methods

3.1. Research design

As the proposed study aims to investigate time-related individual and group level change in tone accuracy and possible word-related effects in relationship to such changes, a longitudinal design will be used. The dependent variable is tone accuracy and the predictors will be grouped into lexical and phonetic variables. Specific variables will be described in the section 3.6. *Definition and Measurement of Variables*.

3.2. Participants

Participants were four typical developing functionally monolingual Mandarin-speaking children, three boys and one girl. The participants were located by informal referral from the surrounding community in the Austin, Texas area. All children were taken care of by Mandarin-speaking parents, grandparents, or nannies, but were not going to English daycare centers during the time period when the data were collected. Mandarin Chinese was the primary language spoken at home**.** Since the phonological characteristics of Taiwan Mandarin are substantially different from Mandarin spoken in Mainland China as described above, all families were originally from the north part of Mainland China. According to parent report and medical records shared by the parents, the children had no reported sensory, speech, motor or cognitive disorders. The study was approved by the Institutional Review Board of the University of Texas at Austin for proper involvement of human subjects prior to the start of data collection. Informed consents were obtained from parents according to IRB regulations.

3.3. Data collection

Children were observed at home when interacting with caretakers and/or the author at bi-weekly intervals from 12 months to 24 months. A digital recorder (Roland Edirol R-09HR) was used to record spontaneous speech in wave format at a resolution of 24 bit and a sampling rate of 44.1 kHz. A mini microphone (SONY ECM-C15) was attached to the child's clothing approximately 2 inches below the mouth to ensure adequate signal to noise characteristics in the home environment. The recorder was placed in a baby belt-bag carried by the child in order not to limit mobility. Child and caretakers interacted as in their usual daily routine with no purposeful elicitation. The interaction episodes were also video-recorded in case audio data alone were not sufficient to determine the intent and content of productions during transcription. Audio- and audio-visual recordings lasted sixty minutes for each session under normal circumstances.

3.4. Data extraction, inclusions and exclusions

Continuous speech data were segmented and recognizable words were extracted into separate files using the software Adobe Audition. According to Vihman (1996), a recognizable word is considered as either (1) a form recognized as a word by caretaker(s) involved in the interaction, i.e. adult repetition of child's vocalization with the corresponding adult form; or (2) a form suggested by the context, i.e. the context strongly suggested the corresponding adult form of the child's vocalization and no other. Unclear words, exclamatory words, and onomatopoeic words were not extracted for analysis. The following instances of productions were extracted and transcribed for further studies but were excluded from the analysis due to production variability of adult input. (1) 2,969 utterance tokens with reduplicated words (e.g. "鸟鸟" [niau3] [niau 0/2/4] *birdie*) were excluded and (2) 227 token of the word "好" [xau], meaning "OK, alright". Both spontaneous and imitated vocalizations were analyzed.

3.5. TRANSCRIPTION AND RELIABILITY CALCULATION

The recognizable words were perceptually transcribed by the author, a native speaker of Mandarin. For each word token, International Phonetic Alphabet (IPA, 2005) was used to transcribe segments. Tones were transcribed in terms of pitch contours (Level, Rising, Fall-Rise or Falling). Within each contour category, pitch height will also be specified as High, Mid or Low. See Appendix 1 for sample transcription page.

15% of word tokens in each session for each child were selected randomly to be transcribed by a second transcriber, a native speaker of Mandarin Chinese familiar with transcription of child speech but blind to the hypotheses of the present study. Transcriber agreement was calculated by using a point-to-point agreement method (Shriberg and Lof, 1991). Transcriber agreement of each word token was calculated as the percentage agreement (P) between the two transcribers (T1 and T2).

$$\text{Percentage agreement } (P\,(T1-T2)) = \frac{\text{Number of agreements}}{\text{Number of disagreements} + \text{number of agreements}} \times 100\%$$

Transcription reliability is the average of the transcriber agreement on each word token analyzed. Table 3 illustrates the transcription format. Transcriber 2 will not be able to see the transcription of Transcriber 1.

Table 3: Illustration of Transcription of Segments and Tones.

Word token (gloss)	Transcriber 1			Transcriber 2		
	Segments	Tone		Segments	Tone	
		contour	height		contour	height
妈 (*mother*)	[ma]	L	H	[ma]	L	H
抱 (*hug*)	[pau]	F	H	[p^hau]	F	H
包 (*bag*)	[p^hau]	L	H	[p^hau]	F	H

Based on the examples in Table 3, the percentage agreement of the segments in each word token was calculated as below:

Word tokens	Segment	Tone
ma (mother):	$PS(T1\text{-}T2) = \frac{2}{2} \times 100\% = 100\%$	$PT(T1\text{-}T2) = \frac{2}{2} \times 100\% = 100\%$
bao (*hug*):	$PS(T1\text{-}T2) = \frac{2}{3} \times 100\% = 67\%$	$PT(T1\text{-}T2) = \frac{2}{2} \times 100\% = 100\%$
bao (*bag*):	$PS(T1\text{-}T2) = \frac{3}{3} \times 100\% = 100\%$	$PT(T1\text{-}T2) = \frac{1}{2} \times 100\% = 50\%$

$$\text{The segment transcription reliability} = \frac{100\% + 67\% + 100\%}{3} = 89\%$$

$$\text{The tone transcription reliability} = \frac{100\% + 50\% + 100\%}{3} = 83.3\%$$

The transcription reliability for segment is 85.77%. The reliability for tone is 89.47%.

3.6. DEFINITION AND MEASUREMENT OF VARIABLES

3.6.1. Dependent variable: Tone Accuracy (TA)

Pitch realization is dependent on the phonetic or phonological contexts, especially for the phenomenon of Tone 3 sandhi in Mandarin (Chao, 1968). A Tone 3 word becomes a rising contour before another Tone3 word and a low-falling contour before words with any of the other tones in Mandarin. A full fall-rise contour for Tone 3 only occurs when the word is produced in isolation. Most of Tone 3 words are realized as low-falling contours. Therefore, for the calculation of tone accuracy, a low-falling contour for a target Tone 3 word was considered as correct. Both high- and mid-falling contours were considered as accurate realizations of target Tone 4 words. For Tone 1 and Tone 2, however, only the pitch contour but not pitch height is phonemically distinctive. Even though there are variations of F_0 contours due to the carry-over effect of the adjacent syllables, speakers tend to preserve the F_0 contour type for perceptual intelligibility, especially the offset of the time phrase of tone production (Xu, 1997). Therefore, for Tone 1 and Tone 2, no matter the utterance position, if the transcribed tone production matched the tone contour of the intended adult target, it was considered as accurate.

These criteria have been used in Yang (2006), Yang & Lee (2006) and Yang & Davis (2008). Other studies on acquisition of Mandarin Tones also use tone contour as the criterion to determine accuracy (Li & Thompson, 1977; Zhu & Dodd, 2000), though methods have not been explicitly described. As illustration, for a Mandarin Tone 1 word "猫" [mau 1] *cat*, transcriptions of the contour as high-level (LH), mid-level [LM] or low-level (LL) pitch in children's production were considered as accurate. For a given instance of a word, when tone is produced accurately, it will be coded as 1. When the tone is produced inaccurately, it will be coded as 0.

Two types of tone accuracy were calculated. (1)***Within tone accuracy*** was the average tone accuracy across all words within the same tone category in each session (2) ***Within word accuracy*** was the average tone accuracy across all the tokens of the same word type in each session.

Since the effect of utterance position on tone production accuracy in young children is still unknown, it is premature to average tone accuracy scores over all utterance positions. As a result, tone accuracy was calculated separately with regard to utterance positions (isolation, utterance initial, utterance medial and utterance final). Words produced in isolation only apply to mono-syllabic words. Initial and final position refer to syllables that occur respectively at the beginning and end of utterances with more than one syllable. For a tri-syllablic utterance, the syllable at the middle of the utterance was considered as occupying the medial position. For utterance of more than three syllables, utterance medial refers to all syllables that were between the initial and final position.

The average tone accuracy (Within-*tone* or within-*word*) was based on a weighted measure of accuracy, with the weights being the percentage of each type of utterance position over the total number of certain tone or word tokens. The examples below illustrate the calculations.

Within tone accuracy of a given utterance position =

$$\frac{\text{Number of words that is accurately produced in a given utterance position}}{\text{Total number of words in that utterance position}}$$

Table 4 below gives example data illustrating potential distribution of words of a tone category, according to utterance positions and the accuracy of each utterance position.

Table 4: Example Data Illustrating Potential Distribution of Words of a Given Tone.

		Number of Tone 1 words	Tone Accuracy (%)
Monosyllabic utterance	Isolation	90	97
Utterances with more than one syllable	utterance initial	10	89
	utterance medial	20	80
	utterance final	30	95

$$\text{Position Weighted Within - Tone Accuracy} = \frac{90}{150}\times 97\% + \frac{10}{150}\times 89\% + \frac{20}{150}\times 80\% + \frac{30}{150}\times 95\%$$

$$= 93.8\%$$

Within word accuracy of a given utterance position =

$$\frac{\text{Number of tokens of a word that is accurately produced in a given utterance position}}{\text{Total number of tokens of the word in that utterance position}}$$

Table 5 shows example data illustrating a potential distribution of tokens of a given word according to utterance position and the accuracy of each utterance position.

Table 5: Example Data Illustrating Potential Distribution of Tokens of a Given Word.

		Number of tokens of word X	Tone Accuracy (%)
Monosyllabic utterance	Isolation	15	95
utterance with more than one syllable	utterance initial	5	85
	utterance medial	2	83
	utterance final	8	90

$$\text{Position Weighted Within - word Accuracy} = \frac{15}{30}\times 95\% + \frac{5}{30}\times 85\% + \frac{2}{30}\times 83\% + \frac{8}{30}\times 90\%$$

$$= 91.2\%$$

3.6.2. Predictors

3.6.2.1. Phonetic variables

Segmental Accuracy (SA)

Similar to the determination of accurate production of tone, segments (consonants or vowels) was judged as being produced accurately when they match the identified target word. When a certain segment was produced accurately, it was coded as 1. Otherwise, it was coded as 0. Whole-word accuracy (Ingram, 2002) measurement was adapted to quantify segmental level accuracy. Whole-word segmental accuracy was calculated as the average accuracy of segment within a particular token of a word, as in the formula. Whole-word segment accuracy of all attempted words in each session was averaged to obtain the mean whole-word segment accuracy of that session.

$$\textbf{Whole-word Segment Accuracy of a word token X} = \frac{\text{Number of accurately produced segments in X}}{\text{Total number of segments in X}}\times 100\%$$

The example below illustrates the calculation of whole word segment accuracy of the child production of the word "脏" [tsaŋ 1] *dirty).* There are 3 segments in the word, where only the vowel [a] was produced correct as the adult target. The initial consonants and final nasal were not accurate. Therefore the segment accuracy of the whole word is indicated:

	Segment Accuracy
Adult target [tsaŋ] Child production [tan]	$= \frac{1}{3} \times 100\% = 33.3\%$

<u>Phonetic Complexity</u>

We adapted the Word Complexity Measure (Stoel-Gammon, 2010) termed the Mandarin Word Complexity Measure (MWCM) to quantify the phonetic complexity of words in actual production (MWCM_A) and of target word (MWCM_T) in Mandarin. The scoring criteria in the Mandarin version are also reflections of the order of emergence in relation to the complexity of articulation based on available results on phonological development of Mandarin-speaking children (Hsu, 1996; Yang, 2006; Zhu & Dodd, 2001).The description below illustrates the coding schema of MWCM:

Word patterns

(1)Productions with more than two syllables receive 1 point.

Syllable Structure

(2)Productions with a syllable-final consonant (n, ŋ) receive 1 point.

Sound Classes

Consonant

(3) Productions with a velar consonant receive 1 point for each velar.

(4) Productions with fricative receive 1 point for each fricative.

(5) Productions with affricate receive 2 points for each affricate.

(6) Productions with aspirated consonant receive 1 point for each aspirated consonant.

(7) Productions with liquid (l,r), syllabic liquid(ɹ̩) receive 1 point for syllabic consonants.

Vowels

(8) Productions with diphthong receive 1 point for each diphthong.

(9) Productions with triphthong receive 2 points for each triphthong.

(10) Productions with rhotic vowel (ɚ) receive 1 point for each rhotic vowel.

Tone

(11) Productions with Tone 2 (rising) receive 1 point.

(12) Productions with Tone 3 (low fall-rise, low fall) receive 2 points.

Phonetic complexity of lexical target (MWCM_T) Using the MWCM, every word then receives a “complexity value” derived from summing the number of points awarded for the word. Each target word was scored for phonetic complexity using the MWCM. As MWCM_T quantifies the phonetic complexity of the standard adult pronunciation of a word, the MWCM_T value for a specific word stays constant over time. The mean MWCM_T values were then computed for each session. As children’s attempted words may vary across sessions, the mean MWCM_T may also vary. Take the word "七" [tɕʰi 1] *seven* for example. [tɕʰ] is an affricate, thus obtained 2 point for criterion 5. Since it is also aspirated, 1 point was assigned for criterion 6. Therefore, the total MWCM_T for the word "七" is 3. Table 6 demonstrates the scoring procedures for MWCM_T by attempted word targets taken from GXH's data. For the set of 5 words, the total MWCM score for attempted target is 9 with an average of 9/5= 1.8. A sample coding of MWCM_T is provided in Appendix 2 Table 1.

Table 6: Demonstration of MWCM_T Calculations.

Word target	Transcription of segment target	Total MWCM_T	Criteria obtaining points
[不] (*no*)	[pu]	0	0 pt
[猫] (*cat*)	[mau]	1	Diphthong (1 pt)
[棒] (*stick*)	[paŋ]	2	Final ŋ (1 pt)
[六] (*six*)	[liou]	3	liquid (1pt)+ triphthong (2 pts)
[七] (*seven*)	[tɕʰi]	3	Affricate (2pts) +aspiration (1 pt)

Phonetic complexity of actual production (MWCM_A) By using the same procedure as above, children's actual word productions were also analyzed for phonetic complexity using the MWCM. For example, if children produced the target word "看" [kan 4] *look* as [ka 4]. The MWCM_A was calculated as 1 since the only point assigned is for the dorsal place of articulation of the sound [k]. The mean MWCM_A values were computed for each session and all renditions of each target word. Since MWCM_A depends on how the children actually produce the target word, it may change across sessions as children's production varies on different occasions. Table 7 illustrates the scoring procedures for MWCM_A produced by GXH. For this set of words, the total MWCM score for actual production is 4 with an average of 4/5= 0.8. A sample coding of MWCM_T is provided in Appendix 2 Table 2.

Table 7: Demonstration of MWCM_A Calculations.

Word target	Transcription of segment actual	Total MWCM_A	Criteria obtaining points
[不] (*no*)	[pu]	0	0 pt
[猫] (*cat*)	[ma]	0	0 pt
[棒] (*bravo*)	[puo]	1	diphthong (1 pt)
[六] (*six*)	[la]	1	liquid (1 pt)
[七] (*seven*)	[tɕi]	2	Affricate (2 pts)

3.6.2.2. Lexical variables

Child Output Frequency (COF)

Child Output Frequency (COF) was calculated as the number of tokens in which a lexical item is produced in each session. This variable is time variant, as in different sessions, children's output might not be the same.

Productive Tone Neighborhood Density (PTND)

Although neighborhood density of tone is equal across words in Mandarin (Duanmu, 2006; Howie, 1976), in children's productive lexicons, tones may not necessarily exhibit such an even distribution over syllables in lexical items they actually produce. Therefore, we employed Productive Tone Neighborhood Density (PTND) as the counterpart of ambient neighborhood density for segments. For a given word, the number of word that differ from it only in terms of tone in children's production repertoire is defined as PTND of that word. For example, if there were three words “旗” [tɕʰi 2] *flag*, “起” [tɕʰi 3] *rise*, “气” [tɕʰi 4] *air* that differ from the lexical item “七” [tɕʰi1] *seven* only in terms of tone, the PTND of “七” will be 3. As children are constantly expanding their lexicon, PTND will not stay constant over time. In the present study, PTND in each session is calculated to address the time-sensitivity of this variable.

Ambient Word frequency (AWF)

An online Modern Chinese Character Frequency Corpus (Da, 2004) was used to search the frequency of target words attempted by children in the ambient language. http://lingua.mtsu.edu/chinese-computing/statistics/char/list.php?Which=MO. The cumulative frequency in percentile counts were used. This variable is time invariant as it is based on properties of Mandarin that are stable in the input.

All the variables were calculated for each word token before deriving average values. Due to the large number of word tokens, the coding and calculations were performed using a customized computer program. 1% of the data were also manually coded and calculated to ensure accuracy. A sample table with calculated values of phonetic and lexical variables is provided in Appendix 3.

3.7. Data analysis and expected outcomes

Data analysis and statistical procedures to evaluate the hypotheses related to study aims will be described in this section.

Aim 1. Examine the relationship between tone accuracy and segmental accuracy over time.

Hypothesis: Children will produce segments more accurately than tones. Tone and segmental production are relatively independent.

The first aim was to examine the relationship between production accuracy for tones and segments in words attempted by the four children. Three types of analysis were carried out. First, tone and segmental accuracy values were compared over time in these children's words in each tone category. The purpose of the analysis was to evaluate claims in previous studies that tone is produced more accurately than segments before 24 months and to explore whether these relationship for accuracy change over time. Paired Sample T-Tests were used to evaluate the differences between tone and segmental accuracy.

Second, we compared segmental accuracy when tone was produced accurately and when it was produced inaccurately. The purpose of this analysis was to examine whether there was a facilitation or trade-off effect between tone and segmental production. Segmental accuracy might scaffold accurate tone production. On the contrary, tone production accuracy might be achieved at the sacrifice of segmental accuracy. Paired Sample T-tests were used to statistically evaluate the differences between segmental accuracy in these two contexts.

Third, whether and how tone and segmental production were related was further examined by comparing production accuracy of the four tones and segmental accuracy of words with the four tones. If tone and segmental production are inherently related, tone accuracy ranking would be related to rankings of segmental accuracy across tone categories. One-way Analysis of Variances (ANOVAs) were performed to evaluate the effect of tone category on tone and segmental accuracy.

Aim 2. Examine the relationship between tone accuracy and word-level variables over time.

Hypothesis: Tone accuracy of a word will be negatively correlated to the number of contrastive tone minimal pairs in a child's lexicon. Words with a smaller number of contrastive tone minimal pairs will have more chance to be produced accurately in tone than words with a larger number of contrastive minimal pairs. Over time, the coefficients of the correlation will decrease as children expand their lexicon.

The second aim was to explore what factors influence tone accuracy variations among lexical items and how these influences change over time. The dependent variable was within-word tone accuracy. The possible predictors were segmental accuracy (TA), word complexity of the target (MWCM_T) and actual production (MWCM_A) as phonetic variables; and child output frequency (COF), productive tone neighborhood density (PTND), and ambient word frequency (AWF) as lexical variables. Analyses were based on five stages divided relative to the vocabulary size of each child. Specifically, Stage 1: 0-50 words; Stage 2: 50-100 words; Stage 3:100-150 words; Stage 4:150-200 words; Stage 5: more than 200 words. Appendix 1 provides detailed information on sessions and chronological ages corresponding to word stages for each child.

Average variable scores when each tone was produced with high or low accuracy were compared. For each child within each word stage, average tone accuracy values of all lexical items within the same tone category were calculated. With reference to the average tone accuracy, lexical items were grouped as high or low in tone accuracy. Average scores of the phonetic and lexical variables based on high-low tone accuracy for each child were then calculated. If tone accuracy is negatively related to children's Productive Tone Neighborhood Density, we would predict a lower average value of PTND when TA was high and a higher average value of PTND when TA was low.

Pearson Correlation Tests were performed to statistically evaluate the relationship between tone accuracy and word-level variables over time. For each child within each tone category, Pearson Correlation Coefficient and corresponding significance between tone and each word level variable were obtained for each word stage. Analytical procedures were performed in Software SPSS (Version 16). As in Hypothesis 2, we predicted a significant negative correlation between TA and PTND (which characterizes the contractiveness of tone in children's lexicon) for each word. Words with higher PTND values should have low TA. Words with low PTND values should have high TA. The coefficient between TA and PTND should decrease over time as children's lexicon expands.

Table 8 summarizes variables and statistical methods used for each aim.

Table 8: Variables and Statistical Methods.

	Variables	**Statistical methods**
Aim 1 **Tone and segment accuracy**	TA (within tone) SA	Paired Sample T-test One-way ANOVA
Aim 2 **Tone accuracy and word-level variation**	TA (within word) *Phonetic variables* SA , MWCM_T, MWCM_A *Lexical variables* COF, PTND, AWF	Correlation analysis

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

Chapter 4: Results

4.1. CHARACTERISTICS OF THE DATA

Across the four participants, 7,092 utterances with 11,246 word tokens were analyzed. Information about participants including gender, chronological ages, number of sessions and vocabulary size for each child is summarized in Appendix 1 (Table 1-4).

4.1.1. Distribution of tone contours in actual production

Figure 3 shows the percentages of occurrence for the four tone contours in the actual production of participants GXH, LXB, MYK and YKX. The actual number of tokens and their percentages of occurrence are listed in Appendix 3 (Table 1-4).

According to the Figure 3, Tone 4 (falling) contours show the highest percentage in the first several sessions for all four participants. Tone 3 (fall-rise) contours are produced with the lowest percentage of occurrence among the four tones. For GXH and LXB, in the first three months of observation, the majority of pitch contours were Tone 4 (falling). They did not produce Tone 3 (fall-rise) pitch contours during these months. For MYK, Tone 4 pitch contours were predominant and there was no Tone 3 pitch contour in the first month of data collection. For all four children, the predominance of Tone 4 pitch contours diminished as they grew older. However, Tone 3 pitch contours continued to have the lowest percentage of occurrence. Overall, as summarized in Table 9, Tone 4 contours show the highest percentage of occurrence and Tone 3 contours the lowest across all sessions observed.

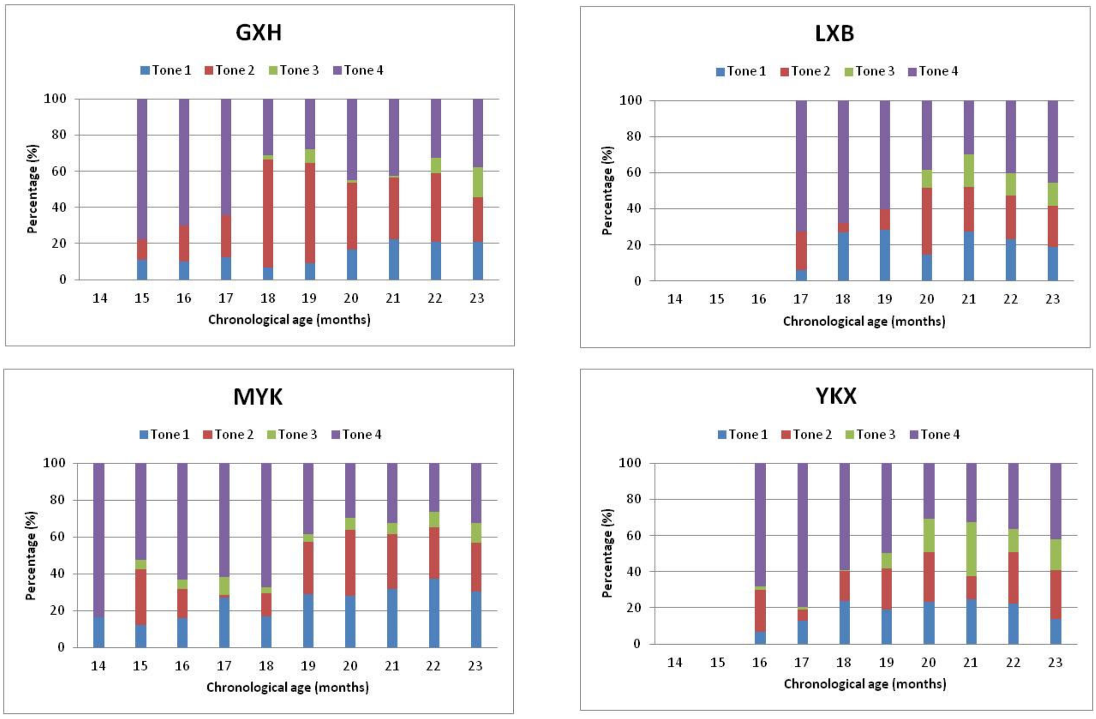


Figure 3: Percentage of Tone 1-4 Contours in Actual Production across Months.

Table 9: Summary of Distribution of Tone 1-4 Contours in Actual Production.

Child	Actual Production								
	Tone 1 (level)		Tone 2 (Rising)		Tone 3 (Fall-rise)		Tone 4 (Falling)		Total Token
	Token	%	Token	%	Token	%	Token	%	
GXH	438	18%	846	35%	196	8%	936	39%	2416
LXB	781	22%	861	24%	444	13%	1461	41%	3547
MYK	1052	31%	948	28%	257	8%	1116	33%	3373
YKX	426	20%	481	23%	323	15%	903	42%	2133

4.1.2. Distribution of tone contours in attempted target words

The percentages of tokens and types of the four tones in participants' attempted target words are illustrated in Figures 4 and 5 respectively. The actual number of word tokens/types and their percentage of occurrence are presented in Appendix 4 and Appendix 5 (Table 1-4).

As in Figures 4 and 5, varied lexical selection patterns were observed in the earlier months of data collection. For GXH, Tone 4 word targets were predominant when he was 15 months of age. During months 15 and 16, Tone 2 word targets were produced at lower percentages in comparison to Tone 1 and Tone 4 word targets. No Tone 3 word targets were attempted. For LXB, the percentage of Tone 4 word targets was initially the highest and decreased as he started to produce more targets from other tonal categories. The percentage of Tone 4 word targets rose again during his last two months. MYK predominantly produced Tone 1 word targets at 14 months of age. Unlike GXH and LXB, MYK did not show a specific preference for Tone 4 word targets. However, his Tone 3 word targets were produced with the lowest percentage among the four tone categories. YKX produced predominantly Tone 2 words compared to words in other tone categories in her 16th month. Tone 3 words are less frequent during the first three months.

Overall, individual variations for selection of word targets with the four tones in the earlier months were found across these four children. All four tended to attempt fewer Tone 3 word targets compared to words in other categories (Tables 12 and 13).

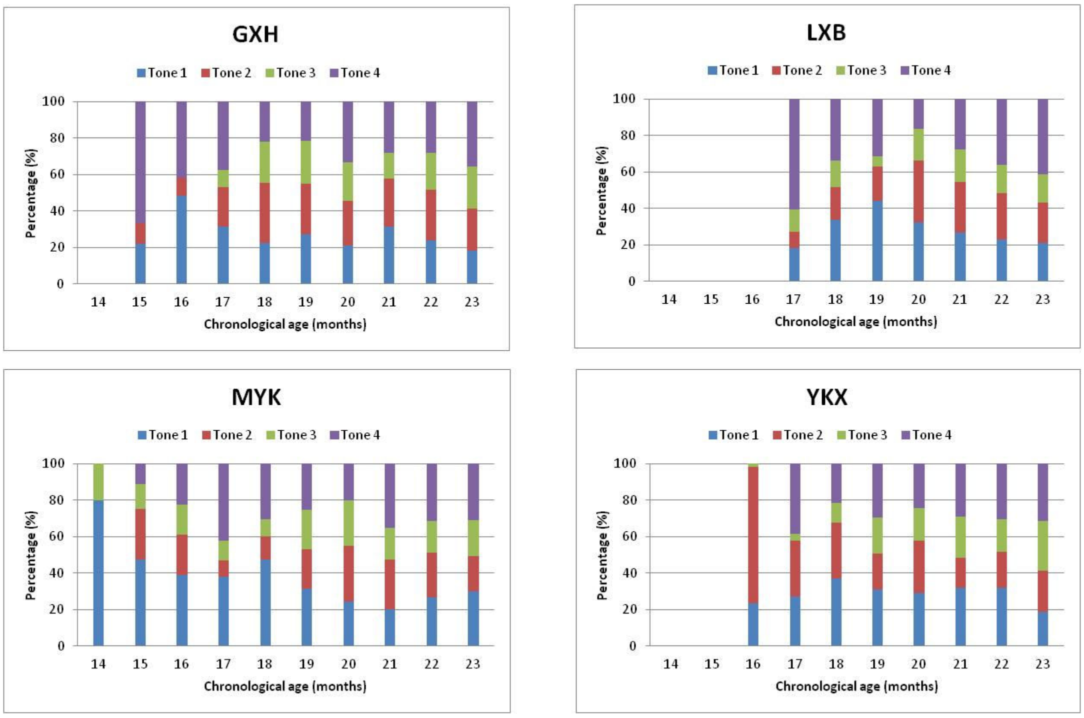


Figure 4: Percentage of Word Tokens of Tone 1-4 in Attempted Targets across Months.

Table 10: Summary of Word Token Distribution of Tone 1-4 in Attempted Targets.

Child	Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Token
	Token	%	Token	%	Token	%	Token	%	
GXH	560	24%	604	26%	470	20%	732	31%	2366
LXB	854	25%	854	25%	545	16%	1147	34%	3400
MYK	907	27%	785	24%	628	19%	981	30%	3301
YKX	626	29%	527	24%	423	19%	609	28%	2185

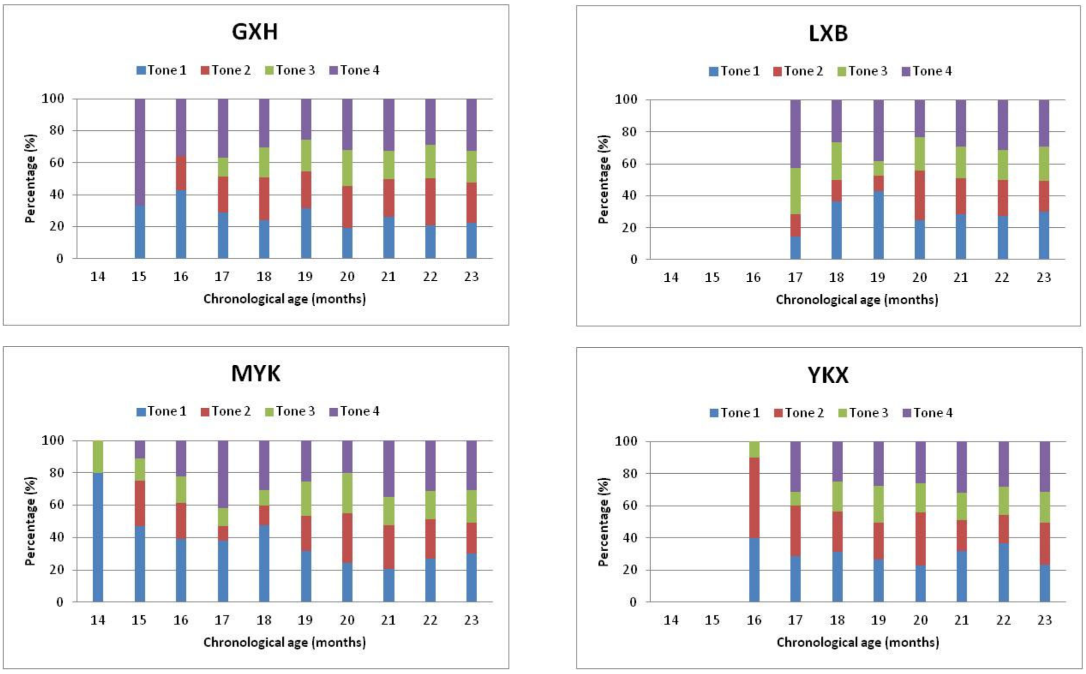


Figure 5: Percentage of Word Types of Tone 1-4 in Attempted Targets across Months.

Table 11: Summary of Word Type Distribution of Tone 1-4 in Attempted Targets.

Child	Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Type
	Type	%	Type	%	Type	%	Type	%	
GXH	168	24%	177	25%	133	19%	221	32%	699
LXB	315	28%	243	22%	221	20%	328	30%	1107
MYK	325	25%	299	25%	222	20%	348	29%	1194
YKX	238	29%	207	25%	153	18%	237	28%	835

Summary of general characteristics of the data

These results indicated that children produce predominantly more words with Tone 4 (falling) pitch contours and less words with Tone 3 (fall-rise) pitch contours. This preference decreased as children grow older. Considerable individual variation existed in the selection of target words across children. Children did not always attempt words with more frequently occurring pitch contours, such as Tone 4, but they tended to avoid words with less frequent pitch contours (Tone 3).

4.2. TONE AND SEGMENTAL ACCURACY

Aim 1 of the present study is to examine the relationship between production accuracy of tones and segments in words attempted. Section 4.2.1 presents results of the comparison of tone and segment accuracy values over time in these children's words in each tone category. Results reported in Section 4.2.2 compared segmental accuracy when tone is produced accurately or inaccurately. Whether and how tone and segmental production are related is further examined by comparing production accuracy of the four tones and segmental accuracy of words with the four tones in Section 4.2.3. The prediction for Aim 1 was that segmental accuracy is higher than tone accuracy. Tone and segmental production are relatively independent.

4.2.1. Comparison of tone and segment accuracy within each tone category

Figures 6, 8, 10 and 12 represent the relationship between segmental and tone accuracy across time for the four participants. Figures 7, 9, 11 and 13 represent the difference between these two accuracy measures over time in the four participants (See Appendix 7 for actual numbers of tone and segmental accuracy and the differences numbers). Each figure contains four panels representing the four tone categories in Mandarin. In the A figures, the Y-axis displays accuracy in percentages and the X-axis displays chronological age in months. Red lines with squares represent segmental accuracy and blue lines with triangles represent tone accuracy. In the B figures, the Y-axis illustrates the difference between segmental and tone accuracy. The X-axis shows chronological age in months. If segmental accuracy is higher than tone accuracy, the orange bar shows a positive value. If tone accuracy is higher than segment accuracy, the orange bar shows a negative value. A taller bar height represents a bigger difference between segment and tone accuracy.

GXH

For GXH, Tone 1 and Tone 3 accuracy is lower than segmental accuracy. Tone 4 accuracy is higher than segment accuracy. The differences between these two accuracy measures decrease as he grows older (Figure 7). However, tone and segment accuracy for Tone 2 words seems close across the period of the study.

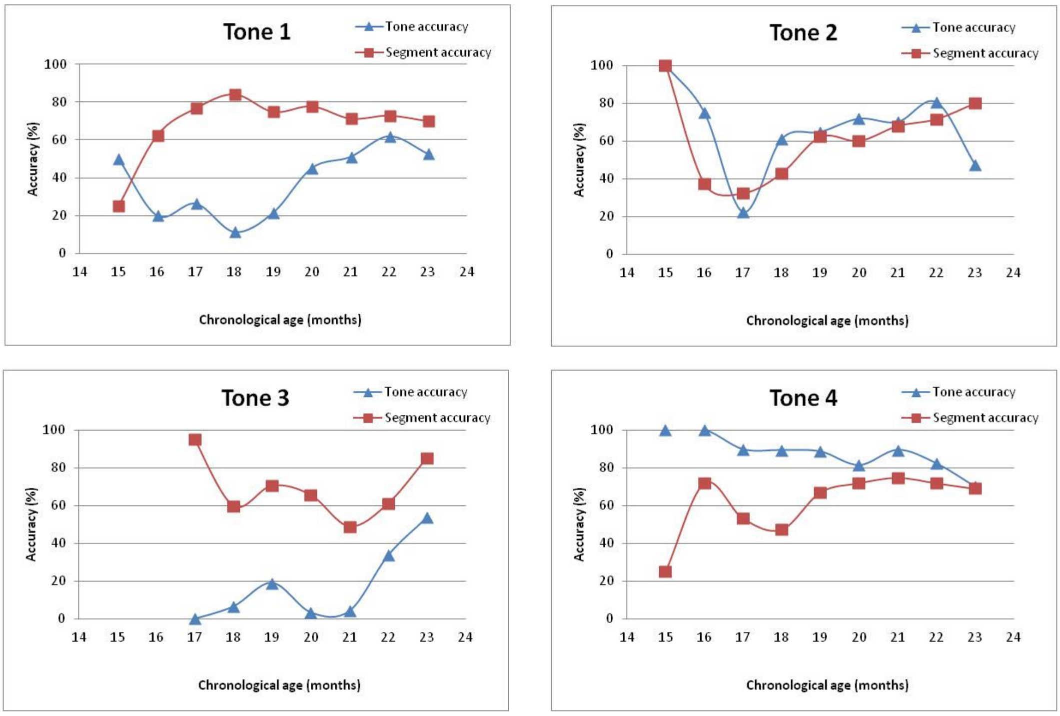


Figure 6: Segment and Tone Accuracy in GXH's Production.

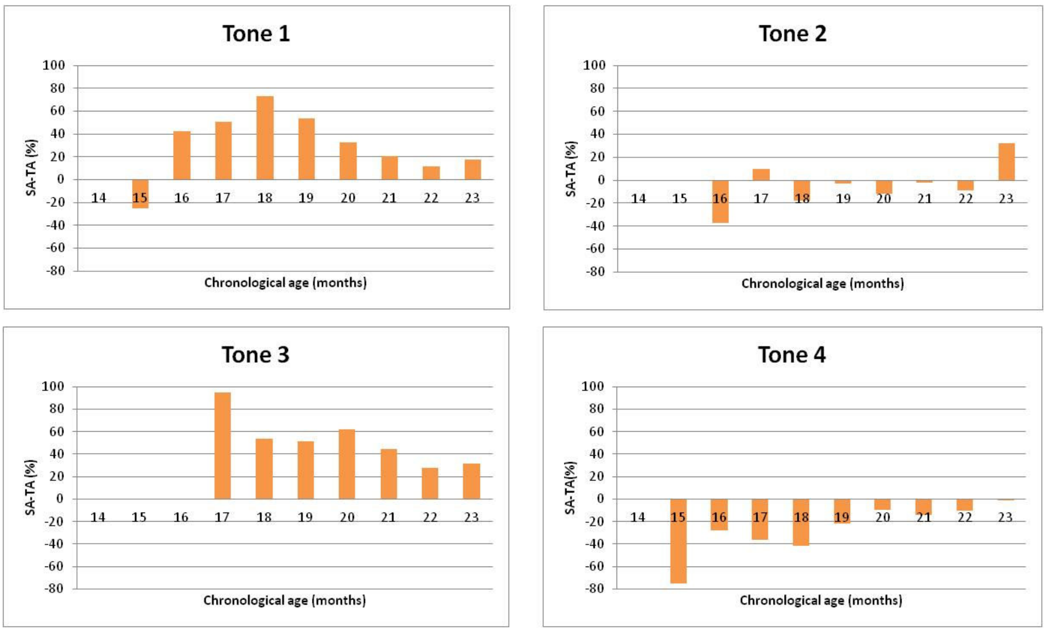


Figure 7: Differences between Segment and Tone Accuracy in GXH's Production.

LXB

For LXB (Figure 8), Tone 2 and Tone 3 accuracy is lower than segmental accuracy while Tone 4 accuracy is higher. The differences between these two accuracy measures also decrease as LXB grows older (Figure 9). For Tone 1 words, there are some months where segmental accuracy is higher than tone accuracy (Months 19 and 20). These two measures did not differ too much for the majority of the months observed.

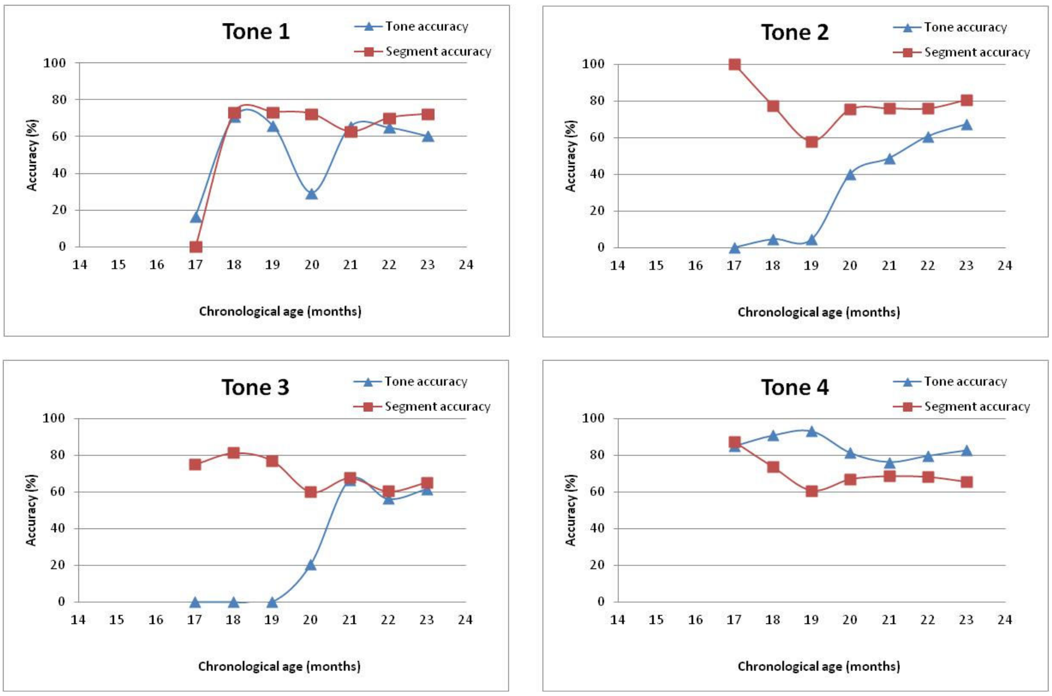


Figure 8: Segment and Tone Accuracy in LXB’s Production.

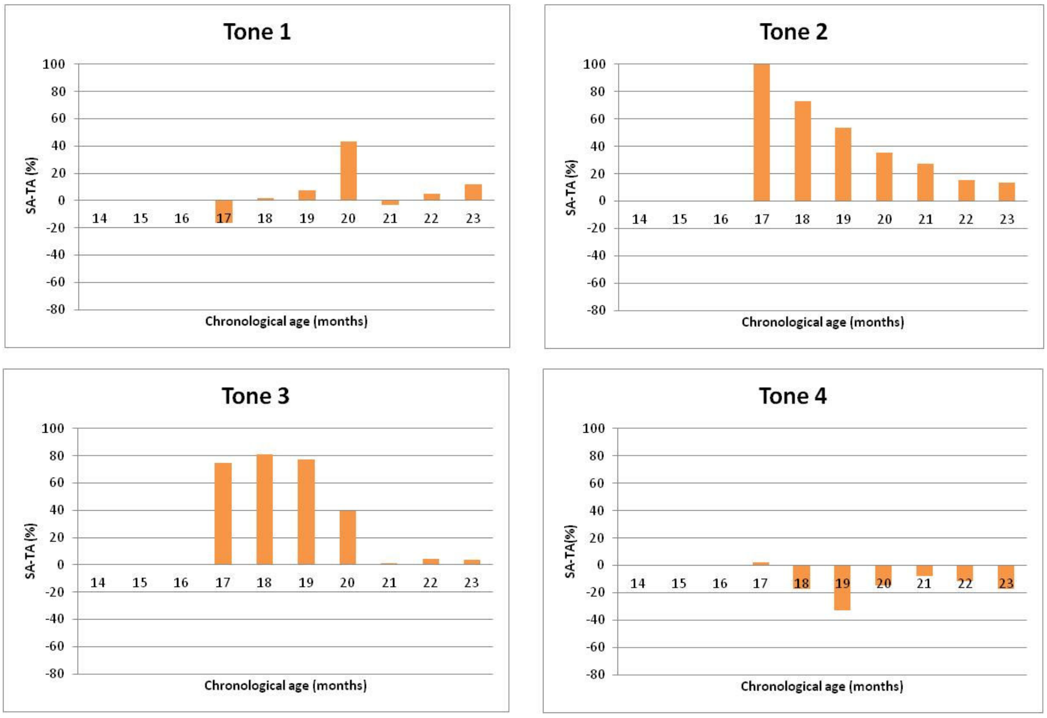


Figure 9: Differences between Segment and Tone Accuracy in LXB's Production.

MYK

Figures 10 and 11 illustrate the developmental changes of segment and tone accuracy in MYK's production and the difference between these two. For Tone 1, segmental accuracy is higher than tone accuracy by Month 19. Tone accuracy is slightly higher than segmental accuracy afterwards. Although the tone and segmental accuracy in Tone 2 words were not as straightforward as in Tone 1 words, their values got closer towards his 20th month. For Tone 3, segmental accuracy was higher than tone accuracy for the entire observation period but the difference between these two get smaller as MYK got older. For Tone 4, tone was higher than segmental accuracy but the differences also decreased over the period of the study.

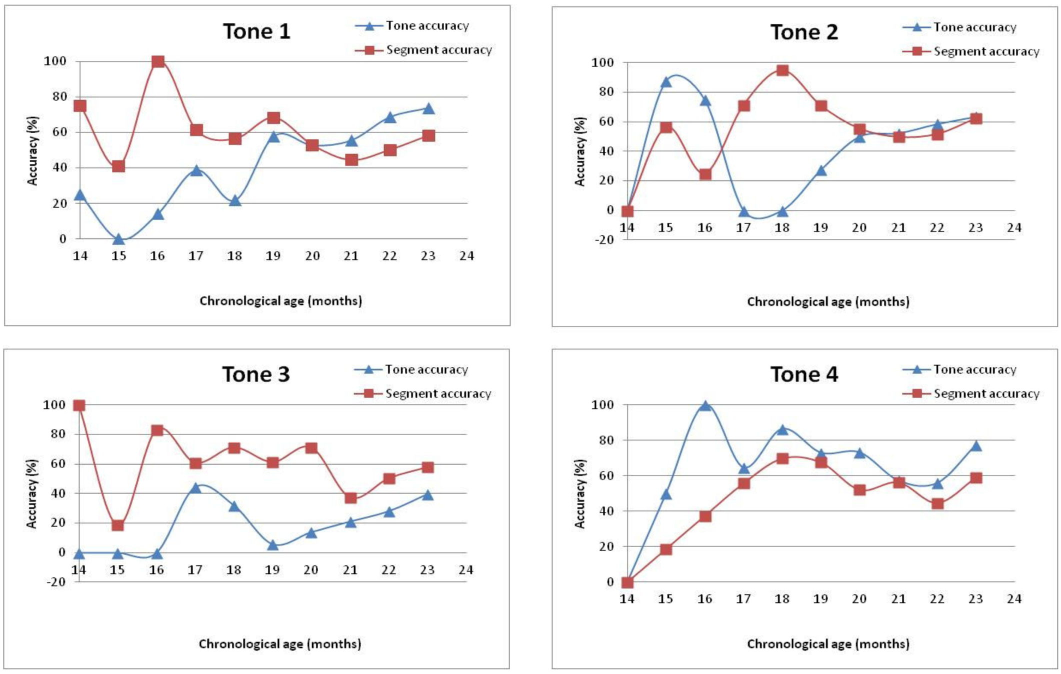


Figure 10: Segment and Tone Accuracy in MYK's Production.

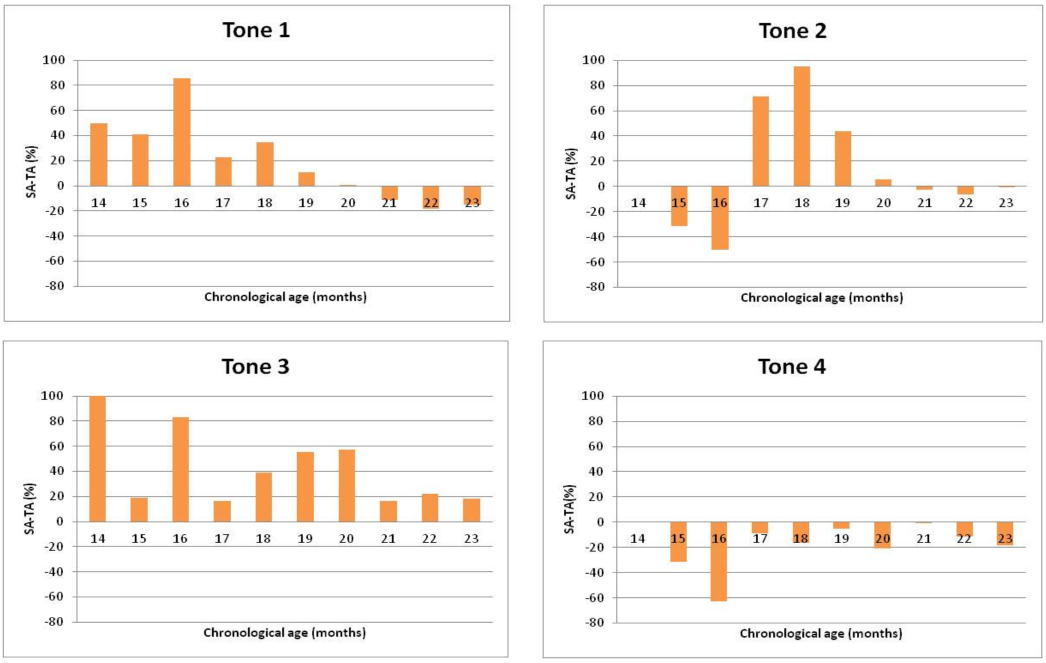


Figure 11: Differences between Segment and Tone Accuracy in MYK's Production.

YKX

The relationship between tone and segment accuracy over time in YKX's production is represented in Figure 12. Figure 13 illustrates the differences between segment and tone accuracy over time. For Tone 1 words, segmental accuracy was higher than tone accuracy by Month 19. For Tones 2 and 3 words, segmental accuracy was higher than tone accuracy before her 20th month. Tone 4 accuracy in YKX's production is higher than segment accuracy over the entire observation period but the differences between these two measures tended to become smaller over time.

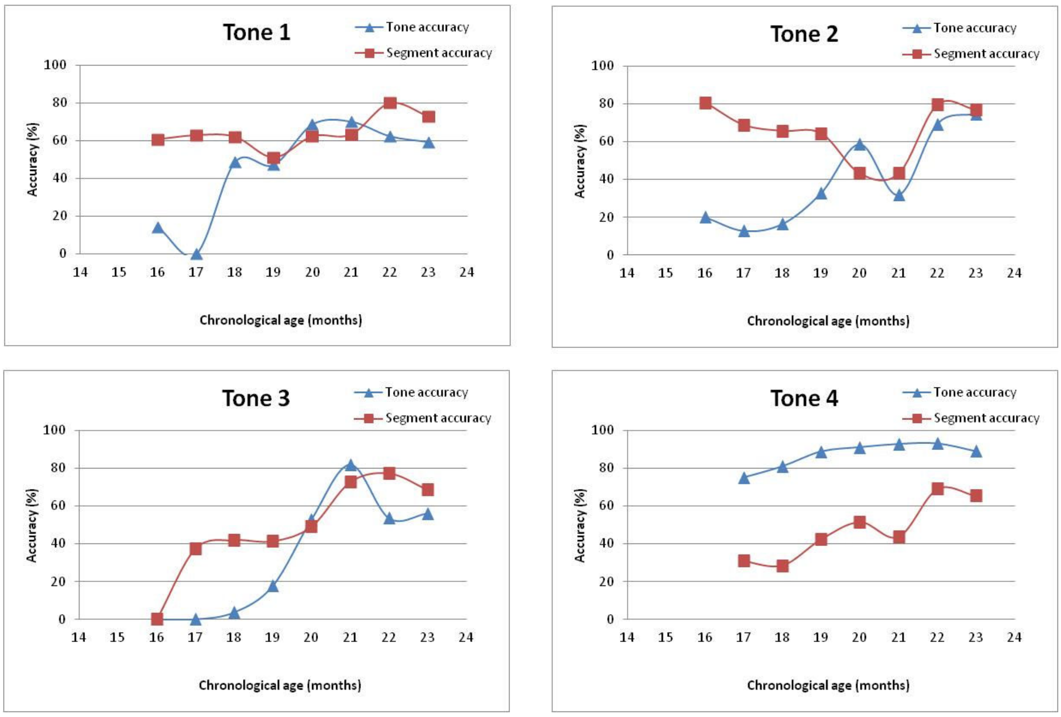


Figure 12: Segment and Tone Accuracy in YKX's Production.

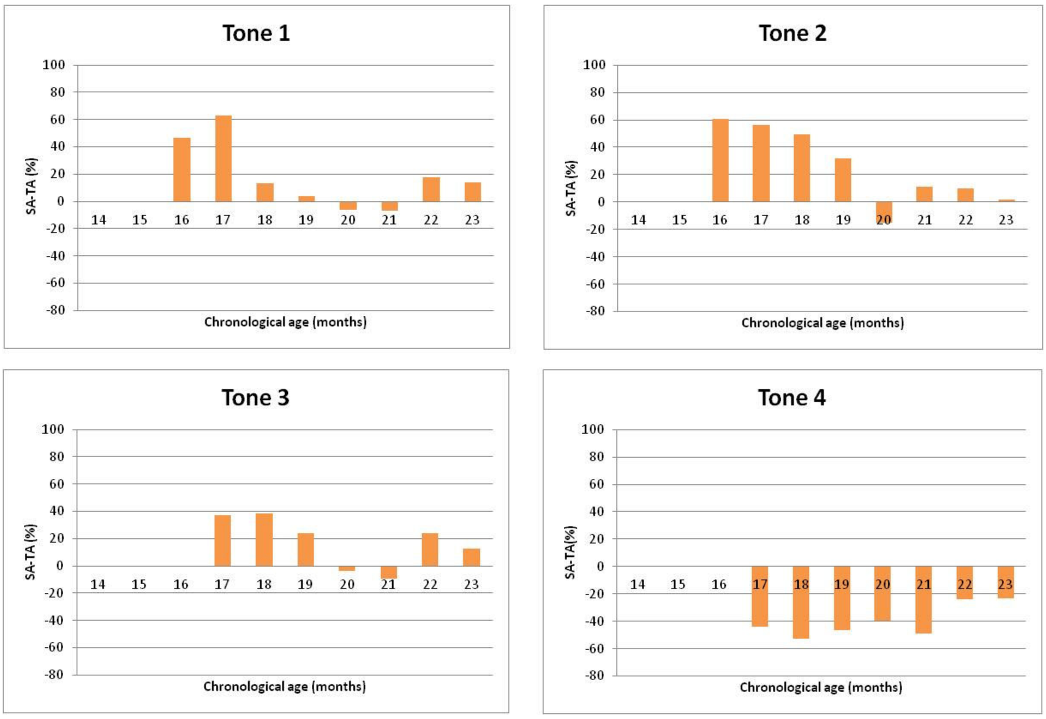


Figure 13: Differences between Segment and Tone Accuracy in YKX's Production.

Results of paired sample t-test to compare segmental and tone accuracy for each child across the period of the study are summarized in Table 12. For all four children, segmental accuracy was significantly higher than tone accuracy for Tone 3 words ($p < .05$). In contrast, for Tone 4 words, tone was significantly higher than segmental accuracy at a 0.05 level. For Tone 1, segmental accuracy is higher than tone accuracy for the four children. The differences in LXB & MYK did not achieve statistical significance. Tone 2 accuracy is significantly higher than segment accuracy for GXH. Tone 2 accuracy is lower than segment accuracy in LXB, MYK and YKX but the difference in MYK does not achieve statistical significance.

Table 12: T-test Results on Segment and Tone Accuracy

Child	GXH		LXB		MYK		YKX	
	t	sig.	t	sig.	t	sig.	t	sig.
Tone 1	5.391	0	1.678	0.119*	1.809	0.089*	2.638	0.019
Tone 2	-0.564	0.581*	4.459	0.001	1.398	0.181*	3.342	0.004
Tone 3	7.936	0	3.072	0.01	5.211	0	2.998	0.009
Tone 4	-2.887	0.011	-4.621	0.001	-3.557	0.003	-10.877	0

* Not significant at .05 level.

4.2.2. Comparison of segment accuracy in cases when tone is produced accurately or inaccurately

Figures 14 to 17 display segmental accuracy when tone was produced accurately (i.e. segment accuracy in the context of accurate tones) or when tone was produced inaccurately (i.e. segment accuracy in the context of inaccurate tones). All comparisons were made across time for the four participants. (See Appendix 8 for segmental accuracy values under these two tone accuracy conditions). The figure layouts are consistent with those described above. Y-axis is segment accuracy in percentage and X-axis is chronological age in months. Red line with squares and blue line with triangles represents segment accuracy when tone is produced accurately and inaccurately respectively. In this case, higher segmental accuracy when tone is produced accurately compared to when tone is produced inaccurately will imply a mutually facilitative relationship between tone and segmental production. On the contrary, if segmental accuracy is lower when tone is produced accurately, it will suggest a trade-off effect of tone and segmental production.

GXH

For GXH's production of Tone 4 (Figure 14, lower right panel), segmental accuracy is higher when tone is produced accurately compared to when tone is produced inaccurately. This result suggests that accurate segmental production may facilitate the accurate production of Tone 4. Such facilitation effect diminishes after 21 months.

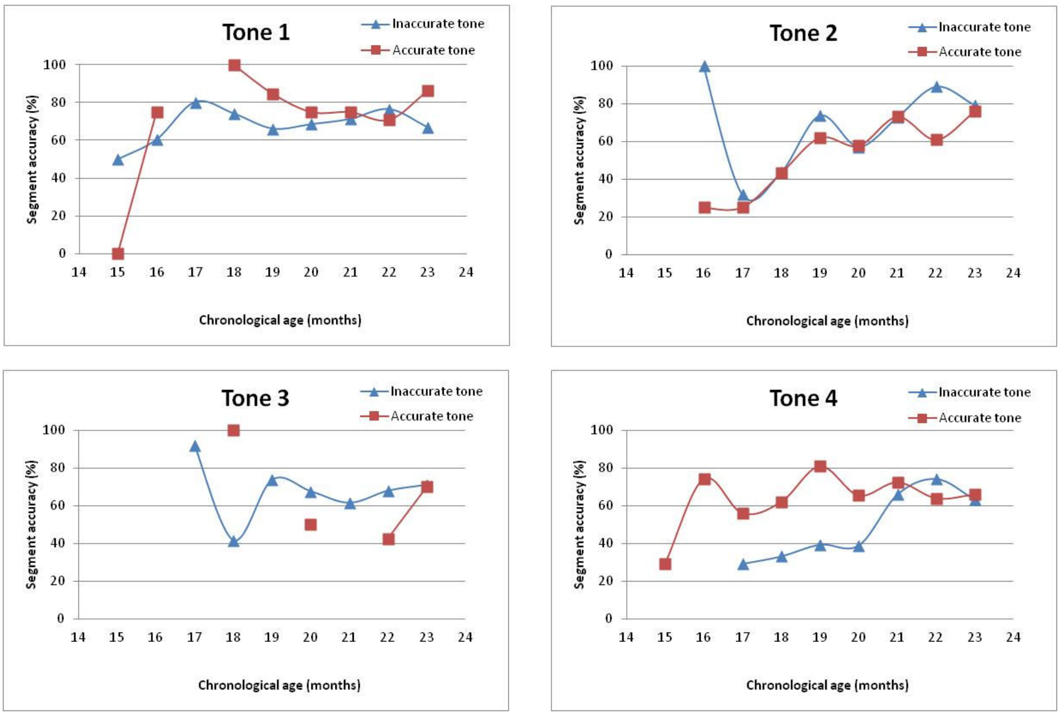


Figure 14: Segmental Accuracy under Two Tone Accuracy Conditions in GXH.

LXB

Figure 15 illustrates the change over time in segmental accuracy under two tone accuracy conditions in LXB. For Tone 3, there seems to be a trade off effect between tone and segment production. As shown in Figure 12 (lower left panel), segment accuracy is higher when tone is produced inaccurately than when tone is produced accurately.

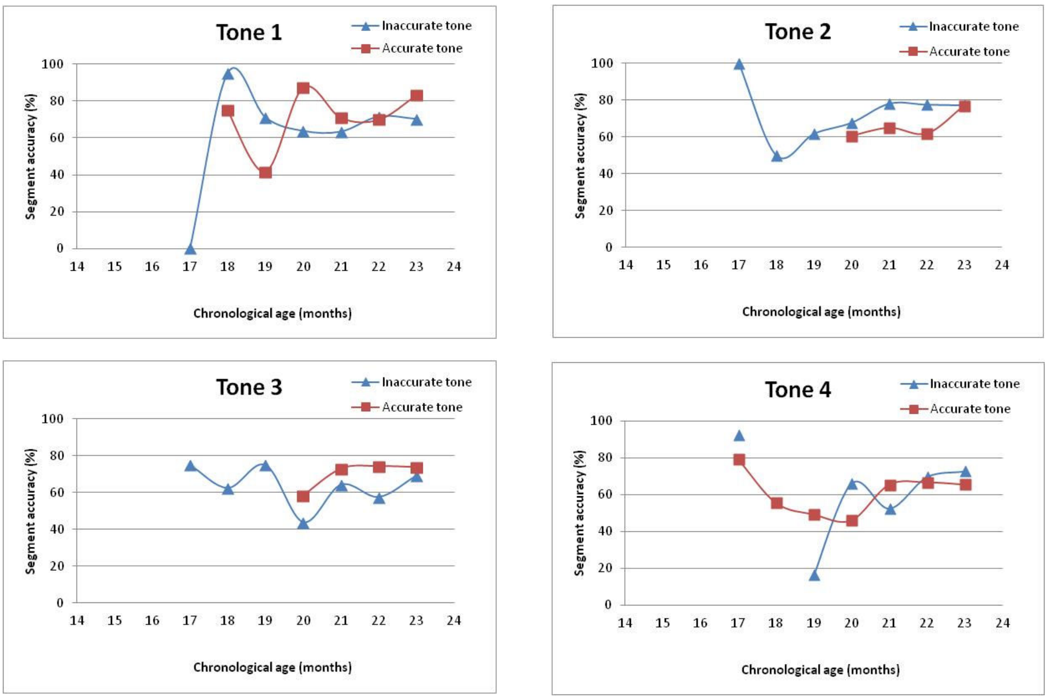


Figure 15: Segmental Accuracy under Two Tone Accuracy Conditions in LXB.

MYK

Similar trade-off effect can be observed in MYK's production of Tone 2 (Figure 16, upper right panel). Before 21 months, segmental accuracy is higher when tone is produced inaccurately compared to when tone is produced accurately. For other tone categories, segment accuracy does not seem to be dependent on tone accuracy.

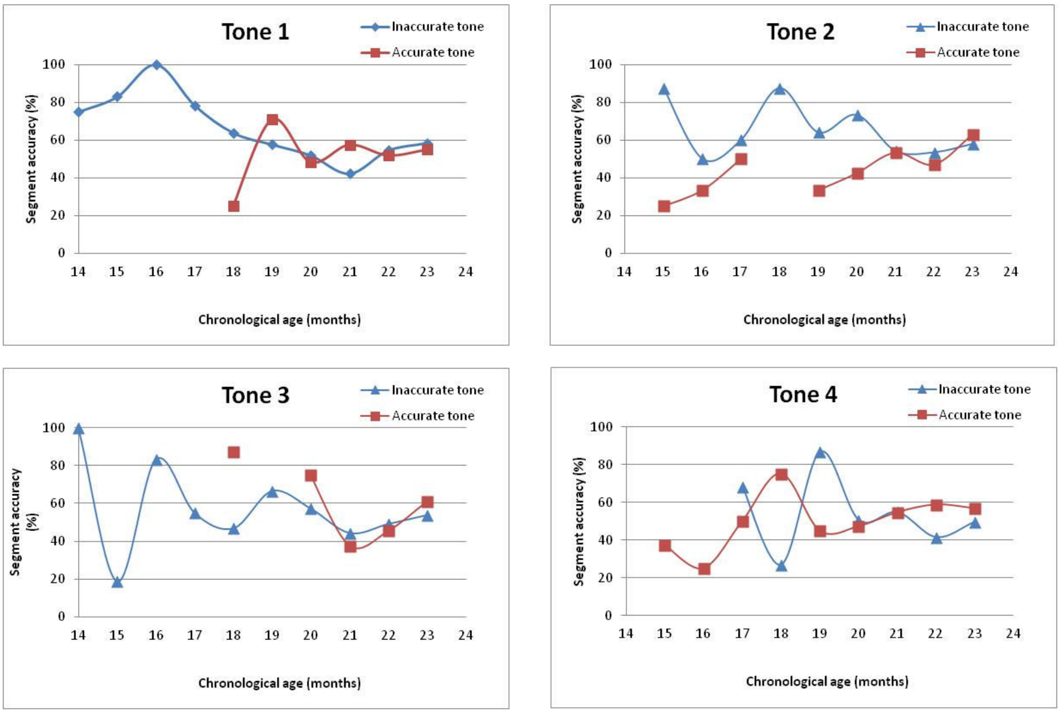


Figure 16: Segmental Accuracy under Two Tone Accuracy Conditions in MYK.

<u>YKX</u>

Figure 17 illustrates the change over time in segmental accuracy under two tone accuracy conditions in YKX's productions. For Tone 4 words (lower right panel), segmental accuracy is higher when tone is produced accurately compared to when tone is produced inaccurately for months 17 and 18. This result may suggest a facilitative effect of segmental accuracy on Tone 4 accuracy in earlier production. For other tonal categories, there does not seem to be defined pattern of segmental accuracy under these two tone accuracy conditions.

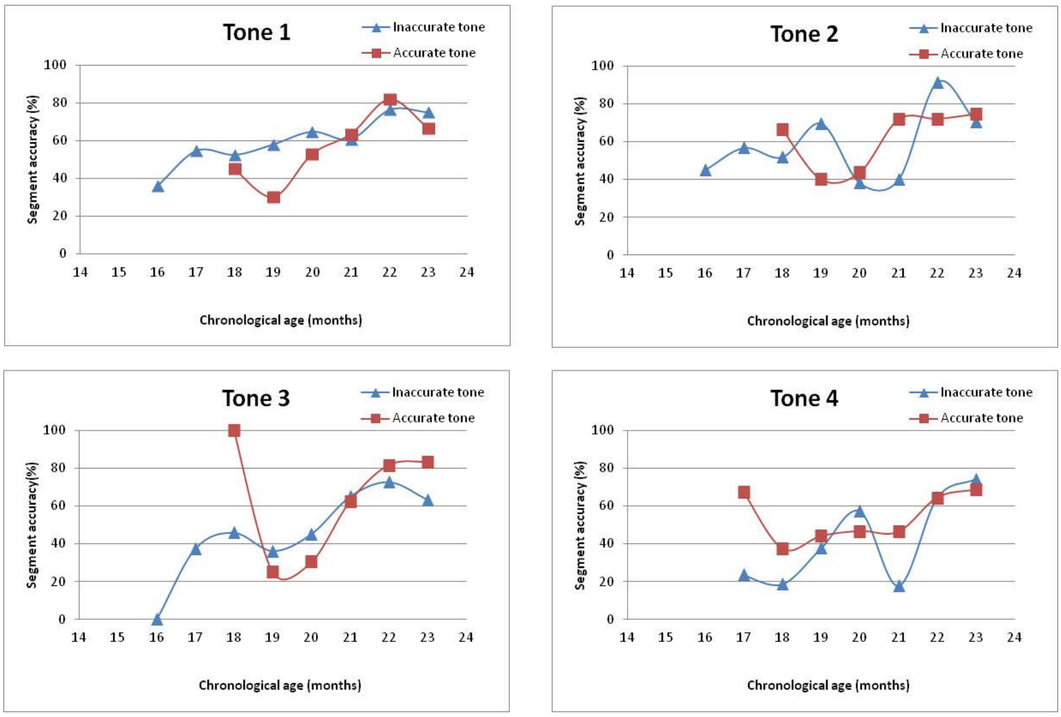


Figure 17: Segmental Accuracy under Two Tone Accuracy Conditions in YKX.

Table 13: Paired Sample T-test Results on Segmental Accuracy under Two Tone Accuracy Conditions.

Child	GXH		LXB		MYK		YKX	
	t	sig.	t	sig.	t	sig.	t	sig.
Tone 1	0.481	0.648	-0.14	0.894	-0.431	0.684	-1.63	0.164
Tone 2	-1.687	0.136	-2.668	0.076	-2.488	0.042*	0.14	0.894
Tone 3	0.194	0.859	4.008	0.028*	1.277	0.271	0.873	0.423
Tone 4	2.534	0.044*	0.03	0.977	0.112	0.915	1.546	0.173

*Significant at 0.05 level

Table 13 summarizes results of paired sample t-test to compare average segmental accuracy in cases when tones were produced accurately or inaccurately. Regardless of tone accuracy, there are no significant differences between segment accuracy except for Tone 4 in GXH, Tone 2 for LXB and Tone 3 in MYK.

4.2.3. Comparison of tone accuracy and segmental accuracy across tone categories.

Figure 15 compares tone accuracy across the four Mandarin tone categories and segmental accuracy of words with the four tones. Tone accuracy for each of the four children is displayed on the left and segmental accuracy is displayed on the right. Y-axis is accuracy (tone or segmental) in percentage and X-axis is chronological age in months.

For all children, Tone 4 has the highest accuracy among the four categories and Tone 3 has the lowest accuracy. However, no such pattern was found for segmental accuracy. Thus, Tone 4 words do not show higher segmental accuracy for any of the children. Segmental accuracy for Tone 3 words were not the lowest compared to segmental accuracy of words in other tone categories.

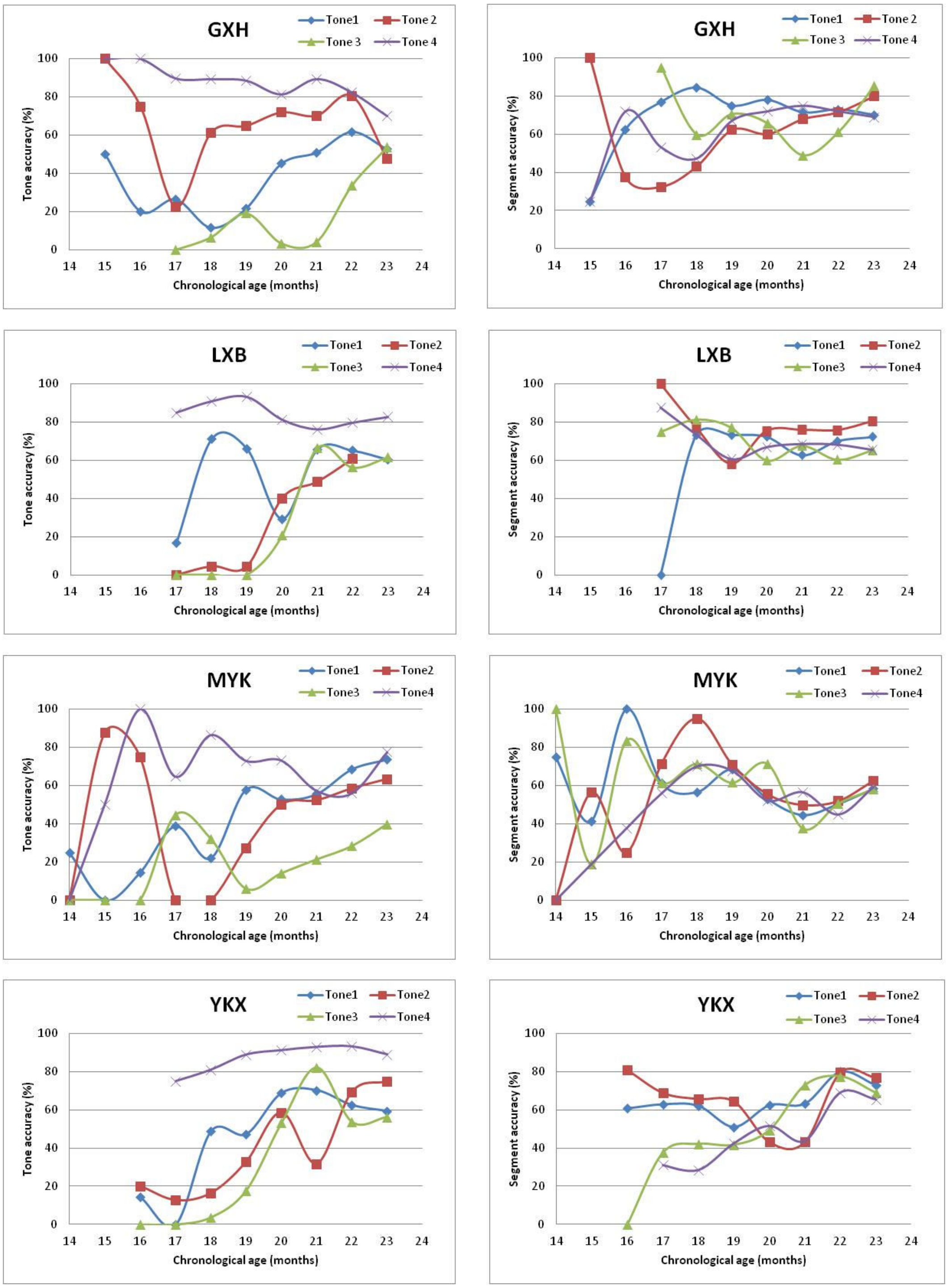


Figure 18: Tone Accuracy and Segmental Accuracy across Tone Categories.

One-way Analysis of Variances (ANOVAs) were performed for each child to evaluate the effect of tone category on tone accuracy and segmental accuracy. Table 14 summarizes the results. For all children, there is a significant effect of tone category on tone accuracy. A post-hoc LSD test suggested that Tone 4 has significantly higher tone accuracy and Tone 3 has significantly lower tone accuracy compared to other tone categories. However, no significant effect of tone was found for segmental accuracy in any of the children. The accuracy ranking observed in tone production accuracy was not present in segmental accuracy. This analysis provided additional support to the relative independence of tone and segmental accuracy observed in Section 4.3.2.

Table 14: ANOVA Results on Tone and Segment Accuracy across Tone Categories.

Accuracy	GXH		LXB		MYK		YKX	
	F	sig.	F	sig.	F	sig.	F	sig.
Tone	23.532	0*	10.349	0*	15.188	0*	10.901	.001*
Segment	1.179	0.327	1.847	0.132	0.334	0.855	3.861	0.007*

*Significant at 0.05 level

Summary of tone-segment relationships

These results indicated that tone category and developmental age influences the relationship between tone and segmental accuracy. For all children during the earlier months, segmental accuracy is higher than tone accuracy for Tone 3 words. Tone accuracy is higher than segmental accuracy in Tone 4 words. Segmental accuracy tends to be higher than tone accuracy in Tone 1 and Tone 2 words but with individual child level variations. Developmentally, the differences between tone and segmental accuracy decrease as these four children grew older. Results also suggest relative independence of tone and segmental accuracy. Within each tone category, similar segmental accuracy was observed whether tone is produced accurately or inaccurately. Across tone categories, Tone 4 words were produced with the highest accuracy and Tone 3 with the lowest accuracy. However, segmental accuracy was not related to the tonal category of the words.

4.3. TONE ACCURACY AND WORD-LEVEL VARIATIONS

4.3.1. Comparison of average variable scores

The average variable values for each participant by vocabulary size are summarized in Appendix 9 (Table 1-4). Tables 15- 18 display the average scores of the phonetic and lexical variables based on high-low tone accuracy for each child. In each table, Parts A-D represent the four tone categories respectively.

The leftmost column displays the word stages represented by vocabulary size. The number in the parenthesis under column "Tone Accuracy (TA)" is the average tone accuracy of lexical items within the same tone category for each word stage. Lexical items with tone accuracy higher than the average value were considered as "High TA." Lexical items with tone accuracy lower than the average value were considered as "Low TA." The number of lexical items that fell within each of the two groups is shown in the column "Number of Lexical Items". Average values for the phonetic variables and lexical variables of these two groups of lexical items are presented in the remaining columns. According to the study hypothesis, if tone accuracy is negatively related to children's Productive Tone Neighborhood Density, we would predict a lower average value of PTND when TA was high and a higher average value of PTND when TA was low. In each of the tables below, corresponding cells are highlighted in green if that pattern occurs.

<u>GXH</u>

Table 15A-D shows average variable scores of Tone 1, 2, 3, and 4 lexical items in GXH's production over the five word stages.

For Tone 1 lexical items, in GXH's 0-50 words stage, the average tone accuracy for all the 11 Tone 1 lexical items is 31%. There are 5 lexical items with tone accuracy higher than 31%, considered as High TA lexical items, and 6 with tone accuracy lower than 31%, considered as low TA lexical items. For High TA lexical items, the average segment accuracy was 0.70. The average MWCM score for target and for child production was 1.20 and 1.08 respectively. For low TA lexical items, the average segment accuracy, MWCM score for target and for actual production was 0.65, 2.00 and 1.49 respectively. With regard to lexical variables, when tone was produced with high accuracy, the average Child Output Frequency of those lexical items was 3.30. The Productive Tone Neighborhood Density was 0. And the average Adult Word Frequency for the lexical items with high TA was 67.22. The corresponding values for lexical items with low TA were 1.96, 0.17 and 86.00. As highlighted in green, when GXH has 0-50 words, 50-100 words and more than 150 words, PTND is negatively related to tone accuracy as high TA corresponds to low PTND and low TA corresponds to high PTND. This relationship between TA and PTND was not shown when GXH had a vocabulary size of 100-150 words. Instead, higher TA corresponded to higher PTND.

For Tone 2, PTND was negatively related to tone accuracy regardless of vocabulary size. Such results indicate that, for Tone 2 lexical items, the more tone contrastive minimal pairs a word has in a session, the more probability it will be produced with lower tone accuracy. For Tone 3 lexical items, with a vocabulary size of 0-50 words, none of them were produced correctly (accuracy equal 0). Therefore no high-low comparison was made in this word stage. When GXH had more than 200 words, high TA corresponded to low PTND and low TA corresponded to high PTND, suggesting an inverse relationship between TA and PTND. With regards to Tone 4 lexical items, for 50-100 words and for more than 200 words, high TA was related to lower PTND values compared to PTND values when TA was low. For other word stages, High TA corresponded to high PTND and low TA corresponded to low PTND.

Table 15: Average Variable Scores of Tone 1-4 Lexical Items by GXH.

(A) Tone 1

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 31%)	5	0.70	1.20	1.08	3.30	0.00	67.22
	low (< 31%)	6	0.65	2.00	1.49	1.96	0.17	86.00
50-100	high (> 26%)	8	0.70	1.63	1.18	1.72	0.07	90.76
	low (< 26%)	12	0.70	2.00	1.59	2.81	0.40	73.36
100-150	high (> 51%)	12	0.70	2.00	1.59	2.81	0.40	73.36
	low (< 51%)	15	0.72	1.73	1.59	3.42	0.14	73.54
150-200	high (> 51%)	8	0.69	2.25	1.62	4.75	0.19	93.29
	low (< 51%)	10	0.76	2.00	1.56	3.00	0.50	71.32
>200	high (> 52%)	14	0.78	2.00	1.49	3.39	0.32	78.10
	low (< 52%)	17	0.69	1.82	1.82	2.53	0.38	89.99

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

(B) Tone 2

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 43%)	5	0.39	2.40	1.27	2.00	0.10	94.24
	low (< 43%)	4	0.75	0.25	0.25	2.75	0.25	78.33
50-100	high (> 72%)	15	0.56	1.67	1.53	2.58	0.07	89.35
	low (< 72%)	9	0.72	2.00	2.11	1.89	0.19	74.87
100-150	high (> 83%)	19	0.63	2.32	1.98	3.24	0.00	88.35
	low (< 83%)	7	0.71	1.25	0.97	2.76	0.45	70.47
150-200	high (> 79%)	19	0.63	2.32	1.98	3.24	0.00	88.35
	low (< 79%)	8	0.81	1.38	1.31	2.25	0.13	69.37
>200	high (> 44%)	18	0.76	2.06	2.01	3.50	0.03	81.98
	low (< 44%)	18	0.79	2.28	1.86	2.61	0.22	81.35

(C) Tone 3

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (=0%)	3	0.83	0.67	0.33	1.83	0.00	72.85
	na							
50-100	high (> 11%)	8	0.59	1.95	1.28	3.10	0.64	66.98
	low (< 11%)	14	0.62	1.47	1.19	2.34	0.13	77.26
100-150	high (> 5%)	2	0.33	3.50	1.75	2.25	0.25	70.78
	low (< 5%)	19	0.70	1.84	1.49	1.92	0.22	73.10
150-200	high (> 37%)	8	0.61	2.25	1.77	2.81	0.38	76.02
	low (< 37%)	12	0.69	1.75	1.56	3.67	0.08	72.21
>200	high (> 57%)	13	0.66	2.38	2.08	3.96	0.00	72.48
	low (< 57%)	15	0.67	1.93	1.54	4.50	0.47	77.19

(D) Tone 4

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 92%)	14	0.51	1.29	0.86	1.96	0.14	75.47
	low (< 92%)	4	0.23	1.50	0.73	3.38	0.00	34.94
50-100	high (> 81%)	24	0.64	1.58	1.33	2.49	0.21	73.15
	low (< 81%)	7	0.45	1.57	1.33	1.86	0.29	52.85
100-150	high (> 85%)	24	0.68	1.54	1.27	2.44	0.24	72.19
	low (< 85%)	8	0.45	1.63	1.31	3.33	0.13	46.35
150-200	high (> 87%)	18	0.62	1.83	1.91	1.78	0.25	68.04
	low (< 87%)	9	0.80	1.67	1.66	8.11	0.11	68.85
>200	high (> 73%)	27	0.66	2.11	1.73	3.43	0.20	72.97
	low (< 73%)	16	0.51	1.25	1.35	4.50	0.31	62.62

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

LXB

Table 16A-D displays average variable scores for Tone 1, 2, 3 and 4 lexical items by LXB.

For Tone 1 lexical items, during LXB's 0-50 word stage, PTND was negatively related to tone accuracy as high TA corresponded to low PTND and low TA corresponded to high PTND. This relationship between TA and PTND was not shown in the subsequent word stages. Instead, higher TA corresponded to higher PTND. For Tone 2 lexical items, previous to the last word stage, PTND was negatively related to tone accuracy. Such results indicate that for Tone 2 lexical items, the less tone contrastive minimal pairs a word has in a session, the more possible it will be produced with lower tone accuracy. For Tone 3 lexical items, with a vocabulary size up to 100 words, tone was not correct in any words. No high-low comparison could be made for these two word stages. Similar to GXH, with a vocabulary size of more than 200 words, LXB produced higher TA when lexical items had lower PTND and lower TA when PTND was high. For Tone 4 lexical items, with a vocabulary size of 0-50, 50-100 words and more than 200 words, a similar negative relationship of TA and PTND were found. For other words stages, high TA corresponded to high PTND and low TA corresponded to low PTND.

Table 16: Average Variable Scores of Tone 1-4 Lexical Items by LXB.

(A) Tone 1

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 55%)	6	0.56	1.00	0.35	2.50	0.33	76.68
	low (< 55%)	7	0.74	0.86	0.47	3.19	0.50	72.43
50-100	high (> 56%)	6	0.58	1.50	1.11	3.17	0.50	67.43
	low (< 56%)	7	0.68	1.71	1.42	8.00	0.14	81.95
100-150	high (> 32%)	6	0.82	1.67	1.63	2.33	0.17	85.49
	low (< 32%)	10	0.63	2.00	1.57	3.40	0.10	67.22
150-200	high (> 72%)	19	0.66	2.42	2.08	1.37	0.47	80.35
	low (< 72%)	13	0.53	2.00	1.54	2.77	0.15	74.22
>200	high (> 64%)	61	0.73	2.06	1.82	2.43	0.43	79.25
	low (< 64%)	45	0.66	2.24	2.06	1.98	0.25	77.38

(B) Tone 2

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 5%)	1	0.45	0.00	0.07	5.00	0.50	86.42
	low (< 5%)	2	0.67	0.00	0.40	4.25	1.00	97.23
50-100	high (> 56%)	6	0.60	1.17	1.00	2.67	0.17	91.46
	low (< 56%)	5	0.57	2.00	2.16	7.80	0.20	63.35
100-150	high (> 52%)	13	0.63	2.46	2.21	2.69	0.00	78.52
	low (< 52%)	12	0.66	2.17	1.96	2.42	0.17	82.07
150-200	high (> 50%)	15	0.63	2.60	2.46	2.67	0.13	80.64
	low (< 50%)	18	0.73	2.11	1.67	3.17	0.39	85.58
>200	high (> 59%)	43	0.67	1.76	1.55	2.17	0.38	77.58
	low (< 59%)	40	0.72	2.15	1.99	1.76	0.20	72.85

(C) Tone 3

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (= 0%)	8	0.68	0.50	0.25	2.31	0.56	75.98
	na							
50-100	high (= 0%)	9	0.63	2.00	1.53	2.00	0.22	69.73
	na							
100-150	high (> 5%)	8	0.60	1.63	1.40	1.88	0.25	74.33
	low (< 5%)	8	0.38	2.25	2.05	3.88	0.13	72.64
150-200	high (> 65%)	12	0.59	2.00	1.85	1.92	0.42	68.41
	low (< 65%)	8	0.63	1.63	1.33	4.00	0.38	78.21
>200	high (> 66%)	43	0.68	2.16	1.84	1.71	0.52	71.60
	low (< 66%)	32	0.60	1.97	1.65	2.71	0.66	71.93

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

(D) Tone 4

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 88%)	8	0.46	1.38	1.00	2.64	0.29	70.06
	low (< 88%)	3	0.43	0.33	0.02	4.56	0.44	67.00
50-100	high (> 79%)	4	0.50	2.50	2.13	1.75	0.25	53.74
	low (< 79%)	3	0.88	1.33	1.46	4.33	0.33	49.48
100-150	high (> 77%)	15	0.49	1.80	1.83	2.27	0.07	63.96
	low (< 77%)	5	0.62	2.60	3.00	1.40	0.00	61.36
150-200	high (> 82%)	17	0.69	2.06	2.27	1.59	0.65	81.81
	low (< 82%)	7	0.64	1.43	1.01	3.14	0.14	66.94
>200	high (> 79%)	76	0.73	2.12	1.94	1.83	0.33	72.78
	low (< 79%)	39	0.58	1.95	1.64	3.30	0.44	62.22

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

<u>MYK</u>

Table 17A-D displays average variable scores of Tone 1, 2, 3 and 4 lexical items in MYK's production.

For tone 1, during the 150-200 words stage, PTND was negatively related to tone accuracy. This relationship between TA and PTND was not shown in other word stages. For Tone 2 lexical items, except for the 50-100 word stage, PTND was negatively related to tone accuracy. Such results indicate that for Tone 2 lexical items, the less tone contrastive minimal pairs a word has in a session, the lower was tone accuracy. For Tone 3 lexical items, negative relationship between TA and PTND were only evident at 150-200 words. Average variable scores of Tone 4 lexical items are shown in Table 19D. Except at the 50-100 word stage, high TA corresponded to low PTND and low TA corresponded to high PTND. This pattern was similar to patterns found in Tone 2 lexical items.

Table 17: Average Variable Scores of Tone 1-4 Lexical Items by MYK.

(A) Tone 1

Vocabulary size (words)	TA	Number of Lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 38%)	8	0.69	0.75	0.57	2.64	0.23	76.11
	low (< 38%)	8	0.74	0.75	0.32	2.88	0.00	82.37
50-100	high (> 38%)	3	0.50	1.00	0.83	1.67	0.00	99.38
	low (< 38%)	5	0.65	0.60	0.27	4.60	0.60	85.97
100-150	high (> 51%)	11	0.53	1.73	1.02	2.18	0.18	81.80
	low (< 51%)	14	0.55	1.43	1.47	2.36	0.29	83.96
150-200	high (> 59%)	12	*0.46*	1.92	1.44	2.08	0.42	79.48
	low (< 59%)	12	0.48	1.50	0.97	2.92	0.50	80.61
>200	high (> 66%)	69	0.50	2.42	2.02	2.14	0.39	78.04
	low (< 66%)	48	0.55	2.10	1.51	2.01	0.38	80.31

(B) Tone 2

Vocabulary size (words)	TA	Number of Lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 26%)	3	0.33	1.00	0.71	2.11	0.00	97.52
	low (< 26%)	8	0.57	0.88	0.61	1.91	0.13	88.92
50-100	high (> 38%)	6	0.67	1.00	0.79	4.00	0.17	94.95
	low (< 38%)	9	0.39	1.11	0.61	1.56	0.11	75.41
100-150	high (> 61%)	14	0.38	2.21	1.74	2.21	0.14	84.48
	low (< 61%)	10	0.71	1.00	0.61	6.60	0.20	78.75
150-200	high (> 60%)	16	0.50	2.13	1.98	1.88	0.25	88.51
	low (< 60%)	13	0.79	1.38	1.05	2.31	0.31	79.70
>200	high (> 61%)	57	0.53	1.98	1.55	1.72	0.30	81.25
	low (< 61%)	52	0.54	2.21	1.98	2.16	0.49	81.43

(C) Tone 3

Vocabulary size (words)	TA	Number of Lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 12%)	2	0.83	0.00	0.16	2.50	0.25	69.62
	low (< 12%)	8	0.59	1.50	0.69	1.71	0.00	76.89
50-100	high (> 10%)	2	0.42	2.00	0.92	8.00	0.50	73.92
	low (< 10%)	6	0.42	2.00	1.28	3.50	0.17	79.17
100-150	high (> 9%)	4	0.74	2.50	2.06	14.25	0.25	69.11
	low (< 9%)	10	0.59	1.50	0.88	1.40	0.10	75.88
150-200	high (> 26%)	6	0.75	1.67	1.25	1.67	0.17	88.06
	low (< 26%)	14	0.48	1.71	1.27	3.71	0.71	66.85
>200	high (> 38%)	31	0.47	1.97	1.58	1.72	0.70	76.98
	low (< 38%)	48	0.50	1.82	1.43	2.54	0.54	74.01

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

(D) Tone 4

Vocabulary size (words)	TA	Number of Lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 84%)	9	0.56	1.78	1.09	2.04	0.00	71.89
	low (< 84%)	3	0.57	1.33	0.97	3.25	0.11	49.62
50-100	high (> 66%)	6	0.29	2.17	1.38	2.50	0.17	71.25
	low (< 66%)	4	0.69	1.00	0.71	4.75	0.00	79.07
100-150	high (> 72%)	8	0.58	2.00	1.83	2.13	0.25	64.15
	low (< 72%)	6	0.55	1.33	0.99	3.67	0.33	57.72
150-200	high (> 70%)	16	0.41	2.06	1.46	2.44	0.19	77.11
	low (< 70%)	11	0.41	1.18	0.81	2.91	0.45	62.15
>200	high (> 64%)	83	0.58	1.86	1.73	2.14	0.39	74.14
	low (< 64%)	61	0.46	1.84	1.26	2.61	0.48	72.19

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

YKX

Average variable scores of Tone 1, 2, 3 and 4 lexical items in YKX's production over the five word stages are displayed in Table 18A-D.

For Tone 1, PTND was negatively related to tone accuracy at 50-100, 100-150 words and beyond 200 words. This relationship between TA and PTND was not shown in other word stages. For tone 2 lexical items, except at the 50-150 word stage, PTND was negatively related to tone accuracy. High TA corresponded to low PTND and low TA corresponded to high PTND. For Tone 3 lexical item, for less than 50 words, tone was incorrect for all of the lexical items and no high-low comparison was made. A negative relationship between TA and PTND was found when YKX had 50-200 words. Beyond 200 words, lexical items with higher TA tended to have lower PTND compared to PTND of lexical items with lower TA. Average variable scores of Tone 4 lexical items are shown in Table 20D. PTND was negatively related to tone accuracy at less than 50 words and at 150-200 words. The opposite pattern was found in other word stages. Higher TA corresponded to higher PTND and lower TA corresponded to lower PTND.

Table 18: Average Variable Scores of Tone 1-4 Lexical Items by YKX.

(A) Tone 1

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 10%)	3	0.73	0.00	0.11	3.83	0.42	62.45
	low (< 10%)	9	0.36	1.89	0.81	1.65	0.00	77.43
50-100	high (> 38%)	11	0.51	1.00	0.82	2.62	0.17	74.13
	low (< 38%)	12	0.46	1.83	1.26	1.79	0.17	81.63
100-150	high (> 63%)	16	0.61	1.56	0.69	2.16	0.09	79.79
	low (< 63%)	10	0.69	1.50	1.07	4.05	0.35	81.49
150-200	high (> 80%)	15	0.59	1.53	1.09	1.93	0.03	71.71
	low (< 80%)	9	0.59	2.44	1.46	3.61	0.00	79.53
>200	high (> 66%)	28	0.68	2.25	1.79	2.07	0.10	79.59
	low (< 66%)	26	0.67	1.93	1.68	2.55	0.11	72.65

(B) Tone 2

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 10%)	2	0.96	0.50	0.02	9.38	0.00	91.75
	low (< 10%)	12	0.30	1.25	0.54	1.29	0.10	94.97
50-100	high (> 11%)	4	0.67	2.00	1.25	1.67	0.25	91.52
	low (< 11%)	18	0.53	1.33	0.91	1.66	0.17	85.52
100-150	high (> 53%)	16	0.40	1.44	0.94	1.97	0.19	90.64
	low (< 53%)	20	0.47	2.00	1.44	2.13	0.15	82.80
150-200	high (> 40%)	8	0.62	1.63	1.03	2.69	0.25	73.49
	low (< 40%)	17	0.33	2.29	1.63	1.71	0.29	80.87
>200	high (> 71%)	29	0.71	2.07	1.83	2.03	0.12	81.00
	low (< 71%)	19	0.62	1.79	1.28	1.88	0.14	82.28

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

(C) Tone 3

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (= 0%)	8	0.25	1.50	0.69	1.25	0.38	78.19
	na							
50-100	high (> 4%)	1	1.00	1.00	1.00	1.00	0.00	98.78
	low (< 4%)	15	0.41	1.27	1.10	1.54	0.18	73.91
100-150	high (> 45%)	12	0.31	2.50	1.70	1.88	0.17	60.57
	low (< 45%)	11	0.44	1.36	0.77	1.36	0.32	78.50
150-200	high (> 70%)	11	0.49	1.64	1.45	2.77	0.00	66.05
	low (< 70%)	6	0.58	2.33	1.91	4.00	0.58	69.68
>200	high (> 67%)	19	0.77	1.79	1.67	2.11	0.32	73.49
	low (< 67%)	19	0.59	2.53	1.77	3.19	0.23	61.59

(D) Tone 4

Vocabulary size (words)	TA	Number of lexical items	Average					
			Phonetic variables			Lexical variables		
			SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	high (> 75%)	9	0.36	1.89	1.44	1.83	0.00	67.73
	low (< 75%)	4	0.19	1.25	0.55	3.04	0.08	36.04
50-100	high (> 85%)	15	0.32	1.93	1.12	1.67	0.02	71.96
	low (< 85%)	6	0.47	1.00	0.51	3.50	0.00	66.94
100-150	high (> 90%)	23	0.47	1.61	1.22	1.87	0.24	66.37
	low (< 90%)	5	0.56	1.60	1.37	2.30	0.20	63.39
150-200	high (> 91%)	26	0.49	1.85	1.48	2.00	0.15	65.24
	low (< 91%)	4	0.21	2.00	1.65	3.38	0.25	57.66
>200	high (> 92%)	49	0.67	2.04	1.73	2.07	0.26	68.23
	low (< 92%)	13	0.56	2.31	2.19	2.81	0.12	51.08

TA= Tone Accuracy; SA= Segment Accuracy; MWCM_T = Mandarin Word Complexity Measure in Target; MWCM_A=Mandarin Word Complexity Measure in Actual Production; COF= Child Output Frequency; PTND= Productive Tone Neighborhood Density; AWF= Ambient Word Frequency.

To summarize, a larger degree of variation was observed for the relationship between Tone Accuracy and Productive Tone Neighborhood Density. For an individual child, there were no consistent findings within each tone category related to vocabulary size. Higher TA did not necessarily correspond to lower PTND whether children had a smaller or larger vocabulary size. Across tone categories, negative TA-PTND relationships found in certain word stages for one tone category did not entail the same relationship in the same word stages for other tone categories. Comparing across the four children, the TA-PTND relationship manifested in Tables 17-20 was also difficult to determine a generalized pattern. Based on this analysis, TA and PTND did not seem to be reliably related to each other.

The two-way division of Tone Accuracy might obscure the variations of tone accuracy even within the High or Low conditions. For example, a lexical item with a TA value of 74% and another one with 37% would both be considered Low TA in YKX's first word stage because the average TA is 75%. On the other hand, a lexical item with a TA value of 74% would be considered as Low TA but another one with a value of 76% would be considered as High TA although they only differ in 2% in accuracy. Therefore, it is necessary to take into consideration gradients of variability in Tone Accuracy. The analysis in the next section will consider the correlation of Tone Accuracy and other word-level variables statistically.

4.3.2. Correlation between tone accuracy and word-level variables

Figures 19-22 illustrate the results of correlation tests for the four children. The four panels represent the four tone categories respectively.

All the variables were arranged horizontally with regard to vocabulary size. Phonetic variables are presented in the top panel and lexical variables in the bottom panel. The lines indicate the relationship between TA and each of the phonetic and lexical variables. The number represents the correlation coefficient between variables connected by the lines. Significant correlations ($p< 0.05$) were marked with "*". A significant positive coefficient (highlighted in green) indicates a direct relationship between tone accuracy and a variable. For a specific word stage, words with higher Tone Accuracy would be associated with higher variable values. On the other hand, a significant negative coefficient (highlighted in pink) indicates a negative relationship between tone accuracy and a variable. Words with higher tone accuracy would be associated with lower variable values. According to our hypothesis, if Tone Accuracy and Productive Tone Neighborhood Density are negatively related, we would predict a significant negative coefficient. As the children's vocabulary expands, the coefficient decreases or even becomes positive, indicating that larger PTND or contrasts of tones facilitate accurate tonal production.

GXH

Figure 19 displays the correlation coefficients of TA and other word-level variables for Tone 1, 2, 3 and 4 lexical items in GXH's production over the five word stages.

For Tone 1 lexical items, TA and PTND were significantly correlated, when GXH produced 100-150 words. The coefficient was positive suggesting an direct relationship between these two. For Tone 2, Adult Word Frequency (AWF) was directly correlated with TA ($p < .05$) when GXH had 50-200 words. For other variables and vocabulary sizes, no significant correlations were observed. For Tone 3, PTND showed significant positive correlation with TA for a vocabulary size of 50-100, 150-200. Over 200 words, the correlation coefficient was negative. When GXH produced less than 50 words, since no words were produced accurately, no correlation was calculated. Several significant correlations were found in Tone 4 lexical items: AWF for vocabulary size of 100-150; MWCM in child actual production, COF and AWF for vocabulary size of 150-200; MWCM in target and COF at more than 200 words. However, over the five word stages, no decreasing tendency in coefficients was observed for any of the variables for the four tone categories. Overall, with regard to word stages, GXH's tonal production accuracy did not seem to be sensitive to any word-level variables when he had less than 50 words. SA is not correlated with Tone accuracy in any word stage, confirming our findings on the relative independence of tone and segmental production reported in Section 4.2. All the significant correlations between TA and AWF are positive, indicating that GXH tends to produce tone more accurately in words that he heard more often in ambient language.

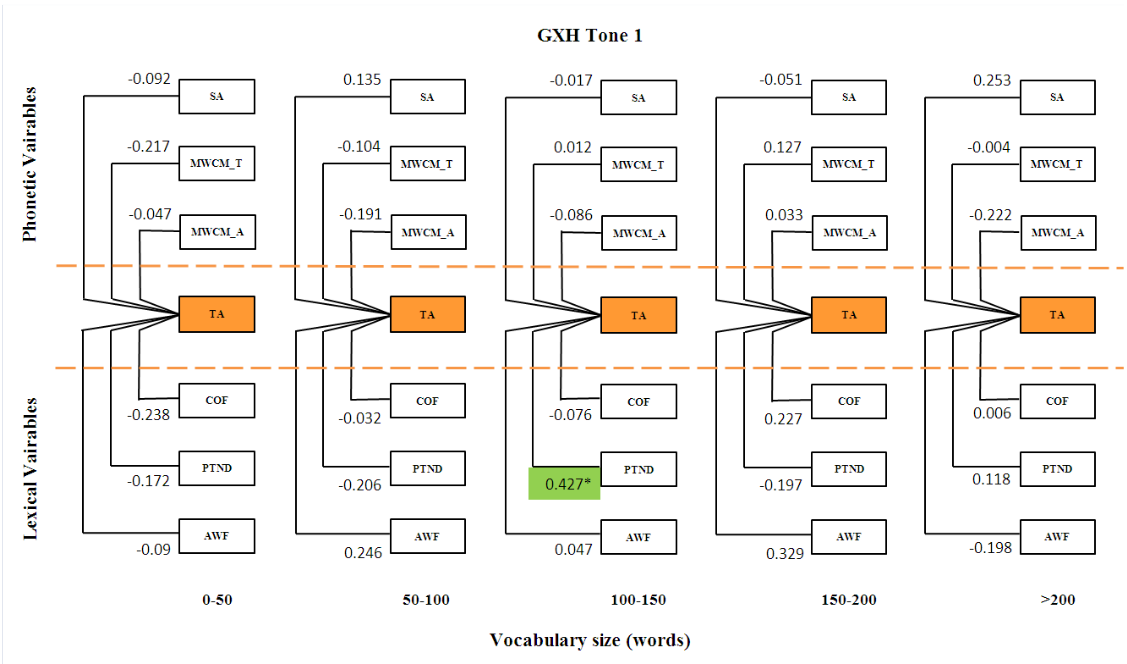
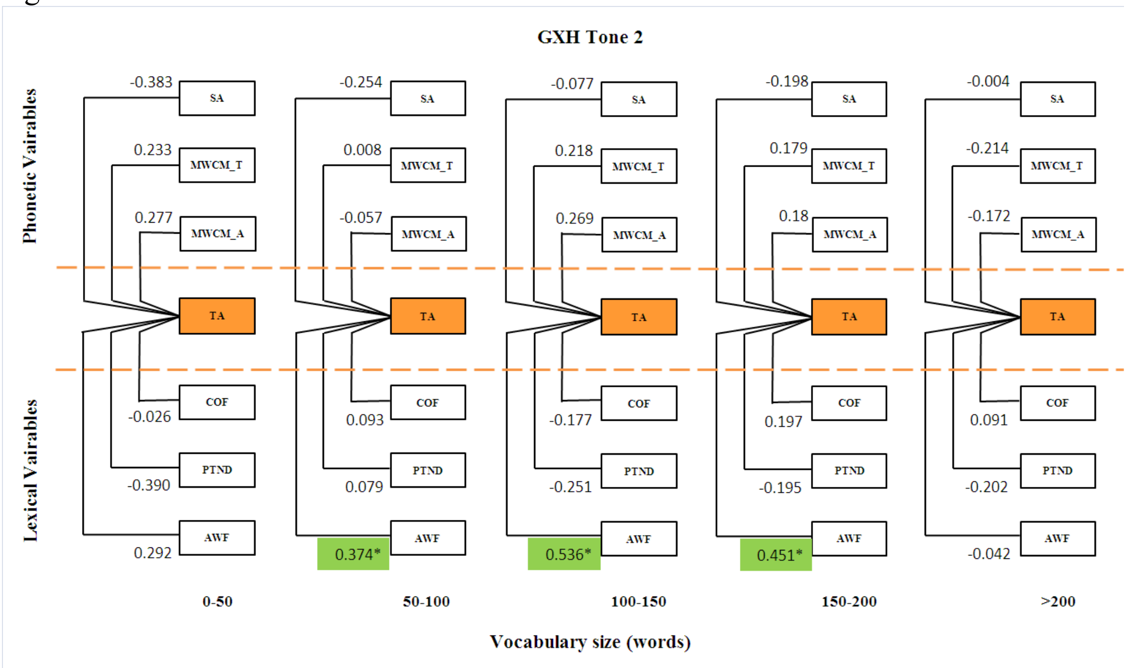


Figure 19: Correlation Coefficients of Tone 1-4 Lexical Items by GXH.

Figure 19: cont.



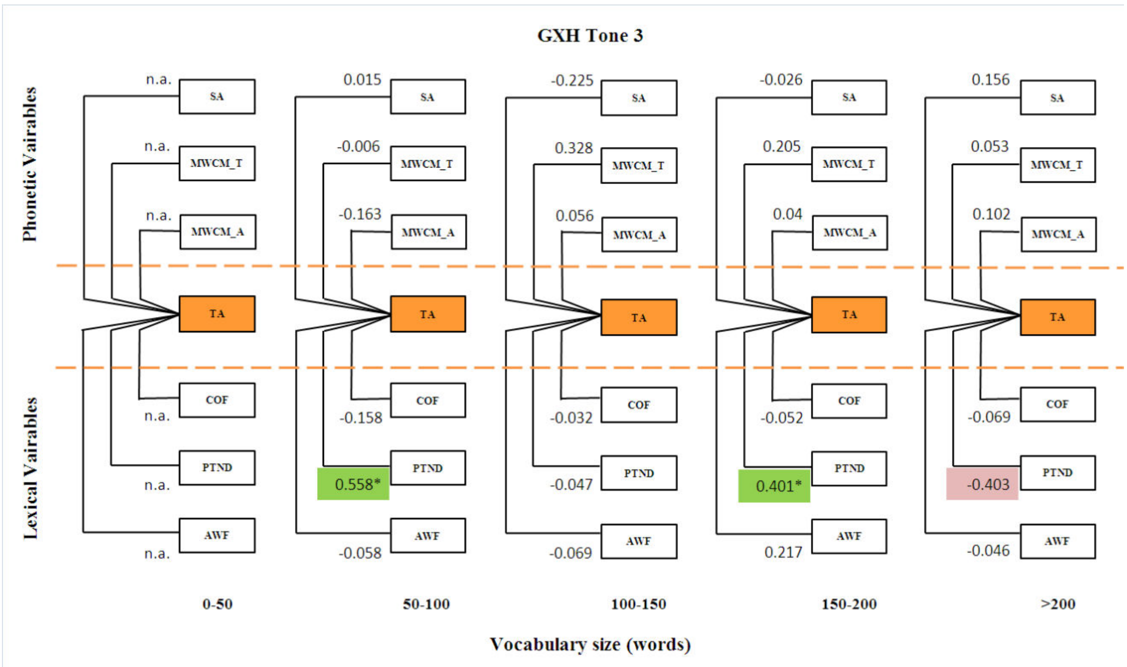
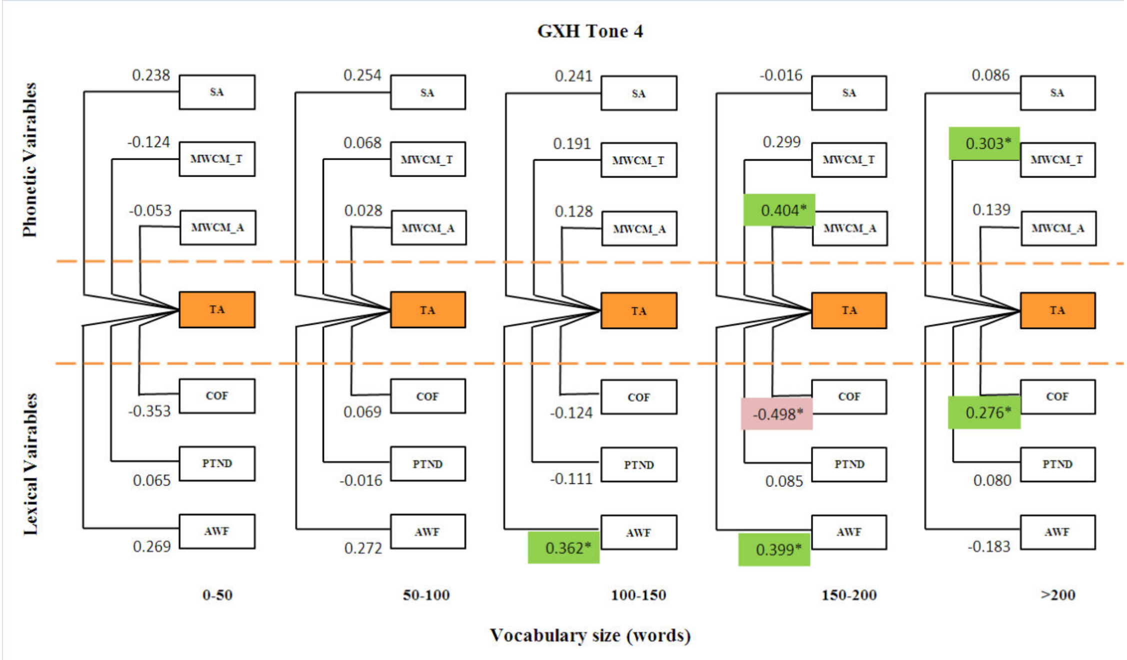


Figure 19: cont.

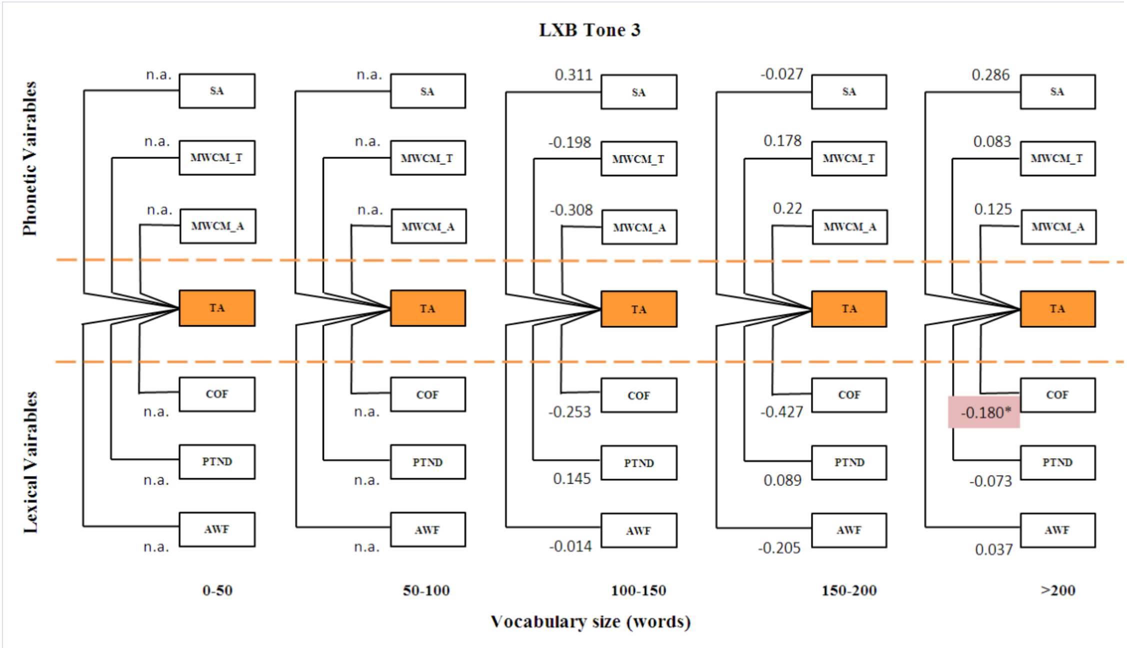


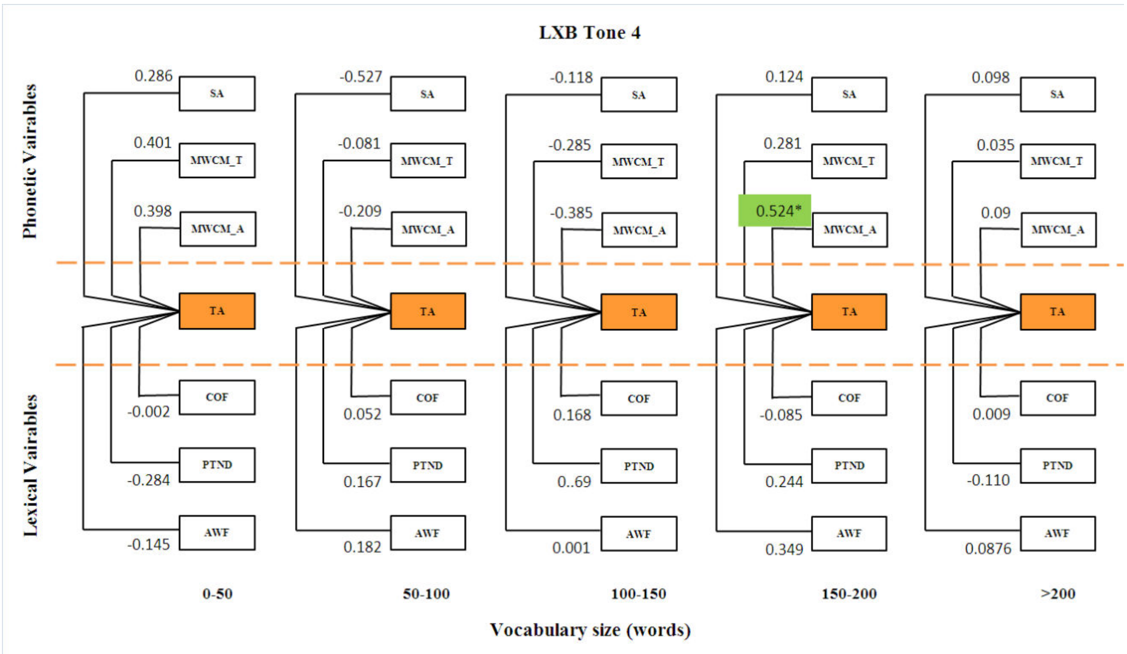
LXB

Figure 20 displays the correlation coefficients of TA and other word-level variables of Tone 1, 2, 3 and 4 lexical items in LXB. For Tone 1, a significant negative correlation was found between TA and COF when LXB had 150-200 words. For Tone 2, MWCM in actual production was significantly correlated with TA at a vocabulary size of 150-200 words. The coefficient was positive. For Tone 3, a significant negative correlation was found between TA and COF when LXB had more than 200 words. Similar to Tone 2, significant positive correlation between MWCM_A and TA were observed in LXB's Tone 4 production. Overall, with regard to word stages, no significant correlations between TA and other word-level variables were found when LXB had less than 150 words. SA was not correlated with tone accuracy in any word stages, providing additional evidence of the relative independence of tone and segment production.

Figure 20: Correlation Coefficients of Tone 1-4 Lexical Items by LXB.

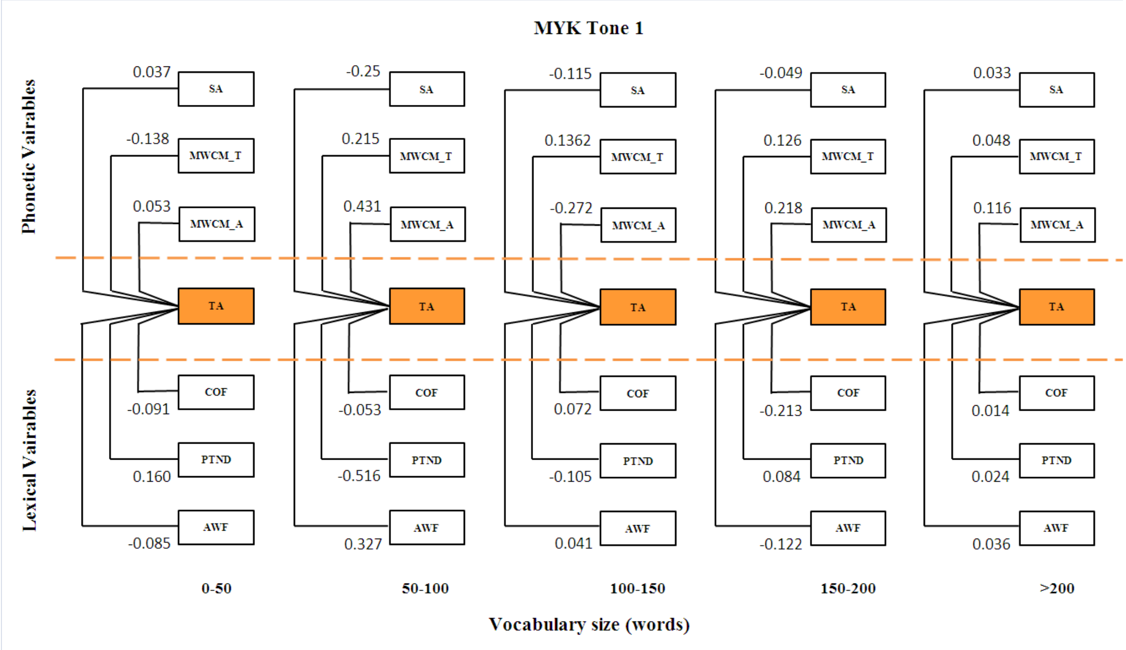
Figure 20: cont.





MYK

Figure 21 displays the correlation coefficients of Tone Accuracy and other word-level variables of Tone 1, 2, 3 and 4 lexical items in MYK's productions. For Tone 1, none of the correlations were significant for any vocabulary size. For Tone 2, when MYK had more than 100 words, two phonetic variables, MWCM in target and MWCM in actual production, were significantly correlated with TA. The coefficient was positive. For Tone 3, a significant negative correlation was found between TA and COF when MYK had more than 200 words, resembling LXB's Tone 2 production. For Tone 4, significant negative correlation between PTND and TA were found when MYK had less than 50 words and 150-200 words. MWCM_T and MWCM_A were significantly correlated with TA during the last two word stages respectively. Both of the coefficients were positive. Overall, with regard to word stages, no significant correlations between TA and other word-level variables were found before MYK had 100 words, except Tone 4. Resembling the other two children, SA was not correlated with tone accuracy in any word stages.



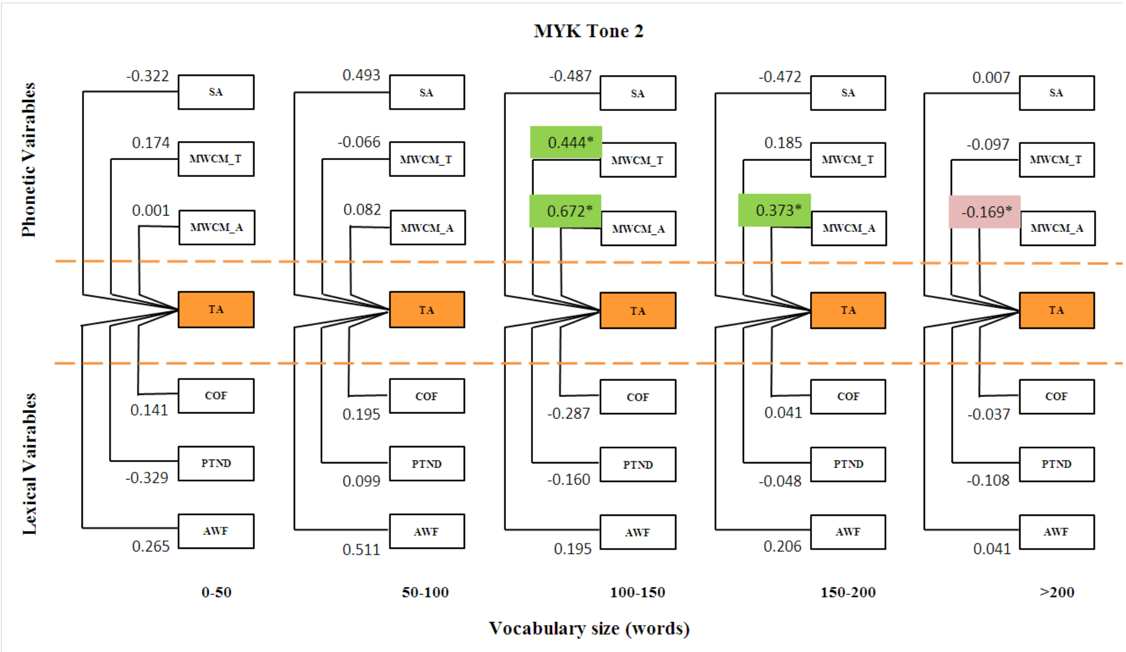
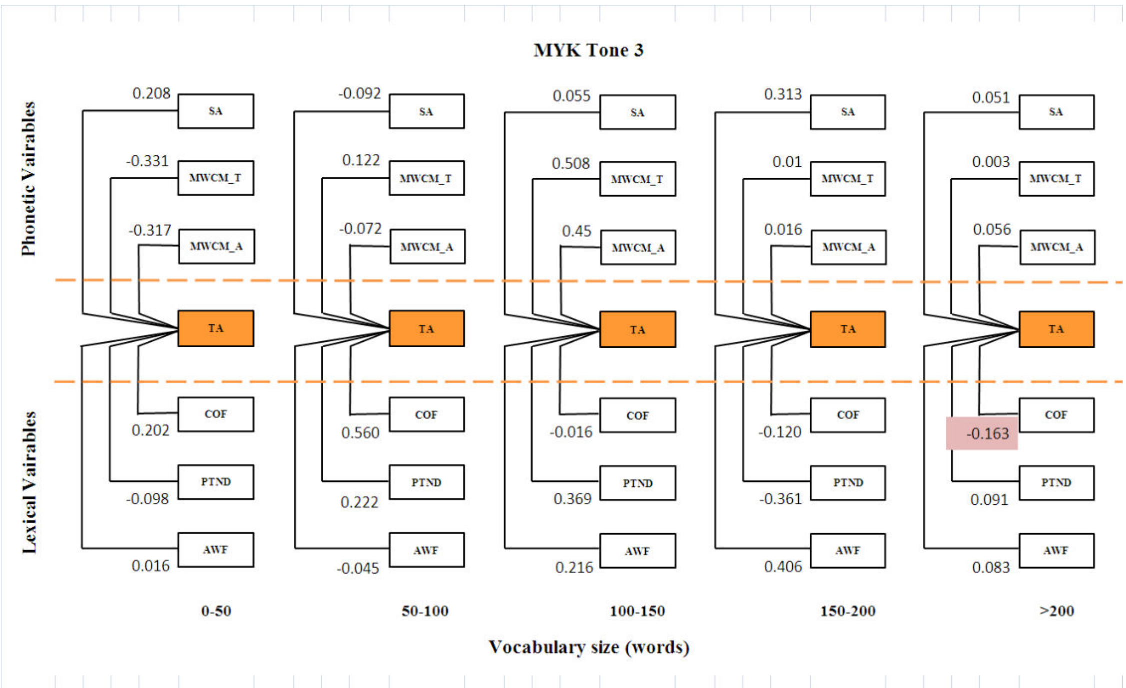
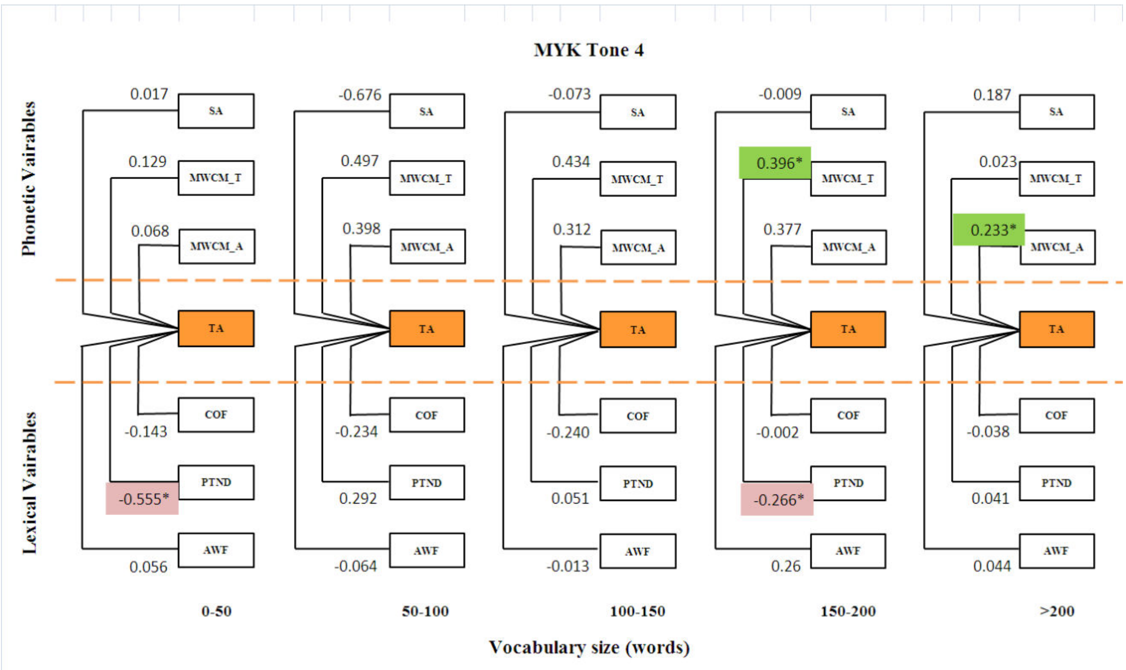


Figure 21: Correlation Coefficients of Tone 1-4 Lexical Items by MYK.

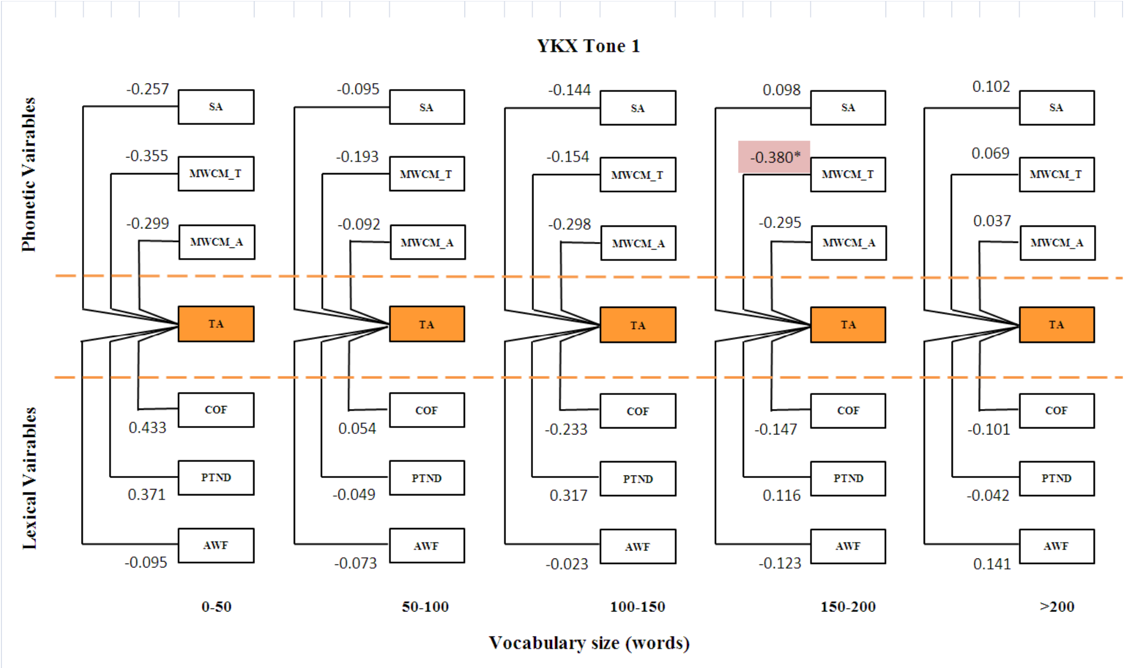
Figure 21: cont.





YKX

Figure 22 shows the correlation coefficients of Tone Accuracy and other word-level variables of Tone 1, 2, 3 and 4 lexical items. For Tone 1, a significant negative correlation was found between TA and MWCM_T when YKX had 150-200 words. For Tone 2, no significant correlations were observed between TA and other word-level variables for any vocabulary size. For Tone 3, all the significant correlations were found after YKX had more than 150 words. Specifically, during 150-200 word stage, PTND was negatively correlated with TA ($p<.05$). Over 200 words, MWCM_T and COF were significantly correlated with TA and the coefficients are negative. During this word stage, another word-level variable, AFW is also significantly correlated with TA. The coefficient was positive. For Tone 4, significant negative correlation between PTND and TA were found when YKX produced less than 50 words. AFW is directly correlated with TA in this word stage as well. Similar significant correlation was also found when YKX had more than 200 words. In general, no significant correlations between TA and other word-level variables were found when YKX produced less than 150 words, except Tone 4. SA was not correlated with tone accuracy in any word stage. Similar to GXH, significant correlations between TA and AWF were all positive in YKX's production. This result may indicate that words with higher ambient frequency tend to have more accurate Tone compared to words with lower ambient frequency.



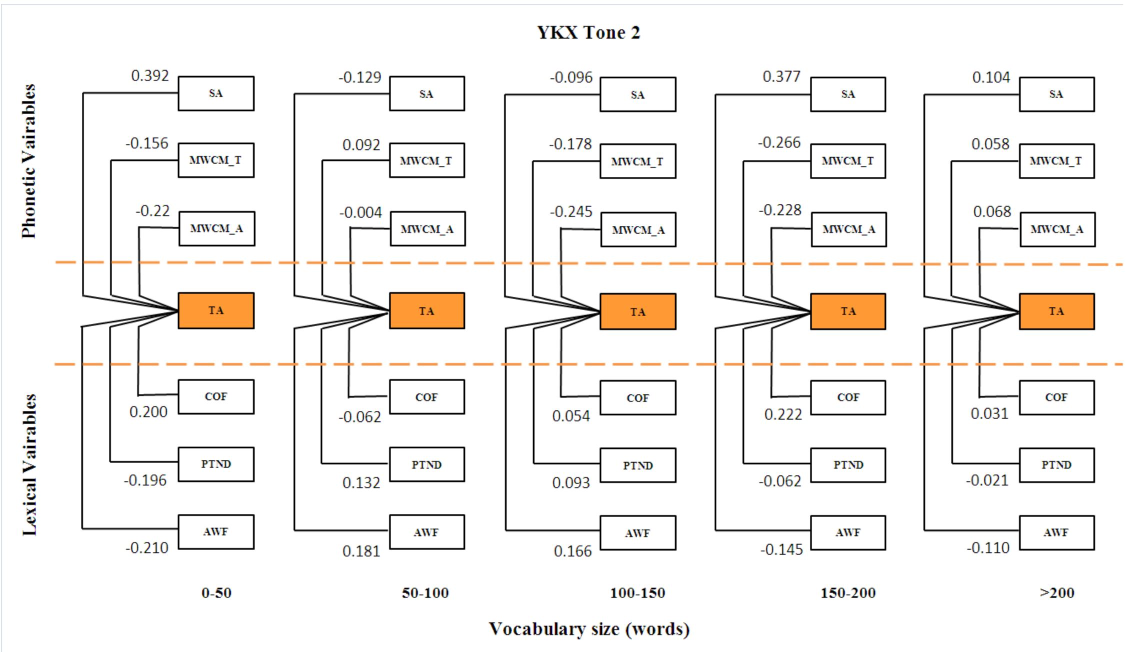
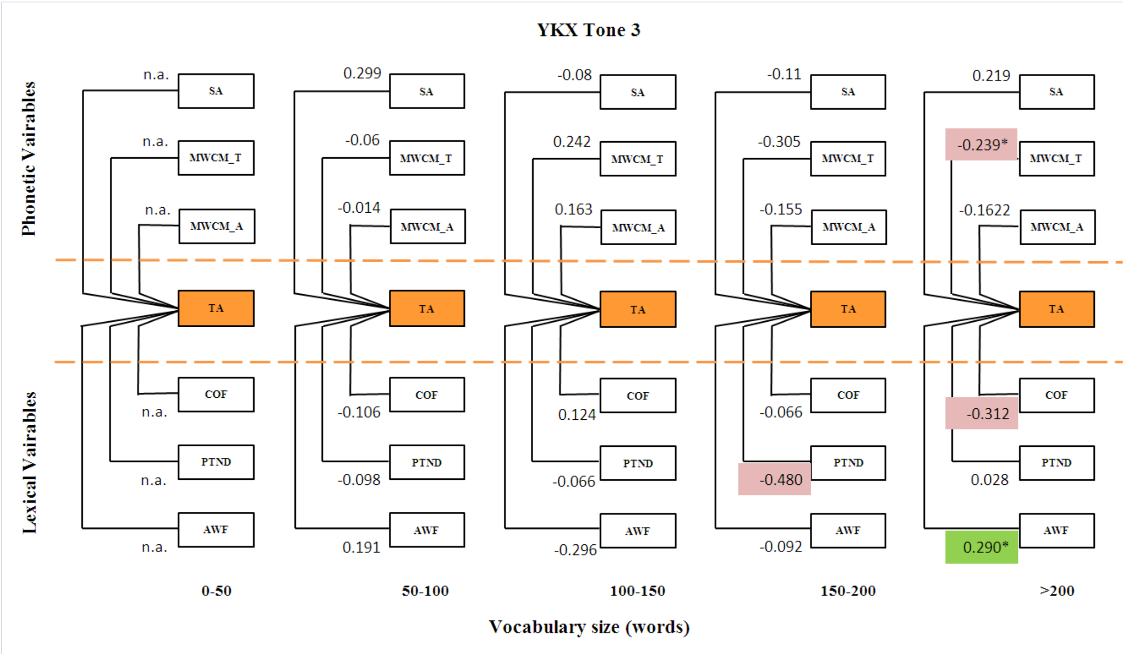
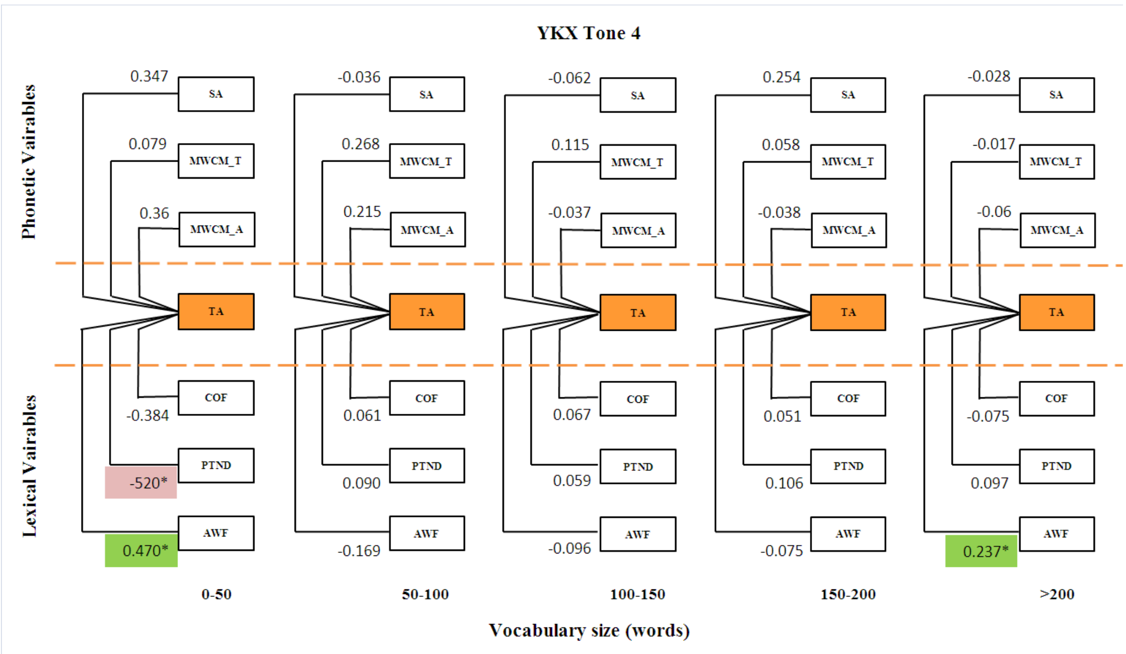


Figure 22: Correlation Coefficients of Tone 1-4 Lexical Items by YKX.

Figure 22: cont.





Summary of tone accuracy and word-level variables

These results revealed a great amount of variation in the relationship between tone accuracy and word-level variables, across tone categories, over word stages and across children. First, word-level variables significantly correlated with tone accuracy were not the same for the four tone categories. For some tone categories, significant correlations were found only with lexical variables but not with phonetic variables. For others, the opposite patterns were observed. For certain tone categories, e.g. Tone 1 in MYK and Tone 2 in YKX, tone accuracy was not related to any of the word-level variables. Second, tone accuracy and word-level variables were related differently at the five word stages. For the same word-level variable, during a given word stage, the relationship was negative but positive in another word stage. Third, children exhibited varied patterns in the relationship between tone accuracy and word-level variables. Tone production in some children tended to be significantly correlated more to lexical level than to phonetic variables (e.g. GXH). For some other children, their tone accuracy was more correlated to phonetic level variables (e.g. MYK).

According to our hypothesis, we would expect negative correlations between TA and PTND, and a decreasing tendency or loosened correlation as children expand their lexicon. But these results did not reliably support this hypothesis.

However, two generalizations can be drawn from the four children: (1) Segmental accuracy was not significantly correlated with tone accuracy for any vocabulary size in any children. This finding lends additional support for the relative independence between segmental and tonal production reported in Section 4.2. (2) Time wise, when children had a vocabulary of less than 50 words, tone accuracy was less likely to be correlated with any word-level variables. For LXB, significant correlations were only found with even larger vocabulary size of more than 150 words. Such a pattern suggests that word-level variables only start to affect tone accuracy when children acquire considerably larger vocabulary which requires internal organization of their mental lexicon.

Chapter 5: Discussion

Lexical tone is a unique phonological feature of tone languages. Children learning a tone language such as Mandarin need to acquire the phonetic ability to produce both tonal and segmental (consonant and vowel) contrasts. They must learn to associate these behavioral capacities with lexical forms to establish phonological knowledge necessary for effective linguistic communication. Behaviorally, this process involves both articulatory control for segmental production and phonatory control for tone production. One goal of the present study was to understand the interaction of these dual-level motor adjustments at the onset of meaningful speech. Relationships between tone accuracy and segmental accuracy were compared longitudinally to evaluate previous claims (e.g. Li & Thompson, 1977; Tuaycharoen, 1977) that tones are acquired earlier than segments.

Conceptually, the establishment of association between tone and words might be inherently related as phonological knowledge is considered as emerging from early word learning. Another goal of the present study was to assess the relationship between tone accuracy and word-level properties at the earliest stages of phonological acquisition. For each lexical item, both phonetic variables and lexical variables were considered as possible sources for variations across lexical items within the same tone category.

5.1. Tone accuracy and segment accuracy

Comparing accuracy of tones and segments across the period of the study, we found that tone accuracy was not always higher than segmental accuracy. This result is contrary to previous studies proposing that tones show error-free production before children are 2 years of age (Chao, 1951; Li & Thompson, 1977; Tuaycharoen, 1977). Chao (1951, pp 278) remarked "the correct tone system is acquired relatively quickly and is mastered well in advance of the segmental system". The claims of these previous studies about the higher accuracy of tone compared to segmental accuracy were made by observations of anecdotal productions of individual children. For example, Li & Thompson (1977) reported that some segments (e.g. affricates) were still pronounced incorrectly when all the tones were produced correctly.

In the present study, we compared the accuracy of tones and segments based on quantitative measures. Since both word position and tone contours influence accuracy of tone production, we used weighted measures to quantify accuracy. The comparisons were also carried out with regard to tone categories. Furthermore, conclusions of previous studies were based on the observations of children's tone and segment production at the end of the study period. The present study considered developmental stages in evaluating the relationships between tone and segmental accuracy. Our results indicate that tone shape and developmental age influence the relationship between tone and segmental accuracy. The production of tone and segment are relatively independent.

5.1.1. Influence of tone shapes

According to our results, tone accuracy is higher than segmental accuracy for Tone 4 words but lower than segmental accuracy in Tone 3 words for all children during earlier periods of development. Segmental accuracy tends to be higher than tone accuracy in Tone

1 and Tone 2 words but with individual child level variations.

One explanation of the effect of tone category would be the difficulty involved in the production of different pitch contours. In Mandarin, Tone 4 is high-falling. Production of Tone 4 requires speakers to manipulate pitch falls by lowering fundamental frequency (F_0). Mandarin Tone 3 is composed of a low falling contour and a rising contour in citation forms. Production of Tone 3 requires phonatory control of a bi-directional pitch change: first lowering F_0 and then raising it.

Physiologically, falling F_0 is considered as a by-product of declining subglottal pressure at the end of utterances produced during a respiratory cycle (Lieberman, 1967). Compared to rising F_0, a falling F_0 requires less motoric efforts in both adults and children (Sunberg, 1979; Xu & Sun, 2002, Ortega-Llebaria, Davis & Yang, in preparation). Falling pitch contours have been reported to predominate in pre-linguistic babbling compared to rising pitch contours in Mandarin and Cantonese infants (Chen, 2005; Chen & Kent, 2009; Lee et al., 2005). The predominance of falling pitch contours during one-word stage was also reported in Li and Thompson (1977). The difficulty of rising versus falling contours has been documented in studies of non-tone languages as well. Snow (1998) reported that 4-year-old children tend to imitate falling intonation contours more frequently than rising contours. This pattern was also observed in the present study as reported in Section 4.1.1. Children produced predominantly more Tone 4 pitch contours regardless of target forms. A recent acoustic study (Wong, 2012) suggested that the Tone 4 produced the most adult-like performance among the four tones in 3-year-old Mandarin-speaking children.

In comparison, phonatory manipulation of Tone 3 requires the speaker to produce a combination of pitch contours. Thus it is motorically more demanding than Tone 4 and other tones in Mandarin. Tone 3 was the least frequent contour in the children's inventory (Yang, Ortega-Llebaria & Davis, 2010). In addition, children were found to have more difficulty with low F_0 compared to higher F_0. The low-falling component in Tone 3 also poses potential challenge for their developing ability of phonatory control.

Therefore, Tone 4 would be more accurately produced compared to Tone 3 if the segmental portions of words with these two tones are similar. Future analysis is needed to compare the complexity of segmental inventory across tone categories.

5.1.2. Influence of developmental stages

Our results on the relationship between tone and segmental accuracy also suggest that the differences between these two decreased as children grew older. The selectivity of words based on segmental properties could be a possible explanation. Take Tone 3 words in LXB for example. During the earlier sessions, LXB selected segments considered easier in speech production of Mandarin children (e.g. "马" [ma3] *horse*, "把" [pa3] *handle*, "打" [ta3] at 17 months). Since Tone 3 is considered motorically more demanding for children, the differences between tone and segmental accuracy would be large at this early stage of development. However, as he attempted productively more challenging segments later in development (“脚” [tɕʰiau 3] *foot*, “桶” [tʰuŋ 3] *bucket*, “嘴” [tsuei 3] *mouth* at 20 months), segmental accuracy decreased to some extent. But LXB started to be more capable of controlling the bidirectional pitch contour of Tone 3, as indicated by an increase in tone accuracy. With an increase in tone accuracy but a decrease in segmental accuracy, the differences between these two measures became smaller. Initial inspection of the Target Word Complexity Measures (Appendix 7) also suggests an increasing tendency toward segmental complexity over time. However, future research is needed to analyze children's segmental inventory over time for a more comprehensive developmental profile.

5.1.3. Relative independence of tone and segmental production

Tonal accuracy does not appear to be interactive with segmental accuracy in these four children. Within each tone category (Section 4.2.2), similar segmental accuracy was observed whether tone is produced accurately or inaccurately. There did not seem to be either facilitative or trade-off effects between tone and segmental accuracy. It is not the case that more accurate production of tone was scaffolded by accurate segmental production. Accurate tone production was not achieved at the sacrifice of segmental accuracy either. Across tone categories (Section 4.2.3), Tone 4 words were produced with the highest accuracy and Tone 3 with the lowest accuracy. However, the accuracy of the segmental portions of Tone 4 words was not higher than the segmental portions of words from other tone categories. Tone 3 words were not produced with the lowest segmental accuracy either.

Tone production is realized through changes in phonation while segmental production of consonants and vowels is realized through changes in articulation. Phonation and articulation are considered as two separate facets of speech production. Anatomically, they are generated by different and independent parts of the speech apparatus. The coordination of various muscles, cartilages and membranes in the laryngeal system are necessary for controlling pitch. Articulators (i.e. tongue, lips, jaw, etc.) in the supralaryngeal system are concerned with the filtering of F_0 to achieve varied vocal qualities of consonants and vowels. Neurologically, the anatomical distinction of phonation and articulation are represented differently in the motor cortex (Brown et al., 2010). Recent neuro-imaging studies have indicated a specific laryngeal area in the human motor cortex which is independent of the area associated with articulation (Brown, Ngan & Liotti, 2008; Simonyan & Horwitz, 2011).

As remarked by Liberman (1984), phonation and articulation must be elaborately and flexibly modified in human speech in order to produce different vocal properties at many different pitches (Fitch, 2000). For Mandarin speakers, the independence of phonation and articulation are also important for meaning generation through pitch changes with the same syllable. The larynx is the physical foundation of phonation and the hyoid provides basis for tongue movements that participate in articulation. Developmental studies on the maturation of the laryngeal and supra-laryngeal systems indicate that the laryngeal skeleton descends rapidly from the hyoid during early infancy (Wind, 1970). Such descent ensures that physical linkage between the laryngeal and the supra-laryngeal systems are weakened (Nishimura, 2003), thus contributing to increased flexibility between phonation and articulation. Because of this biological organization, Mandarin-learning children have already gained relative autonomous control over tone and segment production. Our results furnish additional behavioral evidence, in a language that uses both systems intensively, on the independence of phonation and articulation from a developmental perspective.

In general, due to the lack of systematic research on lexical tone acquisition, people take for granted that the tonal system is mastered well in advance of the segmental system in children learning Mandarin, Cantonese, and other tone languages. Our initial hypothesis that segmental accuracy would always be higher than tone accuracy also oversimplified the developmental progression. These results indicate that the relationship between accuracy of tones and segments, especially at the onset of word learning, is far from a simple high-low dichotomy. It can be influenced by many factors such as the motoric availability of pitch shapes (Tone 4 over Tone 3), preference or selectivity for segments, and the developmental changes in the production repertoire of individual children.

5.2. TONE ACCURACY AND WORD-LEVEL VARIABLES

The second goal of the present study was to evaluate the relationship between tone accuracy and word-level variables in words with the same tone categories in each child. We hypothesized that Productive Tone Neighborhood Density (PTND) which measures tonal contrastivity in children's own production system would be responsible for the within-tone variability observed in Yang (2006). However, tone accuracy was less likely to be correlated with any word-level variables when children had less than 50 words in their vocabulary. No consistent patterns were observed by comparing PTND scores with words with high or low tone accuracy. Statistically, analysis of correlation between tone accuracy and PTND does not suggest consistency in significant correlation outcomes. No decreasing tendency of correlation coefficients between these two was observed either. This result would suggest that children are not sensitive to the contrastivity in their own production at the earliest stage of word learning. For other word-level variables, only Ambient Word Frequency exhibited significant positive correlations with tone accuracy for certain word stages. Other word-level variables did not show consistent correlations with tone accuracy. Tentative explanations are discussed in the following sections.

5.2.1. Tone accuracy and lexical factors

5.2.1.1. Productive Tone Neighborhood Density

The notion of Neighborhood Density was proposed to capture the phonological similarities across words that form "phonological neighborhoods in the mental lexicon (Luce & Pisoni, 1998). Productive Tone Neighborhood Density (PTND) is a measure designed for this study to qualify the contrastivity of tone in children's own productive lexicon. One reason for using this measure is the impossibility of measuring tone neighborhood density in ambient Mandarin. As introduced in Chapter 1, the four tones are evenly distributed in Mandarin syllables (words). As a result, the number of tone neighbors for all words equals 3. However, since children are still in the process of constructing their vocabulary, a certain syllable might not have lexical items illustrating all four tones. It is possible to use PTND to quantify the tone density of different syllables. Another reason for using this density measure based on children's own production is that the vocabulary of a mature adult speaker will be substantially different from children with a smaller vocabulary even if the tonal distribution is not even. Since the present study documented tone acquisition from children's initial word use, using a neighborhood density measure based on children's own productions is more appropriate than an adult-based measure.

Given the appropriateness of the design, one possibility for the unclear relationship between TA and PTND might be holistic representations of word forms in these children's early lexicon. At the onset of word learning, Mandarin children might store and retrieve words as a single unit rather than separate segmental and tonal portions. For example, they might not be conscious that Tone 1word 妈 [ma] "mother" and Tone 3 word 马 [ma] "horse" are similar in segmental components but differ in tone. Therefore, they might not be able to extract patterns of contrastivity as defined by PTND.

This holistic representation in early word learning has been proposed in studies of English-speaking children (e.g. Ferguson & Farewell, 1975; Vihman, 1996, Sosa & Stoel-Gammon, 2006). According to Ferguson & Farewell (1975), children might be sensitive to the detailed phonetic properties of the words, but lack the knowledge of how these phonetic details are organized into words. As our first finding indicates, Mandarin children have already achieved a relative independence of control of segment and tone in their productions. However, the representation of a word as segment and tone may still not be established in their phonological knowledge base. Therefore, tone accuracy might not be sensitive to PTND when segment and tone are stored as a single and unanalyzed unit.

According to our results, significant correlations between TA and PTND were observed as vocabulary increased. But the word stages in which significant correlations started to occur varied across children and across tone categories. Studies in English children suggest that as children expand their lexicon, the number of holistic forms becomes too large to be kept in memory. Children need to move toward a more segmental representation of words (Vihman & Velleman, 1989, Walley, 1993). However, when this transition from whole-word to segmental patterns occurs is debatable. Some studies suggest 50 words (e.g. Vihman, 1996); others suggest 150-200 words (Sosa & Stoel-Gammon, 2006). In the pilot study of Yang & Davis (2008), the relationship between tone accuracy and PTND was not analyzed according to the child's vocabulary size. It is possible that that by 27 months, the child had already begun the process of storing words in a systematic pattern rather than as holistic units. Thus, she may have already begun the process of learning the contrastivity of tone relative to segments in the same word. Therefore, effect of PTND on tone accuracy was observed.

In the case of tone acquisition, it is still unknown whether children follow this holistic to separate segment-tone representation in their word learning. The time when this transition happens in the acquisition process has not been studied either. Li and Thompson (1977) tried to delineate tone acquisition with developmental stages in syntax. They suggested that after the earlier one-word stage and before two-three-word stages, four-way contrast of tones starts to appear but the exact timing was not mentioned. Besides answering these questions, further research is also needed to search for an independent indicator of this representational transition in order to understand whether sensitivity to PTND is caused by the changes in mental representations of tones and segments. A prolonged period of observation is needed to examine whether variations of tone accuracy are related with the contrastivity of tone in children's productive vocabulary.

5.2.1.2. Word frequency measures

In addition to PTND, this study also employed word frequency measures for the purpose of exploring the effect of word properties on production variability of word within the same tone category. Both the effect of ambient language (Ambient Word Frequency) and children's own production (Child Output Frequency) were considered.

Results suggested that a significant correlation between tone accuracy and AWF was only found when children had more than 50 words. One concern about this result is the validity of using frequency counts based on adult corpora. Studies of the difference between child-directed and adult-directed speech indicate that specific vocabularies are often used in conversation with younger children (Snow, 1986). Child-directed speech is more facilitative in younger children with smaller vocabularies than older children who eventually need to learn words presented in adult-directed speech (Ma et al., 2011). Therefore, it might be more appropriate to use frequency corpora based on parental input, especially during the first 50 word stage. Considerable individual variations across children and across tone categories were also observed. Although there are general characteristics of child-directed speech, the specific lexicon in the input to individual children might differ. It would be more informative to use the actual input of each child as a measure for ambient frequency in future studies. Although not directly related to variations in within-tone accuracy, Clumeck (1978) and Clumeck (1980) suggested the influence of actual input on the order of tone acquisition. In Clumeck (1978), high-level and high-falling tones were mastered earlier than other tones. However, for the participant of the 1980 study, rising tone were acquire first, because rising tones occurred more frequently in the input.

With regard to child output frequency, both positive and negative correlations were observed in the four children, indicating that higher frequency of words in children's own production might either facilitate or impede the accurate production of tones. On one hand, it would be logical to assume that high frequency words in each tone category would be more accurate because of more motor practice. On the other hand, vocal motor routines that occur more frequently may be more resistant to change when new sounds or contrasts enter the child's phonological system. Such conflicting findings were also found for children learning English. As reported in Tyler & Edwards (1993), accurate productions of voiceless stops emerged first in words that occurred more often in children's production. On the contrary, Velten (1943) found that words produced most frequently remained inaccurate longest and were more resistant to change.

5.2.2. Tone accuracy and phonetic factors

5.2.2.1. Segmental accuracy

In addition to lexical factors (PTND and word frequency in the ambient language and child output), segmental accuracy was considered as one of the possible phonetic factors influencing within-tone accuracy variations. According to these results, tone accuracy was not influenced by how accurately the child produced the segmental portion of the words. Thus segmental accuracy did not contribute to variations in within-tone accuracy. This lack of relationship was consistently observed for all tone categories in all the children. No significant correlations were found.

Segmental accuracy was considered as the phonetic context for tone production. Its potential influence on tone accuracy variations with the same tone category was predicted based on the findings from studies on segmental accuracy in non-tone languages (Ref. Section 2.4.1.2, Section 2.4.2.2). Davis et al. (2005) reported that production accuracy of English vowels was influenced by the phonetic context of neighboring consonants. However, both production of vowels and consonants were produced by manipulations of relative position of articulators. Their mutual influences were thus expected (MacNeilage & Davis, 1990). In comparison, production of tone and segment involves different speech production systems (articulatory and phonatory) and children seem to gain autonomous control over these two systems (Section 5.1.3) quite early. Although, tone production needs to be realized simultaneously with segmental accuracy in the same syllable, segmental accuracy is not the same type of within system phonetic context that is found in non-tone languages, due to the relative independence of segment and tone production. This finding confirmed pilot results obtained previously for one Mandarin child (Yang and Davis, 2008).

Instead of considering segmental level accuracy, two suprasegmental conditioning factors can be considered as alternative dimensions of phonetic context for tone accuracy. One is syllable duration. The duration of Mandarin tone is acoustically related to the shape of pitch contours (Howie, 1976). Tone 3 tends to have the longest duration compared to other tone categories. This duration pattern has been reported in both infant babbling (Ortega-Llebaria, Davis & Yang, in preparation) and children's speech (Yang, Davis & Diehl, 2008). Therefore, it is possible that a Tone 3 word might be produced with higher accuracy if it occurs in a longer syllable than a word that occurs in a shorter syllable. Yang (2006) found that Tone 3 occurred more frequently on syllables with more than one vowel in the rime portion. Acoustic analysis is needed to further evaluate this hypothesis.

The other possible factor related to phonetic context is the type of tone found on the adjacent syllable (or word). Wong (2008) suggested that bi-syllabic words with compatible tonal sequences are produced more accurately than those with non-compatible sequences. By 'compatible', she meant whether the ending F_0 height of the first tone is the same as the beginning F_0 height of the second. In Mandarin, for example, Tone 1 (55)

- Tone 4 (51) sequences are considered as compatible. In contrast, a Tone 1 (55) – Tone 3 (214) sequence is not. Accordingly, the Tone 1 word in the compatible sequence would be produced more accurately than the Tone 1 word in the non-compatible sequence. Both "吃" [[tʂʰɻ̩]] *eat* and "喝" [xɤ] "drink" are Tone 1 words. 吃 in 吃 (Tone 1)-饭(Tone 4) (eat food) would have higher accuracy than 喝 in 喝 (Tone 1)-水(Tone 3) (drink water). In the present study, accuracy calculation and analysis were based on single syllable morphemes without considering this sequencing effect. Further analysis is needed to evaluate whether tonal shapes in an adjacent syllable can account for accuracy variation of word in the same tone category.

5.2.2.2. Mandarin Word Complexity Measures

Two phonetic factors that could influence with-in tone accuracy were considered; word complexity in the segmental portion of both target words and child actual productions of those words. The relationship between tone accuracy and word complexity measures were mixed. Both positive and negative correlations were observed in different tone categories and varied across the four children. In some cases, within the same tone category, a word with more complex segments in the target would be produced more accurately than word that contained less complex segments. The opposite pattern happened as well. In other cases, if children produced a word in a more complex way regardless of target, it showed a higher accuracy than a word that contained less complex segments. Nevertheless, out of these unclear relationships, for all children, word complexity measures were not significantly correlated with tone accuracy before 100 words. As suggested in Stokes & Surendran (2005), articulatory complexity plays a role in accuracy of consonant production in English at 25 months of age but not before. We could expect more patterned outcomes if later developmental periods are investigated.

The validity of the Mandarin Word Complexity Measure might also relate to these mixed results. In order to quantify production complexity in Mandarin, the Mandarin Word Complexity Measure was designed by adapting a word complexity measure used for English (Stoel-Gammon, 2010). Measurement of phonetic complexity or articulatory complexity has no consensual agreement in the literature. The measurement used in the current study reflects the children's changes required to achieve accurate production for more and more sounds in their ambient language. Later acquired sounds, syllable shapes and word patterns were considered as more complex compared to earlier phonetic units. However, this measure is the first attempt towards a quantitative evaluation of production complexity in Mandarin children. Further large scale studies are needed to test the validity of this measure for capturing developmental trends in Mandarin-learning children.

In summary, it seems that there is no easy answer to the relative importance of Productive Tone Neighborhood Density (PTND), word frequency, and word complexity on word-level variations of tone accuracy within the same tone category based on the present dataset. Hypotheses emerging from studies of segmental-level acquisition in older children do not apply to tone acquisition in Mandarin nor to children aged 24 months and younger. Data beyond 24 months need to be collected to further evaluate what developmental stage is more susceptible to the influence of these word-level variables. Alternative methods of calculating neighborhood density, word frequency and word complexity would be another avenue for further investigation. Cross-linguistic studies from other tonal languages may also provide more information on this issue.

5.3. GENERAL DISCUSSIONS

Physiological basis for tone production in early word learning stage

Tone acquisition in the early stages of Mandarin children's speech development is child-centered. Physiological availability of capacities to produce ambient language tonal forms underlies the distribution patterns of tonal contours in children's actual production and their selection of target word forms. Early tonal production is characterized by the predominance of Tone 4 (falling pitch) which is also found to be more frequent in pre-linguistic babbling and thus appears to be motorically more available to children's production system(e.g. Chen, 2005; Chen & Kent, 2009). This pattern confirmed findings in Li and Thompson (1977). Tone 3 (fall-rise pitch), which occurs least frequent in babbling and appears to be motorically more demanding was produced with the lowest frequency in early sessions by these four children.

Similar to findings in Yang and Lee (2007), GXH and LXB in the present study tended to select more words with falling tones in their early word productions. The other two children didn't show such selectivity. According to a cognitive approach to phonological acquisition, some children first attempt to produce words that are more suitable for their production abilities and others do not (Ferguson & Farwell, 1975; Menn, 1976; Stoel-Gammon & Cooper, 1982; Vihman, 1993). Although different children may follow idiosyncratic approaches to achieving production accuracy, the central tendency is that physiological availability of tonal forms accounts for the higher accuracy of Tone 4 words compared to words with other tonal shapes (Section 4.2.3).

With regard to accuracy variations in words with the same tone, children seemed to store early words holistically in their mental lexicon, as they were not sensitive to the contrastivity of tones in their own production when they had a vocabulary of less than 50 words. Segmental and tonal components were not treated as separate representations in children's knowledge. Then what accounts for the within-tone accuracy variations at earlier stages?

Although it is not possible to address this issue with the present data set, within the same tone category, target words with segment-tone combinations that mirror a child's pre-linguistic vocalization output would be acquired earlier and more accurately compared to sequences that are not present in children's babbling repertoire. Also segment-tone combinations that occur more frequently in babbling might be produced more accurately compared to combinations that occur less frequently. For example, a segment-tone combination [pa] + Tone 4 (falling pitch) might be produced more often in a child's babbling compared to another combination [da] + Tone 4 (falling pitch). When children start to attach meaning to these two combinations in early word learning, 爸 [pa [pa 4] "father" would be more accurate than 大 [ta 4] "big" although they are both Tone 4 words. The idiosyncratic production patterns found in individual children might also be due to the difference in the pre-linguistic forms they use. If another child produced more [da] + Tone 4 combination in his babbling compared to [pa] + Tone 4, the word “大” might be produced with higher accuracy than “爸” at the onset of word learning.

Findings on the relationship between babbling and first words in English support this assumption of pre-linguistic availability for tone accuracy (MacNeilage & Davis, 2000; Oller, Wieman, Doyle, & Ross, 1976; Vihman, 1996). According to MacNeilage (2011), “first words are produced almost exclusively by drawing on the existing babbling repertoire … words are babbling episodes pre-empted for communicating a concept”. Child-specific pre-linguistic vocal patterns are carried forward to the production patterns found in first words (Stoel-Gammon & Cooper, 1984; Vihman, 1993). Previous studies on the continuity of babbling and first words in tone languages have been carried out only at the tone level. Whether a holistic combination of tonal and segmental components in non-meaningful vocalizations evolves into early meaningful speech forms has not been documented. Pre-linguistic babbling data in acquisition of tone languages is needed to evaluate this proposition.

Establishment of tonal knowledge at the increase of vocabulary

As children expand their vocabulary, their early production preference and avoidance for certain tone shapes decreased. Children start to explore words with varied tone and segmental combination in preparation for a more mature lexicon. As suggested by our findings on the relationship between tone accuracy and productive tone neighborhood density, in some children significant correlation between these two were only found beyond 50-word stage. For other children, significant correlations were found even later with a vocabulary size larger than 150 words. Children only start to be sensitive to the contrastivity among tones with similar segmental features when they have a larger vocabulary. Similar to difficulties found in the study of phonological representation in English, only measurable speech behaviors can provide indirect inference to the organization of tonal knowledge in children learning Mandarin and other tone languages. The relationship between PTND could serve as one of the markers of emergent systematicity (Vihman, 1996) in children's tonal acquisition. If this assumption is correct, the attainment of a certain vocabulary size might be associated with phonological reorganization and the emergence of phonemic representation in individual children.

As discussed earlier, tone accuracy and adult word frequency only become significantly correlated beyond the first 50 words. Although using frequency measures based on adult corpora might still be problematic, this finding would imply that children start to pay attention to ambient language properties in tonal acquisition. As a consequence of the expanding lexicon, tonal production is changing from a more child-centered, motor-based process within 50 words to a process that is sensitive to the statistical features of the ambient language.

Dissociation between phonetic output and phonological knowledge

At we discussed earlier, within the first 50 words, a child may treat segmental and tonal properties of a word as an unanalyzed single unit. The knowledge of tonal shape into phonological categories is not yet established. However, in children's production, the relationship between tone accuracy and segmental accuracy suggests relative independence of the two components. These findings seem conflicting at first. However, the dissociation between phonetic output and phonological knowledge implies a dual-level encoding of Mandarin word representations parallel to the proposal of Beckman, Munson & Edwards (2007) for English. Based on the dual-encoding theory in Pierrehumbert (2003), Beckman et al. (2007) posited two levels of phonological representations suggesting that each word is encoded in at least two ways: a "fine-grained" representation of the auditory and articulatory patterns at an item-based level and a "coarser-grained" generalization over phonological patterns at an abstract level. In production, Mandarin children have already acquired the ability to control tone and segment production separately resulting from different anatomical, physiological and neurological basis. Such autonomy offers children the flexibility to add new words with similar segmental component which differ in tone or the same tone which differs in segments into their vocabulary. As the consequence of increased vocabulary size, there is higher-level abstraction of tonal categories thus the separation of tone and segment in phonological knowledge are established. The establishment of categorization of tone would in turn lead to the reorganization of mental lexicon from holistic to systematic (Hoff, Core & Bridges, 2008; Metsala & Walley, 1998).

In summary, the acquisition of phonatory complexity, as represented by tone acquisition in Mandarin, is not a passive process where innate phonological knowledge is revealed through children's maturation. Rather, phonological knowledge emerges on the basis of children's pre-linguistic motor capacities and is established in concert with the expansion of their lexicon.

5.4. METHODOLOGICAL CONSIDERATIONS AND FUTURE RESEARCH

Position-weighted measurement vs. by-position analysis

In considering the effect of word position on production, weighted measures were used in the calculation of tone accuracy, segment accuracy and word complexity (Chapter 3, Section 3.6.1). An alternative method is to calculate accuracy in each word position and analyze separately. For example, comparison of tone and segmental accuracy could be done separately for words that occur in isolation, in the initial or final position of a two-word utterance, etc. Such analysis would permit the investigation of positional effects on the relationship between tone and segmental accuracy. Considering by-position analysis would be even helpful in understanding effect of word-level variables on tone accuracy. For example, the accuracy of a word occurring in isolation might be associated with word-level variables in a different way compared to a word occurring in the medial position of the three-word utterance. Therefore, position-separated analyses are needed in future analysis.

Validity of statistical models

The present longitudinal study of spontaneous speech provided an opportunity to understand phonological development during early word-learning stages when manipulated production experiments are difficult to execute in younger children. However, unlike controlled experimental paradigms in cross-sectional studies, longitudinal data has some unique properties that pose challenges for proper statistical analysis.

First, repeated measurements of any given variable are correlated because they are measured from the same participant over time. A traditional regression models are therefore not suitable due to the violation of the assumption of independence. Secondly, for studies involving home visits of children, unexpected circumstances such as child sickness, family emergency, can interrupt the data collection schedule. Therefore, unbalanced or uneven sampling of time points also makes modeling of temporal correlation challenging. Thirdly, the amount of samples is insufficient. For earlier sessions, children may produce less than 10 word tokens in a session. It is also hard to find words that occur more than 10 times in every session. Such small numbers of word types and tokens prevent statistical analysis. Last, the asymptotic growth pattern prevailing in human growth data requires nonlinear models to study longitudinal data in speech development.

Although statistical analyses moved the field of (psycho)-linguistics beyond qualitative analyses to use quantitative methods, these huge amounts of variability and unpredictability violate the assumptions of many statistical models, thus undermining the validity of these tests. Future efforts are needed to search for a mathematical solution that is most appropriate for the description of longitudinal data involving spontaneous speech collected in younger children.

CONCLUSIONS

Tonal acquisition at the onset of speech development is not a passive process where innate phonological knowledge is revealed solely through children's maturation. Rather, phonological knowledge is established on the basis of children's pre-linguistic motor capacities in concert with cognitive learning occurring via the socially embedded expansion of their lexicon. Tone acquisition at the earliest onset of word learning is more child-centered in that availability of tonal forms to the child's production system underlies accuracy. The relationship between tone and segmental accuracy is influenced by physiological complexity of tonal forms and the child's developmental age. Children gain autonomous control of phonatory and articulatory goal directed behaviors in their production of tones and segments as they learn the nature of precise word targets. Accurate production of tone was not scaffolded by accurate segmental production. It was not achieved at the sacrifice of segmental accuracy either. Accuracy ranking across the four tones didn't entail accuracy ranking of the segmental portions of these four tones. Results indicate that conceptually, tone and segments may be represented holistically in early word forms. Children were not sensitive to the contrastivity (characterized by Productive Tone Neighborhood Density) involved in tonal categories with a vocabulary of less than 50 words. The influences from word-level variables were not consistently found in the four children beyond 50 words to the end of the observation period (i.e. 24 months). Influences from lexical properties of word would only be apparent when phonological knowledge of tonal categories is established with vocabulary expansion. Some research has asserted a lexical driving force for acquisition of phonology for children older than 4 years (e.g. Beckman et al., 2007). This assertion should be tested in Mandarin to establish the generality of this principle across languages.

Appendices

Appendix 1: Participants and Session Information.

Table 1: GXH, Male.

Vocabulary size (Words)	Session number	Age (Year; Months, Days)	Age (Months)	Number of word types	Number of new words	Cumulative number of new words
0-50	1	1;3,15	15.5	6	6	6
	2	1;3,29	16	9	9	15
	3	1;4,11	16.4	5	3	18
	4	1;4,30	17	8	5	23
	5	1;5,17	17.6	27	15	38
	6	1;6,0	18	13	3	41
50-100	7	1;6,25	18.8	46	24	65
	8	1;7,1	19	35	5	70
	9	1;7,12	19.4	35	12	82
	10	1;7,25	19.8	52	27	109
100-150	11	1;8,8	20.3	58	16	125
	12	1;9,4	21.1	36	5	130
	13	1;9,18	21.6	71	22	152
150-200	14	1;10,8	22.3	55	17	169
	15	1;10,26	22.9	69	20	189
>200	16	1;11,10	23.3	63	12	201
	17	1;11,28	23.9	111	42	243

Table 2: LXB, Male.

Vocabulary size (Words)	Session number	Age (Year; Months, Days)	Age (Months)	Number of word types	Number of new words	Cumulative number of new words
0-50	1	1;5,25	17.8	7	7	8
	2	1;6,12	18.4	13	8	15
	3	1;6,23	18.8	17	11	26
	4	1;7,7	19.2	21	9	35
50-100	5	1;8,5	20.2	40	28	63
100-150	6	1;8,19	20.6	77	51	114
150 -200	7	1;9, 4	21.1	109	70	184
> 200	8	1;9, 17	21.6	143	75	259
	9	1;10,0	22	121	41	300
	10	1;10,14	22.5	136	46	346
	11	1;10,28	22.9	130	29	375
	12	1;11,11	23.4	164	44	419
	13	1;11,24	23.8	129	25	444

Table 3: MYK, Male.

Vocabulary size (Words)	Session number	Age (Year; Months, Days)	Age (Months)	Number of word types	Number of new words	Cumulative number of new words
0-50	1	1;2,28	14.9	3	3	4
	2	1;3,19	15.6	6	4	7
	3	1;3,26	15.9	3	2	9
	4	1;4,10	16.3	9	5	14
	5	1;5,14	17.5	18	11	25
	6	1;6,1	18	17	9	34
	7	1;6,13	18.4	14	2	36
	8	1;7,0	19	14	6	42
	9	1;7,9	19.3	25	7	49
50-100	10	1;7,22	19.7	41	17	66
100-150	11	1;8,18	20.6	77	55	121
150-200	12	1;8,28	20.9	100	51	172
>200	13	1;9,18	21.6	91	56	228
	14	1;9,26	21.9	146	62	290
	15	1;10,10	22.3	145	63	353
	16	1;10,24	22.8	194	91	444
	17	1;11,10	23.3	140	24	468
	18	1;11,23	23.8	151	38	506

Table 4: YKX, Female.

Vocabulary size (Words)	Session number	Age (Year; Months, Days)	Age (Months)	Number of word types	Number of new words	Cumulative number of new words
0-50	1	1;4,17	16.6	10	10	10
	2	1;5,1	17	19	14	24
	3	1;5,17	17.6	16	9	33
	4	1;5,29	18	23	14	47
50-100	5	1;6,12	18.4	39	20	67
	6	1;6,27	18.9	34	13	80
	7	1;7,10	19.3	46	20	100
100-150	8	1;7,24	19.8	63	23	123
	9	1;8,7	20.2	82	34	157
150-200	10	1;8,22	20.7	65	20	177
	11	1;9,5	21.2	58	15	192
>200	12	1;9,17	21.6	77	18	210
	13	1;10,17	22.6	56	11	221
	14	1;10,24	22.8	69	18	239
	15	1;11,3	23.1	72	19	258
	16	1;11,14	23.5	106	30	288

Appendix 2: Samples of MWCM Coding.

Table 1: MWCM Coding of Attempted Targets.

Child	Word	Segment target	Segment actual	Tone target	Tone actual	Criteria 1	2	3	4	5	6	7	8	9	10	Total segment MWCM_T	Total tone MWCM_T
GXH	[飞]	[fei]	[tei]	[LH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[机]	[tɕi]	[ti]	[LH]	[FH]	0	0	0	0	2	0	0	0	0	0	2	0
GXH	[机]	[tɕi]	[ki]	[LH]	[RH]	0	0	0	0	2	0	0	0	0	0	2	0
GXH	[飞]	[fei]	[vi]	[LH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[机]	[tɕi]	[ki]	[LH]	[FH]	0	0	0	0	2	0	0	0	0	0	2	0
GXH	[饿]	[ɤ]	[xuə]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[猫]	[mau]	[mau]	[LH]	[RH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[猫]	[mau]	[ma]	[LH]	[LH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[咪]	[mi]	[mi]	[LH]	[LH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[鼻]	[pi]	[pi]	[RH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	1
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[抱]	[pau]	[pau]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[in]	[LH]	[FRH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[i]	[LH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[xi]	[LH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[i]	[LH]	[RFH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[棒]	[paŋ]	[puo]	[FH]	[FH]	0	1	1	0	0	0	0	0	0	0	2	0
GXH	[猫]	[mau]	[mɛn]	[LH]	[LH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[大]	[ta]	[ta]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0

Table 2: MWCM Coding of Actual Productions.

Child	Word	Segment target	Segment actual	Tone target	Tone actual	Criteria 1	2	3	4	5	6	7	8	9	10	Total segment MWCM_A	Total tone MWCM_A
GXH	[飞]	[fei]	[tei]	[LH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[机]	[tɕi]	[ti]	[LH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[机]	[tɕi]	[ki]	[LH]	[RH]	0	0	1	0	0	0	0	0	0	0	1	1
GXH	[飞]	[fei]	[vi]	[LH]	[FH]	0	0	0	1	0	0	0	0	0	0	1	0
GXH	[机]	[tɕi]	[ki]	[LH]	[FH]	0	0	1	0	0	0	0	0	0	0	1	0
GXH	[饿]	[ɤ]	[xuə]	[FH]	[FH]	0	0	0	1	0	0	0	1	0	0	2	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[猫]	[mau]	[mau]	[LH]	[RH]	0	0	0	0	0	0	0	1	0	0	1	1
GXH	[猫]	[mau]	[ma]	[LH]	[LH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[咪]	[mi]	[mi]	[LH]	[LH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[鼻]	[pi]	[pi]	[RH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[抱]	[pau]	[pa]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[抱]	[pau]	[pau]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[饿]	[ɤ]	[ɤ]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[in]	[LH]	[FRH]	0	1	0	0	0	0	0	0	0	0	1	2
GXH	[一]	[i]	[i]	[LH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0
GXH	[一]	[i]	[xi]	[LH]	[FH]	0	0	0	1	0	0	0	0	0	0	1	0
GXH	[一]	[i]	[i]	[LH]	[RFH]	0	0	0	0	0	0	0	0	0	0	0	1
GXH	[棒]	[paŋ]	[puo]	[FH]	[FH]	0	0	0	0	0	0	0	1	0	0	1	0
GXH	[猫]	[mau]	[mɛn]	[LH]	[LH]	0	1	0	0	0	0	0	0	0	0	1	0
GXH	[大]	[ta]	[ta]	[FH]	[FH]	0	0	0	0	0	0	0	0	0	0	0	0

Appendix 3: A Sample of Calculated Variable Values.

Child	Word	Age (Y; M; D)	Word stages	Target Segment	Target Tone	TA	SA	MWCM_T	MWCM_A	COF	PTND	AWF
GXH	[猴]	1;7,12	50-100	[xou]	2	0.00	1.00	2.00	2.00	2	0	97.73
GXH	[星]	1;7,12	50-100	[ɕiŋ]	1	0.00	0.67	3.00	2.00	1	0	77.32
GXH	[亲]	1;7,12	50-100	[tɕʰin]	1	0.00	1.00	4.00	4.00	1	0	68.80
GXH	[想]	1;7,12	50-100	[ɕiaŋ]	3	0.00	1.00	4.00	4.00	2	1	41.58
GXH	[胖]	1;7,12	50-100	[pʰaŋ]	4	1.00	0.50	3.00	4.00	2	0	96.70
GXH	[红]	1;7,12	50-100	[xuŋ]	2	0.00	1.00	3.00	3.00	1	0	75.89
GXH	[里]	1;7,12	50-100	[li]	3	0.00	0.33	1.00	0.00	3	0	30.52
GXH	[高]	1;7,25	50-100	[kau]	1	0.00	0.00	2.00	1.00	1	0	47.45
GXH	[后]	1;7,25	50-100	[xou]	4	0.00	0.50	2.00	2.00	1	1	29.96
GXH	[前]	1;7,25	50-100	[tɕʰien]	2	1.00	0.67	5.00	4.00	1	0	40.43
GXH	[盖]	1;7,25	50-100	[kai]	4	1.00	0.50	1.00	0.00	3	0	89.94
GXH	[上]	1;7,25	50-100	[ʂaŋ]	4	1.00	0.33	3.00	4.00	1	0	17.61
GXH	[右]	1;7,25	50-100	[iou]	4	0.00	0.00	2.00	1.00	1	1	84.90
GXH	[左]	1;7,25	50-100	[tsuo]	3	0.00	0.00	3.00	2.00	2	1	84.87
GXH	[来]	1;7,25	50-100	[lai]	2	0.50	0.75	1.00	2.00	2	0	17.06
GXH	[牙]	1;7,25	50-100	[ja]	2	1.00	1.00	0.00	0.00	4	0	89.09
GXH	[下]	1;7,25	50-100	[ɕia]	4	0.91	1.00	2.00	2.00	11	0	28.12
GXH	[笔]	1;7,25	50-100	[pi]	3	0.00	1.00	0.00	0.00	1	0	88.40
GXH	[鹿]	1;7,25	50-100	[lu]	4	1.00	0.50	1.00	0.00	5	0	97.32
GXH	[倒]	1;7,25	50-100	[tau]	4	0.00	1.00	1.00	1.00	1	0	79.91
GXH	[看]	1;7,25	50-100	[kʰan]	4	0.00	0.00	3.00	4.00	1	0	36.88
GXH	[烫]	1;7,25	50-100	[tʰaŋ]	4	1.00	1.00	3.00	3.00	2	0	99.08
GXH	[伞]	1;7,25	50-100	[san]	3	0.00	0.67	2.00	1.50	2	0	97.78

Appendix 4: Distribution of Tone 1-4 Contours in Actual Production across Months.

Table 1: GXH.

Age (months)	GXH Actual Production								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
15	1	11.11	1	11.11	0	0.00	7	77.78	9
16	3	10.00	6	20.00	0	0.00	21	70.00	30
17	12	12.24	23	23.47	0	0.00	63	64.29	98
18	14	6.93	120	59.41	5	2.48	63	31.19	202
19	14	9.03	86	55.48	12	7.74	43	27.74	155
20	55	16.72	121	36.78	5	1.52	148	44.98	329
21	72	22.43	110	34.27	3	0.93	136	42.37	321
22	106	20.74	194	37.96	44	8.61	167	32.68	511
23	161	21.16	185	24.31	127	16.69	288	37.84	761
Overall	438	18.13	846	35.02	196	8.11	936	38.74	2416

Table 2: LXB.

Age (months)	LXB Actual Production								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
17	2	6.06	7	21.21	0	0	24	72.73	33
18	33	27.05	6	4.92	0	0	83	68.03	122
19	20	28.57	8	11.43	0	0	42	60	70
20	53	14.89	131	36.8	36	10.11	136	38.2	356
21	189	27.67	167	24.45	123	18.01	204	29.87	683
22	291	23.24	305	24.36	152	12.14	504	40.26	1252
23	193	18.72	237	22.99	133	12.9	468	45.39	1031
Overall	781	22.02	861	24.27	444	12.52	1461	41.19	3547

Table 3: MYK.

Age (months)	MYK Actual Production								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
14	1	16.67	0	0	0	0	5	83.33	6
15	5	12.5	12	30	2	5	21	52.5	40
16	3	15.79	3	15.79	1	5.26	12	63.16	19
17	22	27.16	1	1.23	8	9.88	50	61.73	81
18	16	16.84	12	12.63	3	3.16	64	67.37	95
19	63	29.03	61	28.11	10	4.61	83	38.25	217
20	141	27.87	182	35.97	33	6.52	150	29.64	506
21	223	31.86	209	29.86	41	5.86	227	32.43	700
22	317	37.29	239	28.12	68	8	226	26.59	850
23	261	30.38	229	26.66	91	10.59	278	32.36	859
Overall	1052	31.19	948	28.11	257	7.62	1116	33.09	3373

Table 4: YKX.

Age (months)	YKX Actual Production								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
16	4	6.67	14	23.33	1	1.67	41	68.33	60
17	10	12.66	5	6.33	1	1.27	63	79.75	79
18	48	24	33	16.5	1	0.5	118	59	200
19	49	19.22	58	22.75	21	8.24	127	49.8	255
20	79	23.1	94	27.49	64	18.71	105	30.7	342
21	86	24.5	45	12.82	105	29.91	115	32.76	351
22	85	22.19	109	28.46	49	12.79	140	36.55	383
23	65	14.04	123	26.57	81	17.49	194	41.9	463
Overall	426	19.97	481	22.55	323	15.14	903	42.33	2133

Appendix 5: Distribution of Word Tokens of Tone 1-4 in Attempted Targets across Months.

Table 1: GXH.

Age (months)	GXH Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
15	2	22.22	1	11.11	0	0.00	6	66.67	9
16	15	48.39	3	9.68	0	0.00	13	41.94	31
17	31	31.63	21	21.43	9	9.18	37	37.76	98
18	46	22.44	68	33.17	46	22.44	45	21.95	205
19	43	27.39	43	27.39	37	23.57	34	21.66	157
20	68	21.25	77	24.06	68	21.25	107	33.44	320
21	102	31.58	85	26.32	45	13.93	91	28.17	323
22	122	24.06	140	27.61	102	20.12	143	28.21	507
23	131	18.30	166	23.18	163	22.77	256	35.75	716
Overall	560	23.67	604	25.53	470	19.86	732	30.94	2366

Table 2: LXB.

Age (months)	LXB Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
17	6	18.18	3	9.09	4	12.12	20	60.61	33
18	42	33.87	22	17.74	18	14.52	42	33.87	124
19	31	44.29	13	18.57	4	5.71	22	31.43	70
20	119	32.51	123	33.61	64	17.49	60	16.39	366
21	186	26.76	193	27.77	124	17.84	192	27.63	695
22	271	23.16	293	25.04	185	15.81	421	35.98	1170
23	199	21.13	207	21.97	146	15.5	390	41.4	942
Overall	854	25.12	854	25.12	545	16.03	1147	33.74	3400

Table 3: MYK.

Age (months)	MYK Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
14	4	80	0	0	1	20	0	0	5
15	17	47.22	10	27.78	5	13.89	4	11.11	36
16	7	38.89	4	22.22	3	16.67	4	22.22	18
17	31	38.27	7	8.64	9	11.11	34	41.98	81
18	45	47.37	12	12.63	9	9.47	29	30.53	95
19	69	31.65	47	21.56	47	21.56	55	25.23	218
20	126	24.47	158	30.68	128	24.85	103	20	515
21	141	20.35	189	27.27	121	17.46	242	34.92	693
22	224	26.96	203	24.43	144	17.33	260	31.29	831
23	243	30.04	155	19.16	161	19.9	250	30.9	809
Overall	907	27.48	785	23.78	628	19.02	981	29.72	3301

Table 4: YKX.

Age (months)	YKX Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Tokens
	Token	%	Token	%	Token	%	Token	%	
16	14	23.33	45	75	1	1.67	0	0	60
17	22	27.16	25	30.86	3	3.7	31	38.27	81
18	76	37.25	62	30.39	22	10.78	44	21.57	204
19	82	30.83	53	19.92	53	19.92	78	29.32	266
20	105	29.09	103	28.53	66	18.28	87	24.1	361
21	120	31.83	62	16.45	86	22.81	109	28.91	377
22	121	32.1	74	19.63	67	17.77	115	30.5	377
23	86	18.74	103	22.44	125	27.23	145	31.59	459
Overall	626	28.65	527	24.12	423	19.36	609	27.87	2185

Appendix 6: Distribution of Word Types of Tone 1-4 in Attempted Targets across Months.

Table 1: GXH.

Age (months)	GXH Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Types
	Type	%	Type	%	Type	%	Type	%	
15	2	33.33	0	0.00	0	0.00	4	66.67	6
16	6	42.86	3	21.43	0	0.00	5	35.71	14
17	10	28.57	8	22.86	4	11.43	13	37.14	35
18	14	23.73	16	27.12	11	18.64	18	30.51	59
19	22	31.43	16	22.86	14	20.00	18	25.71	70
20	21	19.09	29	26.36	25	22.73	35	31.82	110
21	28	26.17	25	23.36	19	17.76	35	32.71	107
22	26	20.97	36	29.03	26	20.97	36	29.03	124
23	39	22.41	44	25.29	34	19.54	57	32.76	174
Overall	168	24.03	177	25.32	133	19.03	221	31.62	699

Table 2: LXB.

Age (months)	LXB Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Types
	Type	%	Type	%	Type	%	Type	%	
17	1	14.29	1	14.29	2	28.57	3	42.86	7
18	11	36.67	4	13.33	7	23.33	8	26.67	30
19	9	42.86	2	9.52	2	9.52	8	38.1	21
20	29	24.79	36	30.77	25	21.37	27	23.08	117
21	71	28.17	57	22.62	50	19.84	74	29.37	252
22	106	27.39	86	22.22	73	18.86	122	31.52	387
23	88	30.03	57	19.45	62	21.16	86	29.35	293
Overall	315	28.46	243	21.95	221	19.96	328	29.63	1107

Table 3: MYK.

Age (months)	MYK Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Types
	Type	%	Type	%	Type	%	Type	%	
14	2	66.67	0	0	1	33.33	0	0	3
15	4	44.44	2	22.22	2	22.22	1	11.11	9
16	3	33.33	3	33.33	1	11.11	2	22.22	9
17	5	27.78	4	22.22	4	22.22	5	27.78	18
18	13	41.94	3	9.68	6	19.35	9	29.03	31
19	26	32.5	21	26.25	13	16.25	20	25	80
20	49	27.68	53	29.94	34	19.21	41	23.16	177
21	59	24.89	58	24.47	39	16.46	81	34.18	237
22	88	25.96	81	23.89	66	19.47	104	30.68	339
23	76	26.12	74	25.43	56	19.24	85	29.21	291
Overall	325	27.22	299	25.04	222	18.59	348	29.15	1194

Table 4: YKX.

Age (months)	YKX Attempted Target								
	Tone 1		Tone 2		Tone 3		Tone 4		Total Types
	Type	%	Type	%	Type	%	Type	%	
16	4	40	5	50	1	10	0	0	10
17	10	28.57	11	31.43	3	8.57	11	31.43	35
18	30	31.25	24	25	18	18.75	24	25	96
19	29	26.61	25	22.94	25	22.94	30	27.52	109
20	34	23.13	48	32.65	27	18.37	38	25.85	147
21	43	31.85	26	19.26	23	17.04	43	31.85	135
22	46	36.8	22	17.6	22	17.6	35	28	125
23	42	23.6	46	25.84	34	19.1	56	31.46	178
Overall	238	28.5	207	24.79	153	18.32	237	28.38	835

Appendix 7: Tone Accuracy (TA), Segment Accuracy (SA) and the Differences between Segment and Tone Accuracy (Diff.) across Months (in Percentage).

Table 1: GXH.

GXH	Tone 1			Tone 2			Tone 3			Tone 4		
Age (months)	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.
15	50.00	25.00	-25.00	100.00	100.00	0.00	na	na	na	100.00	25.00	-75.00
16	20.00	62.50	42.50	75.00	37.50	-37.50	na	na	na	100.00	71.97	-28.03
17	26.33	76.89	50.56	22.50	32.50	10.00	0.00	95.00	95.00	89.71	53.10	-36.60
18	11.42	84.36	72.94	61.16	43.04	-18.13	6.52	59.78	53.26	89.29	47.26	-42.02
19	21.54	75.00	53.46	64.80	62.28	-2.52	19.01	70.54	51.54	88.54	67.07	-21.47
20	45.09	77.95	32.86	71.85	60.08	-11.77	3.33	65.57	62.24	81.30	71.97	-9.33
21	50.97	71.44	20.46	70.11	67.97	-2.14	4.17	48.93	44.76	89.38	74.86	-14.52
22	61.82	73.03	11.21	80.50	71.61	-8.90	33.64	61.17	27.53	82.25	71.92	-10.33
23	52.66	70.07	17.41	47.56	79.93	32.37	53.66	85.18	31.53	69.91	68.97	-0.93

Table 2: LXB.

LXB	Tone 1			Tone 2			Tone 3			Tone 4		
Age (months)	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.
17	16.67	0.00	-16.67	0.00	100.00	100.00	0.00	75.00	75.00	85.00	87.50	2.50
18	71.14	73.07	1.93	4.55	77.27	72.73	0.00	81.25	81.25	90.91	73.52	-17.39
19	66.06	73.35	7.29	4.55	58.04	53.50	0.00	77.08	77.08	93.18	60.64	-32.54
20	29.36	72.39	43.03	40.18	75.40	35.23	20.65	60.04	39.39	81.25	66.88	-14.38
21	65.73	62.70	-3.02	48.72	75.90	27.18	66.59	67.63	1.04	76.22	68.52	-7.69
22	65.03	70.13	5.10	60.64	75.77	15.13	56.26	60.38	4.13	79.67	68.14	-11.53
23	60.37	72.31	11.94	67.18	80.49	13.31	61.64	65.30	3.65	82.65	65.45	-17.20

Table 3: MYK.

MYK	Tone 1			Tone 2			Tone 3			Tone 4		
Age (months)	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.
14	25.00	75.00	50.00	0.00	0.00	0.00	0.00	100.00	100.00	0.00	0.00	0.00
15	0.00	41.18	41.18	87.50	56.25	-31.25	0.00	18.75	18.75	50.00	18.75	-31.25
16	14.29	100.00	85.71	75.00	25.00	-50.00	0.00	83.33	83.33	100.00	37.50	-62.50
17	38.71	61.29	22.58	0.00	71.43	71.43	44.44	61.11	16.67	64.71	55.88	-8.82
18	21.85	56.51	34.66	0.00	95.00	95.00	32.14	71.43	39.29	86.43	69.92	-16.51
19	57.78	68.42	10.64	27.35	70.97	43.62	5.88	61.50	55.61	73.02	67.70	-5.32
20	52.68	52.92	0.24	50.01	55.45	5.44	13.99	71.38	57.39	73.17	52.18	-20.99
21	55.47	44.50	-10.98	52.24	49.80	-2.44	21.20	37.55	16.35	57.33	56.47	-0.85
22	68.51	50.25	-18.26	58.48	51.83	-6.64	28.21	50.52	22.30	55.98	44.84	-11.13
23	73.64	58.32	-15.32	63.39	62.46	-0.93	39.78	58.08	18.31	77.44	59.12	-18.32

Table 4: YKX.

YKX	Tone 1			Tone 2			Tone 3			Tone 4		
Age (months)	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.	TA	SA	Diff.
16	14.29	60.71	46.43	20.00	80.74	60.74	0.00	0.00	0.00	na	na	na
17	0.00	62.88	62.88	12.70	68.85	56.15	0.00	37.50	37.50	75.00	31.02	-43.98
18	48.75	62.05	13.31	16.43	65.69	49.27	3.70	42.06	38.36	80.98	28.48	-52.49
19	47.19	50.99	3.80	32.78	64.42	31.64	17.74	41.53	23.79	88.89	42.39	-46.49
20	68.67	62.34	-6.33	58.60	43.23	-15.38	53.03	49.25	-3.77	91.16	51.54	-39.62
21	70.18	63.19	-6.99	31.77	43.21	11.44	82.12	72.80	-9.32	92.88	43.71	-49.18
22	62.44	79.89	17.45	69.31	79.45	10.14	53.69	77.43	23.74	93.14	68.97	-24.17
23	59.32	72.93	13.61	74.77	76.85	2.08	56.22	68.81	12.59	89.00	65.42	-23.58

Appendix 8: Segment Accuracy (SA) When Tone is Produced Accurately (Acc.) or Inaccurately (Inacc.) across Months (in Percentage).

Table 1: GXH.

GXH	Tone 1		Tone 2		Tone 3		Tone 4	
Age (months)	Acc.	Inacc.	Acc.	Inacc.	Acc.	Inacc.	Acc.	Inacc.
15	0.00	50.00	na	na	na	na	29.17	na
16	75.00	60.42	25.00	100.00	na	na	74.17	na
17	na	80.29	25.00	31.79	na	91.67	55.91	29.17
18	100.00	74.17	43.18	43.57	100.00	41.64	61.96	33.13
19	84.72	66.22	62.08	73.63	na	73.61	80.83	39.35
20	75.00	68.58	57.75	57.12	50.00	67.52	65.36	39.00
21	75.00	71.31	73.21	72.72	na	61.46	72.39	66.07
22	70.83	76.51	60.96	88.95	42.50	67.86	63.87	74.24
23	86.31	66.93	76.18	79.08	70.14	70.94	65.95	63.25

Table 2: LXB.

LXB	Tone 1		Tone 2		Tone 3		Tone 4	
Age (months)	Acc.	Inacc.	Acc.	Inacc.	Acc.	Inacc.	Acc.	Inacc.
17	na	0.00	na	100.00	na	75.00	79.17	92.31
18	75.00	95.24	na	50.00	na	62.50	55.42	na
19	41.67	71.03	na	61.90	na	75.00	49.07	16.67
20	87.50	63.75	60.61	67.86	58.33	43.81	45.83	66.20
21	71.00	63.38	65.00	78.19	72.99	64.29	65.31	52.52
22	69.95	71.60	61.92	77.63	74.23	57.59	66.60	69.57
23	83.21	70.27	76.92	77.27	73.72	69.23	65.39	72.67

Table 3: MYK.

MYK	Tone 1		Tone 2		Tone 3		Tone 4	
Age (months)	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**
14	na	75.00	na	na	na	100.00	na	na
15	na	83.04	25.00	87.50	na	18.75	37.50	na
16	na	100.00	33.33	50.00	na	83.33	25.00	na
17	na	78.33	50.00	60.00	na	55.00	50.00	68.33
18	25.00	63.84	na	87.50	87.50	46.88	75.30	27.08
19	70.83	57.72	33.33	64.05	na	66.67	44.91	86.94
20	48.33	51.97	42.39	73.30	75.00	57.26	47.32	50.75
21	57.50	42.32	53.32	54.28	37.50	44.37	54.72	54.97
22	51.90	54.59	47.01	53.41	45.37	49.20	58.84	41.64
23	54.91	58.48	62.84	57.85	61.11	53.90	56.83	49.84

Table 4: YKX.

YKX	Tone 1		Tone 2		Tone 3		Tone 4	
Age (months)	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**	**Acc.**	**Inacc.**
16	na	35.94	na	45.00	na	0.00	na	na
17	na	54.58	na	56.73	na	37.50	67.36	23.61
18	45.00	52.39	66.67	51.84	100.00	45.83	37.19	18.75
19	30.21	58.01	40.00	69.58	25.00	36.24	44.05	37.75
20	52.98	64.70	43.72	38.09	30.42	45.24	46.66	57.29
21	63.45	60.71	71.83	39.92	62.50	65.00	46.31	17.92
22	81.73	76.34	72.08	91.35	81.67	72.65	64.37	64.45
23	66.50	74.97	74.65	70.47	83.44	63.34	68.67	74.37

Appendix 9: Average Variable Values by Vocabulary Size.

Table 1: GXH.

A). Tone 1

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.31	0.70	1.32	1.12	2.77	0.05	75.70
50-100	0.26	0.72	1.93	1.53	2.43	0.28	81.65
100-150	0.51	0.72	2.00	1.71	3.63	0.24	76.21
150-200	0.51	0.75	2.15	1.70	4.58	0.38	81.14
>200	0.52	0.73	1.90	1.71	3.33	0.33	85.07

B). Tone 2

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.43	0.44	1.64	0.80	2.21	0.14	89.05
50-100	0.72	0.61	1.91	1.70	2.77	0.16	85.25
100-150	0.83	0.64	1.74	1.46	3.00	0.18	81.61
150-200	0.79	0.69	1.97	1.77	3.47	0.06	82.99
>200	0.44	0.77	2.20	1.93	3.57	0.11	81.08

C). Tone 3

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.43	0.44	1.64	0.80	2.21	0.14	89.05
50-100	0.72	0.61	1.91	1.70	2.77	0.16	85.25
100-150	0.83	0.64	1.74	1.46	3.00	0.18	81.61
150-200	0.79	0.69	1.97	1.77	3.47	0.06	82.99
>200	0.44	0.77	2.20	1.93	3.57	0.11	81.08

D). Tone 4

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.43	0.44	1.64	0.80	2.21	0.14	89.05
50-100	0.72	0.61	1.91	1.70	2.77	0.16	85.25
100-150	0.83	0.64	1.74	1.46	3.00	0.18	81.61
150-200	0.79	0.69	1.97	1.77	3.47	0.06	82.99
>200	0.44	0.77	2.20	1.93	3.57	0.11	81.08

Table 2: LXB.

A). Tone 1

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.55	0.65	0.81	0.43	3.76	0.38	74.54
50-100	0.56	0.63	1.62	1.28	5.77	0.31	75.25
100-150	0.32	0.70	1.88	1.59	3.00	0.13	74.07
150-200	0.72	0.61	2.25	1.86	1.94	0.34	77.86
>200	0.64	0.71	2.06	1.84	2.74	0.39	74.28

B). Tone 2

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.05	0.68	0.00	0.26	5.43	0.86	94.47
50-100	0.56	0.58	1.55	1.53	5.00	0.18	78.68
100-150	0.52	0.65	2.32	2.09	2.56	0.08	80.22
150-200	0.50	0.69	2.33	2.03	2.94	0.27	83.33
>200	0.59	0.72	1.98	1.80	2.51	0.26	73.35

C). Tone 3

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.00	0.77	0.36	0.18	2.36	0.64	74.94
50-100	0.00	0.63	2.00	1.53	2.00	0.22	69.73
100-150	0.51	0.49	1.94	1.72	2.88	0.19	73.49
150-200	0.65	0.60	1.85	1.64	2.75	0.40	72.33
>200	0.66	0.68	2.02	1.69	2.73	0.55	65.02

D). Tone 4

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.88	0.55	1.00	0.61	4.42	0.26	66.86
50-100	0.79	0.66	2.00	1.84	2.86	0.29	51.92
100-150	0.77	0.52	2.00	2.13	2.05	0.05	63.31
150-200	0.82	0.67	1.88	1.91	2.04	0.50	77.47
>200	0.79	0.67	1.97	1.72	3.78	0.34	59.09

Table 3: MYK.

A). Tone 1

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.38	0.74	0.75	0.50	3.20	0.14	79.41
50-100	0.38	0.60	0.75	0.48	3.50	0.38	91.00
100-150	0.51	0.54	1.56	1.27	2.28	0.24	83.01
150-200	0.59	0.47	1.71	1.21	2.50	0.46	80.05
>200	0.66	0.53	2.19	1.71	2.65	0.42	77.28

B). Tone 2

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.26	0.57	0.78	0.57	2.50	0.22	91.28
50-100	0.38	0.50	1.07	0.68	2.53	0.13	83.23
100-150	0.61	0.52	1.71	1.27	4.04	0.17	81.99
150-200	0.60	0.63	1.79	1.56	2.07	0.28	84.26
>200	0.61	0.55	1.95	1.64	2.31	0.33	79.95

C). Tone 3

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.12	0.68	0.95	0.64	2.05	0.05	69.44
50-100	0.10	0.42	2.00	1.19	4.63	0.25	77.86
100-150	0.09	0.63	1.79	1.22	5.07	0.14	73.95
150-200	0.27	0.56	1.70	1.27	3.10	0.55	73.21
>200	0.38	0.51	1.85	1.42	2.78	0.53	69.12

D). Tone 4

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.84	0.60	1.30	0.94	3.30	0.04	62.16
50-100	0.66	0.45	1.70	1.11	3.40	0.10	74.38
100-150	0.72	0.57	1.71	1.47	2.79	0.29	61.40
150-200	0.70	0.41	1.70	1.19	2.63	0.30	71.01
>200	0.64	0.52	1.88	1.60	2.87	0.43	63.32

Table 4: YKX.

A). Tone 1

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.10	0.46	1.45	0.72	2.60	0.25	64.20
50-100	0.38	0.51	1.39	0.98	2.33	0.17	79.36
100-150	0.63	0.67	1.51	0.87	3.17	0.23	80.11
150-200	0.80	0.58	1.82	1.19	2.76	0.03	70.10
>200	0.66	0.71	2.06	1.68	2.65	0.12	74.91

B). Tone 2

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.10	0.56	0.81	0.34	4.48	0.24	94.08
50-100	0.11	0.56	1.58	1.08	1.77	0.16	87.36
100-150	0.53	0.44	1.84	1.31	2.25	0.16	85.28
150-200	0.40	0.43	2.13	1.50	2.16	0.26	74.94
>200	0.71	0.71	2.01	1.68	2.48	0.19	77.80

C). Tone 3

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.00	0.25	1.50	0.69	1.25	0.38	78.19
50-100	0.04	0.46	1.30	1.05	1.65	0.17	74.34
100-150	0.45	0.42	1.94	1.31	1.88	0.22	70.62
150-200	0.70	0.50	1.95	1.66	3.80	0.20	64.39
>200	0.67	0.72	1.93	1.56	3.41	0.23	62.43

D). Tone 4

Vocabulary size (words)	TA	Phonetic variables			Lexical variables		
		SA	MWCM_T	MWCM_A	COF	PTND	AWF
0-50	0.75	0.32	1.58	0.97	2.58	0.05	48.24
50-100	0.85	0.38	1.55	0.91	2.34	0.03	70.55
100-150	0.90	0.50	1.68	1.29	2.06	0.24	63.30
150-200	0.91	0.44	1.92	1.50	2.61	0.13	60.30
>200	0.92	0.62	2.04	1.78	2.70	0.21	57.57

References

Allen, G. D., & Hawkins S. (1980) Phonological rhythm: definition and development. In *Child Phonology. Volume 1:Production*. G.Yeni-Komshian, J. Kavanagh and C. Ferguson (eds.). New York: Academic Press.

Barlow, J. A., & Gierut, J. A. (1999). Optimality theory in phonological acquisition. *Journal of Speech, Language, and Hearing Research, 42*(6), 1482–1498.

Beckman, M. E., & Edwards, J. (2000). The ontogeny of phonological categories and the primacy of lexical learning in linguistic development. *Child Development, 71*(1), 240-249.

Beckman, M. E., Munson, B., & Edwards, J. (2007). Vocabulary growth and the developmental expansion of types of phonological knowledge. *Laboratory Phonology, 9*, 241-264.

Bernhardt, B. H., & Stemberger, J. P. (1998). *Handbook of phonological development from the perspective of constraint-based nonlinear phonology*. San Diego: Academic Press.

Boysson-Bardies, B. de, & Vihman, M. M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language, 67*(2), 297-319.

Boysson-Bardies, B. de, Vihman, M.M., Rough-Hellochius, L., Durand, C., Landberg, I., & Arao, F. (1992) Material evidence of infant selection from target language: A cross Linguistic phonetic study. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications* (pp. 369-391).

Timonium, MD: York Press.

Brown, S, Laird, A. R., Pfordresher, P. Q., Thelen, S. M., Turkeltaub, P., & Liotti, M. (2010). The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain Cognition*, 70(1), 31-41.

Brown, S., Ngan, E., & Liotti, M. (2008). A larynx area in the human motor cortex. Cerebral Cortex, *18*(4), 837-845.

Chao, Y. R. (1930). A system of tone letters. Reprinted in *Fangyan* (Dialects), 1980 (2), 81-83.

Chao, Y. R. (1951). The Cantian idiolect: an analysis of the Chinese spoken by a 28-months-old child. In Y. R. Chao (Ed.), *Aspects of Chinese sociolinguistics: essays*. Stanford, CA: Stanford University Press.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chen, F. Y. (2007). The Development of Aspiration in Mandarin: A longitudinal case study. MA thesis. Hunan University, Changsha, Hunan.

Chen, L.-M., & Kent, R. D. (2009). Development of prosodic patterns in Mandarin-learning infants. *Journal of Child Language, 36*, 73-84.

Chen, Y.H. (2005). *The pitch contours of infant babbling adapting to the target language*. MA thesis. Hunan University, China.

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

Clumeck, H (1977). *Studies in the Acquisition of Mandarin Phonology.* University of

California, Berkeley.

Clumeck, H. (1980). The acquisition of tone. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson (Eds.), *Child phonology: production* (pp. 257-275). New York: Academic Press.

Da, Jun. 2004. Chinese text computing. http://lingua.mtsu.edu/chinese-computing.

Davis, B. L., MacNeilage P.F., & Matyear, C. (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production. *Phonetica*, *59*, 75-107.

Davis, B. L., MacNeilage, P.F., & Matyear, C.M. (2005) Continuing studies in emergence of accuracy from pre-linguistic vocal patterns. Presented at Child Phonology Conference, June 16-17, 2005, Ft. Worth, TX.

Davis, B., & MacNeilage, P. (1995). The articulatory basis of babbling. *Journal of Speech and Hearing Research*, *38*(6), 1199-1211.

Davis, B.L., & MacNeilage, P.F. (1990). The acquisition of vowels: a case study. *Journal of Speech and Hearing Research, 33*, 16-27.

Duanmu, S. (2006). Chinese (Mandarin): phonology. In K. Brown (Ed.), *Encyclopedia of Language and Linguistics* (2nd ed.) (pp. 351-355). Oxford, UK: Elsevier Publishing House.

Edwards, J., Beckman, M., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in nonword repetition. *Journal of Speech, Language and Hearing Research, 47*, 421-436.

Ellis, N. (2002). Frequency effects in language processing : a review with implications of implicit and explicit theories of language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*(2), 419-439.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences, 4*(7), 258-267.

Gierut, J. A., & Morrisette M.L. (1998). Lexical properties in implementation of sound change. In A. Greenhill, M. Hughes, & H. Littlefield (Eds.), *Proceedings of the 22nd Annual Boston University Conference on Language Development* (pp. 257-268). Massachusetts: Cascadilla Press.

Gierut, J. A., Morrisette, M. L., & Champion A. H. (1999) Lexical constraints in phonological acquisition. *Journal of Child Language, 26*(2), 261-294.

Gordon, R.G., Jr. (Ed.). (2005). *Ethnologue: languages of the world*. (15th ed.). Dallas, TX.: SIL International.

Heisler, L. (2004) *The influence of neighborhood density and phonotactic probability on representational and production factors in word learning*. Doctoral Dissertation, Purdue University.

Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20-24-month olds. *Journal of Child Language, 35*, 903-916.

Hollich, G, Juscyk, P.W. and Luce, P.A. 2002. Lexical neighborhood effects in 17-

month-old word learning. In B. Skarabela, S. Fish, and A.H.-J. Do (Eds.), *Proceedings of the 26th Annual Boston University Conference on Language Development* (pp. 314-323). Somerville, MA: Cascadilla Press.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. New York: Cambridge University Press.

Hsu, J (1996). A study of the stages in the development and acquisition of Mandarin Chinese by children in Taiwan. Taipei: The Craine Publishing Co., Ltd.

Ingram, D. (2002). The measurement of whole-word productions. *Journal of Child Language, 29*(4), 713–733.

Jakobson, R. (1968). *Child language, aphasia and phonological universals*. (A. Keiler, Trans.). The Hague: Mouton. (Original work published 1941)

Lee, T., Lee, B., & Chen, Y.H. (2005). The pitch patterns of infants acquiring tonal languages: the perspective from Chinese. Paper presented at *ELA2005*, December 8-10, 2005, Lyon, France.

Lee, W.S., & Zee, E. (2003). Beijing Mandarin - an IPA illustration. *Journal of the International Phonetic Association, 33*(1), 109-112.

Lehiste, I. (1970) *Suprasegmentals.* Cambridge, MA: MIT Press.

Levitt, A.G., & Utman, J. G. (1992). From babbling towards the sound systems of English and French: a longitudinal two-case study. *Journal of Child Language*,*19*(1), 19-49.

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

Lieberman, P. (1967). *Intonation, perception, and language*. Cambridge, MA: MIT Press.

Lieberman, P. (1984). *The biology and evolution of language*. Cambridge: Harvard University Press.

Lindblom, B. (1992). Phonological units as adaptive emergents of lexical development. In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications.* (pp. 131–163). Timonium, MD: York Press.

Lowenstein, J. H., & Nittrouer, S (2008). Patterns of acquisition of native voice onset time in English-learning children. *Journal of the Acoustical Society of American*, *124*(2), 1180–1191.

Luce P. A., & Pisoni, D. B. (1998). Recognizing spoken words: the neighborhood activation model. *Ear & Hearing, 19*(1), 1-36.

Ma, W., Golinkoff, R. M., Houston, D., & Hirsh-Pasek, K. (2011). Word learning in infant- and adult-directed speech. *Language Learning and Development, 7*, 209-225.

Macken, M. A., & Barton, D. (1980). The acquisition of the voicing contrast in English: A study of voice onset time in word-initial stop consonants. *Journal of Child Language*, 7, 41-74.

Macken, M. A. (1992) Where is phonology? In C.A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: models, research, implications.* (pp. 367–397). Timonium, MD: York Press.

MacNeilage, P. F. (2011). Lashley's Serial order problem and the acquisition/evolution of speech. *Cognitive Critique, 3*, 49-84.

MacNeilage, P. F. (In press). Sound patterns of first words and how they became linked with concepts. In C. Lefebre (Ed.), *The evolution of language*. Cambridge: Cambridge University Press.

MacNeilage, P. F., & Davis, B. L. (2000). On the Origin of the Internal Structure of Word Forms. *Science, 288*, 528-531.

McCune, L., & Vihman, M. M. (2001). Early phonetic and lexical development: a productivity approach. *Journal of Speech, Language, and Hearing Research*, *44*(3), 670-684.

Menn, L. (1976). *Pattern, control and contrast in beginning speech: A case study in the development of word form and word function.* PhD dissertation, University of Illinois at Urbana-Champaign.

Metsala, J., & Walley, A. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). Mahwah, NJ: Lawrence Erlbaum.

Nishamura, T. (2003). Comparative morphology of the hyo-laryngeal complex in anthropoids: two steps in the evolution of the descent of the larynx. *Primates, 44*(1), 41-49.

Oller, D.K., & Eilers, R. E (1975). Phonetic expectation and transcription validity. *Phonetica, 31*, 288-304.

Oller, D.K., Wieman, L., Doyle, W., & Ross, C. (1976). Infant babbling and speech. *Journal of Child Language, 3*, 1-11.

Pierrehunmbert, J. B. (2003). Probabilistic phonology:discrimination and robustness. In

R.Hay Bod, J. & S. Jannedy (Eds.), *Probabilistic linguistics*. Cambridge, MA: MIT Press.

Prather, E., Hedrick, D., & Kern, C. (1975). Articulation development in children aged two to four years. *Journal of Speech and Hearing Disorders, 40*, 179-191.

Prince, A., & Smolensky, P. (1993) *Optimality theory: constraint interaction in generative grammar*. Blackwell.

Robb, M., & Bleile, K. (1994). Consonant inventories of young children from 8 to 25 months. *Clinical Linguistics and Phonetics, 8*, 295-320.

Simonyan, K., & Horwitz, B. (2011). Laryngeal motor cortex and control of speech in humans. *Neuroscientist, 17*(2), 197-208.

Snow, C. (1986). Conversations with children. In P. Fletcher & M. Garman (Eds.), *Language Acquisition* (2nd ed., pp. 365-375). New York: Cambridge University Press.

Snow, D. (1998). Children's imitations of intonation contours: Are rising tones more difficult than falling tones? *Journal of Speech, Language, and Hearing Research, 41*, 576-587.

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

Sosa, A. V. (2008). *Lexical effects in typical phonological acquisition.* Unpublished PhD Dissertation, University of Washinton, Seattle.

Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33*, 31-50.

Sosa, A. V., & Stoel-Gammon, C. (In press). Lexical and phonological effects in early word production. *Journal of Speech and Hearing Research*.

Stoel-Gammon, C., & Cooper, J. (1984). Patterns of early lexical and phonological development. *Journal of Child Language*, *11*(2), 247-271.

Stoel-Gammon, C. (2010). The Word Complexity Measure: description and application to developmental phonology and disorders. *Clinical Linguistics and Phonology, 24*(4-5), 271-282.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38*, 1-34.

Stokes, S. F., & Surrendran, D. (2005) Articulatory complexity, ambient frequency, and functional load as predictors of consonant development in children. *Journal of Speech, Language and Hearing Research, 48*(3), 577-591.

Storkel, H. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics, 25* (2), 201–221.

Sundberg, J. (1979). Maximum speed of pitch changes in singers and untrained subjects. *Journal of Phonetics, 7*, 71-79.

Thelen, E., & Smith, L. B. (1996). *A Dynamic systems approach to the development of cognition and action*. Cambridge: MIT Press.

Tsay, J. (2001). Modeling tone acquisition. Joint seminar talk given to Department of Chinese, Translation and Linguistic and Institute of Chinese, May 11, 2001, Hong Kong.

Tse, A.C.Y. (1993). The development of a phonological system in Cantonese: A case report. In Eve Clark (Ed.), *Proceedings of The 25th Annual Child Language Research Forum. Stanford*, CA: Center for the Study of Language and Information.

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

Tuaycharoen, P. (1977). The phonetic and phonological development of a Thai baby: From early communicative interaction to speech. Unpublished PhD dissertation, University of London.

Tyler, A., & Edwards, J. (1993). Lexical acquisition and the acquisition of initial voiceless stops. *Journal of Child Language, 20*, 253-273.

Velleman, S. L., & Vihman, M. M. (2002). Whole-word phonology and templates: Trap, bootstrap, or some of each? *Language, Speech, and Hearing Services in the Schools*, *33*(1), 9-23.

Velten, H. V. (1943). The growth of phonemic and lexical patterns in infant language. *Language, 19*, 281-292.

Vihman, M. M. (1996). *Phonological development: The origins of language in the child*. Cambridge, MA: Blackwell.

Vihman, M., & Velleman, S. (1989). Phonological reorganization: a case study. *Language, Speech and Hearing Services in Schools, 33*, 9-23.

Vihman, M.M. (1993). Variable paths to early word production. *Journal of Phonetics*, *21*(1), 61-82.

Vitevitch, M.S., & Sommers, M. (2003). The facilitative influence of phonological similarity and neighborhood frequency in speech production. . *Memory & Cognition, 31*, 491-504.

Walley, A.C. (1993). The role of vocabulary development in children's spoken word recognition and segmentation ability. *Developmental Review, 13*, 286-350.

Wind, J. (1970). *On the phylogeny and the ontogeny of the human larynx: A morphological and functional study*. Groningen: Wolters-Noordhoff.

Wong, P.-S. (2008). *Development of lexical tone production in disyllabic words by 2- to 6-year-old Mandarin-speaking children.* Unpublished PhD Dissertation, City University of New York, New York City.

Wong, P.-S. (2012). Acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions. *Journal of Phonetics, 40*(1), 141-151.

Wong, P.-S., Schwartz, R. G., & Jenkins, J. J. (2005). Perception and production of lexical tones by 3-year-old Mandarin-speaking children. *Journal of Speech, Language and Hearing Research, 48*(5), 1065-1079.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics, 25*, 61-83.

Xu, Y., & Sun, X. (2002). Maximum speech of pitch change and how it may relate to speech. *Journal of the Acoustical Society of America, 111*(3), 1399-1413.

Yang, J. (2006) *The Acquisition of Tone in Mandarin-speaking Children*. MA thesis, Hunan University, China.

Yang, J., & Davis, B. L. (2008). Lexical/segmental influence on tone accuracy in

Mandarin-speaking Children: A Longitudinal Case Study. In *Abstract volume of the 16th Biennial International Conference on Infant Studies.* P102, March 27-30, 2008, Vancouver, Canada.

Yang, J., & Davis, B. L. (2010). *Contextual influences on the excursion size of fundamental frequency in infant babbling*. Paper presented at the 2nd PanAmerican/Iberian Meeting on Acoustics.

Yang, J., & Lee, T. H-T. (2007). *Lexical Selectivity and Avoidance: Evidence from Tonal Acquisition in Mandarin Children*. Paper presented at the Texas Research Symposium on Language Diversity.

Yang, J., & Lee, T.H.-T. (2006). Lexical variation and rime-tone correlation in early tonal acquisition: a longitudinal study of Mandarin Chinese. In A. Belotel-Grenié, & M.l Grenié (Eds.) *Proceedings of Second International Symposium on Tonal Aspects of Languages* (pp.133-138), April 27-29, 2006, La Rochelle, France.

Yang, J., Diehl, R., & Davis, B. L. (2008). The development of tonal duration in Mandarin-speaking children. [abstract]. *Journal of the Acoustical Society of America, 124*(4), 2498.

Yang, J., Ortega-Llebaria, M., & Davis, B. L. (2010). *A comparative study of pitch patterns in the babbling of English and Chinese infants*. Paper presented at the 2010 International Conference of Child Phonology.

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

Zamuner, T. (2009). The structure and nature of phonological neighborhoods in children's early lexicons. *Journal of Child Language, 36*, 3-21.

Zhu, H. (2002) P*honological Development in Specific Context: Studies of Chinese-speaking Children*. Clevedon: Multilingual Matters Ltd.

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (Modern Standard Chinese). *Journal of Child Language* , 27, 3-42.