THE PRODUCTIVITY OF TONE SANDHI IN MANDARIN-SPEAKING CHILDREN AND ITS DEVELOPMENTAL FEATURES

HUANG XUNAN

PhD

The Hong Kong Polytechnic University

2021

The Hong Kong Polytechnic University

Department of Chinese and Bilingual Studies

The Productivity of Tone Sandhi in Mandarin-speaking Children and Its Developmental Features

Huang Xunan

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

08.2019

CERTIFICATE OF ORIGINALITY

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it reproduces no material previously published or written, nor material that has been accepted for the award of any other degree or diploma, except where due acknowledgment has been made in the text.

______________________(Signed)

_____Huang Xunan_____(Name of student)

Abstract

Children's productive knowledge of morphological alternations has received substantial attention in the literature. However, little is known about how and when the tonal alternations, such as tone sandhis in Chinese dialects, become productive. The productivity of tone sandhi is thought to be influenced by two operational mechanisms, i.e., lexical and computation mechanism. It has been well established that Mandarin Tone 3 sandhi in adults is mediated via a computation mechanism. However, the development of Tone 3 sandhi in children is still unclear. Besides, these two mechanisms relate to the issues of mental representation and decomposition of Mandarin Tone 3 sandhi. There are different theoretical accounts for how tone sandhi is represented and implemented in adult speakers, including the underlying representation view, the surface representation view, the multi-variant account, or the underspecification view. It is still unknown how they are represented in the mental lexicon and decomposed in speech production in children. This thesis examines the productivity of Mandarin Tone 3 sandhi in Mandarin-speaking children, and those developmental features, which have received very little attention so far. It aims to explore how such productive sandhi is applied in pseudo- and novel words in children's speech production from both perceptual and acoustic analysis, how they are processed online or internalized in the mental lexicon in children, and how they are decomposed

from the speech productions from childhood to adulthood. Explorations on these research questions are expected to further provide new knowledge into language acquisition.

To this end, I first explored the productivity of the sandhi pattern in speech production by examining children's performances in real words, pseudo-words, and novel words by observing the phoneticians' speech production judgment and the acoustic realizations. I found that children's application of Mandarin Tone 3 sandhi undergoes a developmental trajectory. Children showed relatively low productivity at three- to five-years-old, especially in novel words. This indicates the involvement of the lexical mechanism at three-to-five-years old. Seven years old children reached the adult-like application in those novel words, which indicates the full development of the computation mechanism. Second, I conducted an auditory-auditory priming lexical decision task to determine children's mental representation of tone sandhi. I found that four-and five-year-old children represented sandhi items in their surface forms, while the children from six to eight years old have both surface and underlying forms represented in their mental lexicon. Compared with the result of adults that only exhibited a significant underlying priming effect when deciding Tone 3 sandhi real words, children have not yet reached an adult-like underlying representation even between seven and eight years old. The relatively delayed underlying representation in children compared with the speech production task may be explained by the children's literacy effect and the frequency effect of the tone sandhi words. These results are also of great value because they illustrate how children process real sandhi words online at different ages, which informs the development of theoretical models for language acquisition from a different perspective. Third, I explored the tone awareness of the sandhi items with the explicit syllable deletion task and the implicit tone matching that were widely used for examining phonological awareness. I found that children also undergo a similar developmental trajectory in line with their productivity features in the way of reaching the adult-like underlying Tone 3 recognition. Children reached the adult level of explicit processing at seven years old. This experiment enhances our understanding of how the phonological alternations of Mandarin Tone 3 sandhi are recognized by Mandarin-speaking adults and children.

Taken together, this thesis offers a comprehensive understanding of how Mandarin-speaking children learn to apply a sandhi pattern and represent the phonological alternation pattern in the mental lexicon, and how to internalize and decompose the pattern from the surface form during the language development. The overall results reflect children's general developmental timeline of language processing from the lexical mechanism to the computation mechanism. Previous studies primarily focus on children's operational mechanisms in language components that involve phonological or morphological alternations with regular and irregular forms. The tone sandhis in Chinese dialects offers a new perspective to observe children's development of operational mechanisms and language processing.

Acknowledgments

I cannot believe that the clock has ticked down to this stage: writing the acknowledgments. Three years may seem long, but they feel so short. I have enjoyed these three years of Ph.D. life spent devoting all my attention to such an exciting area. I am grateful to my supervisors, the department, my friends, and my family, who have all made my Ph.D. life so enjoyable. This thesis would not have been possible without their valuable support throughout my Ph.D. studies.

First and most importantly, I would like to give my sincerest gratitude to my chief supervisor Dr. Caicai Zhang. She is so intelligent and efficient. I am grateful for her constructive comments on my research and for her generosity in sharing her inspiring ideas and experience in academia. It has been a great honor to have her as my supervisor. My deepest gratitude also goes to my co-supervisor, Professor William Shi-Yuan Wang. He is full of wisdom. He has always created stimulating viewpoints for me to think about my project and has always encouraged me to broaden my horizons and pursue a higher level of scientific achievement. His encouragement has been a great force in keeping up my progress.

I want to thank Professor Jie Zhang, who has offered me valuable comments on my project. His insightful views on phonological processing have opened a new window for me to think about the phonological implications of my study. I would like to thank my external reviewers, Professor Hua Shu and Professor Yiya Chen, for their very constructive comments. I am also indebted to my master's supervisor, Professor Feng Shi, who first ignited my interests in this area. He is a role model of mine. Moreover, I am indebted to my team members from our group meeting, who offered me many insightful comments and suggestions. I am thankful for the companionship of my friends. Without those friends, Jing Shao, Haitao Guan, Wenchun Yang, and Cheungwai Zhang, Shuang Huang, and the classmates in my office, my life over the past three years would not have been so enjoyable. Besides, I want to thank Shang who shared the happiest time of the past three years with me.

Finally, I extend my heartfelt thanks to my parents Haiyan Li and Junshan Huang who have offered me their continued support and love. The research for this project included around 300 children in total. It was my parents who helped me to arrange data collection rooms and who supported me in accomplishing the whole data collection process. I will never forget the days when my parents helped me take care of the children in the waiting room who were going to take part in the experiments and then drove me home after a long day's work. Their support and selfless love have encouraged me to focus all my attention on my research.

Contents

Chapter 1 Introduction

Productivity is a fundamental human cognitive ability in language learning. During the process of first language acquisition, individuals do not merely memorize words or sentences, but often try to generalize over linguistic patterns statistically and apply them productively to new items (Jenny & Ann, 2001; Pinker, 1999; Schuler et al., 2005). For example, during the process of learning the English past-tense inflections, English-speaking children demonstrated a generalization of '-ed' productively to both regular and irregular verbs at around three years old (e.g., *Judy goed to the room*; see Pinker, 1999). However, it turns out that not all linguistic patterns are equally productive.

Yang (2016) explains productivity with a learning model called the Tolerance Principle (e.g., Yang, 2016 & 2018; Schuler et al., 2016), which states that a pattern is productively computed only when it is computationally efficient enough to do so, i.e., the percentage of regular forms was greater than that of irregular forms by a certain degree. In this model, he quantifies the precise proportion of exceptions (e.g., irregular-form items: go-went) that a productive pattern (e.g., a regular past tense in English, -ed) can tolerate. The calculated formula has been verified by a variety of acquisition literature, corpus data, and reports of children's acquisition of novel words (Yang, 2016; Schuler et al., 2016).

However, language processing is far more complicated than an analysis of productivity based on the efficiency of pattern generation. Some phonological patterns show different degrees of productivity in their application to novel words even though they are almost regular. The tone sandhi patterns in Chinese dialects are an example of this phenomenon. Tone sandhi refers to tonal alternations in speech production conditioned by linguistic context (Chen, 2000; Yip, 2002). Tone sandhis are manifested in different ways among Chinese dialects. For example, in Mandarin Chinese, when a Tone 3 syllable is followed by another tone 3 syllable, the first Tone 3 becomes tone 2 (213+213→35+213; Chao, 1948; Lee, 2002) or a tone 2-like syllable (Wang & Li, 1967; Myers & Tsay, 2003; Yuan & Chen, 2014; Zhang et al., 2015). For instance, the Tone 3+Tone 3 word /mei213ny213/ 美女 "beauty" is ultimately pronounced "mei35ny213". In Shanghainese, the tone of the first syllable is spread across the disyllabic word, which neutralizes the tone of the second syllable (Xu, Tang, & Qian, 1981; Zee, and Maddieson 1980; Zhu 1999; Zhang, 2018). For example, the surface tones of the compounds /sã55foŋ31/ 伤风 "to catch a cold" are derived by spreading the underlying tones of the initial syllables "sã53" over the disyllables. Whereas the tone sandhi of Wuxi Wu is formed by pattern substitution, that the tone of the first syllable is firstly substituted by another tone, and the substituted tone is then spread across the disyllabic word (e.g., Chan, & Ren 1989; Yan and Zhang, 2016; Zhang, 2018). In Southern Min, every tone undergoes sandhi, unless it occurs in a final word position (Hsieh, 1976; Hsieh, 1970; Zhang, & Turnbull-Sailor, 2006). For example, the word /pai24kai51/ 排解 "to eliminate" is pronounced as /pai33kai51/ in speech production. These tone sandhi patterns are all regular in the input. While some sandhi patterns, like Mandarin Tone 3 sandhi and Shanghainese tone sandhi, demonstrate high productivity in novel words (Zhang & Peng, 2013; Zhang, Xia, & Peng, 2015; Zhang & Lai, 2010; Zhang & Meng, 2016). Other sandhi patterns, especially those that involve a series of sandhi phenomena in a circular chain shift, like in Southern Min and Wuxi Wu, often fail to generalize to novel words (Hsieh, 1970; Yan & Zhang, 2016; Zhang et al., 2011). It will enrich our knowledge of language acquisition and phonology to understand how such productive and unproductive types of sandhi are applied in speech production, how they are processed online or internalized in the mental lexicon, and how they are decomposed from the speech productions from childhood to adulthood.

The productivity of sandhi patterns in adults has been suggested to be influenced by two types of operational mechanisms, that is, rote-memory-based lexical mechanisms (hereafter, **lexical mechanisms**) and computation-based productive mechanisms (hereafter, **computation mechanisms**) (e.g., Zhang et al., 2015). If an item is processed by the lexical mechanism, it indicates that the item is memorized and retrieved from the mental lexicon in its surface lexical representation; whereas, if it is processed by the computation mechanism, it means that the item is processed by generalizing a pattern productively on the underlying representations. The two mechanisms were derived from the "word and rule" principle (Pinker, 1998), which was later developed as the "dual-route model" (e.g., Pinker, 1998; Marcus, Brinkmann, Clahsen, Wiese, & Pinker, 1995; Ullman & Pinker, 2002; Ullman, 1999, 2001). Targeting the English inflections that include both regular and irregular forms, research has found that the two forms are distinct categories, one of which is relatively rule-like and the other one less so. This has resulted in a processing model whereby they are also assumed to be processed differently. The regular items are processed by the computation mechanism, by which patterns are applied to items of abstract categories (e.g., Verb, Noun), whereas the morphologically complex, irregular items are processed by rote memory. Different sandhi patterns in Chinese dialects, even though they are regular patterns, have also been found to be processed differently in adult speakers. For some complex sandhi items (e.g., the sandhi circle in Southern Min), the patterns were listed in the mental lexicon by the lexical mechanism (e.g., Hsieh, 1970; Zhang et al., 2011). However, the Mandarin Tone 3 sandhi pattern, for example, was processed productively by a computation mechanism (Zhang et al., 2015; Zhang & Lai, 2010; Zhang & Meng, 2016). Detailed information will be given in the literature review.

Another interesting question is how the tone sandhi items are represented or internalized in the mental lexicon. There are three well-attested theoretical accounts for how tone sandhi is represented in adult speakers, which includes the underlying representation view, the surface representation view, the multi-variant account. Based on the empirical studies of tone sandhis in Chinese dialects with high and low productivity, Zhang (2018) speculated that adult speakers tend to process the tone sandhi items with high and low productivity differently in the mental lexicon. When items participating in a sandhi pattern (henceforth, **sandhi items**) show less productivity in novel-words, they may be processed from the surface representation, which indicates that the form is represented just as it appears in spoken language. For example, one Southern Min sandhi pair with low productivity in sandhi application (in which the tone 51 appears as 55 in non-word-final position) was found to be represented in the mental lexicon in its surface form (Chien et al., 2017; Zhang, 2018). If a sandhi pattern shows high productivity in non-words, it is more likely to be processed from the underlying representation, suggesting that the sandhi word is represented as the concatenation of its underlying (canonical) forms of morphemes along with the governing pattern. Some tone pairs participating in productive sandhi applications, such as the Mandarin Tone 3 sandhi, have been determined to be represented in their underlying forms in adult Mandarin speakers (Chien et al., 2016). For example, the speech sound of the sandhi pair /fu35tau315/ 辅导 "to counsel" is likely to be represented in "fu213" plus "tau213" in the mental lexicon. Given the above introduction in spoken sandhi word comprehension, only Tone 3 is represented in the mental lexicon in the lexical decision of Tone 3 sandhi targets. However, the phonological representations of T2-like T3 sandhi variant and Tone 3 are both activated for the production of Tone 3 sandhi words (Chen et al., 2011; Nixon et al., 2015). The multi-variant account (activation of both Tone 2-like T3 sandhi variant and Tone 3) was also verified in the mental representation of the Tone 3 monosyllabic words (Li & Chen, 2015). According to the psycholinguistic models of Levelt et al. (Indefrey, 2011; Levelt, 1999; Levelt et al., 1999), lexical and phonological processing occur before the initiation of speech production. Therefore, the overall picture for the representation of Mandarin Tone 3 sandhi in adult speakers should be: the underlying Tone 3 may be activated when hearing the Tone 3 sandhi word, while both Tone 2 and the abstract underlying tones are activated when producing the sandhi word.

This project focuses on one productive sandhi pattern, i.e., Mandarin Tone 3 sandhi, which has been verified to be processed via the computational mechanism in adult Mandarin speakers (Zhang et al., 2015; Zhang & Peng, 2013). Importantly, however, little is known about when the pattern becomes productive and how it is processed by Mandarin-speaking children of different ages. These studies incorporate tonal alternations into speech production, speech representation, and speech decomposition models. The thesis contributed to a better understanding of the acquisition of tone sandhi. More importantly, this study has the potential to inform the development of theoretical models for language acquisition, speech production, speech representation, phonological awareness, phonology, and morpho-phonology.

First, I explored the productivity of the sandhi pattern in speech production during language acquisition in the project. I found that children's application of Mandarin Tone 3 sandhi undergoes a developmental trajectory. Children showed relatively low productivity at three- to five-years-old, especially in novel words. This indicates the involvement of the lexical mechanism at three-to-five-years old. Seven years old children reached the adult-like application in those novel words, which indicates the full development of the computation mechanism. Understanding the developmental features of the sandhi pattern is crucial because it sheds direct light on the acquisition of morphological and phonological features. Second, I conducted an auditory-auditory priming lexical decision task on tone sandhi items to determine children's mental representation of tone sandhi. The initial Tone 3 syllable of the sandhi words have two allophones, one of which is the underlying Tone 3, and the other is the Tone 2-like form (surface form in the sandhi context). It has been observed that adult Mandarin speakers tend to represent sandhi words in their underlying forms (Chien et al., 2016; Zhang, 2016, 2018). However, no study has provided empirical data to investigate children's developmental trajectory before reaching adult representation. I found that four-and five-year-old children represented sandhi items in their surface forms, while the children from six to eight years old have both surface and underlying forms represented in their mental lexicon. Children also tended to rely more and more on underlying representations as their age increases. The priming lexical decision tasks primarily examined individuals' immediate processing of the linguistic phenomenon. These results are also of great value because they illustrate how children process real sandhi words online at different ages, which informs the development of theoretical models for language acquisition from a different perspective. Third, I explored the tone awareness of the sandhi items, particularly the allophones of the first syllable of the sandhi items. Tone awareness is one of the crucial subsets of phonological awareness (i.e., the ability to detect and manipulate speech sounds in a language; Anthony & Francis, 2005; Chen et al., 2016; Shu et al., 2008). The first syllable of the sandhi items is underlyingly represented as Tone 3 (213) in adult Mandarin speakers, while it surfaces with an identical contour to Tone 2 (35) in the flow of speech in tone sandhi environment (Chien et al., 2016). With the picture-matching task and the traditional paradigm for explicit phonological awareness (see Shu et al., 2008), I examined children's preference for linking the sandhi tones to surface or underlying tones. This research shows that children also undergo a similar developmental trajectory in line with their productivity features in the way of reaching the adult-like underlying Tone 3 identification. Mandarin Tone 3 sandhi is a type of phonological alternation, which is central to phonological and morphological theory, but little is known about how it is recognized or internalized by speakers. This project enhances our understanding of how the phonological alternations of Mandarin Tone 3 sandhi are recognized by Mandarin-speaking adults and children.

The project is presented in eight chapters. Chapter 2 provides background information for this study by reviewing previous investigations of Mandarin tone sandhi, its operational mechanisms, and its representation in the mental lexicon of adults and children. As a theoretical background, the chapter also reviews previous studies of Mandarin tone systems, as well as developmental features in Mandarin-speaking children which contribute to the acquisition of sandhi patterns. Chapter 3 reports an analysis of the productivity of Mandarin Tone 3 sandhi in children's speech production. Chapter 4 presents the acoustic features of the realization of Mandarin Tone 3 sandhi in Mandarin-speaking children. Chapter 5 reports the representation of Mandarin Tone 3 sandhi in children's mental lexicons via the lexical decision task. Chapter 6 examines the tone awareness of the initial allophones of sandhi items. Chapter 7 is a general discussion, where all studies reported in this project are discussed at a higher level within a broader theoretical context to inform the productivity theories and it's driving force for the children to reach the adult-like productivity. Chapter 8 presents a short conclusion regarding the findings, limitations, and an outlook for future research.

Chapter 2 Literature Review

This chapter provides a comprehensive review of studies on the productivity of tone sandhi. The review is organized into five parts: (1) Mandarin tones, tonal phonology, and Tone 3 sandhi; (2) the productivity of tone sandhi and its operational mechanisms; (3) theories regarding the representation of tone sandhi items; (4) developmental features in Mandarin-speaking children which contribute to the acquisition of the sandhi pattern; and (5) the productivity of tone sandhi in novel words and the mental representation of tone sandhi in Mandarin-speaking children.

2.1. Mandarin tones, tonal phonology and Tone 3 sandhi in Mandarin Chinese

Tone is employed in most parts of the world (Wang, 1967). It is reported to exist in 60–70% of languages (Yip, 2002), including many Asian, African and indigenous American languages, as well as some European and South Pacific languages (Wang, 1967). Tones, however, serve different functions in different languages. In the Sino-Tibetan family of languages in Asia, which contains Mandarin Chinese, tone serves a lexical function. This section focuses on the tone system in Mandarin Chinese. Tones in Mandarin Chinese are in a contour-based system with two contour tones in the tonal inventory. Tones are typically associated with syllables and serve to distinguish lexical meaning (Wan & Jaeger, 1998; Wang, 1967; Zhang, 2014). There are four tones which distinguish semantic meanings in Mandarin Chinese. Tone 1 has a high-level pitch, Tone 2 a rising pitch, Tone 3 a low-dipping pitch, and Tone 4 a high-falling pitch (Chao, 1948; Brito, 1980; Wang, 1967). For example, the syllable /ma/ refers to "mother" when it is pronounced with Tone 1, means "hemp" with Tone 2, and means "horse" and "scold" respectively with Tones 3 and 4 in Mandarin Chinese. F0 contours for each of the four Mandarin Chinese tones can be seen in Figure 2.1.

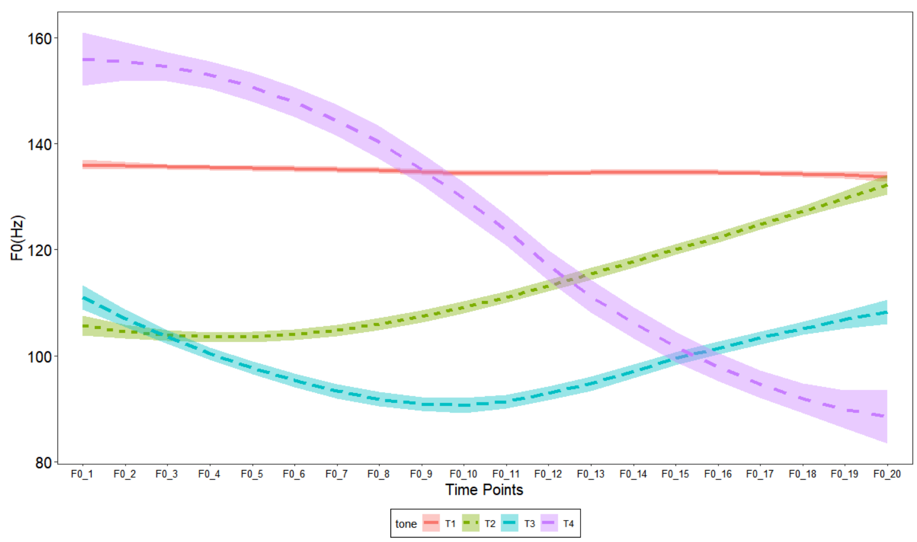


Figure 2.1 *F0 contours (mean and ± standard error) for each of the four Mandarin Chinese tones spoken in isolation by a male Mandarin speaker*

The four lexical tones are also widely labeled using a five-level tone notation system, which uses numbers to indicate pitch height and trajectory (Chao, 1948, 1968;

Lee, 2007). In five-level tone marking, a speaker's lexical pitch level range is divided into five levels, of which 1 is used to describe the lowest pitch level, and 5 the highest (Lin, 2007). For example, Tone 1 is labeled as '55', which indicates a pitch trajectory that starts at the highest pitch level and remains there. Tone 3 is '213', which indicates a relatively low starting pitch, followed by a drop to the lowest pitch and then an increase to a higher pitch than the initial pitch. The Mandarin tone inventory is summarized in table 2.1.

Table 2.1 *a summary of the four lexical tones in Mandarin Chinese*

Tone category	Descriptive naming	Five-level tone marking	Examples
Tone 1	high-level tone	55	/ma55/ 妈 "mother"
Tone 2	rising tone	35	/ma35/ 麻 "hemp"
Tone 3	low-dipping tone	213	/ma213/ 马 "horse"
Tone 4	falling tone	51	/ma51/ 骂 "scold"

When produced in isolated contexts, the four lexical tones are quite stable and seem well-defined (Chao, 1968). When produced in connected speech, however, the tonal contours undergo certain variations depending on preceding and following tones

(Chao, 1968; Kratochvil, 1984; Shen, 1992; Xu, 1997). This is particularly true of Tone 3, which is the focus of this project. It undergoes different types of tonal variations when it is produced in actual speech. Specifically, it becomes a rising tone (Tone 2-like) when preceding another Tone 3 (213+213→35+213), and a low falling tone when preceding any other tones (213+55/35/51→21+55/35/51) (e.g., Duanmu, 2007; Zhang et al., 2015). For instance, the combination of /mei213/ and /ny213/ is pronounced /mei35ny213/ 美女 "beauty"; while the combination of /mei213/ and /sə51/ would be /mei21 sə51/ 美色 "good looks". The acoustic analysis of sandhi Tone 3 has shown that it has an overall lower F0 height (Peng, 2000), a lower F0 peak (Peng, 2000; Yuan & Chen, 2014), as well as a later F0 rising onset (Yuan & Chen, 2014) when compared to Tone 2. However, the perception results have demonstrated that native speakers are usually not able to consciously distinguish the sandhi Tone 3 from Tone 2 (Peng, 2000; Wang, & Li, 1967).

The origin of Mandarin Tone 3 sandhi can be traced back to at least the 16th century based on philological evidence (Mei, 1977). Comparing lexical tones between Mandarin and Middle Korean at that time according to 16th-century Chinese language textbooks for Korean speakers, Mei reconstructed the tonal values of the four tones in the 16th century Mandarin. And he detected that Tone 3 sandhi was in place in the 16th century and that it was phonetically similar to the Tone 3 sandhi of modern Mandarin.

The force driving Mandarin Tone 3 sandhi is not fully understood. Previous works have described it as a dissimilating process, by which the two low-dipping tones are changed into a sequence of a high tone followed by a low dipping tone (Mei, 1977). However, this view cannot explain why the initial syllable changed to a rising tone rather than another high tone, such as the high-level tone. Meanwhile, the sandhi pattern has not yet been found in other languages with similar low-dipping tone combinations, such as in Thai, or in Cantonese where Tone 4 /21/ is quite similar to half Tone 3. By reviewing the productivity and the mechanisms of operation of tone sandhi, and by observing its developmental features, we will have a better understanding of its nature.

2.2. The productivity of tone sandhi and its operational mechanisms

Extensive work has been done on the productivity of Mandarin Tone 3 sandhi, especially with the widely known "wug test." In this test, "wug" represents a non-word. Gleason (1958) designed the wug-like non-words to test the productivity of English morphological patterns. Later the paradigm has been widely used to examine the productivity of morphological patterns (e.g., Selby, 1972; Albright & Hayes, 2003; Joan & Bybee, 1981; Pierrehumbert, 2006). Hsieh (Hsieh, 1970, 1976, 1975) was the first to bring the paradigm into the morpho-phonological area, by examing the productivity of tone sandhi in Southern Min. Southern Min sandhi occurs when a tonal syllable is in a non-word-final position, in a circular chain-shift where the Tone 51 becomes 55, and the Tone 55 becomes 33…, and the Tone 21 becomes 51 (see Figure 2.2). It has been characterized as being phonologically "opaque" (Kiparsky, 1973; Zhang et al., 2009), because the output of one sandhi pattern is also the input of a second sandhi pattern (thus being a non-surface-true phonotactic generalization; Zhang, 2016). Hsieh found that the overall productivity of sandhi patterns in Southern Min is low in pseudo-words. Moreover, different individuals showed high variation among different lexical items and tone categories in the application of tone sandhi to those pseudo-words. Such low productivity and high variability led Hsieh to conclude that sandhi forms are primarily stored in memory as surface-form allomorphs of a word, which can be explained by appealing to lexical mechanisms, rather than being computed productively.

51 → 55 → 33 ← 24

↖ ↙

21

Figure2.2 *tone sandhi circle in Taiwanese Southern Min*

Researchers have repeatedly examined the opaque Southern Min tone circle with wug tests from different perspectives (Wang, 1993; Zhang et al., 2006, 2009, 2011, 2018; Zhang & Lai, 2008). For example, to test the productivity of the different tone sandhi pairs in Southern Min, Zhang and his collaborators (Zhang, 2018) used the elicited the production of two types of disyllabic words from native speakers: (i) real words (actually occurring words that all participants were familiar with); and (ii) novel words (which were composed of an accidental gap (AG) syllable and an existing noun). The AG syllable refers to a type of novel syllable whose combination between consonant and vowel occurs in the target language, but which yields no real morpheme. Results have confirmed the unproductivity of the pattern in general but have observed a different degree of productivity of different sandhi pairs. The 24→33 sandhi has been found to have a lower non-application rate in novel words compared with other tone pairs. Besides, the results showed variations associated with individual speakers, lexical items, and tone categories in the generalization of tone sandhi to novel words in Southern Min.

However, previous studies have consistently shown that the application of Mandarin Tone 3 sandhi is highly productive in novel disyllabic sequences without exception (Xu, 1997; Zhang and Lai, 2006; Zhang & Peng, 2013; Zhang et al., 2015; Zhang, 2016; Zhang & Lai, 2010). Different types of non-words were investigated in the literature, including pseudo-words and novel words. A pseudo-word is defined as a non-occurring sequence of two actually occurring morphemes (e.g., /tʂʰʅ213 sa213/ 尺洒 "ruler sprinkle"), or the actually occurring morpheme followed by a syllable which never occurs in Mandarin phonology (e.g., /tʂʰuɑŋ213 tsəŋ213/ 闯 zeng213 "rush zeng213"). Even though the pseudo-word seldom occurs in daily life, there still exists the sandhi allomorphs of the initial Tone 3 syllables (e.g.,[tʂʰ ʅ35] 尺 in /tʂʰ ʅ213 pan213/ 尺板 "ruler") for the participants to take advantage of. Therefore, the output of pseudo-words can lean upon either the lexical mechanism (by retrieving previously heard sandhi allomorphs) or the computation mechanism. The novel word refers to the combination of novel syllables, each of which yields no meaning (e.g., /pʰiŋ213 tsəŋ213/), or a novel syllable followed by an actually occurring syllable in Mandarin phonology (e.g., /pʰiŋ213 ma213/ ping213 马 "ping213 horse"). Since no lexical entry of the initial syllable can be retrieved from the mental lexicon, participants could only rely on the computation mechanism in this type of non-word.

Firstly, when examining acoustic variations of tones in Mandarin in different tonal contexts using the sequence /mama/ with all 16 possible two-tone combinations of the four Mandarin tones, Xu (1997) found that Mandarin Tone 3 sandhi applies to the pseudo disyllabic word (/ma213 ma213/ 马马 "horse horse").

In addition to this, Zhang and Lai (2010) systematically exploited the productivity of Mandarin Tone 3 sandhi in adult speakers in real words, pseudo-words, and novel words using acoustic analysis. The results indicated that the sandhi pattern could be applied to all three types of words without exception, which provided crucial evidence for its productivity. However, the results also showed that the contour shapes of the sandhi form of Tone 3 in novel words were significantly different from those in real and pseudo-words, exhibiting a lower and later turning point and longer duration.

Zhang and Peng (2013) scrutinized the productivity of Mandarin Tone 3 sandhi with both acoustic analysis and perceptual experiments. They compared the realization of sandhi Tone 3 with Tone 2 in minimally contrastive contexts, both in real words (such as /əɹ213 y213/ 耳语 "whisper" & /əɹ35 y213/ 儿语 "baby talk") and novel words (such as /fia213 sua213/ & /fia35 sua213/). The overall results suggested that Mandarin Tone 3 sandhi applies without exception in the production of real words and novel words. However, their results of the acoustic analysis showed that the production of the sandhi Tone 3 items exhibited significantly lower F0 than Tone 2 in novel words. It is not clear whether this phonetic difference has phonological implications, for the reason that the Mandarin-speaking participants could not reliably distinguish between Tone 2 and sandhi Tone 3 in real or novel words in the perception task. Altogether, these findings provide compelling evidence for the productivity of Tone 3 sandhi in adult speakers and may indicate the presence of a phonological computation mechanism.

Using the same paradigm, Zhang et al. (2015) adopted an ERP study to observe the neural activities of the on-line encoding of the productive Mandarin Tone 3 sandhi pattern in speech production. They found a significant main effect of sandhi encoding in the P2 time-window, where Tone 3 + Tone 3 sequences elicited greater P2 amplitude than did Tone 2 + Tone 3 sequences, irrespective of whether they were real or novel words. As P2 is sensitive to phonological processing (Crowley & Colrain, 2004; Landi, Crowley, Wu, Bailey & Mayes, 2012), this suggests that phonological processing of the second Tone 3 syllable was more effortful when it was preceded by Tone 3 than when it was preceded by Tone 2. The result likely indicates that the encoding of Mandarin Tone 3 sandhi involves greater phonological processing than the encoding of Tone 2.

To summarize, different types of phonological alternation may exhibit different productivity behavior when tested with pseudo- or novel words in the language. The studies of Southern Min tone sandhi pairs have been repeatedly shown to be less productive, to different degrees. Moreover, there is a large amount of variation associated with tone categories, individual differences, and lexical items in the generalization of tone sandhi to pseudo-words in Southern Min. For Mandarin Tone 3 sandhi, by contract, the results from the Mandarin Tone 3 sandhi confirmed that the sandhi environment (Tone 3 + Tone 3) alone triggers the sandhi, no matter how unfamiliar the carrying syllables are. These may indicate that sandhis with high and low productivity are internalized and governed in different ways by speakers.

As for the operational mechanism that governs the productivity of the tone sandhi items, both lexical mechanisms and computation mechanisms could play a role, as mentioned in the previous chapter (also see Zhang et al., 2015). If the items are processed by lexical mechanisms, the surface allomorphs were likely memorized as a frozen chunk; the pattern thus only applies to real words. In contrast, if the items are operated via the computation mechanism, it indicates that they are produced via applying the phonological pattern productively to the underlying representation of the syllables; therefore, the pattern can be applied productively to both real and pseudo-words. Speakers internalize the productive and unproductive sandhi patterns in different ways. The unproductive tone sandhi patterns, for instance, in Southern Min (Hsieh, 1970; 1976; 1975; Wang, 1993; Zhang et al., 2006, 2009, 2011; Zhang & Lai, 2008), operate via the lexical mechanism. The productive sandhi patterns, by contrast, such as the Mandarin Tone 3 sandhi (Zhang & Peng, 2013; Xu, 1997; Zhang et al., 2015; Zhang & Lai, 2010), are internalized through the computation mechanism.

2.3. Theories regarding the representation of tone sandhi items

Speech variations are ubiquitous in different languages. Diverse models have been proposed to deal with the representation of speech variations, among which three have become more representative: the episodic storage account, the abstract representation account, and the hybrid model. The episodic theory suggested that perceptual details are stored in memory as the basic substrate (see Goldinger, 1998, for a review). It has been supported by different studies (Goldinger, 1996; Connine et al., 2008; Pitt, 2009). For example, Pitt (2009) showed that variant forms of newly learned words are recognized only when the form has been exposed before. However, some other literature supports the abstract representation account which assumes that a word in the mental lexicon is associated with a highly abstract phonological representation (e.g., Norris & McQueen, 2008; Pierrehumbert, 2016). According to the account, non-distinctive phonological features are unspecified or represented in an underlying form. Between the two extreme spectrums, it is increasingly acknowledged that the mental lexicon contains a hybrid which entails both episodic and abstract representations (Mitterer et al., 2011; Sumner et al., 2014; Pierrehumbert, 2016). When limiting the topic to the representation features tone sandhi items, some related accounts have also been put forward. There are four generally recognized accounts of how tone sandhi words are represented in the mental lexicon: the surface representation view, the underlying representation view (also called canonical representation view), the multi-variant account, and the underspecification account (Chen et al., 2011; Li & Chen, 2015; Chien et al., 2016, 2017; Nixon et al., 2015; Wewalaarachchi & Singh, 2016; Zhang, 2018; Politzer-Ahles et al., 2016).

The surface and underlying representation views were primarily tested from the perspective of spoken word recognition, which concerned how tonal alternations are represented in the mental lexicon in spoken language comprehension. According to the surface representation view, words are stored and represented just as they appear in spoken language, i.e., in their surface form (the post-sandhi form). For example, the word /tsuo213 tɕiɑu213/ 左脚 "left foot" is represented in the mental lexicon in the form of /tsuo35 tɕiɑu213/, even though the underlying tone of the initial phoneme is a Tone 3 syllable /tsuo213/ 左 "left". The surface representation has been found in the sandhi items of Southern Min (Chien et al., 2017; Zhang, 2018), especially those participating in the 51→55 pattern, which was rather unproductive in pseudo-words (Zhang et al., 2009; Zhang & Lai, 2008; Zhang et al., 2011). The surface representation may be episodic and is stored in memory as a frozen chunk.

The underlying representation view suggests that all the disyllabic words are listed in their pre-sandhi form (i.e., canonical form) despite that they surface in the post-sandhi form in communication. For example, the representation of the tone sandhi word /tsuo213 tɕiɑu213/ 左脚 "left foot" would be /tsuo213/ plus /tɕiɑu213/ in the mental lexicon. According to this view, the allophonic Tone 2-like forms are abstracted away from the mental lexicon. Therefore, it overlaps with the abstract representation account. The underlying representation account can be seen from the data of an early study concerning how Mandarin Tone 3 sandhi is represented in the mental lexicon (Zhou & Marslen-Wilson, 1997). In the study, Zhou & Marslen (1997) designed three types of disyllabic prime words: (i) the Tone 3-initial non-sandhi words (Tone 3 + Non-Tone 3 syllable; e.g., / ts'ai213 xuŋ35/ 彩虹 "rainbow"), (ii) the Tone 2-initial words (e.g., /ts‘ai35 xua35/ 才华 “talent”), and (iii) the control words (e.g., /t‘iæn55 ɣ35/ 天鹅 “swan”). They then used an auditory-auditory priming lexical decision task to examine the effect of the different types of primes on the lexical decision of tone sandhi words (e.g., /ts‘ai213 tɕ’y213/ 采取 “adopt”). The Tone 3-initial non-sandhi words are in the same tonal category with the sandhi words in their underlying form, whereas the Tone 2 initial words share similar pitch contours with the sandhi words in their surface form. Control primes do not share anything in common in terms of lexical tones or segments. The reaction latency showed that the processing of Tone 3 sandhi words was significantly facilitated by Tone 3-initial non-sandhi words, which belong to the same lexical tonal category in the underlying form. Meanwhile, they found an inhibitory effect from Tone 2-initial disyllabic primes. The facilitatory effect from the Tone 3-initial prime words could indicate an underlying representation of Mandarin Tone 3 sandhi. However, when they examined the priming effect of the Tone 2 initial disyllabic words, Tone 3 initial disyllabic words, and the Tone sandhi disyllabic words on the decision latency of Tone 2 initial disyllabic targets (e.g., /ts‘ai213 p’an51/ 裁判 “referee”) in the other experiment, the result showed that even the Tone 2 initial disyllabic prime words showed a significantly more inhibitory effect. According to an empirical study of Sereno & Lee (2015), a facilitatory effect can be found in the priming lexical decision task when a prime word and a target word overlapped in both tonal and segmental information. While in Experiment 2 of Zhou and Marslen-Wilson, no priming effect was found when the initial syllables of Tone 2 prime words and Tone 2 target words overlapped in terms tones and segments, which may result from the distraction of the second syllable of the prime words in the experiment.

Chien and his colleagues (2016) verified the underlying representation view of Mandarin Tone 3 sandhi in Mandarin adult speakers with an updated paradigm. In the research, they adopted the paradigm of Zhou & Marslen-Wilson's study (1997), with the only crucial difference being that they used monosyllabic primes in examining the representation of tone sandhi, which eliminates any priming effect arising from the second syllable of the disyllabic word. Specifically, they examined the reaction time to tone sandhi targets with an auditory-auditory priming lexical decision experiment. Each disyllabic sandhi target (e.g., /fu213 tau213/ 辅导 "to counsel") was preceded by any of three types of monosyllabic primes, namely, a Tone 2 prime (e.g., /fu35/ 服 "be convinced"), a Tone 3 prime (e.g., /fu213/ 辅 "to assist"), or a control prime (e.g., /fu55/ 敷 "to apply"). The Tone 3 prime matched the sandhi initial syllable in its underlying form, whereas the Tone 2 shared a similar pitch contour with the sandhi initial syllable in its surface form. Control primes did not share anything in common in terms of lexical tones. They found that Tone 3 primes facilitated participants' lexical decision latency, while Tone 2 primes did not show any facilitatory effect relative to the control primes. In addition to this, they added word frequency as a factor to examine the nature of the representation of Mandarin Tone 3 sandhi. If the sandhi words were represented in their underlying form, these underlying priming effects would occur regardless of the target word frequency. The results showed that the frequency of the sandhi words did not affect the effect of priming. As a result, they concluded that Mandarin Tone 3 sandhi words were represented in their underlying form (Tone 3 + Tone 3) in Mandarin adult speakers' mental lexicon.

The multi-variant account assumes that the underlying tone and its tonal variants can co-exist in the mental lexicon for the Mandarin Tone 3, possibly together under the node of the abstract underlying Tone 3. In other words, if the account applies to the encoding of Mandarin Tone 3 sandhi items, the representation form of the initial word would be both the underlying Tone 3 and the allophonic Tone 2-like form. This account has been verified by a series of studies (Chen et al., 2011; Li & Chen, 2015; Nixon et al., 2015). For example, Chen et al. (2011) designed three experiments with the implicit priming paradigm (also called form preparation paradigm; see Cholin & Levelt, 2009; Meyer, 1990, 1991, for details), to seek evidence from the phonological encoding of Mandarin Tone 3 sandhi word. The results of vocal reaction time demonstrated that the mental lexicon contains both the underlying- and the surface-form representations. However, the surface-form representations in this study may not necessarily be episodic. The surface form Tone 2-like representations may also be abstracted from the sandhi exemplars. The finding is also compatible with results of a behavioural study with the picture-word interference experiments (Nixon et al., 2015) and a mismatch negativity (MMN) study investigating the effect of allophonic variation on the mental representation and neural processing of Mandarin Tone 3 (Li & Chen, 2015). For example, Li and Chen (2015) used the MMN paradigm in which Tone 1, Tone 2, and Tone 3 syllables were designed as stimuli with the carrier segment [ma]. Four standard/deviant conditions were designed: Tone 1/Tone 3, Tone 3/Tone 1, Tone 2/Tone 3, and Tone 3/Tone 2. Results showed that Tone 1 standard /Tone 3 deviant and Tone 3 standard/Tone 1 deviant conditions yielded comparable MMN effects. However, Tone 2 standard/Tone 3 deviant and Tone 3 standard/Tone 2 deviant conditions showed asymmetrical MMN effects. In Tone 3 standard/Tone 2 deviant condition, the MMN effect was significantly reduced compared to the reversed Tone 2 standard/Tone 3 condition. Li and Chen (2015) claimed that the reduced MMN effect in the Tone 3 standard/Tone 2 deviant condition may be due to the fact that when participants heard a Tone 3 syllable, both Tone 3 and sandhi Tone 3 representations were activated. This, consequently, affected the MMNs to T2 deviant. Based on these results, Li and Chen (2015) proposed that Tone 3 is stored as both Tone 3 and sandhi Tone 3 in the mental lexicon. The multi-variant account for Mandarin Tone 3 sandhi has extended the hybrid model targeting at the segmental and prosodic levels (e.g. Mitterer et al., 2011; Pierrehumbert, 2016)

Besides, it should be noted that MMN studies of Li and Chen (2015)focused on how the Tone 3 monosyllables are represented in the mental lexicon rather than the

Tone 3 sandhi words. Concerning the same question regarding the representation features of Mandarin Tone 3, the experiments of Politzer-Ahles et al. (2016) showed a different result. To be specific, they designed three experiments investigating the representation of Mandarin Tone 3 by looking at the MMN responses to Tone 3 in contrast with the other three Mandarin tones. In the experiments, the participants were presented with monosyllabic Tone 3 words and the other three Mandarin tones which alternatively served as standards and deviants. The results showed that when Tone 3 contrasted with other tones, asymmetrical MMNs were observed, while no asymmetry was observed in conditions when the other tones were contrasted. Since they (2016) also found an asymmetry in the contrast between Tone 3 and Tone 4 in the experiments in addition to the contrast between Tone 3 and Tone 2. They argued that the putatively underspecified nature of Tone 3 may weaken the contrast between standards and deviants. Taken together, further research may be required to confirm the mental representation of Mandarin Tone 3 in monosyllabic conditions.

In addition to the representation features of the productive Mandarin Tone 3 sandhi, it is also worth mentioning that speakers may store unproductive sandhi patterns differently in the mental lexicon in spoken language comprehension when compared to the productive sandhi pattern (Zhang, 2018). He suggested that unproductive alternation patterns (e.g., those involved in a circular chain shift) call for an analysis of surface-form lexical listings (facialized by the lexical mechanism), whereas the productive Mandarin Tone 3 sandhi pattern requires computational processing involving the underlying representation. While the multi-variant account possibly relates to both lexical and computation mechanisms where the abstract underlying form connecting to the computation mechanism and the storage of Tone 2 allophones relating to the lexical mechanism.

2.4. Children's acquisition of tones and tone sandhi

Convergent results indicate that Mandarin Tone 3 sandhi is highly productive and is manipulated by means of the computation mechanism in adult speakers. However, it is still unclear when the pattern becomes productive during children's language development. Many studies have investigated children's first language acquisition and have tried to describe children's learning process (Cazden, 1968; Pinker, 1999; Jenny & Ann, 2001). Child language acquisition and cognitive development are reported to take the course of a U-shaped learning process (Case, 2005). Take the development of plural inflection in English as an example (Cazden, 1968). The acquisition of English plural inflection goes through four stages: (i) The inflection is absent, e.g., "What are those?" "Shoe." (ii) Occasional productions with plural inflections occur, but with no error overgeneralizations. e.g., "Girls are singing." (iii) Children start to produce more plurals but tend to make more overgeneralizations at the same time. "e.g., Two mans are talking." "You pull my pantses down." (iv) Children start to use the plural ending correctly nearly all the time. The four stages can be summarized as a gradual development from the lexical mechanism to the computation mechanism, and finally to a combination of both. At stages (i) and (ii), children memorize the words via the lexical mechanism regardless of whether or not they added the plural-form. When they try to use the computation mechanism at stage (iii), overgeneralization occurs. Although more errors appear at stage (iii), it does not necessarily mean that their ability has decreased. Instead, it indicates that children have started to become aware of the morphological pattern and have begun actively to apply the pattern to new words. At the final stage, the precise use of plural forms requires the involvement of both the computation mechanism and lexical mechanism. The plural inflection (e.g., shoe-shoes) is manipulated through the computation mechanism, whereas for the irregular nouns (e.g., man-men), the lexical mechanism is presumably adopted.

The acquisition of Mandarin tone sandhi may not necessarily undergo all the stages because the application of the sandhi pattern was rather regular in adults' speech production. However, it is still highly likely that they undergo development from the lexical mechanism to the computation mechanism. In particular, children might start producing tone sandhi words based on adult utterances, that is, memorizing the words in their surface forms at an early age. However, the acquisition of Tone 3 sandhi pattern goes beyond the storage of memorized tone sandhi words at a certain age. They can apply the pattern to words they have not heard. For example, they may learn to pronounce the word /ɕiau213 kou213/ 小狗 "puppy" by memorizing the heard surface-form [ɕiau35 kou213] as a frozen chunk. That is, they may acquire the Tone 3 sandhi words by the lexical mechanism at the beginning, but the computation mechanism becomes dominant after they are exposed to enough sandhi and non-sandhi words.

An adult-like performance in the application of the productive tone sandhi requires collaborative efforts in many aspects. It is not difficult to imagine that during the process of language acquisition, both the surface and the underlying forms of Tone 3 words are provided to children in language input. For example, the word /ɕiɑu213/ 小 "small" may be frequently heard by children in different forms, such as /ɕiɑu213 xua55/ 小花 "flower", /ɕiɑu35 kou213/ 小狗 "dog", or single word /ɕiɑu213/ 小. At the very beginning, children may only memorize the different forms as frozen chunks. To achieve this, it first requires children's acquisition of tones and tone sandhi words in both production and perception. While if a child is asked to link the surface allophones of an initial syllable (e.g., /ɕiɑu213/) appearing in different tonal contexts (e.g., /ɕiɑu213 xua55/ and /ɕiɑu35 kou213/ to the same morpheme, some more efforts are required. It involves children's ability to manipulate the disyllabic tone sandhi words into syllables first, which is usually called syllable awareness (e.g., Shu et al., 2008). In speech productions, as seen in the example, the speech flow is not always neatly presented in single words but is rather showed as a surface-form speech flow.

To discover the initial syllable of the sandhi words, children need to decompose the disyllabic words into syllables. Besides, it possibly further requires children's development of morphological awareness (Hao et al., 2013; Chang, et al., 2003), which refers to the explicit or implicit manipulation of morphological units to morphemes. To identify the surface-form allomorphs of the initial syllable (/ɕiɑu35/) in /ɕiɑu35 kou213/ as the underlying Tone 3, children are expected to aware the word meaning regardless of whether it is in surface or underlying form. Besides, children's maturation, language input, vocabulary, literacy, or Pinyin learning may influence the progress of sandhi application. Pinyin refers to a romanized translation of Chinese characters (see https://www.pinyin-guide.com/ for details). In Mainland China, all character pronunciations can be denoted by the alphabetic coding system. For example, 小 is transcribed as 'xiao3' in Pinyin. The letter-sound correspondence is highly regular. Children typically receive systematic explicit Pinyin instructions when they enter Grade 1 of primary school. Previous literature has demonstrated that explicit learning of Pinyin enhances children's phonological awareness (e.g. Siok, & Fletcher, 2001; Yin et al., 2011).

Plenty of studies have examined children's productions of Mandarin tones. They observed children's productions of tones in both longitudinal and synchronic perspectives. The convergent results showed that children's acquisition of tones showed a tremendous amount of individual differences. The first longitudinal study can be traced back to Chao's observations (Chao, 1973) of phonological development in a 28-month-old girl who acquired Mandarin Chinese as her first language in the U.S. He recorded and analyzed the child's vocabulary and phonology during a one-month observation period. The tracking result showed that the child produced tones in isolation correctly at the beginning of the study (i.e., two years and four months; hereafter 2;4, in which 2 represents the year, and 4 represents month).

Later, Clumeck (1977) reported longitudinal data for three children. Two of them were raised in monolingual families. Of these, one child, a boy, was studied from 2;3 to 3;5 years old, and the other child, a girl, from 1;0 to 2;10. Their tone productions in isolation and utterance-final position were analyzed in imitated, elicited, and spontaneous production. The results showed that both children acquired the level tone and the falling tone "almost" entirely, but that they were still undergoing development in producing the rising and low-level tones until three years old (Clumeck, 1977).

Another longitudinal study (Zhu, 2002) recorded four children's productions from 10 to 14 months in Beijing, China. The data collection ended when the children aged 2, except one who was only recorded to 20 months old. Results found that the children's productions of the four lexical tones in Mandarin Chinese were stably acquired between the age of 1; 2-1; 9.

Li and Thompson (1977) observed tone acquisition features in Mandarin-speaking children from both longitudinal and synchronic perspectives. They recruited

17 children from 1;5 to 3 years in Taipei and traced their tone production for seven months. In their results, they did not explicitly illustrate the timeline of tone acquisition; instead, they reported the order of tone acquisition in the Mandarin-speaking children, namely that high and falling tones (Tone 1 and Tone 4) were acquired before the rising and dipping tones, and that the acquisition of lexical tones was expected to be earlier than the acquisition of segmental elements.

Zhu and her collaborator Dodd also conducted a comprehensive large-scale cross-sectional study on children's production of Mandarin (Zhu & Dodd, 2000). This study gave a better picture of the acquisition of tones, based on 129 children (1;6-4;6). They found that the children in the youngest age group (aged 1;6 – 2;0) had acquired the production of tones in different linguistic contexts.

According to the papers reviewed above, it remains challenging to determine the critical age for children's acquisition of the four lexical tones in Mandarin Chinese. Several results have shown that tone production was acquired before two years old (e.g., Chao, 1973; Zhu & Dodd, 2000), whereas other studies have reported that tone production was not fully acquired even at three years old. For example, three-year-old children were found in one study to continue to have difficulties in distinguishing between rising and dipping tones (Clumeck, 1977). The discrepancies in the findings mainly result from different rating criteria and different sample sizes. The production accuracy of tones was determined merely by one or two transcribers' subjective transcriptions in some studies (e.g., Chao, 1973; Zhu & Dodd, 2000). Besides, children's production of tones could also be judged with the support of lexical and other contextual cues (e.g., Chao, 1973; Clumeck, 1977), which created expectations for the target tone. The effect may lead to the tones produced by children being accepted as correct by the judge, even though there might be subtle phonetic discrepancies in the tone production (e.g., Zhu & Dodd, 2000).

Wong and her collaborators (Wong, 2005; Wong, 2012; Wong & Strange, 2017) conducted a series of experiments to examine Mandarin-speaking children's acquisition of lexical tones in both monosyllabic and disyllabic environments. Their results showed that children's acquisition of tones was not acoustically accurate as early in age as previous studies had suggested. In a speech perception task, they recruited children as well as some of their mothers to produce a set of words that carried different tones in a picture reading experiment. They also recruited Mandarin-speaking judges to identify the productions of tones. To eliminate the influence of lexical cues, they filtered out the segmental information from the produced words before conducting the tone production judgment. The results suggested that children's production of the four lexical tones in monosyllables was not yet adult-like at three years old. Among the four lexical tones, the children had the most difficulties with the low dipping tone. In the disyllabic environment, children had not reached the adult level in the production of Mandarin tones even at six years old. In lexical tone perception, the results revealed that three-year-old Mandarin-speaking children perceived the four lexical tones with high accuracy (Wong, Schwartz, Jenkins, 2005). Therefore, the conclusion might be that the children may have a good recognition of the four lexical tones by three years old. However, possibly due to their immature speech motor control, children's tone production was still undergoing development even at six years old in the disyllabic environment.

For the acquisition of tone sandhi items, previous results were also inconsistent. In general, children who speak Mandarin as their mother tongue could produce lexical tones and Tone 3 sandhi words at or before age 3 (Chao, 1973; Tang, 2013; Wang, 2011; Wang et al., 2015). However, their productions were still not adult-like at five years old (Rattanasone et al., 2018). The earliest studies can also be traced back to Chao's paper (1973) in which the girl was reported to produce some tone sandhi words from age 2;4 to age 2;5. In a more recent study, Wang et al. (2015) recruited children from 2;6 to 3;6. By observing the acoustic features of pronounced sandhi words, such as /ʂ'ou213 piau213/ 手表 "watch," they concluded that children had acquired the production of tone sandhi words at two years and seven months old. However, the conclusion made in the paper was based only on a comparison between the children's productions in the different age groups and a description of the figure derived from the acoustic pitch trajectory. When compared with the acoustic pitch realization of sandhi words in adults' speech production (Rattanasone, Tang, et al., 2018), the results revealed that children from three to five years of age produced the sandhi words with a more flattened rising curve than adults.

2.5. The productivity of tone sandhi in novel words and its mental representation feature in Mandarin-speaking children

It is worth noting that the correct production of Tone 3 sandhi real words does not imply the productivity of Tone 3 sandhi in novel words. Previous studies on the acquisition of productive voicing neutralization and alternations in Dutch have revealed that children still exhibited errors in applying the patterns productively to novel words even when they were able to produce the alternations correctly in real words (e.g., Zamuner, et al., 2006). Likewise, children's acquisition of Mandarin Tone 3 sandhi may also undergo developmental stages in reaching an adult level of productivity in novel words.

As far as children's productive knowledge of tone sandhi in pseudo-and non-words is concerned, however, less is investigated. Tang and his collaborators (Tang, Yuen, et al., 2019) explored children's productive use of Tone 3 sandhi pattern in pseudo-words in three- to five-year-old children. For those pseudo-words, the output of which can lean upon either the previously heard sandhi allomorphs or the computation mechanism, Mandarin-speaking children were reported to have reached an adult level of productivity at three years old (Tang, et al., 2019). To be specific, the authors examined the children's productions of disyllabic Tone 3 sandhi pseudo-words (called novel words in that paper) such as /ma213 ku213/ 马鼓 "horse drum" with adults as controls. The acoustic analysis showed that the children from three to five years of age applied the pattern productively to those pseudo-words, just as the adults did. However, they only used the pseudo-words (e.g., /ma213 ku213/) in their experiment. Since the pseudo-words still have lexical entries for participants to retrieve, it is unknown when the Mandarin-speaking children would reach adult-like productivity in the production of Tone 3 sandhi non-words, thereby demonstrating a developed computation mechanism. Meanwhile, considering the possible acoustic difference between children and adults, it is more credible to take advantage of children's productions of real tone sandhi words as controls. Second, they only included four items in the experiments. Some variations may, therefore, be hidden in this case. In order to take a closer observation of the productive knowledge, this thesis included 20 test items, and made use of both tone judgment and acoustic analysis.

Regarding children's representation of tone sandhi in the mental lexicon, a previous study (Wewalaarachchi & Singh, 2016) indicated that when children were presented with both surface-form Tone 3 sandhi words (e.g., /iy35 san213/) and the underlying-form sandhi words (e.g., /iy213 san213/), they only recognized the surface-form at ages three and five. This possibly suggests that children who speak Mandarin Chinese as their mother tongue still process the tone sandhi words as frozen chunks (subserved by the lexical mechanism) up to age five. However, since the underlying forms are inaccurate productions, and the experiment did not explain whether adults accept the underlying form words. If even adults don't accept such forms as okay examples, then we can't pin it down as underdevelopment in children.

While considering the adult production and mental representation of Mandarin Tone 3 sandhi, the sandhi form is suggested to be computed productively from the underlying form (Chien et al., 2016; Zhang & Peng, 2013; Zhang & Lai, 2008). Given the aforementioned results, children may undergo a transformation from the lexical mechanism to the adult-like computation mechanism. To find out when Mandarin children reach the adult-like pattern, I conducted a series of experiments that examined the performance of children from three to eight years old, as well as that of adults, who acted as controls. Moreover, the factors that may contribute to the development of tone sandhi were discussed. These studies are elaborated in the following chapters.

Chapter 3 The Productivity of Mandarin Tone 3 Sandhi in Mandarin-speaking Children

3.1. Introduction

As discussed in the previous chapter, two operational mechanisms might influence the productivity of tone sandhi in Chinese dialects: a rote-memory lexical mechanism and a computation-based productive mechanism (originated from Pinker, 1998; Zhang & Peng, 2013; Zhang, et al., 2015; Zhang et al., 2011). Studies have reached the consensus that the application of the Mandarin Tone 3 sandhi pattern is highly productive in novel words in adult Mandarin speakers (Zhang & Lai, 2010; Zhang et al., 2015; Zhang & Peng, 2013). However, it is still uncertain when this phonological pattern becomes productive in novel words in children learning Mandarin as their mother tongue.

Children's acquisition of some morphological patterns, such as the English past tense inflections, has been demonstrated to proceed according to developmental stages, progressing from unanalytic memorization of lexical items (implying the lexical mechanism; Jenny & Ann, 2001), to a regularized computation and productive application of the linguistic patterns to new items (implying the computation mechanism), and finally to a combination of both mechanisms (lexical: go-went & computation: watch-watched; Pinker, 1999).

Unlike the morphological processing such as the English past tense or plural infections, which includes both regular and irregular forms, the Mandarin Tone 3 sandhi pattern is highly regular. It is thus expected to undergo development through the first two stages only, from the lexical mechanism to the computation mechanism. The hypothesis has been tested and verified in part by some empirical studies. As has been shown, children were able to produce some recognizable sandhi words at or before three years old (Wang et al., 2015; Li & Thompson, 1977; Chao, 1973; Hua & Dodd, 2000), although their productions were not yet acoustically accurate enough even at five years old (Rattanasone et al., 2018). For the pseudo-words, the output of which can lean upon either the previously heard sandhi allomorphs or the computation mechanism, Mandarin-speaking children were reported to have reached an adult level of productivity at three years old (Tang et al, 2019). The accumulative results imply that the lexical mechanism is dominantly utilized at around three years old. To uncover when Mandarin-speaking children reach adult-like productivity in the production of Tone 3 sandhi novel words, thereby demonstrating a developed computation mechanism, this study examined the speech productions of both Mandarin-speaking children from three to eight years old and adults, using real words, pseudo-words, and novel words. It is expected that Mandarin-speaking children acquire the adult-like pattern in real and pseudo-word conditions at three to five years old, and reach the adult-like pattern in novel words, where the computational phonological processing is required.

3.2. Methods

3.2.1. Participants

Seventy-four children (age range: 3;0–8;10) and thirteen adults were recruited for the study. Among the children, nine were aged 3 (between 3;0-3;11), thirteen were aged 4 (between 4;5-4;11), thirteen were aged 5 (between 5;2-5;11), thirteen were aged 6 (between 6;0-6;11), thirteen were aged 7 (between 7;2-7;11), and thirteen were aged 8 (between 8;3-8;9). All the adults have received systematic explicit Pinyin instructions. The seven- and eight-year-old children were all recruited from primary school and had thus received explicit Pinyin instructions. The children aged three to six were all recruited from kindergartens. The three and four-year-old children have never received any explicit Pinyin instructions. While five and six-year-old children either had a few exposures to explicit Pinyin instructions beforehand (partially trained) or were systematically trained. Table 3.1 listed the specific information about the participants in each age group. All children were monolingual speakers of Mandarin Chinese who had not been exposed to any other languages, such as English or other Chinese dialects. All children passed the hearing screening at 1000 Hz, 2000 Hz, and 4000 Hz at 30 dB on both left and right ears using the headphone (Interacoustic AS608 Screening Audiometer). All children were instructed to accomplish intelligence tests (WISC-R; Zhang & Lin, 1979), which were revised versions of Wechsler's test. The tests included two versions, one of which was for preschoolers below six and a half years old and the other one for children above six and a half years old. The IQ scores were comparable between different age groups (except for the 3-year-old children, who were not tested when the first three participants failed to complete the tests). The experimental procedure was approved by the Human Subjects Ethics Sub-committee of The Hong Kong Polytechnic University. Informed written consent was obtained from all of the adult subjects and children's caregivers in compliance with experimental protocols.

Table 3.1 *Information of participants in each age group*

age groups	**number**	**mean age (age range)**	**gender ratio**	**whether received Pinyin instructions**
3y	9	3;8 (3;0-3;11)	5M/4F	9 (no training experience)
4y	13	4;9 (4;5-4;11)	8M/5F	13 (no training experience)
5y	13	5;5 (5;2-5;11)	8M/5F	5 (systematically trained); 8 (partially trained)

6y	13	6;7 (6;0-6;11)	3M/10F	7 (systematically trained); 6 (partially trained)
7y	13	7;9 (7;2-7;11)	7M/6F	13 (systematically trained)
8y	13	8;8 (8;3-8;9)	7M/6F	13 (systematically trained)
adult	13	27;6 (23.2-31.3)	6M/7F	13 (systematically trained)

3.2.2. Stimuli

To examine whether the Mandarin Tone 3 sandhi pattern was applied productively, real words, pseudo-words, and novel words were designed for the experiment. Each type of word included 20 test items (Tone 3 + Tone 3 words) and 20 control items (Tone 3 + Non-Tone 3 words). For the items in the test and control groups, the initial syllables within each pair were always the same. Table 3.2 listed the specific information about the three types of words. Considering the association between vowel height and F0 values (e.g., Lehiste, 1970), the experiment design covered as many consonants and vowels as possible in each lexicality condition.

Table 3.2a *a list* of *the real words where the onset and rhyme of the initial syllable for the Tone 3 +Tone 3 words and Tone 3+ Non_Tone3 words are provided.*

	Tone 3 +Tone 3	Tone 3 + Non_Tone 3	onset	rhyme

1	老鼠	老鹰	l	ao
2	雨伞	雨衣	y	u
3	粉笔	粉鞋	f	en
4	海马	海豚	h	ai
5	小鸟	小鸡	x	iao
6	老虎	老牛	l	ao
7	手表	手机	sh	ou
8	小草	小花	x	iao
9	水果	水杯	sh	ui
10	水桶	水枪	sh	ui
11	手指	手枪	sh	ou
12	洗脚	洗牙	x	i
13	洗手	洗头	x	i
14	蚂蚁	马路	m	a
15	洗脸	喜糖	x	i
16	洗澡	喜字	x	i
17	左脚	左边	z	uo
18	左手	左轮	z	uo
19	左脸	左面	z	uo
20	左腿	左键	z	uo

Table 3.2b *a list of the pseudo-words where the onset and rhyme of the initial syllable for the Tone 3 +Tone 3 words and Tone 3+ Non_Tone3 words are provided.*

	Tone 3 +Tone 3	Tone 3 + Non_Tone 3	onset	rhyme
1	马草	马树	m	a
2	笔狗	笔猪	b	i
3	草雨	草云	c	ao
4	打海	打山	d	a
5	枣眼	枣眉	z	ao
6	纸手	纸牙	zh	i(ʅ)
7	狗酒	狗烟	g	ou
8	眼腿	眼头	y	an
9	饼跑	饼哭	b	ing
10	雪脚	雪毛	x	ue
11	买打	买爬	m	ai
12	洗海	洗冰	x	i
13	跑笔	跑书	p	ao
14	碗鸟	碗牛	w	an
15	海火	海花	h	ai
16	鸟雪	鸟楼	n	iao
17	腿马	腿猫	t	ui
18	火枣	火姜	h	huo
19	脚米	脚葱	j	iao
20	酒伞	酒门	j	iu

Table 3.2c *a list of the novel words where the onset and rhyme of the initial syllable for the Tone 3 +Tone 3 words and Tone 3+ Non_Tone3 words are provided.*

	Tone 3 +Tone 3	Tone 3 + Non_Tone 3	Fillers
1	zhei3 笔	zhei3 书	zhen2 笔
2	mie3 草	mie3 树	ming1 草
3	hei3 手	hei3 牙	huai1 手
4	sen3 果	sen3 花	shun1 果
5	cou3 脚	cou3 头	cong4 脚
6	diu3 酒	diu3 烟	dang2 酒
7	nou3 表	nou3 钟	nuo1 表
8	ce3 碗	ce3 盆	cui2 碗
9	miu3 跑	miu3 哭	mian1 跑
10	kuo3 伞	kuo3 球	kao2 伞
11	nen3 手	nen3 毛	nun4 手
12	shuo3 枣	shuo3 姜	shua2 枣
13	chuo3 饼	chuo3 鱼	chua2 饼
14	chua3 雪	chua3 冰	chuo2 雪
15	te3 纸	te3 针	tie2 纸
16	ne3 嘴	ne3 鞋	nei1 嘴
17	den3 狗	den3 鸡	diu4 狗
18	suan3 马	suan3 猪	sa2 马
19	eng3 海	eng3 山	er1 海
20	zhui3 雨	zhui3 云	zhun2 雨

To further explain the three types of words in this study, the real words are actual disyllabic words or phrases which occur in daily life. They are all frequently used or heard and all have concrete meanings, for example, /ɕiau213 ts‘au213/ 小草 “grass” vs. /ɕiau213 xua55/ 小花 “flower”. The pseudo-words are non-occurring combinations of real monosyllables, that is, both their initial and final syllables are actual morphemes occurring in Mandarin Chinese, but their combination is rarely found in everyday communication. In the pseudo-word condition, the two syllables also have concrete meanings which would be highly familiar to children, such as animals, plants, or body parts. For example, /ma213 ts‘au213/ 马草 “horse grass” vs. /ma213 ʂ‘u51/ 马树 “horse tree”. The novel words were the composition of an AG word in Mandarin phonology and a real monosyllable. In this experiment, the AG word refers to a type of novel syllable whose combination between consonant and vowel occurs in Mandarin Chinese, but which yields no real morpheme when combined with either Tone 2 or Tone 3, such as the syllable /miɛ213/. The syllables /miɛ213/ and /miɛ35/ correspond to no morpheme in Mandarin Chinese, but the combination between the consonant /m/ and vowel /iɛ/ does in fact exist in Chinese phonology (/miɛ51/ is a real word “to destroy” in Mandarin).

3.2.3. Procedure

To ensure that the child would take the initiative in speaking the specified words, this study adopted an elicited picture-naming task, usually called a “wug test” (Gleason, 1958). The experiment started with an elicited word-guessing game for the pseudo- and novel words. The real word condition was conducted last, with a picture-naming task. For the two types of pseudo-sequences, the experimenter presented slides of pictures on a laptop to the subjects and instructed them that the alien girl "Xiao Yu" wanted them to learn words on her planet. In the pseudo-word condition, two pictures were shown on one page to indicate the two syllables, which were introduced syllable by syllable in their underlying forms by an experimenter, and the children were required to combine the two syllables into a new word. In the novel word condition, the experimenter assigned each AG syllable a lexical meaning (e.g., /miɛ213/ is a verb which means "(for two men) to carry something") and asked the children to combine the AG word and the real monosyllabic word into a new word. Take the speech elicitation of the novel word /miɛ213 ts‘au213/ miɛ3 草 "to carry grass by means of two men" as an example: the experimenter firstly showed a picture indicating that ‘miɛ213’ is a verb which refers to "to carry something by means of two men." The experimenter then familiarized the children with the word by giving examples, like so: "If two men carry a book (/ʂ‘u55/ 书 in Chinese), we would say /miɛ213 ʂ‘u55/. Then, if two men carry grass (/ts‘au213/), what we should say?" It is worth emphasizing that the syllables in the pseudo- and novel word conditions were introduced in their underlying form. The task is not easy, especially for three-year-old children. Before the experiment, I offered children training with enough practice trials to ensure that they understood the tasks and could combine the two syllables into new words in pseudo- and novel words. For the real word condition, one picture was presented in each trial. The children were asked to tell the experimenter what was in the picture. The children were instructed to produce each target item twice. The first token would be preserved for analysis unless it was not ideal due to environmental noise or prolonged productions. Figure 3.1 is a list of sample materials for each type of word.

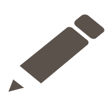
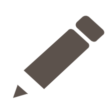
	Test items	Control items
real words e.g./ʂ‘ou213 piau213/ &/ʂ‘ou213 tɕi55/		
pseudo-words e.g./pi213 kou213/ &/pi213 tʂu55/	 	 
novel words (AG) e.g./miɛ213 ts‘au213/ &/miɛ213ʂ‘u51/	 	 

Figure 3.1 *Sample materials for the three types of words*

3.2.4. Recording and segmentation

The interactions of the experimenter and participants were all recorded by digital recorders (Sony, ICD-PX470) at 16 bits with a 44100 Hz sampling rate. The experiment lasted 30-60 minutes for a child and approximately 25 minutes for an adult. To avoid fatigue and to encourage the children to be engaged, the experiment was divided into at least six parts. The children were allowed to rest between two parts for at least ten minutes. A child would be awarded some toys or snacks every time he/she actively engaged himself/herself in the task for more than 5 minutes. If a child lost interest in the experiment on that day, she/he would be invited to complete the experiment in the following few days.

Participants' productions were further labelled and extracted from the recordings using Praat (Boersma & Weenink, 2018) for further analysis. Before the analysis, the trials were scrutinized. Trials would be discarded if they met the following criteria: (i) The AG syllable was mistakenly produced as an actual-occurring syllable (e.g., miɛ213, produced as mei213); (ii) The interval between the two syllables in the sandhi context was longer than 200 ms (this criterion was used for excluding trials that did not form a disyllabic word which would affect the application of Tone 3 sandhi);

and (iii) For items that were produced twice, the first production was preferentially used, and the second token would be used for analysis only when the first production was unsuitable due to noise or other factors. Preserving one token for each item in each speaker, there were 10440 tokens (87 subjects * 120 items) in total, of which 8385 of them were eventually preserved for analysis.

3.3. Data analysis

Participants' production was analyzed from two perspectives: (i) judgment of tone production by phoneticians and (ii) acoustic analysis of the sandhi words. For the phoneticians' judgments, the production of each disyllabic sequence (e.g., /miɛ213 ts'au213/), and the initial Tone 3 syllable extracted out of that sequence (e.g., /miɛ213/), were segmented. For the acoustic analysis, the rime of each initial Tone 3 syllable of the disyllabic sequence (e.g., /iɛ213/ in /miɛ213/) was segmented. This chapter focuses on the result of the phoneticians' judgment.

3.3.1. Participants for the tone production judgments

For the tone production judgment, eight Mandarin-speaking phoneticians who were postgraduate students majored in experimental phonetics participated in the task. They were from Northern dialect-speaking areas. None of them was reported to have any hearing impairment. All participants gave informed consent in compliance with a protocol approved by the Survey and Behavioral Research Ethics Committee of The

Hong Kong Polytechnic University. Given the difficulty of detecting the Tone of the initial syllable in the sandhi context, phoneticians' productions showed some variabilities. The results of five of the eight phoneticians were preserved for final analysis because they met the following criteria. (i) Their rating accuracy of the adult Tone 3 + Non-Tone 3 trials was higher than 95% (i.e. consistently, they classified the first syllable as Tone 3 in Tone 3 + Non-Tone 3 trials); and (ii) they reached a significant intra-judge agreement on tone sandhi application/non-application in the trials of the participants (i.e., in rating the data of seven subjects whose productions were repeated three times in the speech production task); and (iii) they reached a significant inter-judge agreement with others on tone sandhi application/non-application in the trials of the adult participants according to the results of the Fleiss' kappa test ($k = 0.826$, $p < .001$) (Gwet, 2014). In the analysis with Fleiss' kappa, the k value gives a measure for how consistent the ratings are among raters more than three, and the p-value expresses the extent to which the observed amount of agreement among raters exceeds the random result. The kappa value is expected to be above 0.75 if the results were in good agreement.

3.3.2. The procedure of tone production judgment

The tone production judgment was presented in E-Prime 2.0 (Psychology Software Tools, Inc., 2017). To eliminate any distracting effects from semantic meaning and context on the tone judgment of the initial syllable, a disyllabic sequence was presented first in each trial, followed by the initial syllable of that disyllabic sequence. The phoneticians were instructed to identify whether or not the initial Tone 3 syllable underwent tone sandhi. If tone sandhi was applied, they were asked to press 2 (a sandhi Tone 3 becomes Tone 2 or Tone 2-like), and if it was produced in the underlying form of Tone 3, they were instructed to press 3 (indicating Tone 3 or a Tone 3-like tone); if the production was neither the underlying form nor the surface form of Tone 3 (indicating a mispronunciation), they were asked to press 0. After the tone production judgment task, a five-degree rating task was presented to evaluate how similar the production they heard was to their chosen tone, on a scale of 1-5, in which 5 represents remarkably similar and 1 very different. If they chose 0 in the identification session, they would also be instructed to press 0 in the rating session. The specific procedure of the task is described in Figure 3.2. The procedures were first shown to the participant in a written instruction and then explained orally to them again to make sure that they understood the experiment.

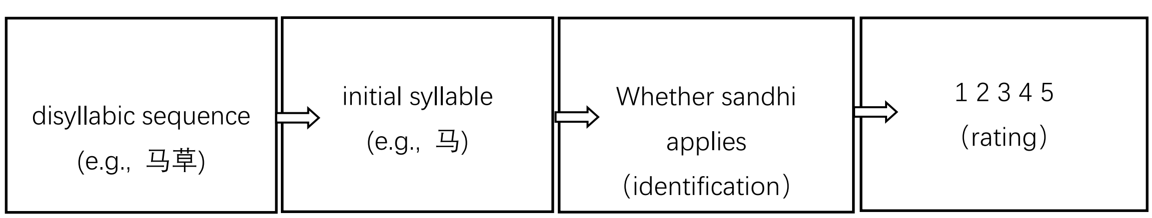


Figure 3.2 *The procedure of the production judgment test*

Both the test items (Tone 3 + Tone 3 sequences) and the control items (Tone 3 + Non-Tone 3 sequences) were included in the production judgment task. The results of the test items were primarily used for data analysis. The whole task is rather long and requires a high degree of concentration. To ensure the accuracy of the judgments, the task was divided into 30 blocks in which the three types of stimuli were randomly mixed. Each block took the raters for about 30 minutes. They were allowed to take breaks at any time they desired. Moreover, if they found that they had made a mistake, they could correct it at any time.

3.4. Results

3.4.1. Result of the tone production judgment

Figure 3.3 summarizes the sandhi application rate of Tone 3 production in the three types of stimuli in each age group. Boxplots, which display the distribution of data within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual points of sandhi application rate, ranging from 30% to 100%, are displayed.

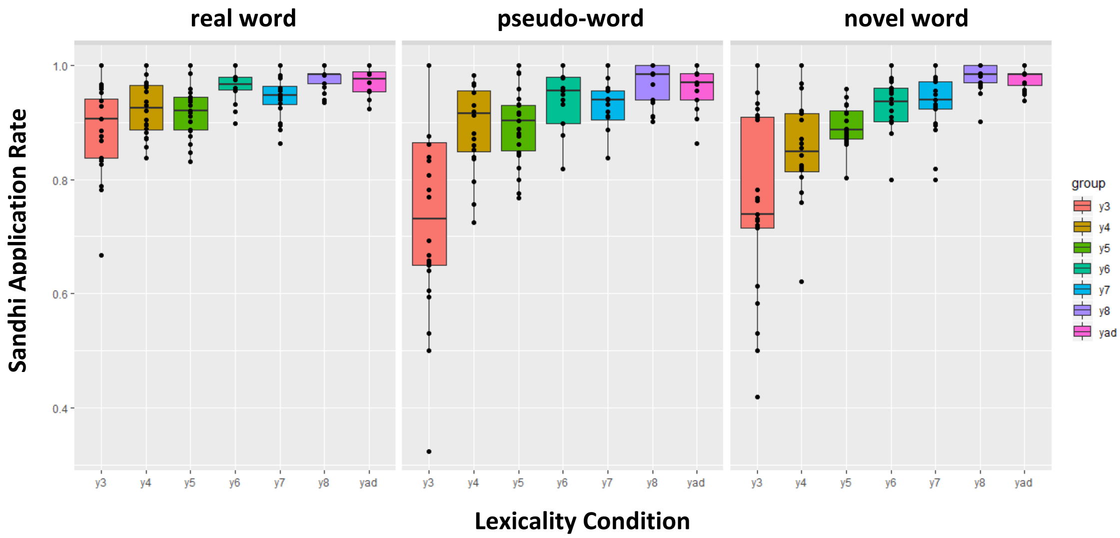


Figure 3.3 *the individual features of the sandhi application rate for the test items. The dots represent individuals' sandhi application rate, grouped by age.*

First, statistical analyses were conducted on the sandhi application rate of the Tone 3 sandhi items across different age groups. We measured the sandhi application rate in Tone 3 sandhi words according to the five phoneticians' judgments. It was labelled as 1 if the initial Tone 3 syllable was judged to be Tone 2-like (indicating the application of the sandhi pattern), and as 0 if the item failed to exhibit a tone sandhi pattern (indicating underapplication). Mispronunciations were deleted for analysis. There is currently a debate in the literature concerning the appropriate method for constructing statistical models. Barr et al. (2013) has argued that random effects structure should be kept maximal, in that model comparisons should be made by starting first with a full random structure and then eliminating non-significant predictors. The advantage of this method is that it circumvents the possibility of falsely reporting a fixed effect which originates from the misspecification of an unknown true random-effect structure. While convergence problems would occur if model models were too complex. If the maximum likelihood estimation of a model failed to reach convergence, one correlation parameter would be removed. Following the method of Barr et al. (2013), I adopted backward algorithms for model comparison. To be specific, data were analyzed by generalized linear mixed-effects models (Baayen, 2008), using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The binomial distribution and the logistic link function with family = "binomial" were selected for the analysis. The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included sandhi application (1 indicates application; 0 indicates underapplication) as the dependent variable, the lexicality condition (real words, pseudo-words, and novel words) and age groups (3y, 4y, 5y, 6y, 7y, 8y, and adult) as independent variables, subject, and item as random factors. Random slopes were also added for each of the fixed effects. Models were compared using log-likelihood ratio tests to determine if given terms significantly improved the model. For the age group, the adult group is treated as the baseline first to which the other age groups are compared.

The results showed that the best model included an interaction between lexicality conditions and age groups, and lexicality condition|item, subject as random factors. To compare whether children performed differently from adults in each lexicality condition, data were analyzed separately in the three lexicality conditions. The analyses demonstrated that the age group variable contributed significantly to the model across all the three lexicality conditions. In real words, the application rates of three- to five-year-old children were significantly lower than adults (*p*s<.001). In the pseudo-word condition, the results were the same: three- to five-year-olds exhibited significantly lower sandhi application rates than did the adults (*p*s<.001). In the novel word condition, however, the three- to six-year-old children showed a significant difference in the application of sandhi pattern from the adults, indicating a significantly lower application rate in three- to six-year-olds than the adults (*p*s<.001), while the difference between six-year-olds and adults was marginally significant (*p*=.05). Detailed information is reported in Table 3.3.

Table 3.3 *Fixed effect estimates of children's sandhi application across different age groups compared to adults in the three lexicality conditions.*

Lexicality condition	age group	Estimate	Std. Error	z value	*P*-value
Real word	(Intercept)	4.0831	0.3756	10.870	*p<.001****
	y3-adult	-1.4754	0.5278	-2.795	*p<.01 ***
	y4-adult	-1.0448	0.468	-2.231	*p<.05 **
	y5-adult	-1.3121	0.4687	-2.799	*p<.01 ***

	y6-adult	0.0754	0.5657	0.133	*p=0.89385*
	y7-adult	-0.5530	0.5222	-1.059	*p=0.2896*
	y8-adult	0.0175	0.5028	0.035	*p= 0.97212*
Pseudo-word	(Intercept)	4.358	0.707	6.166	*p<.001 ****
	y3-adult	-2.848	0.813	-3.502	*p<.001 ****
	y4-adult	-1.840	0.658	-2.796	*p<.01 ****
	y5-adult	-1.366	0.656	-2.083	*p<.05 ***
	y6-adult	-0.402	0.748	-0.538	*p= 0.59090*
	y7-adult	-0.261	0.787	-0.331	*p= 0.74065*
	y8-adult	-0.520	0.543	-0.958	*p= 0.33784*
Novel word	(Intercept)	4.0061	0.3753	10.675	*p<.001 ****
	y3-adult	-2.4903	0.5536	-4.498	*p<.001 ****
	y4-adult	-1.6693	0.5391	-3.097	*p<.01 ***
	y5-adult	-1.4526	0.5123	-2.835	*p<.01 ***
	y6-adult	-0.9613	0.5199	-1.849	*p=.05·*
	y7-adult	-0.5576	0.5757	-0.969	*p=0.33274*
	y8-adult	0.6642	0.5827	1.14	*p=0.25435*

Secondly, another analysis was conducted adding Pinyin instructions as a variable in addition to the lexicality conditions and age that examined the influence of explicit Pinyin instructions on sandhi application. The analysis focused on the five and six-year-olds where there was a mixture of children with and without systematic Pinyin training. The factor Pinyin instruction was treated as a categorical variable (systematically trained, partially trained). Data were analyzed by a generalized linear mixed-effects model (Baayen, 2008), using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included sandhi application (1 indicates application; 0 indicates underapplication) as the dependent variable, lexicality condition (real words, pseudo-words, non-words) and Pinyin instruction (systematically trained, partially trained) as independent variables, subject, and item as random factors. Random slopes were also added for each of the fixed effects. The statistical analysis started with a full random structure. Models were compared using log-likelihood ratio tests to determine the best model. The results showed that adding the variable 'Pinyin instructions' did not significantly improve the model. The analysis, therefore, demonstrated that the prior experience on explicit Pinyin instructions does not significantly influence children's operation mechanism of tone sandhi. However, it is noteworthy that only 5 five- and six- year -old children among the 26 have received systematic explicit Pinyin instructions. The null effect of Pinyin instruction in this part may result from the limited number of sample size.

It is worth noting that, although the three- to five-year-old children showed significantly lower sandhi application rates than the adults in both real and pseudo-word conditions, the degree of underapplication was different in the two lexicality conditions. The difference between three- to five-year-olds and adults was much smaller in the real words than in the pseudo-words. The evidence can be found by the

Estimate (effect size) in Table 3.2, which represents the difference between the child age *group* and the adults in each lexicality. To be explicit, the larger the absolute value of *effect size*, the larger the difference between the children's age group and the adults. Positivity and negativity indicate the difference between the target group and the baseline (adult). Therefore, a negative *effect size* suggests lower application rates in children than in adults. As can be seen in the table, the absolute *effect size* in the pseudo- and novel words was rather large in three-year-old children, suggesting a more underapplication rate in the pseudo- and novel words of three-year-olds.

Thirdly, an analysis was conducted comparing the application rate of children in the seven-year-old group to the younger age groups, since the seven-year-old children had reached the adult-level sandhi application in all three lexicality conditions. The results showed that there was no significant difference in children across those age groups in the real word condition. In the pseudo-word condition, the three- and four-year-olds were found to demonstrate significant differences compared with seven-year-old children (3y: estimate = -1.998, S.E. = 0.633, t = -3.153, $p < .001$; 4y: estimate = -1.491, S.E. = 0.574, z = -2.598, $p < .01$). In the novel words, only the three-year-olds were significantly lower in sandhi application than the seven-year-olds (estimate = -1.998, S.E. = 0.578, z = -3.456, $p < .001$). This analysis with seven-year-olds as the baseline also somewhat supports the argument above that there is a larger difference among the younger children in pseudo-words compared to real words.

It is also worth noting that, the individuals showed a wide range in the sandhi application rate among different age groups. For adult participants, the individual application rates were densely distributed above 90%. Individual differences became larger in younger groups. To recapitulate the results of the tone production judgment task regarding the general rate of sandhi application, children's application of tone sandhi was not adult-like in three- to five-year-olds in all three lexicality conditions. We can see from the individual data (Figure 3.3), however, that the pattern was not consistent among the subjects. As can be seen in the case of the three-year-old children, there is a large number of individual differences, where some individuals performed at a very high sandhi application rate, even in novel words, whereas some individuals performed rather poorly.

A further pattern was observed in the data, concerning the sandhi application rate for each test item (Figure 3.4). This pattern is similar to that of the individual participants, in that the adults and the older children showed more or less consistent sandhi applications, regardless of the items. Younger children, by contrast, showed a wider range in the rate of their application of sandhi among different items in pseudo- and novel words.

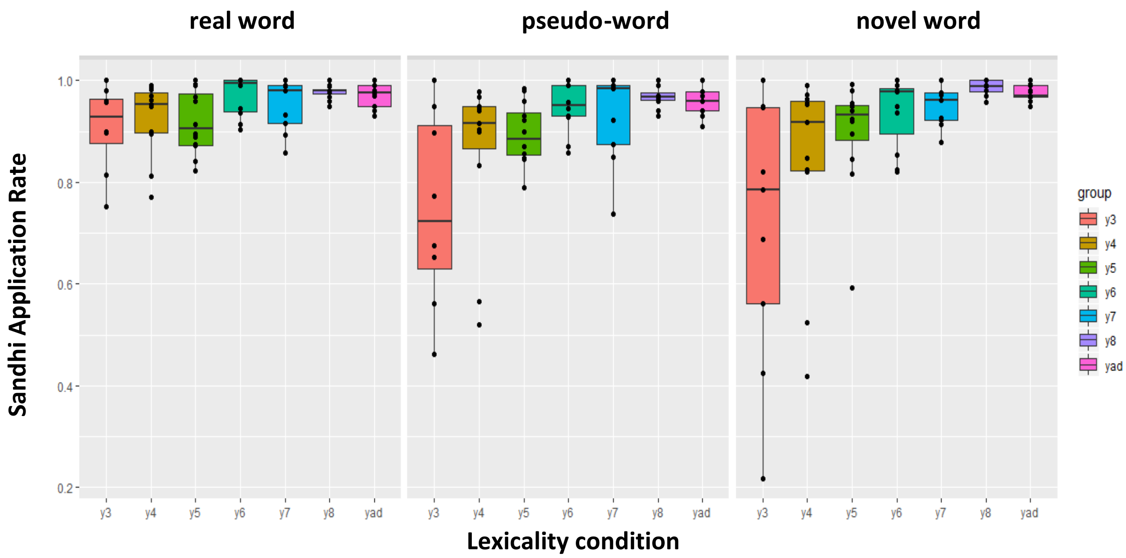


Figure 3.4 *illustrates the sandhi application rate of individual test items. The dots represent the sandhi application rate for each test item in each age group.*

To sum up, the Mandarin-speaking children reached the adult level of productivity at six years old for real and pseudo-words and reached the adult-like level of application at seven years old for novel words. Moreover, even though the results in the real and pseudo-word conditions appeared to be similar, in that the three- to five-year-old children exhibited significantly poorer performance than adults, it was evident that the underapplication rate was different to a degree. The three-year-old children showed a more obvious trend of underapplication in pseudo- and novel words.

Last, analyses were conducted to pinpoint the operational mechanism in the pseudo-word condition. As mentioned in the introduction, two possible operational mechanisms underly the correct application of Tone 3 sandhi for the pseudo-words.

First, children may take advantage of previously heard sandhi allomorphs of the initial syllables, which are listed in the mental lexicon. However, to make use of the sandhi allomorph, children would have to extract those syllables from previously heard disyllabic words that undergo tone sandhi. For example, children may make use of the sandhi allomorph /ɕi35/ 洗 "to wash" in /ɕi35 ʂʻou213/ 洗手 "to wash hands" when producing the pseudo-word /ɕi35 xai213/ 洗海 "to wash sea". The second possibility is that children may merely take advantage of the phonological pattern and apply the pattern computationally. A possible way to segregate these two explanations is to check the probability of the sandhi-initial allophones occurring in the sandhi context in the language input to children, and examine whether their probability of occurring in the sandhi context influences the sandhi application rate. The lexical mechanism (e.g., sandhi allomorphs stored in memory) predicts that the sandhi application rate in pseudo-words would be influenced by the probability of word-initial morphemes occurring in the sandhi context, whereas the computation mechanism predicts no relationship between the two. The ideal way to test the above two hypotheses would be to examine the occurrence of word-initial morphemes occurring in the sandhi context in a corpus of children's vocabulary or speech input to the children. However, there is only one toddler's checklist available (Hao et al., 2008), which targeted children before three years old. It can be seen from this checklist that children have very little vocabulary with sandhi allophones before three years of age (Hao et al., 2008). To glean some information about the operational mechanisms in pseudo-words, this study made use of an existing Mandarin adult corpus (Cai & Brysbaert, 2010), which was subtitle-based. Given that the children's primary input was from adults' speech (including possible TV influence), it may be worth taking advantage of the adult corpus.

The sandhi probability of the initial syllables occurring in the sandhi context was treated as a continuous variable. The operational mechanism was discerned by examining whether the variable *sandhi probability* significantly influenced the sandhi application rate in each age group. Data were analyzed by a generalized linear mixed-effects model (Baayen, 2008), using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included sandhi application (1 indicates application; 0 indicates underapplication) as the dependent variable, age groups (3y, 4y, 5y, 6y, 7y, 8y, and adult), and sandhi probability as independent variables, subject, and item as random factors. Random slopes were also added for each of the fixed effects. Statistical data analyses started with a full random structure. Models were compared using log-likelihood ratio tests to determine the best model. The results showed that the variable 'probability' did not significantly influence the model. This suggests that it is the computation mechanism that may be active in the application of the sandhi pattern in pseudo-words. The sandhi application results in this condition suggest that children reach adult-like productivity in the computation-based pseudo-word processing at six years old.

3.4.2. The result from the rating score

As was explained earlier, the rating score indicates how similar productions were in terms of the tone chosen: 5 represents remarkably similar and 1 extremely different. Therefore, the rating score of an instance of sandhi "application" can be interpreted as the degree to which it sounds like Tone 2; for the "underapplication" category, it can be interpreted as the degree to which it sounds like Tone 3. It would be intriguing to examine the rating scores of the sandhi sequences (Tone 3 + Tone 3 combinations) that were categorized as "underapplications" in the tone judgment task. In this condition, the higher the rating score, the more underlying Tone 3 like the item is. Figure 3.5 summarizes the rating score of these test items in the three lexicality conditions in each age group. In the figure, the distribution of rating scores within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual points of the rating score, ranging from 2 to 5, are displayed.

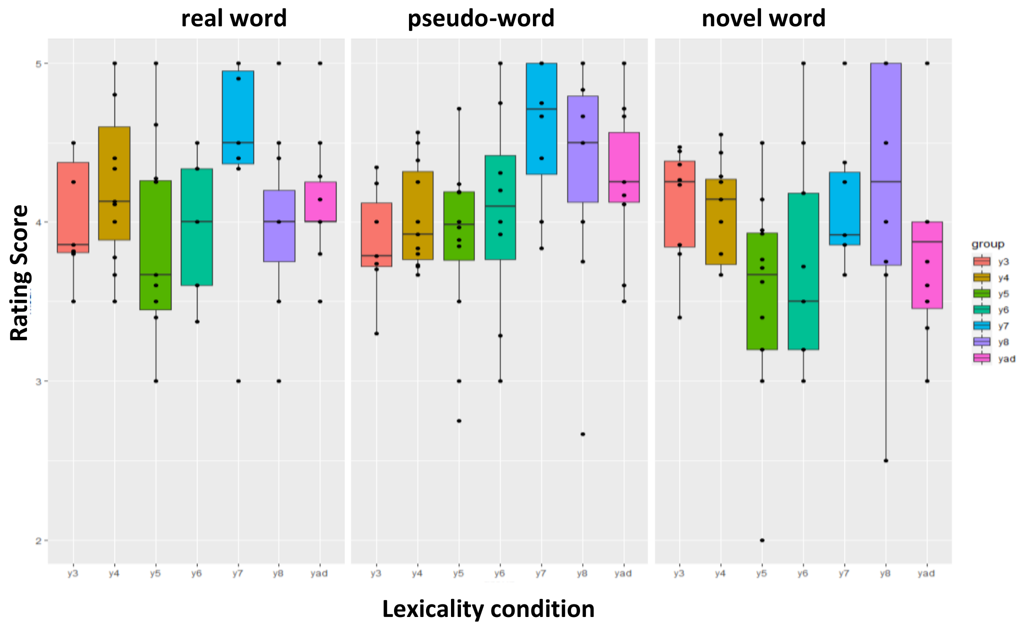


Figure 3.5 *The rating score of Tone 3 productions in the "underapplication" category (Tone 3-like) in each age group.*

Data were statistically analyzed by a linear mixed-effects model using the package lme4 (Bates et al., 2014) in R (R CoreTeam, 2014). The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included participants' rating scores for the tone sandhi items in the underapplication category as the dependent variables, and the lexicality conditions (real word, pseudo-word, novel word) and age groups (y3, y4, y5, y6, y7, y8, and adult) as independent variables. It also included individual subjects, items as random factors, and the random slopes of each independent variable. The analysis began with a full random structure. Terms were dropped off individually, and successive models were compared with likelihood ratio tests to determine if a given term significantly contributed to the fit of the model.

The results showed that the best model included interactions between the lexicality conditions and age groups, with a random intercept of individual subjects. The data were then analyzed separately in each of the three conditions. For age groups, the rating score for the adults was treated as the baseline. The analyses demonstrated that the age group variable did not significantly contribute to the model in either real word, pseudo-word, or novel word conditions, which suggests no significant difference in rating scores between children and adults.

3.5. Discussion

The study examined the productivity of Mandarin Tone 3 sandhi in children learning Mandarin Chinese as their mother tongue with adults as controls across the three lexicality conditions, i.e., real words, pseudo-words, and novel words. The results of the tone production judgment and the rating score demonstrate a developmental trajectory in gaining adult-like high productivity. Firstly, the result confirmed high productivity in adult Mandarin speakers in sandhi application across the three lexicality conditions. By comparing children's performance to that of the adults, the results suggest that children showed relatively low productivity to the three types of words from three to five years old, this is especially true for the three-year-old children when producing pseudo- and novel words. This indicates that the lexical mechanism is involved in these age groups under such conditions. Besides, children reached the adult-like application in pseudo-words at six-year-old, and in novel word condition at seven years old, indicating that children started to transit from using the lexical mechanism to applying the computational mechanism at the age of six, and had fully developed the ability to tone sandhi via the computational mechanism at the age of seven.

Tang et al. (2019) examined the productivity of Mandarin Tone 3 sandhi in the disyllabic pseudo-words (e.g., /ma213ku213/ 马鼓 "horse drum") in children aged three-to-five with adults as controls. They found that three-year-olds have already applied the sandhi pattern productively to the pseudo-words as well as adults. The discrepancies between the findings of the current study and Tang et al. (2019) may be due to the difference in the stimulus complexity. In their study, there were only four pseudo-words and they were all designed with noun-noun combinations. In contrast, the current study contained twenty tone sandhi pseudo-words that covered different types of combinations such as noun-noun-combinations, noun-verb combinations, verb-noun combinations, in addition to twenty novel words. With a larger number of items and richer grammatical structures in the stimuli, the current study provided compelling evidence that only children aged six and older have reached the adult level of productivity in pseudo-words. In fact, another study which measured the acoustic features of Mandarin Tone 3 sandhi (Rattanasone, Tang, et al., 2018), showed that children haven't fully reached the adult-like acoustic realizations at five years old even in the real words, which is partially consistent with the findings here.

It is also noteworthy that, although Mandarin-speaking children had not reached an adult level from three to five, their tone sandhi application rates in pseudo- and novel words were not particularly low. Their sandhi application rate was above 80% in real words across all the age groups. Even in novel words, the application rate was still above 80% for Mandarin-speaking children from four years old onwards. Besides, when looking at the individual features, it is evident that participants showed substantial individual differences, especially in the younger age groups. For instance, some three-year-old children exhibited rather high sandhi application rates in novel words. This may be influenced by children's vocabulary size and language inputs, etc. Moreover, previous Pinyin instructions are also expected to influence the progress of sandhi application, even though the current statistical analysis concerning the five to six-year-old children did not found a significant influence on them. Future studies should examine the possible reasons for the development of tone sandhi and the individual variations systematically to shed further light on the productivity of tone sandhi and the factors influencing it.

Chapter 4 The Acoustic Realization of Mandarin Tone 3 Sandhi in Mandarin-speaking Children

4.1. Introduction

The realization of a phonological pattern can be reflected in its phonetic basis. However, even though there existed acoustic differences between the sandhi Tone 3 and Tone 2 in novel words, these differences were not detected in adult speakers' perception (Zhang & Peng, 2013). Given the possible discrepancies between how Tone 3 sandhi is perceived and its actual acoustic features, it is valuable to observe the acoustic features of the test items based on the tone production judgment.

Previous studies have explored the productivity of Mandarin Tone 3 sandhi by observing its acoustic realization in different types of pseudo- and novel words in adult Mandarin speakers. For example, the results from Zhang and Lai (2010) showed that speakers' production of pseudo- and novel Tone 3 sandhi words had lower and later turning points than did their production of real sandhi words. Tone 3 sandhi pseudo- and novel words were produced more like an underlying Tone 3 than real words. This indicates an acoustic incomplete application of the sandhi pattern in pseudo- and novel words in adults. Zhang and Peng (2013) further compared the acoustic features of sandhi Tone 3 syllables with Tone 2 syllables in real and novel word conditions. If the tone sandhi pattern was realized in the sandhi context, the initial Tone 3 syllable was expected to be Tone 2-like. The results showed that sandhi Tone 3 exhibited significantly lower F0 than Tone 2 syllables in novel words, but no significant difference was found in real words. This further confirmed the acoustic incomplete application of the sandhi pattern in adults' production of novel words. It is essential to differentiate the two terms "underapplication" and "incomplete application" here. The incomplete application of the pitch trajectory cannot be detected perceptually in adult speakers, while the "underapplication" has its perceptual basis.

Despite accumulating interest in the acoustic realization of Mandarin Tone 3 sandhi in pseudo- and novel words in adult speakers (Zhang and Lai, 2010; Zhang and Peng, 2013), no consensus has been reached on the acoustic features of the application of Mandarin Tone 3 sandhi in novel combinations over the course of children's development. Tang et al. (2019) examined three- to five-year-old children's productions of disyllabic Tone 3 sandhi pseudo-words with adults as controls, and acoustic analysis showed that the children from three to five years of age applied the pattern productively to the pseudo-words, just as the adults did. However, this requires further verification because previous literature has found that children's production of Tone 3 sandhi real words was still not adult-like from three to five years of age (Rattanasone et al., 2018). According to the acoustic analysis of children's production of Tone 3 sandhi real words (Rattanasone et al., 2018), children in these age groups were found to produce the sandhi words with a more flattened rising curve than adults.

To gain better insight into children's production of Tone 3 sandhi, this study examined the acoustic characteristics of the production of real, pseudo- and novel sandhi words in both three- to eight-year-old children and adults. This study further divided children's performances into three categories: application, underapplication, and mispronunciation. By exploring the acoustic realizations of items that were identified as "application", I can gain a better understanding of the acoustic realization features for the items that were perceived as sandhi Tone 3. According to the result of the speech production judgment, the three to five years old children showed a significant trend of underapplication in pseudo- and novel words. It is highly possible that trends of incomplete application may still exist in children especially from three to five years old when calculating the production of sandhi Tone 3 syllables in the "application" category. While it is also possible that the acoustic differences no longer exist for those who have been judged as applying the sandhi pattern. The acoustic features were analyzed in terms of pitch trajectory to test the predictions.

4.2. Method

The subjects and stimuli were identical to those of the experiment in Chapter 3. Specifically, 74 children, whose ages ranged from three to eight years old, and 13 adults were recruited for the study. All children were monolingual speakers of Mandarin Chinese and had not been exposed to any other languages or other dialects. All children passed the hearing screening at 1000 Hz, 2000 Hz, and 4000 Hz at 30 dB on both left and right ears using headphones (Interacoustic AS608 Screening Audiometer). IQ scores were comparable among the different age groups (except for the 3-year-old children). Informed written consent was obtained from all adult subjects and children's caregivers, in compliance with the experimental protocols.

To examine whether the Mandarin Tone 3 sandhi pattern was applied productively, real words, pseudo-words, and novel words were designed for the experiment. Each type of word included 20 test items (Tone 3 + Tone 3 words) and 20 control items (Tone 3 + non-Tone 3 words). The study adopted a picture-naming task that elicited spontaneous productions of the three types of words.

4.3. Data analysis

As mentioned in the previous chapter, participants' productions were labeled and extracted from the recordings in Praat (Boersma & Weenink, 2016) for both tone production judgment and acoustic analysis. Before the data analysis, the experimenter scrutinized the data and preserved 8385 tokens out of the original 10440 (87 subjects * 120 items) for analysis. Among the preserved tokens, 4505 were test items. The results of the acoustic analysis of the tone were based on the tone production judgment.

Specifically, data were further classified into three subgroups according to the judgments of five phoneticians: application of tone sandhi (Tone 2-like), underapplication (Tone 3-like), and misproduction. Precisely, a production was classified into the "application" subgroup if it was identified as exhibiting applied tone sandhi by three or more phoneticians. It was categorized into the "underapplication" group if identified as underapplied by three or more of the five phoneticians. Similarly, a production was categorized into the "mispronunciation" group if it was identified as a mispronunciation by three or more phoneticians. The number of test items in each condition is shown in Table 4.1.

Table 4.1 *The number of preserved test items in each lexicality condition*

	adult	8y	7y	6y	5y	4y	3y
real word	260	239	200	196	260	200	134
pseudo-word	260	253	200	206	260	223	133
novel word	258	246	190	200	256	197	134

The duration and F0 of the rhyme of each initial syllable in a Tone 3 disyllabic sequence (e.g., /iɛ213/ in /miɛ213/) were measured with Praat (Boersma & Weenink, 2016) from the trials within each speaker. F0 was measured at 20 sampling points evenly spaced across the rhyme duration (0%, 5%, 10%, ... 95%, and 100%). The extracted F0 values were then log-transformed from Hz to semitones. The log-transformed F0 values were then z-score normalized against the mean pitch across all tokens for each speaker to minimize inter-speaker variations in pitch (e.g., Zhu, 2005). The formula for normalizing the F0 values of each token is illustrated below.

Z-score Normalized Pitch = [Observed Pitch - Mean Pitch (Individual)] / Standard Deviation of Pitch (Individual)

4.4. Results

4.4.1. Acoustic realization of Mandarin Tone 3 sandhi items that were classified as "application"

Figure 4.1 depicts the pitch trajectories for the initial syllables of the test items that were classified as "application". In the Figure, mean and S.E. of the pitch values of test items identified as sandhi "application" were presented.

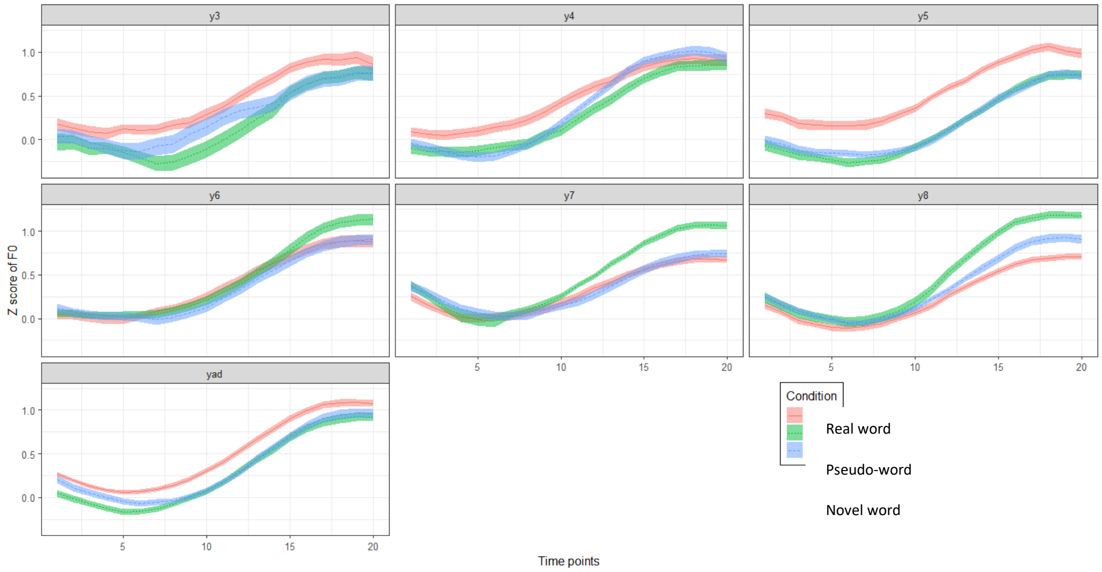


Figure 4.1 *Mean and S.E. of the pitch values of test items identified as sandhi "application"*

Growth curve analysis (Mirman, 2014) was used to characterize the pitch trajectories of the three types of words (real words, pseudo-words, and novel words) that were classified as "application" over the course of the 21 time points. The overall learning curves were modeled with second-order orthogonal polynomials, following previous studies on Mandarin tones or tone sandhi (Li & Chen, 2016; Rattanasone et al., 2018; Tang, 2019), with fixed effects of lexicality conditions and age groups on all time terms. The model also included random effects of participants and test items on all time terms. The fixed effects of lexicality condition and age group were added step by step to the pitch points and their effects on model fit were evaluated using model comparisons. Improvements in model fit were evaluated using -2 times the change in log-likelihood, which is distributed as χ^2 with degrees of freedom equal to the number of parameters added. Parameter-specific *p*-values were estimated using the normal approximation. All analyses were carried out in R (R Core Team, 2014) using the lme4 package (Bates et al., 2014).

The analysis generated three parameters for each pitch contour, i.e., (1) the intercept, (2) the linear trend, (3) the quadratic trend. According to Mirman (2014), the three parameters capture, respectively, the overall F0 mean (the higher the intercept, the higher the average F0), pitch slope (as reflected in the linear trend: a positive value indicates a rising pitch and a negative value a falling pitch; a larger estimate indicates a steeper slope); and curvature (a positive quadratic trend indicates a concave F0 contour, a negative quadratic trend a convex contour, and a larger quadratic trend a more curved F0 contour). These parameters were used to evaluate any differences in the overall F0 contour among the groups and lexicality conditions.

Table 4.2 *Results of the linear mixed regression model with second-order polynomials on pitch points of the initial tone sandhi syllable that was classified as "application" in disyllabic sequences across age group and lexicality condition.*

Pitch parameter	Factor	$\chi 2$	Chi Df	*p*
Intercept	Condition	1.55	2	<.01

	Group	0.0000	2	1
	Group × Condition	**2252.3650**	**2**	**< .001**
	Condition	**10.0031**	**2**	**<.01**
linear	Group	0.000	6	1.00
	Group × Condition	**234.6197**	**20**	**< .001**
	Condition	**15.1980**	**0**	**< .001**
Quadratic	Group	6.3686	6	.383
	Group × Condition	**37.4582**	**20**	**<.05**

As summarized in Table 4.2, there was a significant main effect of lexicality and interactions between group and lexicality on all the three observed trends. Post-hoc analyses were carried out with second-order orthogonal polynomials to build separate linear mixed-effects models for each age group, analyzing the effect of lexicality condition on the F0 trajectories. For the lexicality condition, the real word condition was treated as the baseline. The parameter estimates for analyzing the effect of lexicality on pitch trajectory are summarized in Table 4.3 where only significant findings are listed. No acoustic comparison between children and adults was made due to the potential distinctions arising from age difference.

Table 4.3. *Parameter estimates for analysis of the effect of lexicality on pitch trajectory*

Age group		Lexicality condition	Estimate	S.E.	df	*p*
adult	Intercept	Pseudo-real	-.21	.05	53.16	<.001
		Novel-real	-.17	.05	54.33	<.001
	Quadratic	Novel-real	.14	.06	110.95	<.05
8y	Intercept	Pseudo-real	.24	.07	51.61	<.05
		Novel-real	.16	.07	52.12	<.001
	Linear	Pseudo-real	.83	.14	52.12	<.001
		Novel-real	.37	.14	51.73	<.001
	Quadratic	Pseudo-real	.25	.07	60.47	<.001
		Novel-real	.18	.07	62.50	<.05
7y	Intercept	Pseudo-real	.26	.10	57.30	<.05
	Quadratic	Pseudo-real	.23	.11	57.54	<.05
		Novel-real	.24	.11	61.79	<.05
6y	Linear	Pseudo-real	.30	.15	55.82	<.05

	Quadratic	Pseudo-real	.20	.09	84.11	<.05
5y	Intercept	Pseudo-real	-.40	.06	63.97	<.001
		Novel-real	-.39	.06	63.35	<.001
4y	Intercept	Pseudo-real	-.17	.06	57.43	<.001
	Linear	Novel-real	0.46	.19	57.54	<.05
	Quadratic	Pseudo-real	.23	.12	59.59	<.05
		Novel-real	.23	.12	62.94	<.05
3y	Intercept	Pseudo-real	-.25	.08	72.10	<.001
		Novel-real	-.25	.08	68.51	<.001
	Quadratic	Pseudo-real	.30	.11	161.71	<.05

In the adult group, the results showed that both pseudo-words and novel words exhibited a significantly lower mean F0 than the real words. The results also found a significantly larger quadratic trend in the novel words than the real words, indicating a more concave F0 contour. Compared to the adult results, children's pitch trajectories showed different trends that developed from underapplication in the three-year-old group to overapplication or over articulation in the six- and eight-year-old groups.

More specifically, in the three-year-old group, a significant effect of lexicality on the intercept was found in both the pseudo-words and the novel words, which suggested that the pseudo- and novel words exhibited a significantly lower mean F0 than the real words. The pseudo-words also exhibited a significantly larger quadratic trend than real words did, indicating a more concave F0 contour in this type of words. In the four-year-old group, the results showed significantly lower mean F0 and more concave F0 contours in the pseudo-words than in the real words, in addition to a significantly more rising trend and a more concave F0 contour in novel words. In the five-year-old group, the results showed significant effects of lexicality on the linear trend, with a lower mean F0 in both pseudo- and novel words compared to the real words.

For the six-year-old children, the results exhibited significant effects of lexicality on both the linear and the quadratic trends in pseudo-words, indicating more rising F0 trends and more curved F0 contours. In the seven-year-old group, the results showed a significantly higher mean F0 and a significantly larger quadratic trend in pseudo-words than did the real words. There was also a significantly larger quadratic trend in novel words than in the real words in seven-year-olds, indicating a more concave F0 contour. In the eight-year-old group, the effect of lexicality was significant on all three observed trends in both pseudo-and novel words, which showed a significantly higher mean F0 than the real words did, as well as more rising pitch contours and more concave F0 contours compared to those of the real words.

Two primary scenarios can be summarized from the results above. 1) a significant negative intercept alone or a significant interaction between the intercept term (negative) and quadratic trend (positive) would indicate a trend of incomplete application (i.e., the sandhi Tone 3 produced with a significantly lower mean F0 or more low-dipping/concaved F0 contour); 2) a significant positive linear trend alone or a positive linear trend plus a significant positive quadratic trend would provide evidence of a more steeply rising tone contour, that is, an over-application or over articulation of the sandhi Tone 3. Based on the principles, the three and five age groups showed significant trends of underapplication in both pseudo-words and novel words than in real words. For the children aged four, the results revealed a trend of underapplication in pseudo-words than in real words, similar to three- and five-year-olds; however, the interpretation of the trend in novel words for the four-year-old children is uncertain. Although the results showed a more rising F0 and a more concave F0 contour, which might suggest overapplication, a closer examination of the F0 trajectories in Figure 4.1 indicated that the significant rising trend in the linear term may be due to the low F0 value in F0 onset. Therefore, it may be better-suited to classify the patterns as underapplication. Intriguingly, the six- and seven-year-olds exhibited a significant trend of overapplication in pseudo-words. And the eight-year-olds exhibited trends of overapplication in both pseudo- and novel words than did the real words. Last, the results of adults revealed a significant trend of incomplete application in both pseudo- and novel words than real words.

4.4.2. Underapplication and mispronunciation

In addition to the analyses of the pitch trajectories that were classified as "application", the pitch trajectories of test items that were identified as "underapplication" in different lexicality conditions were also depicted (see Figure 4.2). For the tokens categorized as "underapplication", the pitch trajectories of real and pseudo-words generally showed flatter pitch contours, with larger standard errors. In the novel words, the pitch trajectories generally showed later turning points.

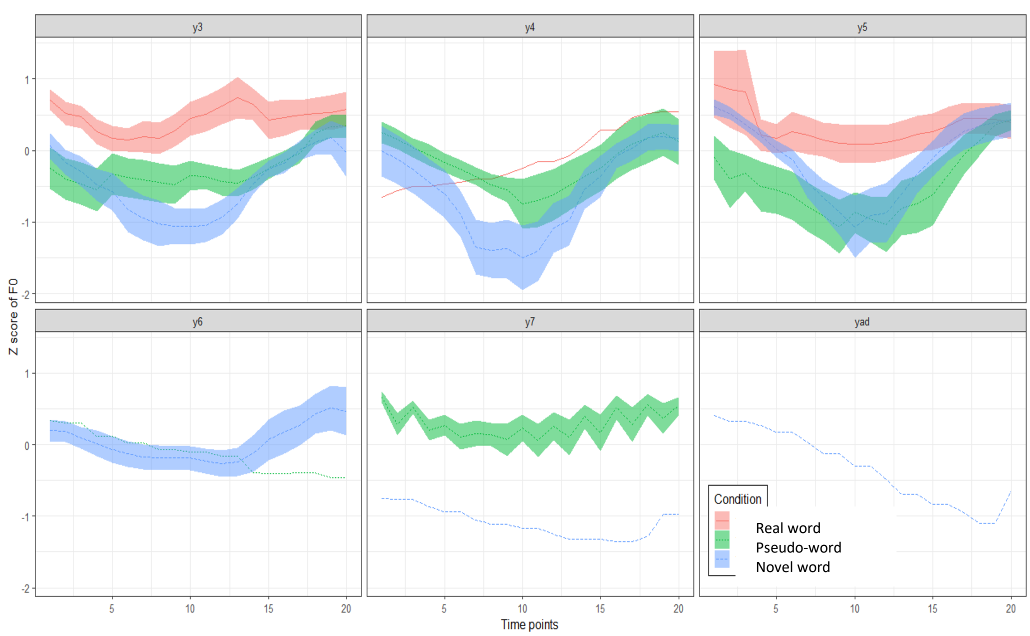


Figure 4.2 *Mean and S.E. of the pitch values in the "underapplication" category*

The study also calculated the percentage of mispronunciations in each age group, as well as in each lexicality condition. Detailed information is given in Table 4.4. As can be seen from the table, the three-year-old children exhibited a relatively larger percentage of inaccurate productions in pseudo- and novel words. The rate of mispronunciations for the other age groups was not large in general.

Table 4.4 *The rate of mispronunciations of test items in each age group*

	3y	4y	5y	6y	7y	8y	adult
real word	1.13%	0.97%	0.29%	0.65%	1.12%	0.25%	0.00%
pseudo-word	5.80%	0.42%	0.00%	0.53%	0.22%	0.16%	0.00%
novel word	3.34%	1.22%	0.28%	0.42%	1.04%	0.49%	0.23%
total	3.47%	0.86%	0.19%	0.53%	0.78%	0.30%	0.08%

4.5. Discussion

This chapter has examined the acoustic realization of children's productions of Tone 3 sandhi real words, pseudo-words, and novel words in Mandarin-speaking children and adults. As mentioned before, real words are expected to be listed in children's mental lexicon and can, therefore, be memorized via the lexical mechanism. Pseudo-words have been shown to operate primarily by the computation mechanism. For novel words, children are expected to lean solely on the phonological pattern for computational processing. To understand the acoustic features of the sandhi items in different age groups as well as in different lexicality conditions, pitch trajectories were analyzed. The results showed that the three to five age groups showed significant trends of incomplete application in both pseudo- and novel words that were classified as "applied" tone sandhi. This may further suggest children's low productivity in pseudo- and novel words in these age groups. Contrastively, children aged six-to-eight showed significant trends of overapplication in pseudo-words compared to real words; the eight-year-olds also exhibited a significant trend of overapplication in novel-words. The overapplication trends in the results may be treated as an indicator of children's awareness of the sandhi pattern and their attempts to apply the sandhi pattern, exhibiting a trend of overapplication or over articulation.

The results of adults revealed a significant trend of incomplete application in both pseudos- and novel words than real words. This appears to be consistent with the previous experiments of Zhang and Lai (2010). The result could be because the pseudo- and novel words were presented syllable by syllable in their underlying forms in the experiment, while the real words were elicited via a word-guessing game in which no underlying tones were shown. The sandhi tones in pseudo- and novel words may, therefore, have been more likely to bear more resemblance to the underlying tone than to the sandhi tone. Zhang and Lai (2010) have also linked the incomplete application of tone sandhi in adults to a weaker phonetic motivation for this type of sandhi pattern in nature.

The integrated results of the tone production judgment and the acoustic analysis offered us a comprehensive and broader scale picture of children's productivity of tone sandhi, showing the developmental trajectory of Mandarin-speaking children toward gaining adult-like productivity: three- to five-year-old children were not able to apply the sandhi pattern as well as adults, which implies that children of this age cohort primarily take advantage of mental listings in the lexicon and rely on the lexical mechanism to produce Tone 3 sandhi words. After a year of development and transition at the age of six, the seven-and-eight-year-olds have fully acquired the adult-like computation mechanism. Besides, as the sandhi pattern and the computation mechanism were gradually acquired, children tended to apply it actively, sometimes possibly showing certain trends of over articulations in speech production.

Chapter 5 The Representation of Mandarin Tone 3 sandhi in the Mental Lexicon of Mandarin-speaking Children

5.1. Introduction

Acquiring the sound system of one's native language is a rather complicated process. During this process, a child may take advantage of not only a well developed articulatory control but also the capacity to process and represent the components of the language in the mental lexicon. Till now, children's sensitivity to segments and tones has gained great attention (Ma et al., 2017; Singh & Fu, 2016; Tong et al., 2015; Shen et al., 2013). However, there are some more complicated linguistic phenomena, such as tonal alternations, that have been less investigated in children. This chapter focuses on one such tonal alternation: Mandarin Tone 3 sandhi. It is fascinating to learn how Mandarin Tone 3 sandhi are represented in the mental lexicon and how children's representations of tone sandhi evolve into adult representations during the course of language development.

As summarized in the literature review, there are three well-attested accounts of how tone sandhi words are represented in adults' mental lexicon: the surface representation view, the underlying representation view, and the multi-variant account

(Chen et al., 2011; Li & Chen, 2015; Chien et al., 2016, 2017; Zhang, 2018). Converging evidence (Chien et al., 2016, 2017; Zhang, 2016; Zhang, 2018) has demonstrated that the Mandarin Tone 3 sandhi, which involves phonological computation mechanisms, were retrieved with their underlying tone in the spoken language comprehension of adult speakers (supporting the underlying representation view). While in speech production, both the underlying tone and its Tone 2-like tonal variants are activated (Chen et al., 2011; Li & Chen, 2015). It is worth noting that both the lexical and computation mechanisms may be involved in this condition where the storage of instantiated surface-form representation relating to the lexical mechanism and the abstract underlying-form representation connecting to the computation mechanism. According to the psycholinguistic models of Levelt et al. (Indefrey, 2011; Levelt, 1999; Levelt et al., 1999), lexical and phonological processing occurs before the initiation of speech production. Therefore, the overall picture for the representation of Mandarin Tone 3 sandhi in adult speakers should be: the underlying representation or different forms of exemplars may be activated when hearing the Tone 3 sandhi word, while both the allophonic representation and the abstract underlying forms are activated when producing the word. Moreover, previous literature (Chien et al., 2016, 2017; Zhang, 2016; Zhang, 2018) has shown that adult speakers treat sandhi patterns with low productivity differently in their mental lexicons compared with the patterns with high productivity. Unproductive types of sandhi, which involve lexical memorization of allomorph listings of existing morphemes or syllables, were more likely to be listed in the surface tone (supporting the surface representation view). According to the results of Chapters 3 and 4 concerning the productivity of Mandarin Tone 3 sandhi in the speech production of Mandarin-speaking children, the productivity of Tone 3 sandhi undergoes a gradual transformation from relatively low productivity in pseudo- and novel words in three- to five-year-old children to more and more adult-like productivity—even some exaggerations—by six years old. This suggests that, at first, children primarily rely on the lexical mechanism to process Tone 3 sandhi words, while they take more and more advantage of the computation mechanism from the age of six. Children may make full use of the computation mechanism at seven years old, because, in novel words, where participants could solely rely on the computation mechanism productively, seven-year-olds demonstrated adult-like productivity. Based on results from adult speakers, which found an association between productivity and features of representation (Chien et al., 2016, 2017; Zhang, 2016; Zhang, 2018), it is likely that children also undergo a gradual developmental trajectory to reaching an adult-like underlying representation of Mandarin Tone 3 sandhi. Between three to five years of age, they may primarily take advantage of surface forms when making lexical responses to tone sandhi items. By seven years of age, they are expected to approach the adult mental representation, which primarily makes use of the underlying forms according to the underlying representation view.

As one previous study of children's representation of tone sandhi words in the mental lexicon (Wewalaarachchi & Singh, 2016) has found, when children were presented with surface-form Tone 3 sandhi words (e.g., /iy35 san213/) and underlying-form sandhi words (e.g., /iy213 san213/) between three and five years old, they only recognized the surface-form. This may suggest that children who speak Mandarin Chinese as their mother tongue still process tone sandhi words in their surface form (facilitated by the lexical mechanism) up to age five. However, since the underlying-form words never occur in daily life, and there are no data to support that such forms were acceptable even in adults, we can hardly attribute this to the underdevelopment of underlying representations in children. Therefore, in order to find how and when children reach the adult underlying representations of Mandarin Tone 3 sandhi, the current study observed children from three to eight years old with adults as controls as they were found to be presented in their underlying Tone 3 forms (Chien et al., 2016; Zhang, 2018).

5.2. Method

The study used an auditory-auditory priming lexical decision task, following the paradigm of Chien et al. (2016). This is an efficient way to examine the features of the mental representation of morphological or phonological alternations. In the task, participants judge whether an experimental item they hear is a word or not. The latency with which the participants respond "yes" to this question is influenced by the exposure to an earlier item (prime). If a prime word shares some overlapped modalities, such as form overlap, with the target (teacup-cup), the reaction latency of the target is expected to be shorter. In previous studies using this method, priming effects were found either (a) when the primes were the same as the targets, both in terms of segments and tones, or (b) when the primes were productive phonological alternations of the targets (such as how gree**n**back primes gree**n** in English; e.g., Sereno & Lee, 2015; Gow, 2001; Lahiri et al., 1990). Using this method, the experiment manipulated three types of primes: (i) primes that corresponded to sandhi-initial syllable on only the segmental level (**control**), (ii) primes that overlapped with the sandhi-initial syllable on both the segmental level and the tonal level in its underlying representation (**Tone 3 prime**), and overlapped its initial syllable on the segmental and the tonal level in its surface form (**Tone 2 prime**). If sandhi items are represented in their underlying forms, a Tone 3 priming effect will occur, but if the sandhi items are represented in their surface forms, a Tone 2 facilitatory effect is expected. If the sandhi items are represented in both Tone 2 and Tone 3, the representation of sandhi words may be explained by the multi-variant account.

5.2.1 Participants

Children aged between three to eight years old were recruited for the experiment, although only the age groups from four to eight were preserved for analysis. The reason for excluding the participants in the three-year-old group was that, in this experiment, reaction time, which requires a relatively long attention span, served as a crucial measurement. The three-year-olds, however, were found to be unable to focus their attention on the experiment, as their very low response accuracy rates in the practice trial. They were therefore excluded from further data analysis. In total, 90 children aged from four to eight were retained for the experiment out of 107 children, according to the following criteria. All children were monolingual Mandarin speakers who had seldom been exposed to other languages, like English or other dialects. All children were instructed to accomplish the revised versions of Wechsler's test for intelligence tests (WISC-R; Zhang & Lin, 1979), which included two versions, one of which was for preschoolers below six and a half years old and the other one for children above six and a half years old. None of them were reported to have any hearing impairments. Among the children, 18 were aged four (mean: 4;5), 33 were aged five (mean: 5;3), 28 were six-year-olds (mean: 6;8), and 25 were seven or eight years old (mean: 8;1). Since the results of the speech production task showed that children reached an adult-like application of tone sandhi at seven and eight years old, this experiment combined the two age groups into one for analysis. In addition to the children, we recruited seventeen adult speakers as controls. Their ages ranged from 18 to 32 years old. Detailed information for the participants is listed below in Table 5.1. Experimental procedures were approved by the Human Subjects Ethics Sub-committee of The Hong Kong Polytechnic University. Informed written consent was obtained from all adult subjects and children's caregivers in compliance with the experimental protocols.

Table 5.1 *The number of subjects preserved for analysis*

age group	subject number	item number	target item number	correct target number
4y	15	2070	540	355
5y	25	3405	886	708
6y	25	3308	863	719
78y	25	3447	900	819
Adult	18	2462	648	636

5.2.2. Stimuli

Fifteen Tone 3 + Tone 3 words (e.g., /tsuo35 tɕiɑu213/ 左脚 "left foot") were included as critical targets. They were selected from 22 sandhi words according to a subjective familiarity rating score (ranging from 1, representing "never heard or said," to 9, representing "very often heard and said") given by 12 Mandarin-speaking adults. To minimize the priming effect from previously heard words, these 12 participants did not take part in the lexical decision task. The average subjective familiarity rating was 8.02 (SD = 0.63). To make sure that these words were also familiar to children, especially the youngest ones, we showed the words to children one by one and asked the children to describe the meaning by explaining, giving examples, or performing, etc. If children were unacquainted with any of the items, the experimenter would explain them to the children.

As delineated earlier in the thesis, in order to examine the effect of different prime words on the recognition of Tone 3 sandhi words, each test item (e.g., /tsuo35 tɕiɑu213/ 左脚 "left foot") was preceded by one of three corresponding monosyllabic primes: a Tone 3 prime (e.g., tsuo213 左 "left"), a Tone 2 prime (e.g., /tsuo35 t'iæn55/ 昨 "yesterday"), and a control prime (either Tone 1 or Tone 4, e.g., /tsuo51/ 做 "to make"). Each target's first syllable overlapped with its corresponding Tone 3 prime both on the segmental and the tonal level in its *underlying representation*. Each target's first syllable overlapped its Tone 2 prime on the segmental and the tonal level in its *surface form*. For the control target condition, the prime matched the target's initial syllable only in terms of segmental phonology. Meanwhile, the sandhi targets and the primes also matched in terms of their familiarity rating scores. Detailed information is shown in Table 5.2.

Table 5.2 *Target words, their corresponding primes, and their familiarity rating scores*

No.	Control prime	familiarity	Tone 2 prime	familiarity	Tone 3 prime	familiarity	target words	familiarity
1	捞	7.18	牢	6.73	老	7.91	老虎	6.55
2	样	7.91	阳	7.64	养	7	养鸟	7.09
3	力	8.18	梨	8.09	礼	6.91	礼品	7.18
4	屋	8.45	无	7.64	舞	7.82	舞蹈	7.82
5	兽	6.36	熟	8.27	手	8.64	手表	7.91
6	做	8.75	昨	8.18	左	8.09	左手	8.09
7	妹	8.82	没	8.27	美	8.55	美女	8.18
8	智	7.36	直	7.82	纸	8.64	纸笔	8.27
9	玉	7.45	鱼	8.55	雨	8.36	雨伞	8.36
10	做	9	昨	8.73	左	8.73	左脚	8.36
11	细	8	媳	7.64	洗	7.45	洗脸	8.45
12	叶	7.55	爷	8.73	也	7.91	也许	8.45

13	睡	8.91	谁	8.27	水	8.64	水果	8.55
14	细	8.09	媳	7.73	洗	8.91	洗脚	8.64
15	那	9	拿	8.64	哪	8.91	哪里	8.73
mean		8.07		8.06		8.16		8.04
sd		0.78		0.54		0.65		0.63

In addition, fifteen Tone 2 + Tone 3 words (e.g., tsuo35 t‘ iæn55 昨天 “yesterday”) were included as controls. They matched the sandhi target words in familiarity according to the 12-person rating score and in the segmental phonology of the target’s initial syllable. An equal number of pseudo-words were designed to match the real words. The pseudo-words included 15 pairs of Tone 3 + Tone 3 sequences and 15 pairs of Tone 2 + Tone 3 tone sequences (e.g., tsuo213 ɕy213 左许 vs. tsuo35 p‘u51 昨不). Both the initial and final syllables of the pseudo-words are actually occurring syllables in Mandarin Chinese phonology, but their combination yielded no meaning. The same primes preceded them as preceded the real words.

5.2.3. Stimulus recording

A female native Mandarin speaker was instructed to read the designated words four times. The recordings were accomplished in a sound-proof chamber at the Hong

Kong Polytechnic University with a cardioid microphone (Electrovoice, model N/D767a) using Praat (Boersma & Weenink, 2016). The sampling rate was 22050 Hz. We selected the most clearly pronounced tokens for generating the stimuli. The selected stimuli were further normalized in terms of intensity and duration using Praat. To minimize the influence of other factors, such as pitch length, on the lexical decision, all stimuli were normalized. The monosyllabic prime words were normalized to 400 ms, and the disyllables were normalized to 550 ms. The intensity was normalized to 75 dB for all stimuli.

5.2.4. Procedure

The auditory-auditory priming lexical decision task was presented with E-prime. The participants were instructed to sit in front of the computer screen at a distance of approximately 60 cm. To avoid fatigue and to keep the children rigorously engaged, the experiment was divided into two blocks with a rest between the two parts. Moreover, a child would be given gifts every time he/she actively engaged himself/herself in the experiment for one block. The stimuli were presented randomly to the participants through headphones.

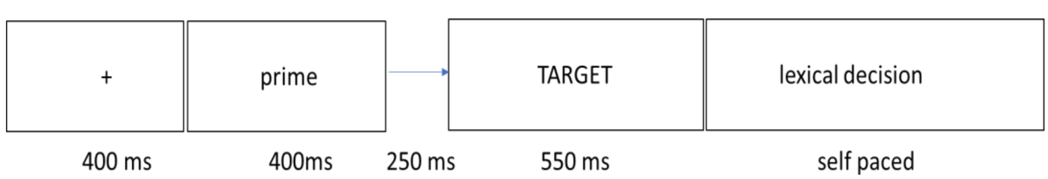


Figure 5.1 *Event order for each trial*

In each trial, participants heard a monosyllabic prime first, followed by a disyllabic target after a 250 ms interval (see Figure 5.1). The target was either a word or a pseudo-word. All targets were presented in their surface form. Children were instructed to listen to each target and then decide whether or not it was a real word. They were instructed to press " √ " if they gave the "yes" answer, or to press "×" representing "no", as quickly and accurately as possible. After hearing the instructions, children participated in twelve practice trials to familiarize themselves with the task. One hundred eighty test trials followed the practice phase. They were presented in a randomized order within the two blocks. The inter-trial interval was set to 400 ms.

5.3. Data analysis

Before conducting analyses, the data were processed at different levels to ensure that they were clean. First, at the individual subject level, good performance on the practice items across the trial indicated that the children quickly understood the lexical decision task procedures. In contrast, if a child's accuracy was still low (below 60%) when conducting the practice block even for the second time, he/she would not proceed to the experiment. A child's data would also be removed for analysis if he/she showed consecutive bias, i.e. he/she consistently made " √ " choices or "×" choices for more than 30 trials in at least one block. Secondly, at the item level, the accuracy of the children for each disyllabic item was screened. If an item was identified with low accuracy across children, it was removed from the analysis (Haebig & Weismer, 2015). As a result, eight real words and eight pseudo-words were deleted from the analysis (the lexical decision accuracy of deleted items was 2 standard errors below the mean in both real and pseudo-words). Three of the deleted items were tone sandhi target words. Thirdly, at the trial level, trials were excluded where reaction times were beyond 2 standard deviations from the mean. To be precise, mean reaction times across different lexicality conditions were calculated for each participant, and any trials with lexical decisions that were 2 standard deviations above or below that participant's mean were removed. This led to the removal of another 1.69% of the data for the four-year-olds, 3.25% of the five-year-olds' data, 3.34% of the six-year-olds' data, 2.44% of the seven & eight-year-olds' data, and 2.33% of the adults' data. The statistical analyses of reaction times only included trials where participants correctly identified the target as a real word. Data from 90 children and 18 adults were included in the statistical analyses. Detailed information for the subjects and items is listed in Table 5.3.

Table 5.3. *Item and accuracy information within the age groups*

age group	item	target	correct target	acc	acc of target	Noutlier	delete rate

4y	2070	540	355	65%	66%	6	1.69%
5y	3405	886	708	69%	80%	23	3.25%
6y	3308	863	719	79%	83%	24	3.34%
78y	3447	900	819	89%	91%	20	2.44%
adult	2331	612	600	96%	98%	14	2.33%

5.4. Results

The analysis focused on the reaction times for lexical decision of the Tone 3 sandhi targets in the Mandarin-speaking children and adults on the presentation of three types of primes. Participants' reaction times were log-transformed prior to further analysis. The reason why raw reaction time data being log-transformed is to eliminate or reduce skewing in the data distribution. This elimination or reduction is necessary because just a few extreme outliers might dominate the result, obscuring the main trends characterizing the majority of data points without the logarithmic transformation (Baayen, 2008). Boxplots, which displays the distribution of log-transformed reaction times (hereafter, log-RT) within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual points of the log-RT are displayed in Figure 5.2. The reaction times to the Tone 3 sandhi targets are given different colors according to the type of prime: the red column represents the control primes, the green column represents the Tone 3 primes, and the indigo column indicates the Tone 2 primes. Mean values, standard deviations, and standard errors of the log-RT within each priming condition and age group are summarized in Table 5.4.

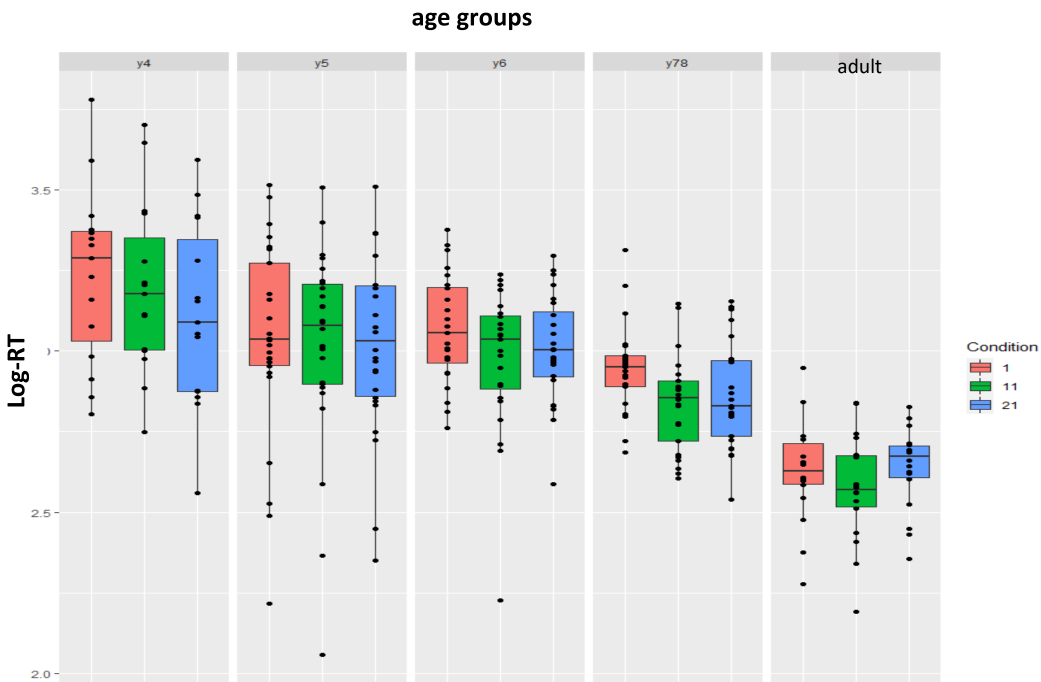


Figure 5.2. *The effect of prime conditions of Tone 3 (11), Tone 2 (21), Control (1) on the lexical decision of Tone 3 sandhi targets. The distribution of log-transformed reaction times (hereafter, log-RT) within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual points of the log-RT.*

Table 5.4 *Means, standard deviations, and standard errors of the Log-RT in each lexicality condition*

Age group	Condition	Mean Log-RT	SD	SE
y4	control	3.224470	0.3486084	0.03236748

	T3 prime	3.169419	0.3645723	0.03460367
	T2 prime	3.107782	0.3914852	0.03544341
y5	control	3.067034	0.3506935	0.02380663
	T3 prime	3.060377	0.3428396	0.02241213
	T2 prime	3.016819	0.3934178	0.02571852
y6	control	3.069305	0.2470604	0.01677155
	T3 prime	2.987413	0.3031372	0.01940637
	T2 prime	3.013685	0.2561737	0.01674660
y78	control	2.937989	0.2293090	0.01395530
	T3 prime	2.845580	0.2637497	0.01611108
	T2 prime	2.866846	0.2669745	0.01652531
adult	control	2.625621	0.2814493	0.02097799
	T3 prime	2.569989	0.3224568	0.02416916
	T2 prime	2.636404	0.2633254	0.02007838

The log-RTs were statistically analyzed by linear mixed-effects analysis using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The R package lmerTest was used to provide *p*-values for models (Kuznetsova et al., 2016). The model included participants' log-RT to the tone sandhi items as the dependent variable, and the prime conditions (Tone 2 prime, Tone 3 prime, Control), age groups (y4, y5, y6, y7&8, adult), and trial order (1st, 2nd, 3rd) as independent variables. The study also included subjects and items as random factors. To explain the trial order variable here, it represents the sequential order of the target that occurred in the experiment. Namely, each target item occurred three times, once with a control, once with a Tone 2 prime, and once with a Tone 3 prime. When the target appeared the first time in the experiment, its response would be labeled with trial order = 1^{st}, and likewise for the second and third appearances, which were labeled trial order = 2^{nd} and trial order = 3^{rd}, respectively. I adopted backward algorithms for model comparison. To be specific, I began the analysis with a full random structure including all the independent variables, and the random intercepts of subjects and target items, as well as the slopes of them for all the three fixed factors. I then eliminated non-significant predictors using model comparisons. When a model failed to reach convergence, I solved the problem by removing the minimum correlation parameters (Barr et al., 2013). The results showed that the best model included interactions between prime conditions, age groups, and trial order, as well as subject and item as the random intercepts.

To compare whether participants in each age group performed differently, the data were analyzed separately within each age group. For the prime condition, the log-RT of the control primes was chosen as the baseline to compare that of the Tone 2 and Tone 3 primes with. I then reset the reference level to that of the Tone 2 primes to compare the Tone 3 primes with. For the trial order condition, the first appearance (trial order = 1st) was chosen as the baseline first to compare the second and the third appearances. I then reset the reference level to 2nd to compare with the other order if necessary. The results are summarized below.

For the four-year-old children, the final model that best fit the data only included a fixed effect of condition. The effects of trial order and the interaction term between condition and order were not found to improve the model significantly. The results of the post hoc analysis revealed a significant facilitatory effect from the Tone 2 prime (estimate = -0.130, SE = 0.038, t = -3.338, p < .001), where the log-RT with the Tone 2 prime was significantly shorter than the controls (log-RT with Tone 2 as prime = 3.11 < log-RT with control as prime = 3.22).

For the five-year-old children, the facilitatory effect from the Tone 2 prime (surface prime) approached significance prime (estimate = -0.046, SE = 0.026, t = -1.59, p= .071), with the log-RT being shorter when it carried a Tone 2 prime than it was for controls (log-RT with Tone 2 as prime =3.02 & log-RT with control as prime =3.07). Similarly, there was also a general order effect with the trials presented the first time being identified significantly slower than when presented a second and third time (2nd: estimate = -0.066, SE = 0.025, t = -2.473, p< .05; 3rd: estimate = -0.055, SE = 0.026, t = -2.086, p< .05).

For the children aged six years old, the final model that best fits the data included fixed effects of condition and trial order as well. As in the case of four-year-olds, adding the interaction term between condition and trial order did not improve the model significantly here either. Looking at the prime effect from the six-year-old children, the results revealed significant facilitation effects by both Tone 2 primes (estimate = -0.054, SE = 0.020, t = -2.692, p < .05) and Tone 3 primes (estimate = -0.076, SE = 0.020, t = -3.797, p < .001). Of the facilitatory effects from Tone 2 and Tone 3 primes, the Tone 3 facilitation was more prominent (2.99 for Tone 3 primes vs. 3.07 for controls) than that of the Tone 2 primes (3.01 for Tone 2 primes vs. 3.07 for controls). However, the log-RT was not significantly different between the Tone 2 and the Tone 3 primes. The trial order effect was still similar to that found for the younger children, in that the trials presented the first time were identified somewhat more slowly than when presented the second and third times in six-year-old children (2nd: estimate = -0.064, SE = 0.020, t = -2.231, p< .005; 3rd: estimate = -0.047, SE = 0.020, t = -2.363, p< .05).

For the seven- and eight-year-old children, only the condition variable was found to contribute to the model significantly. The post hoc analysis showed both a significant facilitation effect from both Tone 2 primes (estimate = -0.075, S.E. = 0.018, t = -4.125, p < .001) and Tone 3 primes (estimate = -0.095, S.E. = 0.018, t = -5.259, p < .001). We then set the Tone 2 primes as the reference level to compare the log-RT of the Tone 2 prime condition to that of the Tone 3 primes. The results were not significantly different between the Tone 2 and the Tone 3 primes (2.85 for Tone 3 primes, 2.87 for Tone 2 primes, and 2.94 for controls).

For the performance of the adult Mandarin speakers, the final model that best fits the data included fixed effects of condition and trial order. Adding the interaction between condition and trial order was not found to improve the model significantly here either. Post hoc analysis found a significant facilitatory effect from Tone 3 primes, with the log-RT for Tone 3 primes being significantly shorter than for control primes (estimate = -0.063, S.E. = 0.261, t = -2.441, p < .05; log-RT of Tone 3 primes = 2.57, log-RT of controls = 2.63). This indicates a significant underlying priming effect on the representation of tone sandhi items. I then set the Tone 2 prime as the reference level to compare the log-RT of the Tone 2 prime condition to that of the Tone 3 primes. The results also showed a significant difference between the Tone 2 and the Tone 3 primes, namely that the log-RT of Tone 3 primes was significantly shorter than that of Tone 2 primes (estimate = -0.068, S.E. = 0.026, t = -2.578, p < .05; log-RT of Tone 3 primes = 2.57, log-RT of Tone 2 primes = 2.64). There was not any significant difference between the Tone 2 prime condition and the controls. The trial order effect in the adult group was similar to that found in younger children, namely that the trials presented the first time were identified significantly more slowly than when they were presented the second time (estimate = -0.073, S.E. = 0.026, t = -2.794, p < .01) and 3rd time (estimate = -0.105, S.E. = 0.026, t = -4.031, p < .001). The results are summarized in Table 5.5.

Table 5.5 *Results of model comparisons of reaction time (likelihood ratio tests)*

age group	condition	order	condition * order
4y	Yes; significant Tone 2 prime facilitation	no	no
5y	Tone 2 prime approached significance (*p* = .0629)	Yes; longer log-RT the first time than 2nd and 3rd time	no
6y	Yes; both significant Tone 2 prime facilitation and Tone 3 prime facilitation	Yes; longer log-RT the first time than 2nd and 3rd time	no
78y	Yes; both significant Tone 2 prime facilitation and Tone 3 prime facilitation	no	no
adult	Yes; significant facilitatory effect from Tone 3 primes	Yes; longer log-RT the first time than 2nd and 3rd time	no

5.5. Discussion

This study focused on the mental representation of Mandarin Tone 3 sandhi in children who speak Mandarin Chinese as their mother tongue, with data from adults as controls. To explore the features of this representation, the study conducted an auditory priming lexical decision task, in which the prime words matched the initial sandhi tone in either the underlying form (Tone 3), or the surface form (Tone 2), or did not match in terms of tones at all (control). If the children took advantage of the Tone 2 primes, the sandhi words may be represented in their surface form. If they make use of the Tone 3 prime, the sandhi words may be represented in their underlying form. While if they make use of both Tone 2 and Tone 3, both the underlying Tone 3 and the surface Tone 2 are expected to be represented in the mental lexicon. The results of the log-RT from the lexical decision task indicate that the Mandarin-speaking adults exhibited a significant underlying priming effect when deciding Tone 3 sandhi real words. Children primarily took advantage of the surface primes until five years old, after which they started to make use of both surface and underlying primes from six to eight years old. The results suggested that children's representation of Tone 3 sandhi words undergoes a transformation from surface priming between four and five to multi-variant representation at later ages, and that they have not yet reached an adult-like underlying representation even between seven and eight years old.

The results from the adult data are consistent with the previous findings (Chien et al., 2016; Zhang, 2018) that Mandarin-speaking adults tend to list Tone 3 sandhi words, which are productive, in their underlying form, even though the words they encounter daily are in the post-sandhi form. That is to say, Mandarin native speakers represent Tone 3 sandhi words as underlying /tone 3 + tone 3/ in their mental lexicon.

It is also basically in line with the hypothesis that the four- and five-year-old children, who were found to exhibit relatively low productivity in sandhi application in speech production, primarily store and retrieve the sandhi words in their surface form. The finding further confirms children's way of learning sandhi words in the early ages that they primarily take advantage of the lexical mechanism to memorize the surface-form sandhi words.

However, for the children from six to eight, this study showed that they relied on both surface and underlying priming when deciding the Mandarin Tone 3 sandhi words. This suggested that the representation of Tone 3 sandhi words may involve different types in these age groups that both the underlying tone and the context-specific surface tone are activated. The finding does not perfectly match the development of productivity in children's production of sandhi words. In the speech production task, children have already undergone a transformation from the lexical mechanism to the computation mechanism from six to eight, and they have achieved the adults-like computation mechanism at seven. However, the six to eight-year-old children still primarily make use of the lexical exemplar retrieving in this study. The explanation can be made as follows: speech perception requires not only top-down processing that relies primarily on language experience but also bottom-up processing which depends mainly on the instant auditory input. Compared to adults, children may not have sufficient language experience to buildup the underlying Tone 3 category, and merely take advantage of underlying representation. Therefore, they take advantage of both Tone 2 and Tone 3 primes.

To further uncover the reasons for the discrepancies between the results of the two experiments, it may be more helpful to look at the task first. The lexical decision task used real words to test the representation of tone sandhi in the mental lexicon. Individuals' degree of literacy may influence their ability to manipulate the disyllabic words into smaller morphological units. In this experiment, the six to eight-year-old children had only learned some of the written forms of words beforehand. Therefore, due to the literacy effect, children may not reach the adult-like representation even by seven and eight years old. To confirm this hypothesis, however, further study will need to control for the literacy effect. For example, I will recruit both literate and illiterate adults and children, and examine the representation features of Mandarin Tone sandhi in their mental lexicon. If only the literate adults showed a significant underlying priming effect when deciding Tone 3 sandhi real words, while both the illiterate children in older age groups (e.g., seven and eight years) and illiterate adults exhibit both surface and underlying effect, I may predict that the representation features in those children and illiterate adults were due to the literacy effect. Secondly, it is also possible that the magnitude of lexical frequency may affect the operation of Mandarin Tone 3 sandhi. Although Mandarin Tone 3 sandhi has been shown to operate via the computation mechanism, this does not necessarily imply that all Tone 3 sandhi words are manipulated by the same computation mechanism. For unfamiliar words, participants may rely on computation mechanisms to apply the tone sandhi pattern productively. For highly familiar words, however, they may directly take advantage of the lexical mechanism, because the sandhi form may be retrieved from the mental lexicon. Therefore, both the surface and underlying priming effect, which links to the integrative effort of lexical and computation mechanism, were found. However, this pattern was not found in adults (Chien et al., 2016, 2017). It may suggest that the adults have already established the way of computation processing of Mandarin Tone 3 sandhi words, therefore the magnitude of lexical frequency took little effect (Chien et al., 2016, 2017). Unlike the adults, children's acquisition of computation operations was not well established, the lexical frequency may thus affect their operation of Mandarin Tone 3 sandhi.

Chapter 6 Tone awareness of sandhi initial syllables in Mandarin-speaking Adults and Children

6.1. Introduction

As was mentioned in the first chapter, phonological awareness refers to the ability to detect and manipulate speech sounds in a language (Anthony & Francis, 2005; Shu et al., 2008; Chen et al., 2016). It is a general term that includes abilities at different levels, such as tone awareness, syllable, onset, rime, and phoneme awareness (Anthony & Francis, 2005; Høien et al., 1995; Shu et al., 2008; Treiman, 1987). Previous research has uncovered a general sequence of phonological awareness development: children become increasingly sensitive to smaller and smaller units of words as they grow older (Anthony & Francis, 2005; Anthony et al., 2002; Anthony et al., 2003). To be specific, they can detect and manipulate syllables before they can detect or manipulate onsets, rimes, and individual phonemes. When exploring phonological awareness in different languages, scholars tend to analyze the explicit processing and acting ability separately from the implicit discrimination of speech sounds (see Castles & Coltheart, 2004; for a review). Various studies have assumed that implicit phonological awareness primarily increases with age and that explicit awareness relates heavily to prior instructions in how to use pinyin or prior educational experience (Castles & Coltheart, 2004; Shu et al., 2008; Morrison, Smith & Dow-Ehrensberger, 1995; Treiman & Zukowski, 1991). However, the implicit awareness of some linguistic units with more subtle differences such as onsets and tones has also been found to relate to Pinyin instructions (Shu, Peng, & McBride-Chang, 2008; Yin et al., 2011; Chen et al., 2016).

Tone awareness, as one of the crucial components of phonological awareness in tonal languages such as Chinese, is acquired quite late. For example, Shu et al. (2008) examined the implicit tone awareness in Mandarin-speaking children aged three to six, the results showed that the accuracy was only around 60% in five-year-old children (Kindergarten 3), and it was still below 75% in children aged six. Besides, tone awareness was influenced not only by age but also by whether participants have priorly received any explicit Pinyin instructions regardless of whether the "awareness" is implicit or explicit (Shu et al., 2008; Yin et al., 2011; Siok & Fletcher, 2001).

However, no study has yet focused on the awareness of sandhi tones in the Mandarin Tone 3 sandhi context (i.e., tone of the initial syllable in Mandarin Tone 3 sandhi words), where the learning process is more complicated. Considering the nature of sandhi tones that the tone's surface form for an underlying morpheme of the initial syllable varies as a function of the phonological environment, children's tone awareness in the Tone 3 sandhi words may be delayed compared to the awareness of tones in monosyllables/single words. Children's tone awareness of the initial sandhi tone is possibly related to the way speakers internalize tone sandhi words. If the word is internalized via the lexical mechanism, the initial sandhi tone in the sandhi word is more likely to be recognized as a sandhi Tone 3 (Tone 2-like), while if the tone sandhi item is operated by the computation mechanism, the initial sandhi tone in the Tone 3 sandhi word is more likely to be recognized as underlying Tone 3.

The experiments from chapters 3 to 5 explored the operational mechanisms and the mental representational features of Mandarin Tone 3 sandhi from a developmental perspective. The results demonstrated that children primarily took advantage of the lexical mechanism from three to five years old, and reached the adult level of computation mechanism at seven years old. However, due to the possible influences of individuals' literacy effect and the frequency difference of the Tone 3 sandhi words, the representation features of tone sandhi words exhibited some delayed progress in reaching the adult-like underlying representation. Based on these findings, I hypothesize that the initial sandhi tone in the sandhi words is more likely to be recognized as the sandhi Tone 3 (i.e., Tone 2 like) in children from three to five years old, whereas it would be recognized as the underlying Tone 3 in seven and eight-year-old children, as well as adults. For example, when asking a Mandarin speaker to identify the initial syllable of /mei213ny213/ 美女 "beauty" (/mei213ny213/ in the underlying form; /mei35ny213/ in the surface form), which is composed of the morphemes /mei213/ and /ny213/, the three to five-year-olds are expected to identify the initial syllable as /mei35/, while the seven, eight-year-old children and adults are expected to identify that as /mei213/. Besides, some delayed (or accelerated) progress may also occur due to the literacy effect, or some other factors.

Previous work has generally focused on child language development as a process moving from smaller speech units to larger ones (e.g., Bloom, 1994; Goodluck, 1991; Dick et al., 2015; Clark, 2016). These authors have primarily described child language acquisition as a gradual development from single-word utterances, first to multi-word combinations, and eventually to the production of complex sentences. It can not be ignored, however, that there is another learning process, one that moves from larger unanalyzed units to an awareness of smaller units (Peters, 1977; Arnon, 2011). The logic is as follows: children usually learn language from adult input. Adult speech flow is not neatly separated into words, phonemes, and morphemes, but is rather presented to children as a connected speech flow. This chapter offers a new perspective concerning how the Mandarin Tone 3 sandhi words are decomposed (be aware of) during language development.

To examine how adult and child Mandarin speakers analyze or recognize the sandhi tones in Mandarin Tone 3 sandhi words, this chapter conducted two experiments that require both explicit and implicit awareness of the sandhi tones. To examine explicit tone awareness of the sandhi initial syllable, a syllable deletion task was adopted. To measure implicit tone recognition of the sandhi initial syllable, a three-alternative forced-choice task was designed. It is a modified version of the two-alternative forced-choice paradigm (2AFC) which uses three choices for each trial to match the test of Tone 3 sandhi items instead of two. Detailed information will be presented in the research methods section below.

6.2. Methods

6.2.1. Participants

Eighty-six children aged from four to eight who spoke Northern Mandarin dialects were recruited for the study. As with the criteria for subject recruitment in the speech production and mental representation experiments in chapters 3 - 5, all children were monolingual Mandarin speakers who had rarely been exposed to other dialects or languages (including English). All children were instructed to accomplish the revised versions of Wechsler's test for intelligence tests (WISC-R; Zhang & Lin, 1979), which included two versions, one of which was for preschoolers below six and a half years old and the other one for children above six and a half years old. Children's IQ scores were comparable among different age groups. None of them was reported to have any hearing impairment. Eight subjects were excluded from the experiment because they were not able to understand the instructions for the experiment. Seventy-nine subjects, therefore, were retained for data analysis. Among the children, 12 were aged four, 25 were aged five, 19 were six years old, 12 were seven, and 12 were eight years old. The seven- and eight-year-old children were all recruited from primary school and had thus received explicit Pinyin instructions. The children aged from four to six were all recruited from kindergartens. The four-year-old children have never received any explicit Pinyin instructions. While five and six-year-old children had a few exposures to Pinyin instructions beforehand. Twelve adults were also recruited as controls. The information of participants in each age group is summarized in Table 6.1. The experimental procedures were approved by the Human Subjects Ethics Sub-committee of The Hong Kong Polytechnic University. Informed written consent was obtained from all adult subjects and children's caregivers in compliance with the experiment protocols.

Table 6.1 *Information of participants in each age group*

age group	subject number	mean age (age range)	gender ratio	Whether received Pinyin instructions
4y	12	4;6 (4;1-4;11)	5F/7M	no
5y	25	5;5 (5;1-5;11)	11F/14M	5(systematically); 11 (partially); 9 (no)

6y	19	6;5 (6;1-6;11)	10F/9M	10 (systematically); 8 (partially); 1(no)
7y	12	7;6 (7;2-7;10)	6F/6M	12 (systematically)
8y	12	8;7 (8;2-8;11)	6F/6M	12 (systematically)
Adult	12	2462	7F/5M	12 (systematically)

6.2.2. Stimuli

The stimuli included 19 target disyllabic words, seven Tone 3 sandhi words (e.g., /li213 p^hin213/ 礼品 "gift") and 12 non-Tone 3 sandhi words (e.g., /ʂou213 tɕi55/ 手机 "phone"). Each target syllable (e.g., /li213 p^hin213/ 礼品 "gift") was paired with three disyllabic word choices whose initial syllable matched with the initial syllable in the target word, in terms of segmental phonology (consonant and vowel). The choices of tone sandhi targets were: (i) a Tone 3-initial word, whose initial syllable's tone matched that of the initial syllable of the target word in its underlying form (e.g., /li213miæn51/ 里面 "inside") (ii) a Tone 2-initial word, whose initial syllable's tone matched that of the initial syllable of the target word in its surface form (e.g., /li35tʂʅ0/ 梨子 "pear"), and (iii) a word beginning with a syllable whose tone was neither Tone 2 nor Tone 3, which therefore did not match the target word's initial tone in either its surface or underlying form (e.g., /li51tɕʻi0/ 力气 "strength"). Given that semantic relatedness and morphological type affect the accuracy of morphological awareness (Hao et al., 2013), the target words were all composed of free morphemes and the choice items were not semantically related to the initial syllable of the corresponding target. The targets were selected from 30 words according to a subjective familiarity rating score (ranging from 1, representing "never heard or said," to 9 "very often heard and said"), as judged by eight Mandarin-speaking adults. Table 6.2 illustrates the target items and their corresponding choices. The tone sandhi targets and their corresponding choices served as the critical elements for the data analysis.

Table 6.2 *The tone sandhi target items and their corresponding choices*

		Target words	*Words for choices*
test items	Tone 3 sandhi words	礼品	*梨子 里面 力气*
		美女	*眉毛 每次 妹子*
		手表	*熟食 守信 瘦子*
		脚趾	*嚼饭 饺子 浇花*
		舞蹈	*无敌 午饭 污染*
		养鸟	*阳光 仰头 样子*
		笔筒	*鼻子 比赛 壁画*
	Non_Tone 3 sandhi combinations	手机	*熟食 守信 瘦子*
		脚丫	*嚼饭 饺子 浇花*
		笔尖	*鼻子 比赛 壁画*

control items		舞鞋	*无敌 午饭 污染*
		养花	*阳光 仰头 样子*
		礼貌	*梨子 里面 力气*
		美图	*眉毛 每次 妹子*
		眉毛	*梅子 每次 妹子*
		阳光	*养花 羊毛 样子*
		书包	*薯条 梳头 树枝*
		树叶	*书包 数字 薯条*
		中国	*种树 种类 钟表*

6.2.3. Experiment procedure

Children and adults were tested on two different tone awareness tasks individually in a quiet room at their own schools with the experimenter. The two experiments were used to examine the explicit and implicit awareness of the tones of the initial syllable in the Tone 3 sandhi words, respectively. Firstly, in the explicit tone awareness experiment, the syllable deletion task was adopted. Only the 19 target stimuli were used in this task. Powerpoint was used to present the stimuli. Children were first presented via earphones with a target stimulus, which was paired with its picture (e.g., 礼品 /li213 p^hin213/ "gift"). They were then asked to answer the question "What is left if I make [the last word] disappear?" after hearing the stimulus (In Chinese: 礼品：如果我把"品"变不见，还剩下什么？).

Secondly, in the three-alternative forced-choice task that was used to test the implicit tone recognition of the sandhi initial syllable, both target words and their corresponding choice items were employed. The specific procedure of the task was as follows: before the experiment, participants were told that two words that share the same syllable are 'good friends'. Participants were then trained on the task using a sequence of four practice trials, followed by the experiment proper. During each practice trial, the children were first presented via earphones with a target syllable paired with its picture. Immediately after that, they were given three choice items, which were also paired with their pictures. The participants were asked to find the 'good friend' of the initial syllable in each target word. The participants were given feedback of 'correct' or 'incorrect' after they responded. If a child's response was not correct, the experimenter repeated this trial again with explanations. During the experiment proper, children were asked to listen to each word and then identify which of the three choices was the 'good friend' of the target word, ignoring their meanings. The children responded by pointing to the picture they selected, and the experimenters recorded the children's responses. All pictures were semantically unrelated and familiar to the children. Figure 6.1. shows two examples from the experiment, where the characters were not shown on the screen (the left one: target item is a Tone 3 sandhi word; the right one: target item is a non_Tone 3 sandhi word). In each slide, within each trial, the picture of the target word was always presented first.

Figure 6.1. *Examples of the experiment design*

To make sure that the children, especially the youngest ones, understood the task, they were only allowed to proceed when they had correctly answered the questions in the practice task. Crucially, the experiment did not include the tone sandhi items in the practice task, to prevent participants from detecting the purpose of the experiment.

6.3. Results

Participants' responses were transcribed as soon as they were given. For the explicit tone recognition task, if the children produced the syllable correctly as Tone 3, it was labeled as 3; if they produced the syllable as Tone 2, it was labeled as 2. When they were unable to answer, it was labeled as 0. The results were as follows: three trials in the overall trials of the four-year-old children and one trial in those of five-year-old children were labeled as 0. For the remaining tone sandhi trials, the participants produced the initial syllable as either Tone 3 or Tone 2. Detailed information was provided in Table 6.3. For the implicit tone matching task, which tested awareness of sandhi tone in word-initial position, the children exhibited great variability in linking the sandhi tones: some trials were linked to Tone 3 by them, others to Tone 2. There was also a large percentage of trials that were linked to both Tone 2 and Tone 3. Detailed information was provided in section 6.4.2. The coding method for implicit matching was the same as for explicit identification, with one addition: responses were labeled as 3 when matching Tone 3, as 2 when matching Tone 2, and as 0 when matching any other tone, as before. Additionally, responses were labeled as 1 when matching both Tone 2 and Tone 3.

6.3.1. The explicit awareness experiment

In the explicit tone recognition task, 97.62% (S.E.=1.67%) of the trials were produced as Tone 3 by the adult speakers. This indicates that the tone of the initial syllable of the tone sandhi word was generally produced as an underlying Tone 3 in adult speakers, even though the sandhi words they encountered were in their surface form. Figure 6.2 describes the ratio of trials produced as Tone 3 (i.e., the rate of Tone 3 recognition) for the Tone 3 sandhi initial syllable in each age group. The boxplot displays the distribution of Tone 3 recognition ratio within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual points of Tone 3 recognition. As can be seen in Figure 6.2 and Table 6.3, the children aged four showed the least underlying Tone 3 recognition (mean= 73.8%; S.E. = 44.23%), and they showed a gradual increase in their rate of recognition of Tone 3 as the age group increased.

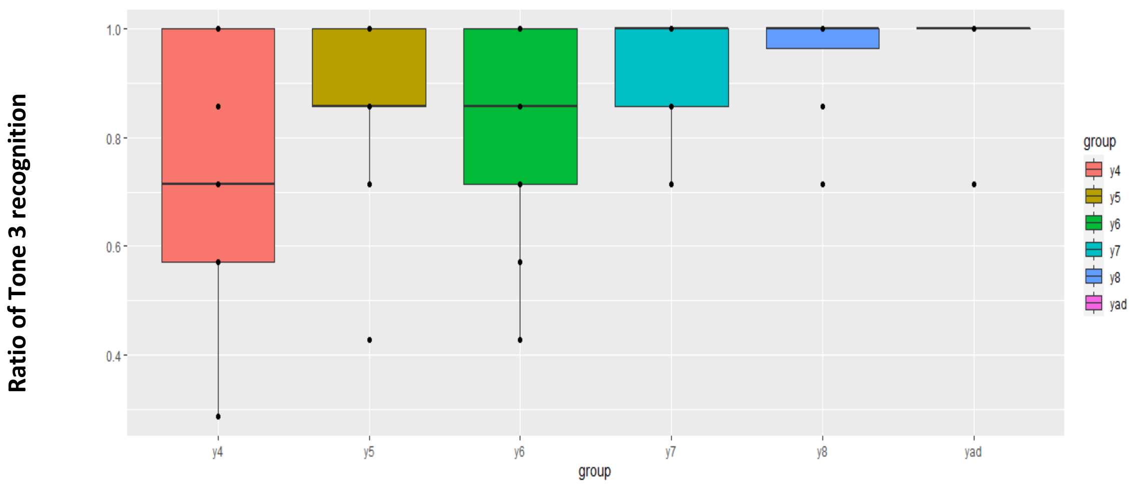


Figure 6.2. *the boxplot displays the distribution of Tone 3 recognition ratio within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual ratio of Tone 3 recognition.*

Table 6.3 *Means, standard deviations, and standard errors of the ratio of Tone 3 recognition in each age group*

age group	*number of test item*	*mean*	*S.D.*	*S.E.*
y4	*84*	*73.81%*	*44.23%*	*4.83%*

y5	*175*	*86.86%*	*33.88%*	*2.56%*
y6	*133*	*83.46%*	*37.30%*	*3.23%*
y7	*84*	*92.86%*	*25.91%*	*2.83%*
y8	*84*	*94.05%*	*23.80%*	*2.60%*
adult	*84*	*97.62%*	*15.34%*	*1.67%*

To find out when children reached adult-like Tone 3 sandhi word recognition, the children's responses were analyzed by generalized linear mixed-effects model (Baayen, 2008), using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The binomial distribution and the logistic link function with family = "binomial" were selected for the analysis. The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included children's responses (1 indicates Tone 3; 0 indicates Tone 2) as the dependent variable, the age group (4y, 5y, 6y, 7y, 8y, and adult), and Pinyin instructions (systematically, partially, no) as an independent variable, and subject and item as random factors. Random slopes were also added for each of the fixed effects. Model comparisons started with a full random structure that included all the fixed factors and random intercepts and slopes. Models were compared using log-likelihood ratio tests to determine if given terms significantly improved the model. The results of the comparison showed that the best model was the one that included age group and random intercepts of subject and item.

For the age group, the adult group is treated as the baseline first to which the other age groups are compared. Comparing children's performance to that of the adults (see Table 6.4), statistical analysis showed that, from ages four to six, children's rates of recognition of the Tone 3 initial as Tone 3 were significantly lower than that of the adults. And children aged seven and eight did not differ significantly from adults. The recognition rate of children was also compared across groups. The results showed that the four-year-olds exhibited a significantly lower rate of Tone 3 recognition than did the eight-year-old children (Estimate = -2.195 , S.E. = 0.748 , z = -2.937, $p<.05$). The rates of underlying Tone 3 recognition shown by children aged four approached significant differences compared to those shown by seven (Estimate = -1.917 , S.E. = 0.711, z = -2.698, $p= .075$). No significant difference was found between other age groups of children.

Table 6.4 *Fixed effect estimates of children's explicit Tone 3 recognition across different age groups compared to the adults*

Fixed effect	Estimate	S. E.	z value	Pr(>|z|)
y4-adult	*-3.2801*	*0.9503*	*-3.452*	*p*<.001 ***
y5-adult	*-2.2051*	*0.9042*	*-2.439*	*p*<.05 *

y6-adult	*-2.5008*	*0.9157*	*-2.731*	*p*<.01 **
y7-adult	*-1.363*	*1.0032*	*-1.359*	*p*=.174
y8-adult	*-1.0849*	*1.0208*	*-1.063*	*p*=.287

6.3.2. The implicit awareness experiment

Table 6.5 summarizes the results of the implicit tone matching task in each age group. The table shows the rates at which participants matched the target words to different tones. Looking at the adult data, 97.4% (S.E.=1.8%) of the targets were matched to Tone 3 words, which suggests that the tone of the initial syllable of the tone sandhi word was generally recognized as an underlying Tone 3 by adult speakers. The children in the younger age groups, by contrast, especially those aged from four to six, tended to match the sandhi initial tone either to Tone 2, or both to Tone 2 and Tone 3. The children aged seven and eight had more or less approached the adult level of Tone 3 matching, although they still exhibited a certain degree of Tone 2 matching. None of the seven- and eight-year-olds matched the sandhi tone to both Tone 2 and Tone 3.

Table 6.5 *Tone matching rate of the sandhi tone in word-initial position by age*

	4y		5y		6y		7y		8y		ad	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
Tone 3 (%)	28.57	24.28	58.39	25.81	66.67	24.98	86.90	18.73	91.67	14.23	97.40	8.61

Tone 2 (%)	18.57	17.88	20.50	21.46	19.84	17.76	10.71	15.08	8.33	14.23	2.60	8.61
Tone 2 and Tone 3 (%)	28.57	26.94	14.29	18.27	7.14	20.35	0.00	0.00	0.00	0.00	0.00	0.00
others (%)	24.29	30.90	6.83	12.08	6.35	10.07	2.38	5.56	0.00	0.00	0.00	0.00

To find out at what age children reach adult-like Tone 3 sandhi word matching, the children's responses were analyzed alongside those of the adults by generalized linear mixed-effects model (Baayen, 2008), using the package lme4 (Joshua et al., 2010) in R (R core team, 2014). The binomial distribution and the logistic link function with family = "binomial" were selected for the analysis. The R package lmerTest was used to provide *p*-values for models (Kuznetsova, et al., 2016). The model included children's responses (1 indicates Tone 3; 0 indicates other tone matching responses) as the dependent variable, the age groups (4y, 5y, 6y, 7y, 8y, and adult), and Pinyin instructions (systematically, partially, no) as an independent variable, and subject and item as random factors. Random slopes were also added for each of the fixed effects. Model comparisons started with a full random structure that included all the fixed factors and random intercepts and slopes. Models were compared using log-likelihood ratio tests to determine if given terms significantly improved the model. The results of the comparison showed that the best model was the one that included age group and Pinyin, as well as the random intercept of the test item.

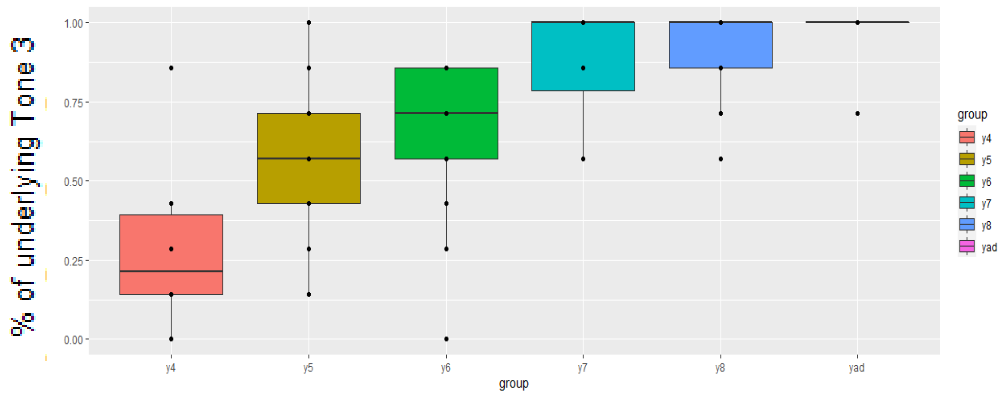


Figure 6.3 *the boxplot displays the distribution of data within each age group at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual ratio of underlying Tone 3 matching.*

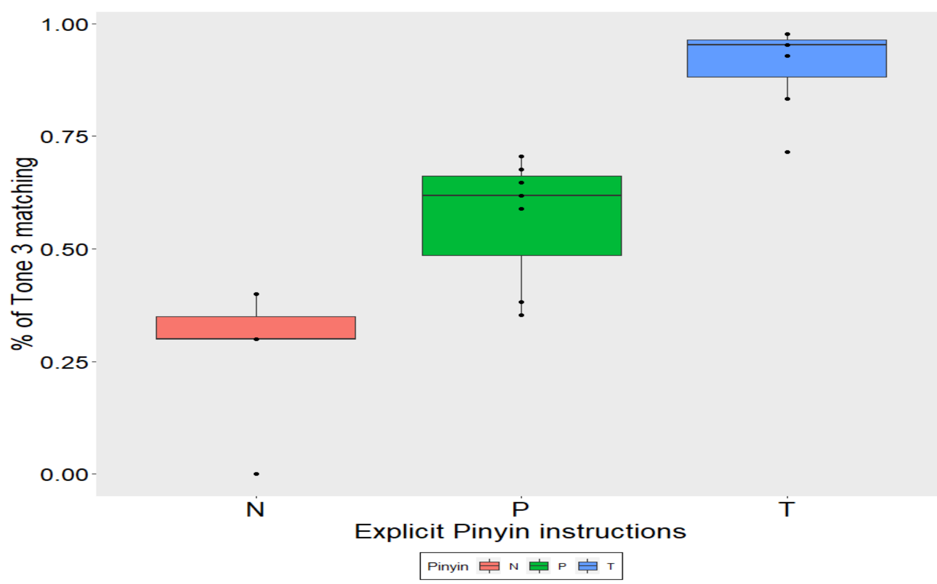


Figure 6.4 *the boxplot displays the distribution of data within each category of Pinyin instructions at the first quartile (Q1), median (Q2), and third quartile (Q3) respectively, and the individual ratio of underlying Tone 3 matching. (N: no Pinyin instructions; P: received partial Pinyin instructions; T: systematically trained in Pinyin instructions)*

Firstly, statistical analysis showed that there was a significant main effect of Pinyin experience (Estimate = -1.4216, S.E. = 0.431, z = -3.301, p<.001), which suggests the influence of explicit Pinyin instructions in implicit tone matching (see Figure 6.4). Secondly, statistical analysis demonstrated that children from four to seven years old exhibited significantly lower underlying Tone 3 matching rates than did adults. And children aged eight did not differ significantly from adults. The underlying Tone 3 matching rates of children were also compared across groups. The results showed that the underlying Tone 3 matching rate of children of four years old was significantly smaller than all the other children's age groups. While no significant difference was found between five and eight in terms of the underlying Tone 3 matching rates. The results of the statistical analysis are provided below in Table 6.6.

Table 6.6 *Fixed effect estimates of child accuracy across different age groups compared to adult accuracy.*

Fixed	Estimate	S.E.	z value	Pr(>|z|)
y4-adult	-4.8002	0.7758	-6.187	p <.001 ***
y5-adult	-2.1598	0.8433	-2.561	p <.05 *
y6-adult	-2.0097	0.8264	-2.432	p <.05 *
y7-adult	-1.7811	0.7923	-2.248	p <.05 *

y8-adult	-1.2538	0.8238	-1.522	*p* =0.128
y4-y5	-2.6404	0.5212	-5.066	*p* <.001 ***
y4-y6	-2.7904	0.4942	-5.646	*p* <.001 ***
y4-y7	-3.019	0.4359	-6.926	*p* <.001 ***
y4-y8	-3.5463	0.4928	-7.197	*p*<.001 ***
y5-y6	-0.15	0.2675	-0.561	*p* =0.574
y5-y7	-0.3786	0.5498	-0.689	*p* =0.491
y5-y8	-0.9059	0.5948	-1.523	*p* =0.127
y6-y7	-0.2286	0.5236	-0.437	*p* =0.662
y6-y8	-0.7559	0.5706	-1.325	*p* =0.185
y7-y8	-0.5272	0.5201	-1.014	*p* =0.310

6.4. Discussion

The chapter examined the tone awareness of the sandhi initial syllable in disyllabic Mandarin Tone 3 sandhi words in Mandarin-speaking children and adults. During the explicit tone recognition task, adult Mandarin speakers were able to recognize the tone of the initial syllable in tone sandhi words as underlyingly Tone 3, which could indicate the involvement of the computation mechanism in internalizing Tone 3 sandhi words. Children reached this adult level of underlying tone recognition at seven years old in the explicit tone awareness task. This result more or less complements the findings regarding the productivity of tone sandhi in Chapter 3, where children employed the computation mechanism at seven years old, not only when combing pseudo- and novel words with the sandhi pattern, but also when decomposing surface-form sandhi words into syllables. Contrastively, the four- to six-year-olds showed significantly less underlying Tone 3 recognition in the explicit tone recognition task compared with adults, this may relate to the involvement of the lexical mechanism at these age groups. However, the results from chapters 3 to 5 revealed that only the three- to five-year-olds primarily took advantage of the lexical mechanism. The children aged six had already been found to use the computation mechanism to a large extent in the speech production task. The significantly lower underlying Tone 3 ratio in children at six than adults in the task can be explained by comparing the differences between six-year-old children and adults. The difference can be found as follows: the six-year-old only received a few prior explicit Pinyin instructions and they only know a limited number of written forms of words. Due to the effect of prior explicit Pinyin instructions and literacy, the six-year-old children showed a relatively delayed underlying Tone 3 awareness.

Furthermore, adult speakers were found to be able to match the tone of the initial syllable in tone sandhi words as an underlying Tone 3 in the implicit tone matching task, which may suggest the involvement of the computation mechanism in memorizing the Tone 3 sandhi words. Statistical analysis demonstrated that children reached this adult level in the implicit tone matching task at eight years old, later than what was found in the explicit tone awareness task. When observing the underlying

Tone 3 matching rates of children across groups, it is obvious that the underlying Tone 3 matching rates were very low in children aged four and six. The underdeveloped tone sensitivity was more obvious in the four year-year-old children, in that almost 30% of trials were linked to both Tone 2 and Tone 3, and 1/4 of the trials were even matched to other tones (neither Tone 2 nor Tone 3) in those four-year-olds. The delayed adult-like underlying Tone 3 matching and the low underlying Tone 3 matching rates in the younger age groups may result from two aspects. Firstly, they might because of the underdeveloped acquisition of tone awareness in children of the younger age groups. Previous empirical experiments (Shu et al., 2008) examining children's (aged three to six) implicit tone awareness in monosyllabic words revealed that the accuracy was not very high (below 75%) even in children aged six. The underdeveloped tone sensitivity may make it more difficult for children in these age groups to be aware of the tones in sandhi contexts. It is also possible that the implicit tone matching task requires a greater memory load than the explicit awareness task, given that individuals were asked to select one out of three items that shared the same initial syllable. Moreover, the visual complexities of the pictures in the experiments were not well controlled. This may also affect the task performances.

When considering the influence of explicit Pinyin instructions on the performance of different age groups in the explicit and implicit tone awareness tasks, one thing stands out that explicit Pinyin instructions contributed significantly to the development of implicit tone awareness of the sandhi initial syllables, rather than the explicit tone recognition. Besides, there might be a developmental trajectory of tone awareness in the sandhi context, which is in line with children's degree of explicit Pinyin instructions. I found that the four-year-old children who have never received any explicit Pinyin instructions performed very poorly in the two tasks. The least underlying Tone 3 recognition and matching ratio at four compared with other age groups may not only because of the age but also the lack of prior explicit Pinyin instructions. However, since the degree of Pinyin instruction is not well controlled in this chapter, further studies are required to further confirm the association with tone awareness in the sandhi context and Pinyin education.

This study is not without its limitations. First, the reliability of the implicit tone awareness tasks was relatively low. Due to the young age of the participants, they were given four pictures when presented with targets and choices. This poses a great challenge for memory load. Second, this study did not control children's levels of literacy and degrees of Pinyin instructions, which may have influenced the results. Despite these limitations, however, this study has extended the literature on tone awareness in the context of Mandarin Tone 3 and has confirmed the development of the computation mechanism during the process of decomposition of Mandarin Tone 3 sandhi words.

Chapter 7 General discussion

There has been widespread and long-lasting interest in the acquisition and processing of morphological and phonological alternations (Bybee, 1995; Bybee & Slobin, 1982; Gleason, 1958; Pinker, 1994, 1999; Pinker & Department, 1998; Ullman & Pinker, 2002; Ullman, 1999; Van Der Lely & Christian, 2000; Kirjavainen, Nikolaev, & Kidd, 2012). This interest has primarily centered on the productivity of the patterns of alternation arising from the regularity of input and how these patterns are applied, internalized, or decomposed by children of different ages (e.g., Gleason, 1958; Pinker, 1994; Steven Pinker, 1998, 1999; Ullman & Pinker, 2002; Ullman, 1999; Van Der Lely & Christian, 2000). According to the dual-route model, regular and irregular morphological alternations (e.g., "jump-jumped" & "go-went" in the past tense of English) are processed in different ways: irregular morphologically complex items are processed item by item via a rote-memory based lexical mechanism, whereas regular items are processed productively via a computational mechanism by which phonological patterns are applied to items in their underlying representation (Steven Pinker et al., 1998; Ullman & Pinker, 2002; Ullman & Pierpont, 2005; Van Der Lely & Christian, 2000). The more productive (e.g., Mandarin Tone 3 sandhi) and less productive sandhis (e.g., the sandhi pairs in Taiwanese) found in Chinese dialects are also manipulated via these two mechanisms, even though all such forms are regular.

For the productive tone sandhis such as Mandarin Tone 3 sandhi, children have also been found to process the sandhi items differently as they get older, thus revealing different levels of productivity over time. By looking at the different forms of tone sandhi from a developmental perspective, this project generated interesting new findings that increase our understanding of the mechanisms by which linguistic productivity operates.

A first step to reveal the complex picture of the developmental features of productivity of different tone sandhi patterns in Chinese dialects, this project has focused on the Mandarin Tone 3 sandhi pattern, which has been demonstrated to be productively processed via the computational mechanism in adult Mandarin speakers (Zhang & Peng, 2013; Zhang et al., 2015). Importantly, however, little is known about when the pattern becomes productive and how it is processed by Mandarin-speaking children of different ages. Besides, the initial Tone 3 morpheme of the sandhi words have two allomorphs: the underlying Tone 3 form and the Tone 2-like form, which surfaces in the sandhi context. It has been observed that adult Mandarin speakers tend to represent sandhi words in their underlying forms (Chien et al., 2016; Zhang, 2016, 2018). However, no study has provided empirical data to investigate the developmental trajectory that children undergo before reaching an adult representation. Furthermore, previous literature mostly focuses on how the Tone sandhi words are acquired (Rattanasone et al., 2018; Wang, 2011) and how the sandhi patterns are applied in pseudo-words (Tang, et al., 2019) and so on, but little is known about how it is recognized or internalized by speakers, even by adults. By discussing the findings of the different experiments which involve the application, representation, and decomposition of phonological and morphological patterns, and comparing them with other studies, this dissertation has provided a fuller and deeper understanding of the productivity of Tone 3 sandhi during language development.

I reported three primary experiments that aimed to address the aforementioned questions about children's phonological acquisition and the operational processing, representation, and tone awareness of the initial syllable in the tone sandhi context during development. The discussion is organized into the following four parts: 1) the productivity of Mandarin Tone 3 sandhi in speech production and its acoustic realization during language acquisition, 2) the representation of Mandarin Tone 3 sandhi in the mental lexicon of Mandarin-speaking children, 3) the tone awareness of sandhi initial syllables in Mandarin-speaking adults and children, 4) the progress from the memory-based lexical mechanism to the computation-based productive mechanism.

7.1. The productivity of Mandarin Tone 3 sandhi in speech production and its acoustic realization during language acquisition

The thesis examined the productivity of Mandarin Tone 3 sandhi in Mandarin-speaking children from three- to eight- years old by observing the application rate and the acoustic features of sandhi in real words, pseudo-words, and novel words. As mentioned in Chapter 3, these three types of words differ in the degree to which they are stored as lexical items and also in the operational abilities and mechanisms they require. Real words are expected to be listed in children's mental lexicon and can, therefore, be memorized via the lexical mechanism. Pseudo-words have been shown to operate by the computation mechanism, even though there may exist certain previously heard sandhi allomorphs that children may take advantage of when producing them. For novel words, since no possible lexical entries can be retrieved from memory, children are expected to lean solely on the phonological pattern for computational processing. The elicited picture naming task (wug test; Gleason, 1958) was used to elicit the spontaneous production of the sandhi words. Children's productions were evaluated based on both the tone production judgment from the phoneticians and the acoustic analysis of the pitch trajectories.

The results of the tone production judgment demonstrate a developmental trajectory in gaining adult-like high productivity. First, the results confirmed the high productivity of sandhi application in adult Mandarin speakers across the three lexicality conditions. Comparing children's performance to that of the adults, the results suggest that children showed relatively low productivity from three to five years old, especially for the three- and four-year-old children when producing pseudo- and novel words. This indicates that the lexical mechanism is involved in these age groups under these conditions. Secondly, children reached adult-like application in pseudo-words at six years old and in the novel word condition at seven years old, indicating that children had made the transition from using the lexical mechanism to applying the computational mechanism at the age of six and had fully developed to processing tone sandhi via the computational mechanism at the age of seven.

Moreover, this study further examined the pitch trajectories of children and adults' productions of Tone 3 sandhi real words, pseudo-words, and novel words. The results revealed that all types of words showed rising trends, i.e., were more Tone-2 like, despite the average F0 of pseudo- and novel words in the three- to five-year-old children appeared lower than the real words. Combined results between statistical analysis and the figures of pitch trajectories in different lexicality conditions suggested the following: when observing the pitch trajectories of the Tone 3 sandhi items that were identified as "application", the three-to-five-year-old children still exhibited significant trends of incomplete application in pseudo-and novel words than in real words, while children aged six-to-eight showed significant trends of overapplication in pseudo-words compared to real words. This may be treated as an indicator to index children's awareness of the sandhi pattern and their attempts to apply the sandhi pattern. When these six- to eight-year-old children tried to apply the sandhi pattern spontaneously, a trend of overapplication or over articulation occurred. Concerning the acoustic realization of the adult Mandarin speakers, they exhibited a trend of incomplete application in pseudo- and novel words compared to real words in the items that were identified as "application". The patterns demonstrated by adults are consistent, more or less, with the previous findings of Zhang and Lai (2010). This could be because the pseudo- and novel words were presented syllable by syllable in their underlying forms in the experiment, the process of which offered the participants more exposures to the underlying tones. In contrast, the participants are expected to encounter more surface-form sandhi real words, especially compared to the novel words in which the sandhi form would never occur in daily life. The sandhi tones in pseudo- and novel words may, therefore, have been more likely to bear more resemblance to the underlying tone than to the sandhi tone. Zhang and Lai (2010) have also linked the underapplication of tone sandhi in adults to a weaker phonetic motivation for this type of sandhi pattern in nature.

To summarize the developmental timeline of gaining the adult-like productivity in the three types of tone sandhi words: three-to-five-year-old children were not able to apply the sandhi pattern like adults, which implies that children at this age may primarily make use of mental listings in the lexicon, relying on the lexical mechanism to produce Tone 3 sandhi words. After a year of development and transition at the age of six, the seven-and-eight-year-olds have fully acquired the adult-like computation mechanism. Besides, as the sandhi pattern and the computation mechanism were gradually acquired, children tended to apply it actively, sometimes showing certain trends of over articulations in speech production.

Tang et al. (2019) conducted a similar experiment that explored the productivity of Mandarin Tone 3 sandhi in the disyllabic pseudo-words (called novel words in that paper; e.g., /ma213ku213/ 马鼓 “horse drum”) in children aged three-to-five with adult performance as controls. However, the result of their study is not in agreement with my findings. They found that three-year-olds have already applied the sandhi pattern productively to the pseudo disyllabic words as well as adults. The discrepancies between my findings and Tang et al. (2019) may be due to the difference in experimental design and stimulus complexity. In their study, there were only four-tone sandhi pseudo-words and they were all designed with two nouns, in contrast, my study contained twenty tone sandhi pseudo-words that covered different types of combinations such as nouns-word-combinations, noun-verb combinations, verb-noun combinations, and twenty novel words. It is likely that children were more likely to apply sandhi patterns in nouns, which led to a higher percentage of correct application in Tang and his collaborators’ study. Furthermore, due to the limited number of stimuli, it is also possible that the pattern is caused by the probability of the sandhi-initial allomorphs that occur in the sandhi context in language input to children. Even though none of the pseudo-words really occurred, individuals could still take advantage of the sandhi allomorphs of the initial syllable. If the allomorph occurs frequently as a sandhi Tone 3, children may easily retrieve the sandhi form from memory, rather than apply the sandhi pattern in the words. In fact, in another study which measured the acoustic features of Mandarin Tone 3 sandhi real words (Rattanasone, Tang, et al., 2018), it showed that children haven't fully reached the adult-like acoustic realizations at five years old even in the real words, which is partially consistent with my observation.

Interestingly, findings from a separate line of inquiry, the developmental timeline of some morphological alternations in Indo-European languages, seem to converge with my findings (e. g., Kerkhoff, 2007; Tomas, Van De Vijver, Demuth, & Petocz, 2017). For example, when investigating children's productive knowledge of Dutch voicing alternation from a developmental point of view, the results revealed that the alternation has not been fully acquired after the age of six (Kerkhoff, 2007). Besides, when asked Russian-speaking children aged four to seven to apply the Russian morphological alternations involving the interacting patterns of vowel deletion and stress shift using a "wug" test with real and novel words, the results showed that even the seven-year-olds exhibited high variabilities in applying the morphological patterns (Tomas et al., 2017). All these studies pointed to a developmental time "seven years old", which is in line with my finding that children at seven have reached an adult level of productivity in applying the sandhi pattern. However, this may not indicate that the "seven years old" is a crucial age. Age 7 broadly coincides with many factors, such as the onset of school education, systematic Pinyin instructions, literacy (learning to read), etc. It'll be interesting to further observe what types of phonological and morphological alternations would be acquired around seven-year-old, and the driven forces for the development.

7.2. The representation of Mandarin Tone 3 sandhi in the mental lexicon of Mandarin-speaking children

The representation features of sandhi items in the mental lexicon related heavily to their operational mechanisms (e.g. Zhang, 2018). Considering the productivity features from low to a higher degree during development, I set about observing how the sandhi items are represented in children's mental lexicon when they approaching the adult-like productivity. The experiment targeted at children from four to eight years old as well as adults as controls to explore the representation features during language development. The study adopted an auditory priming lexical decision task, in which the prime words matched the initial sandhi tone in either the underlying form (Tone 3), or the surface form (Tone 2), or did not match in terms of tones at all (control). If the children took advantage of the Tone 2 primes, the sandhi words may be represented in their surface form. If they make use of the Tone 3 prime, the sandhi words may be represented in their underlying form. While if they make use of both the Tone 2 and Tone 3 primes, the sandhi words are expected to be represented both as Tone 3 Tone3 and Tone 2 Tone 3.

For the adults, the results only exhibited a significant underlying priming effect when deciding Tone 3 sandhi real words, which is consistent with the previous findings (Chien et al., 2016; Zhang, 2018). For the children, the results suggested that children's representation of Tone 3 sandhi words undergoes a transformation from surface priming between four and five to both surface and underlying priming at later ages, and that they have not yet reached an adult-like underlying representation even between seven and eight years old.

To compare the results of the speech production with the mental representation features, four- and five-year-old children, who were found to exhibit relatively low productivity in sandhi application in speech production, primarily store and retrieve the sandhi words in their surface form. The consistent timeline further confirms children's way of learning sandhi words in the early ages that they primarily take advantage of the lexical mechanism to memorize the surface-form sandhi words.

However, for the children from six to eight, this study showed that they relied on both surface and underlying priming words when processing the Mandarin Tone 3 sandhi words. This finding does not perfectly match the development of productivity in children's production of tone sandhi words that the seven and eight-year-old children have already reached the adult-like level of computation mechanism. The results suggested that the processing of Tone 3 sandhi words in these age groups still involves the lexical storage of both the underlying tone and the context-specific surface tone. In the current study, children from six to eight were also found to take advantage of both surface-form and underlying-form exemplars before they reached the adult-like underlying representation in speech comprehension. The reasons may be as follows: firstly, the more abstract underlying-form representation tends to rely more on language experience, while the surface representation depends more on the instant auditory input/acoustic clues. With limited language experiences, children may easily take advantage of both the surface and underlying representations.

In addition, there may be other, more important factors at play here. The lexical decision task used real words to test the representation of tone sandhi in the mental lexicon. Individuals' degree of literacy may influence their ability to manipulate the disyllabic words into smaller morphological units. In this experiment, the six to eight-year-old children had only learned some of the written forms of words beforehand. Therefore, due to the literacy effect, children may not reach the adult-like representation even by seven and eight years old. To confirm this hypothesis, further study should explicitly examine the literacy effect. To be specific, I will recruit both literate and illiterate adults, and examine the representations of Mandarin Tone 3 sandhi in their mental lexicons. If only the literate adults showed a significant underlying priming effect when deciding Tone 3 sandhi real words, while the illiterates exhibit significant surface and underlying priming effects, I may predict that the representation features in those illiterate adults were due to the literacy effect. Secondly, it is also possible that the magnitude of lexical frequency may affect the operation of Mandarin Tone 3 sandhi. Although Mandarin Tone 3 sandhi has been shown to operate via the computation mechanism, this does not necessarily imply that all Tone 3 sandhi words are manipulated by the same computation mechanism. For unfamiliar words, participants may rely on computation mechanisms to apply the tone sandhi pattern productively. For highly familiar words, however, they may not necessarily take advantage of the computation mechanism, because the sandhi form may be retrieved directly from the mental lexicon, which means that those real words could be encoded by the lexical mechanism.

Both the surface and underlying priming effects were found in children aged seven and eight in my thesis. This may be linked to the involvement of the lexical mechanism. However, this pattern was not found in adults (Chien et al., 2016, 2017). It may suggest that the adults' have already established the computation processing of Mandarin Tone 3 sandhi words, therefore the frequency took little effect (Chien et al., 2016, 2017). Unlike the adults, children's acquisition of computation operations were not well established, the lexical frequency may thus affect their operation of Mandarin Tone 3 sandhi.

The results from the adult data are consistent with the previous findings (Chien et al., 2016; Zhang, 2018) that Mandarin-speaking adults tend to list Tone 3 sandhi words, which are productive, in their underlying form, even though the words they encounter daily are in the post-sandhi form. That is to say, Mandarin native speakers represent Tone 3 sandhi words as underlying /tone 3 + tone 3/ in their mental lexicon. The prior presentation of a Tone 2 form, although matching sandhi targets in their surface form, does not seem to facilitate recognition of the tone 3 sandhi words.

7.3. Tone awareness of sandhi initial syllables in Mandarin-speaking adults and children

The thesis observed the awareness of sandhi tones in the Mandarin Tone 3 sandhi context (i.e., tone awareness of the initial syllable in Mandarin Tone 3 sandhi words). In order to trace the development trajectory, the study recruited children from four to eight examining the explicit tone recognitions and the implicit tone matchings of the initial sandhi tones in sandhi words. The results revealed that children reached the adult level of underlying Tone 3 recognition at seven years old in the explicit tone recognition task. The seven-and eight-year-old children could recognize the tone of the initial syllable in the sandhi context as underlying tone 3, like the adults. This may link to the operation of the computation mechanism. When combing the results of the speech production, the mental representation features of Mandarin Tone 3 sandhi and the tone awareness task, we can find that the computation mechanism applies not only when combining the underlying forms with the pattern to new words, represents the surface forms with the underlying representations, it is also required in the decomposition process of the productive phonological alternations. Comparatively, children reached the adult level of underlying Tone 3 matching at eight years old in the implicit tone matching task. Lagging behind, the ratio of underlying Tone 3 matching in the seven-year-olds was still significantly lower than adults in the implicit tone matching task. Integrated the children's performance in younger age groups in the implicit tone awareness task as noted in chapter 6, it is reasonable to check whether the task is too difficult to reveal children's awareness of tones in the tone sandhi context.

Besides, the results revealed that four-to six-year-olds showed significantly less underlying Tone 3 recognition in the explicit tone recognition task compared with adults, this may suggest the involvement of the lexical mechanism at these age groups. However, this is not consistent with the results of chapters 3 to 5 in which children primarily took advantage of the lexical mechanism at three to five-year-olds, and made use of the computation mechanism to a large extent at six years old in the speech production task. The significantly lower underlying Tone 3 ratio in children at six than adults in the task can be explained by literacy effect and whether children had any prior explicit Pinyin instructions. To compare the differences between six-year-old children and adults in this chapter, the six-year-old were found only received a few prior explicit Pinyin instructions and only know a limited number of written forms of words.

When observing the influence of explicit Pinyin instructions on the performance of children in different age groups in the explicit and implicit tone awareness tasks, I found that the Pinyin instructions significantly facilitated the implicit tone awareness. Moreover, the overall results exhibited general developmental trajectories for both the underlying Tone 3 recognition and the matching, which are in line with children's degree of explicit Pinyin instructions. To be specific, the seven- and eight-year-old children, who had received explicit Pinyin instructions like adults, have been found to show larger ratios of underlying Tone 3 recognition and matching. While the four-year-old children who have never received any explicit Pinyin instructions were found to show the least ratio of underlying Tone 3 recognition and matching. The developmental trajectories of underlying Tone 3 recognition and matching ratios may not only because of the age but also the degree of prior explicit Pinyin instructions. To confirm this postulation, further studies will be conducted to associate the tone awareness in the sandhi context and Pinyin education with a better-controlled design.

7.4. The progress from the lexical mechanism to the computation mechanism

The thesis offers a comprehensive understanding of how Mandarin-speaking children learn to apply the Mandarin Tone 3 sandhi pattern, how the Tone 3 sandhi patterns are represented in children's mental lexicon, and how to internalize and decompose the pattern from the surface form during language development. The overall results reflect children's general developmental progress of language processing from the lexical mechanism to the computation mechanism, that is, children aged from three to five years old generally made use of a lexical mechanism to retrieve the surface form sandhi words from memory, while their underlying Tone 3 representations became adult-like via computation mechanism at seven years old. According to the previous literature (e.g., Ullman, 2001; Pinker & Prince, 1991), the lexical/computation distinction may be tied to the distinction between two well-studied brain memory systems, i.e., declarative memory and procedural memory. It has been suggested that the declarative memory is usually specialized for memorizing arbitrarily related information, while the procedural memory may be specialized for computing the regular sequences or grammatical patterns (Ullman, 2001; Ullman & Pierpont, 2005). Further studies will focus more on the relation between the operational mechanisms of tone sandhi and the two memory systems.

These developmental features found in children triggered a new question. What may support the developmental progress of language processing from the lexical mechanism to the computation mechanism? It is not difficult to imagine that during the process of language acquisition, both the surface and the underlying forms of Tone 3 words might be provided to children as input in daily lives. For example, the word /ɕi ɑu213/ 小 "small" may be frequently heard by children in different forms, such as /ɕi ɑu213 xua55/ 小花 “flower”, /ɕiɑu35 kou213/ 小狗 “dog”, or single word /ɕiɑu213/ 小. At the very beginning, children may only memorize the different forms as frozen chunks. At this stage, children already acquired the ability to recognize the lexical tones. For example, previous studies (see Singh & Fu, 2016, for a review) have demonstrated that infants have been able to recognize lexical tones before six months. However, if a child is asked to link the surface allophones of an initial syllable (e.g., /ɕiɑu213/) appearing in different tonal contexts (e.g., /ɕiɑu213 xua55/ and /ɕiɑu35 kou213/) to the same morpheme, the task becomes more complex and some more efforts are required. To be specific, in this process, children are expected to be able to break down the disyllabic word into syllables (involving syllable awareness) and link /ɕiɑu213/ and /ɕiɑu35/ to the same morpheme/word (involving morphological awareness). Syllable awareness was found to be acquired quite early in Mandarin-speaking children, that is around four years old (Shu et al., 2008). The acquisition of morphological awareness was quite late, in which free morphemes emerged between ages 3;10 and 5;0, whereas awareness of some bound morphemes was still not adult-like at seven years old (Hao et al., 2013). Children’s development from lexical mechanism to computation mechanism may be supported by an integrative effort of tone sensitivity, syllable awareness, as well as morphological awareness, which is therefore acquired gradually with age.

It is also worth noting that the developmental trajectory of the morphological awareness in Mandarin-speaking children (Hao et al., 2013) is more or less consistent with the progress of acquisition of tone sandhi. The morphological awareness in the abovementioned study refers to children's ability to detect whether words that have the same pronunciation were the same morpheme or different, homophonous morphemes. For example, participants were asked to judge which of the following three morphemes had a different meaning (/xuŋ35 tʂĄ 35/ 红茶 "black tea", /ly51 tʂĄ 35; 绿茶 "green tea", and /tɕ iæn213 tʂĄ 35/ 检查 "to examine"). If children were able to detect that the /tʂĄ 35/ in /tɕiæn213 tʂĄ 35/ had a different meaning from the other two, they were regarded as having acquired morphological awareness (Hao et al., 2013). Comparatively, Mandarin Tone 3 sandhi also involves morphological awareness which requires the ability to link the two different allophones that occur in different contexts to the same meaning. I found that the application tone sandhi pattern was still undergoing development at seven years old, in terms of the mental representations, which is quite consistent with the developmental trajectory of morphological awareness reported in previous studies (Hao et al., 2013). Again, these findings supported the claim that tone sandhi does not only require tone sensitivity, but it is also a process that demands integrations of other language abilities.

Importantly, great individual variations can be seen in all the experiments of my thesis. For example, in the speech production task concerning the productivity of the Mandarin Tone 3 sandhi, the individual application rates of adults and seven- to eight-year-old children were densely distributed above 85%. While the differences became larger in younger groups, especially in the three-year-olds. The large individual differences are also obvious in the lexical decision and tone awareness tasks. The individual variations may be driven by many factors. In the first place, previous explicit Pinyin instructions and literacy are expected to influence the developmental progress. The effects of explicit Pinyin instructions and literacy have been observed to some extent in the thesis. The seven- and eight-year-old children have received systematic Pinyin instructions. This may be one of the crucial driving forces for the development of tone sandhi. Moreover, children's vocabulary size, language inputs, children's IQ, and their statistical learning performances, may also influence the developing progress. Statistical learning, which is recognized as a foundational mechanism for the successful acquisition of spoken language, refers to the ability to implicitly detect recurring patterns and regularities in environmental input (e.g., Raviv & Arnon, 2018). Previous literature has demonstrated that procedural memory-based learning abilities connect to statistical learning. My future studies will focus more on systematically exploring those possible reasons for the development of tone sandhi.

Moreover, the degree of complexity of tone sandhi also influences the developmental progress. When trying to observe if different productive phonological alternations go through similar developmental progress, I compared the developmental productivity features of Mandarin Tone 3 sandhi in my thesis and Xiamen tone sandhi in Li and Mok (2019). The application of Mandarin Tone 3 sandhi in novel words is found to have reached the adult-like productivity at seven years old in this thesis. However, when investigating the productivity of Xiamen tone sandhi, which involves a complex tone sandhi circle, children as old as ten still have not fully acquired the tone sandhi pattern (Li & Mok, 2019). The Xiamen dialect is a dialectal group of Southern Min, and its sandhi pattern is therefore similar to that of Southern Min, which is more complex than the tone sandhi patterns in Mandarin. The adult native speakers are able to apply the sandhi patterns to pseudo-words of Xiamen tone sandhi items. However, the developmental timeline is delayed in those complex sandhis in Children who speak Xiamen dialect, and they tended to acquire those patterns in a much later stage (Li & Mok, 2019). These findings suggest that children showed different speeds of progress in reaching the adult-like productivity in acquiring the phonological alternations, the degree of complexity of tone sandhi also influences the developmental progress.

The findings that the acquisition progress of phonological alternations is different across languages/dialects and is influenced by the complexity of tone sandhi patterns are consistent with those found in morphological alternations. Previously, plenty of studies in Indo-European languages have provided compelling evidence that the acquisition progress of morphological alternations was not merely modulated by age growth (e.g., Gleason, 1958; Kerkhoff, 2007; Rumelhart & McClelland, 1987;

Slobin, 1985; Steven Pinker, 1999; Tomas et al., 2017; Bybee, 1995; Kellerman, 2019; Kirjavainen et al., 2012; Schuler et al., 2016). It was also influenced by many factors, among which the complexity of the alternation patterns was also found to play an important role. For example, children have already been found to be able to apply the plural noun inflection /-s/, which is relatively simple, to pseudo nouns at around the age three (e.g. Gleason, 1958; Bybee & Slobin, 1982; Steven Pinker, 1999). However, for the Russian morphological alternations, the process of which involves a complex interaction between vowel deletion and stress shift, the results showed that even the seven-year-olds still exhibited high variabilities in applying the morphological patterns (Tomas et al., 2017). Taken together, these findings suggested that the acquisition of both the phonological alternations and morphological alterations do not appear to follow similar developmental timelines across different languages/dialects, and that the complexity of the pattern itself matters. Finally, the developmental progress of phonological alternations and morphological alterations may also be affected by some other factors such as whether the alternation is regular or not, or whether it is phonetically motivated. The two factors have been verified to relate heavily to the developmental progress of some morphological and phonological alternations (e.g., Schuler et al., 2016; Slobin, 1985). My future studies will also focus on these unexplored questions.

Chapter 8 Conclusion

This project examined the productivity of Mandarin Tone 3 sandhi and its representation features. First, I tracked the developmental trajectory in children's application of Mandarin tone sandhi with an elicited picture naming test. I recruited seventy-four children from three- to eight years old and 13 adults as controls for this task. Three types of words, including real words, pseudo- and novel words, which are different in lexicality, were designed for the experiment. As can be seen in chapter 3 and chapter 4, I combined speech production judgment and acoustic analysis to evaluate their sandhi application features. The results of tone production judgment suggest that children showed relatively low productivity from three to five years old, this is especially true for the three-year-old children when producing pseudo- and novel words, which indicates the lexical mechanism is involved in these age groups under such conditions. Besides, children reached the adult-like application in pseudo-words at six-year-old, and in novel words condition at seven years old, indicating that children started to transit from using the lexical mechanism to applying the computational mechanism at the age of six, and fully developed to process tone sandhi via the computational mechanism at the age of seven. The result from the acoustic analysis basically confirmed the developing timeline for children reaching the adult-like computation mechanisms. However, it is noteworthy that children aged from six to eight showed a significant trend of overapplication in pseudo-words compared with real words. This may be due to the children's awareness of the sandhi pattern, and their active efforts to apply the pattern when trying to produce pseudo-words.

In addition to children's performance in speech production, I further observed how the sandhi items are represented in children's mental lexicon as they are approaching adult-like productivity via the computation mechanism. I recruited 107 children from age four to age eight as well as adults as controls to achieve this goal. I adopted an auditory priming lexical decision task, in which the prime words were designed to match either the surface (Tone 2) or underlying form (Tone 3) of the initial syllable of the tone sandhi words, or not match in terms of tones. If the sandhi items are represented in the underlying form, we expected a facilitatory effect from Tone 3. If they are represented in the surface form, we expected a significant facilitatory effect from Tone 2. While if the tone sandhi items are represented as multi-variant exemplars, we expected significant facilitatory effects from both the surface Tone 2 and the underlying forms. The results showed that the representation features of sandhi items in mental lexicon related heavily to their operational mechanisms. That is, children's representation features of tone sandhi were also going through development in line with the development of their operational mechanisms in that tone sandhi was primarily represented in the surface form at four and five-year-old, and went through a gradual development after that. Nonetheless, the results showed that even the seven and eight-year-old children have not established an adult way of underlying representation of tone sandhi. These findings suggested/confirmed that the lexical mechanism is primarily involved in children aged four and five. The seven and eight-year-old children still have not reached the adult-like underlying representation that links to the full use of computation mechanism, which is inconsistent with the findings from my productivity experiments. There are primarily three possible reasons for this discrepancy. First, it may result from children's limited language experiences. Secondly, since the priming task used the real words as targets, individuals' degree of literacy may influence their ability to break down the disyllabic words into smaller morphological units. Indeed, the six to eight-year-old children had only learned some of the words beforehand. Due to this literacy effect, children may not reach the adult-like representation even by seven and eight years old. Thirdly, it is possible that the magnitude of lexical frequency may affect the operation of Mandarin Tone 3 sandhi. Although Mandarin Tone 3 sandhi has been shown to operate via the computation mechanism, this does not necessarily imply that all Tone 3 sandhi words are manipulated by the same degree of computation process. For highly familiar words, they may not necessarily take advantage of the computation mechanism, because the sandhi form may be retrieved directly from the mental lexicon, which means that those real words could be encoded by the lexical mechanism. Therefore, both the surface and underlying priming effect, which links more to the lexical mechanism, were found in children who haven’t fully established the operational mechanism of computation.

My third experiment explored how individuals decompose sandhi items from larger units. The details are elaborated in chapter 7. In order to trace the developmental trajectory, the study recruited 86 children from four to eight, as well as adults as controls, and used the well-established explicit and implicit phonological awareness tasks to examine their tone recognition and tone matching features. The results showed that the gradual development in the explicit and implicit tone awareness of the initial tone sandhi words in children. Children reached the adult level of explicit processing at seven years old, while they reached the adult-level of underlying processing in the implicit task at eight years old. The adult-like underling Tone 3 recognition in seven years old in the explicit tone awareness can also be linked to the operation of the computation mechanism. Further studies should be conducted with an easier implicit tone awareness task to confirm the implicit awareness of the tones with tonal alternations.

The overall picture of the study offered us a general understanding of how native Mandarin children process phonological alternations, particularly tone sandhi, during language development. The phonological Mandarin Tone 3 sandhi pattern has been tested to be highly productive and operated via a computation mechanism in Mandarin-adult speakers. From the current study, it can be seen that children were not as productive as adults from the beginning. They primarily rely on the lexical mechanism to produce real words with Tone 3 sandhi at an early age, and probably have not reached adult-like computational processing until around seven years old.

However, the thesis is not without limitations. First, this thesis did not control individuals' cognitive and linguistic abilities very well, such as children's levels of literacy and degrees of Pinyin instructions, vocabulary size, IQ, and their statistical learning ability, which may have influenced the results to different extents. Secondly, there are different theoretical accounts for the representation features of tone sandhi in the mental lexicon of Mandarin-speaking adults. The thesis only tested part of them currently. Based on the current literature and findings, the overall pictures for the representation of Mandarin Tone 3 sandhi in adult speakers should be: the underlying representation may be activated when hearing the sandhi words, while both the surface and the abstract underlying forms are activated when producing the words (supporting the multi-variant account). The multi-variant account for Mandarin Tone 3 sandhi has been verified in the encoding of tone sandhi words in adult speakers (Chen et al., 2011; Nixon et al., 2015). However, it is still unclear if the account also applies to children in different age groups. It would be meaningful to observe children's progress on the representation of tone sandhi in speech production.

Considering the limitations in the thesis, my future studies will combine the factors that may influence the operational mechanisms of tone sandhi, for example, literacy effect, Pinyin instructions, children's vocabulary size, IQ, and their statistical learning ability, to achieve a better understanding of the acquisition of tonal alternations. In addition to this, I will also explore the developmental features of tone sandhi with low productivity such as that in Southern Min, and a further look at the neuro processing mechanisms of the two types of tone sandhi patterns in Chinese dialects from developmental perspectives. Moreover, more models will be considered and tested in future studies. It is expected that these studies will broaden our comprehension of tone sandhi in nature. More importantly, these studies will provide empirical evidence to the theoretical models of speech production, language acquisition, and phonology.

Reference

Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*(2), 119–161. https://doi.org/10.1016/S0010-0277(03)00146-X

Anthony, J. L. , Lonigan, C. J. , Driscoll, K. , Phillips, B. M. , & Burgess, S. R. . (2003). Phonological sensitivity: A quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly*, *38*(4), 470–487. https://doi.org/10.1598/rrq.38.4.3

Anthony, J. L., & Francis, D. J. (2005). Development of phonological awareness. *Current Directions in Psychological Science*, *14*(5), 255–259. https://doi.org/10.1111/j.0963-7214.2005.00376.x

Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., Phillips, B. M., & Cantor, B. G. (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology*, *82*(1), 65–92. https://doi.org/10.1006/jecp.2002.2677

Arnon, I. (2011). *Units of learning in language acquisition*. *2011*, 167–178. https://doi.org/10.1075/tilar.7.10arn

Baayen, R. H. (2008). Analyzing linguistic data: A practical introduction to statistics using R. Cambridge: Cambridge University Press.

Barr, D., Levy, R., Scheepers, C., & Tily, H. J. (2013). Keep it maximal. *Journal of Memory and Language*, *68*(3), 1–43. https://doi.org/10.1016/j.jml.2012.11.001.Random

Bloom, P. (1994). *Language Acquisition : Core Readings.* 1st MIT Press ed. Cambridge, Mass.: MIT,.

Boersma, P., & Weenink, D. (2016). *Praat: Doing phonetics by computer* (6.0.12).

Bryant, P. E., MacLean, M., Bradley, L. L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology*, *26*(3), 429–438. https://doi.org/10.1037//0012-1649.26.3.429

Bybee, J. (1995). Regular morphology and the lexicon. *Language and Cognitive Processes*, *10*(5), 425–455. https://doi.org/10.1080/01690969508407111

Bybee, J. & E. P. (1981). On lexical and morphological conditioning of alternations : a nonce-probe experiment with Spanish verbs. *Linguistics*, *19*, 937–968.

Bybee, J. L., & Slobin, D. I. (1982). Rules and Schemas in the Development and Use of the English past Tense. *Language*, *58*(2), 265. https://doi.org/10.2307/414099

Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE*, *5*(6). https://doi.org/10.1371/journal.pone.0010729

Case, J. (2005). Theory of U-Shaped Learning. *Lecture Notes in Artificial Intelligence*, *3734*(2), 241–255. https://www.eecis.udel.edu/~case/slides/nugget-ushape.pdf

Castles, A., & Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, *91*(1), 77–111. https://doi.org/10.1016/S0010-0277(03)00164-1

Cazden, C. B. (1968). The Acquisition of Noun and Verb Inflections. *Child Development*, *39*(2), 433–448.

Chan, M. K. M., & Ren, H. M. (1989). Wuxi tone sandhi: From last to first syllable dominance. *Acta Linguistica Hafniensia*, *21*, 35–64.

Chao, Y-R. (1948). *Mandarin Primer*. Harvard University Press.

Chao, Y. R. (1973). The Cantian Idiolect: An analysis of the Chinese spoken by a Twenty-eight-month-old child. In *In C.A. Ferguson & D.I. Slobin (Eds.), Studies of child language development.* (pp. 13–33).

Chao, Yuan-Ren. (1968). *A grammar of spoken Chinese*. CA: University of California Press.

Chao, Yuen Ren. (1948). *Mandarin Primer*. Cambridge: Harvard University Press.

Chen, F., Yan, N., Huang, X., Zhang, H., Wang, L., & Peng, G. (2016). Development of Mandarin onset-rime detection in relation to age and Pinyin instruction. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, *08-12-Sept*(September), 180–184. https://doi.org/10.21437/Interspeech.2016-895

Chen, M. Y. (2000). *Tone Sandhi Patterns across Chinese dialects*. cambridge university press.

Chen, X. (2012). *On the role of the underlying tonal representation in the phonological encoding stage of tone sandhi production of the Mandarin third tone*. *May 2012*. https://doi.org/10.13140/RG.2.2.30813.56802

Chen Y, Shen R, S. N. (2011). Representation of allophonic tone sandhi variants. *Proceedings of the Satellite Workshop to ICPhS: Psycholinguistics Representation of Tone*, 38–41.

Chen, Y., Shen, R., & Schiller, N. (2011). Representation of Allophonic Tone Sandhi Variants. *Proceedings of Psycholinguistics Representation of Tone. Satellite Workshop to ICPhS*, *August*, 38–41.

Chen, Y., & Yuan, J. (2007). A corpus study of the 3rd tone sandhi in standard Chinese. *International Speech Communication Association - 8th Annual Conference of the International Speech Communication Association, Interspeech 2007*, *4*(JANUARY 2007), 2808–2811.

Cheng, R. L. (1968). Tone sandhi in Taiwanese. *Linguistics*, *41*, 19–42.

Chien, Y. F., Sereno, J. A., & Zhang, J. (2016). Priming the representation of Mandarin tone 3 sandhi words. *Language, Cognition and Neuroscience*, *31*(2), 179–189. https://doi.org/10.1080/23273798.2015.1064976

Chien, Y. F., Sereno, J. A., & Zhang, J. (2017). What's in a Word: Observing the Contribution of Underlying and Surface Representations. *Language and Speech*, *60*(4), 643–657. https://doi.org/10.1177/0023830917690419

Cholin, J., & Levelt, W. J. M. (2009). Effects of syllable preparation and syllable frequency in speech production: Further evidence for syllabic units at a post-lexical level. *Language and Cognitive Processes*, *24*(5), 662–684. https://doi.org/10.1080/01690960802348852

Clark, E. V. (2016). *First Language Acquisition*. Cambridge University Press.

Clumeck, H. V. (1977). *Studies in the acquisition of Mandarin phonology (Unpublished doctoral dissertation)*. University of California, Berkeley.

Crowley, K. E., & Colrain, I. M. (2004). A review of the evidence for P2 being an independent component process: age, sleep and modality. *Clinical Neurophysiology*, *115*(4), 732–744.

Cynthiac M. Connine, Larissa J. Ranbom, D. J. P. (2008). Processing variant forms in spoken word recognition: The role of variant frequency. *Perception &*

Psychophysics, *70*(3), 403–411. https://doi.org/10.3758/PP

Dick, F., Krishnan, S., Leech, R., & Curtin, S. (2015). Language Development. *Neurobiology of Language*, 373–388. https://doi.org/10.1016/B978-0-12-407794-2.00031-6

Duanmu, S. (2007). *The phonology of Standard Chinese. 2nd edn*. xford University Press.

Frisch, Stefan A., Janet B. Pierrehumbert, and M. B. (2004). Similarity avoidance and the OCP. *Natural Language and Linguistic Theory*, *22*, 179–228.

Gleason, J. B. (1958). The Child's Learning of English Morphology. *Word*, *14*(3–4), 150–177. https://doi.org/10.1111/j.1467-1770.1969.tb00463.x

Goldinger, S. D. (1998). Echoes of Echoes? An Episodic Theory of Lexical Access. *Psychological Review*, *105*(2), 251–279. https://doi.org/10.1037/0033-295X.105.2.251

Goodluck, H. (1991). *Language Acquisition : A Linguistic Introduction.* Oxford, UK ; Cambridge, USA: Blackwell,.

Gow, D. W. (2001). Assimilation and Anticipation in Continuous Spoken Word Recognition. *Journal of Memory and Language*, *45*(1), 133–159. https://doi.org/10.1006/jmla.2000.2764

Gwet, K. (2014). *Handbook of inter-rater reliability fourth edition*. Advanced Analytics, LLC.

Haebig, E., & Weismer, S. E. (2015). Lexical Processing in School-Age Children with Autism Spectrum Disorder and Children with Specific Language Impairment: The Role of Semantics. *Journal of Autism and Developmental Disorders, July*, 1–26. https://doi.org/10.1007/s10803-015-2534-2

Hao, M., Chen, X., Dronjic, V., Shu, H., & Anderson, R. C. (2013). The development of young Chinese children's morphological awareness: The role of semantic relatedness and morpheme type. *Applied Psycholinguistics*, *34*(1), 45–67. https://doi.org/10.1017/S0142716411000609

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods*, *40*(1), 728–733. https://doi.org/10.3758/BRM.40.3.728

Høien, T., Lundberg, I., Stanovich, K. E., & Bjaalid, I. K. (1995). Components of phonological awareness. *Reading and Writing*, *7*(2), 171–188. https://doi.org/10.1007/BF01027184

Hsieh, H.-I. (1970). The psychological reality of tone sandhi rules in Taiwanese. *6th Meeting of the Chicago Linguistic Society*, 489–503.

Hsieh, H.-I. (1976). On the unreality of some phonological rules. *Lingua*, *38*, 1–19.

Hsieh, Hsin-I. (1975). How generative is phonology? In E. F. K. K. (ed.) (Ed.), *The transformational-generative paradigm and modern linguistic theory* (pp. 109–144). Amsterdam: Benjamins.

Indefrey, P. (2011). The spatial and temporal signatures of word production components: A critical update. *Frontiers in Psychology*, *2*(OCT), 1–16. https://doi.org/10.3389/fpsyg.2011.00255

Jane, J. M. and, & Tsay. (2003). Investigating the Phonetics of Mandarin Tone Sandhi. *Taiwan Journal of Linguistics*, *1*(1), 29–68. https://doi.org/10.6519/TJL.2003.1(1).2

Jenny R. Saffran, Ann Senghas, and J. C. T. (2001). The acquisition of language by children. *PNAS*, *98*(23), 12874–12875. https://doi.org/10.1073/pnas.231498898

Jie, T. (2013). *A Case Study on a Mandarin Chinese Child's Tone Sandhi Rules Acquisition*.

Joshua A. Banta, Martin H. H Stevens, and M. P. (2010). A comprehensive test of the "limiting resources" framework applied to plant tolerance to apical meristem damage. *Oikos*, *119*(2), 359–369. http://dx.doi.org/10.1111/j.1600-0706.2009.17726.x%0D

Kellerman, E. (2019). *Modeling Children ' s Early Linguistic Productivity Through the Automatic Discovery and Use of Lexically*. 782–788.

Kerkhoff, A. (2007). *Acquisition of Morpho-Phonology: The Dutch voicing alternation*. Utrecht University.

Kiparsky, P. (1973). Abstractness, opacity and global rules. In O. Fujimura (Ed.), *Three Dimensions of Linguistic Theory* (pp. 57–86). Tokyo Institute for Advanced Studies of Language.

Kirjavainen, M., Nikolaev, A., & Kidd, E. (2012). The effect of frequency and phonological neighbourhood density on the acquisition of past tense verbs by Finnish children. *Cognitive Linguistics*, *23*(2), 273–315. https://doi.org/10.1515/cog-2012-0009

Kratochvil, P. (1984). Phonetic tone sandhi in Beijing dialect stage speech. *Cahiers de Linguistique - Asie Orientale*, *13*(2), 135–174. https://doi.org/10.3406/clao.1984.1154

Kuznetsova, A., Brockhoff, P., & Christensen, R. *2016. Tests in linear mixed effects models. R software package*.

Lahiri, A., Jongman, A., & Sereno, J. (1990). [笔记]The pronominal clitic [d\r] in Dutch: A theoretical and experimental approach. *Yearbook of Morphology 3*, *3*, 115–127.

Landi, N., Crowley, M. J., Wu, J., Bailey, C. A., & Mayes, L. C. (2012). Deviant ERP response to spoken non-words among ado- lescents exposed to cocaine in utero. *Brain and Language*, *120*(3), 209–216.

Lee, C. Y. (2007). Does horse activate mother? Processing lexical tone in form priming. *Language and Speech*, *50*(1), 101–123. https://doi.org/10.1177/00238309070500010501

Lee, K. A. (2002). *Chinese Tone Sandhi and Prosody*. University of Illinois.

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

Levelt, W. J. M. (1999). Models of word production. *Trends in Cognitive Sciences*, *3*(6), 223–232. https://doi.org/10.1016/S1364-6613(99)01319-4

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*(1), 1–75. https://doi.org/10.1017/S0140525X99001776

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language*, *4*(2), 185–199. https://doi.org/10.1017/S0305000900001598

Li, Q., & Chen, Y. (2016). An acoustic study of contextual tonal variation in Tianjin Mandarin. *Journal of Phonetics*, *54*, 123–150. https://doi.org/10.1016/j.wocn.2015.10.002

Li, Q., Chen, Y., & Xiong, Z. (2019). Illustrations of the ipa: Tianjin Mandarin.

Journal of the International Phonetic Association, *49*(1), 109–128. https://doi.org/10.1017/S0025100317000287

Li, X., & Chen, Y. (2015). Representation and processing of lexical tone and tonal variants: Evidence from the mismatch negativity. *PLoS ONE*, *10*(12). https://doi.org/10.1371/journal.pone.0143097

Li, X. L., Peggy. P. K. M. (2019). The Acquisition of Xiamen Tone Sandhi by Children. *HISPhonCog 2019: Hanyang International Symposium on Phonetics & Cognitive Sciences of Language 2019 May 24-25, 2019, Hanyang University, Seoul, Korea*, 3–4.

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

Luce, P. A., McLennan, C., & Charles-Luce, J. (2003). Abstractness and specificity in spoken word recognition: Indexical and allophonic variability in long-term repetition priming. *Rethinking Implicit Memory*, 197–214. https://doi.org/10.1093/acprof:oso/9780192632326.003.0009

Ma, W., Zhou, P., Singh, L., & Gao, L. (2017). Spoken word recognition in young tone language learners: Age-dependent effects of segmental and suprasegmental variation. *Cognition*, *159*, 139–155. https://doi.org/10.1016/j.cognition.2016.11.011

Marcus, G. F., Brinkmann, U., Clahsen, H., Wiese, R., & Pinker, S. (1995). *German* *inflection: the exception that proves the rule*. *29*(3), 189–256.

McBride-Chang, C., Lam, F., Lam, C., Doo, S., Wong, S. W. L., & Chow, Y. Y. Y. (2008). Word recognition and cognitive profiles of Chinese pre-school children at risk for dyslexia through language delay or familial history of dyslexia. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *49*(2), 211–218. https://doi.org/10.1111/j.1469-7610.2007.01837.x

Mei, T.-L. (1977). Tones and Tone Sandhi in 16th Century Mandarin. In *Journal of Chinese Linguistics* (Vol. 5, pp. 237–260).

Meyer, A. S. (1990). The time course of phonological encoding in language production: The encoding of successive syllables of a word. *Journal of Memory and Language*, *29*, 524–545.

Meyer, A. S. (1991). The Time Course of Phonological Encoding in Language Production: Phonological Encoding Inside a Syllable. *Journal of Memory and Language*, *30*, 69–89.

Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: Evidence from a learning paradigm. *Cognitive Science*, *35*(1), 184–197. https://doi.org/10.1111/j.1551-6709.2010.01140.x

Nixon, J. S., Chen, Y., & Schiller, N. O. (2015). Multi-level processing of phonetic variants in speech production and visual word processing: evidence from

Mandarin lexical tones. *Language, Cognition and Neuroscience*, *30*(5), 491–505. https://doi.org/10.1080/23273798.2014.942326

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian Model of Continuous Speech Recognition. *Psychological Review*, *115*(2), 357–395. https://doi.org/10.1037/0033-295X.115.2.357

Peng, S.-H. (2000). *"Lexical versus 'phonological' representations of Mandarin tones," in Language Acquisition and the Lexicon: Papers in Laboratory Phonology V, eds M. B. Broe, and J. B. Pierrehumbert, (Cambridge: Cambridge University Press)* (pp. 152–167).

Peters, A. M. (1977). Language learning strategies: Does the whole equal the sum of the parts? *Language*, *53*, 560–573.

Pierrehumbert, J. (2006). The statistical basis of an unnatural alternation. In & C. B. (Eds. . L. Goldstein, D. H. Whalen (Ed.), *Laboratory phonology VIII, varieties of phonological competence* (pp. 81–107). Berlin: Mouton de Gruyter.

Pierrehumbert, J. B. (2016). Phonological Representation: Beyond Abstract Versus Episodic. *Annual Review of Linguistics*, *2*, 33–52. https://doi.org/10.1146/annurev-linguistics-030514-125050

Pinker, S. & A. P. (1994). *The Reality of Linguistic Rules.* (R. C. and G. K. I. (eds. . In Susan D. Lima (Ed.); pp. 321–351.). Amsterdam: John Benjamins.

Pitt, M. A. (2009). Test of Generalization To Newly Learned Words. *J Mem Lang*, *61*(1), 19–36. https://doi.org/10.1016/j.jml.2009.02.005.How

Politzer-Ahles, S., Schluter, K., Wu, K., & Almeida, D. (2016). Asymmetries in the perception of Mandarin tones: Evidence from mismatch negativity. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(10), 1547–1570. https://doi.org/10.1037/xhp0000242

Psychology Software Tools, I. (2017). *E-Prime: Documentation Article.* https://doi.org/https://support.pstnet.com.

R CoreTeam. 2014.R: A language and environment for statistical computing.(Version3.1.0)[Computer Program]. Available at: 〈http://www.R project.org/〉. (n.d.).

Rattanasone, N. X., Ping, T., & Yuen, I. (2018). *Mandarin-Speaking 3-year-olds Have Productive Knowledge of Tone Sandhi*. *July*, 141–142. https://doi.org/10.13140/RG.2.2.27938.73926

Rattanasone, N. X., Tang, P., Yuen, I., Gao, L., & Demuth, K. (2018). Five-year-olds' acoustic realization of mandarin tone sandhi and lexical tones in context are not yet fully adult-like. *Frontiers in Psychology*, *9*(MAY), 1–10. https://doi.org/10.3389/fpsyg.2018.00817

Raviv, L., & Arnon, I. (2018). *The developmental trajectory of children ' s auditory* *and visual statistical learning abilities : modality- - based differences in the effect of age*. *May 2017*, 1–13. https://doi.org/10.1111/desc.12593

Rumelhart, D. E., & McClelland, J. L. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Mechanisms of language aquisition* (pp. 195–248).

Schuler, K. D., Yang, C., & Newport, E. L. (2016). Testing the Tolerance Principle: Children form productive rules when it is more computationally efficient to do so. *Proceedings of the 38th Annual Conference of the Cognitive Science Society*, *5*(1958), 2321–2326.

Selby, S. (1972). the Development of Morphological Rules in Children. *British Journal of Educational Psychology*, *42*(3), 293–299. https://doi.org/10.1111/j.2044-8279.1972.tb00722.x

Sereno, J. A., & Lee, H. (2015). The Contribution of Segmental and Tonal Information in Mandarin Spoken Word Processing. *Language and Speech*, *58*(2), 131–151. https://doi.org/10.1177/0023830914522956

Shen, J., Deutsch, D., & Rayner, K. (2013). On-line perception of Mandarin Tones 2 and 3: Evidence from eye movements. *The Journal of the Acoustical Society of America*, *133*(5), 3016–3029. https://doi.org/10.1121/1.4795775

Shen, X. S. (1992). On tone sandhi and tonal coarticulation. *Acta Linguistica*

Hafniensia, *25*(1), 83–94. https://doi.org/10.1080/03740463.1992.10412279

Shu, H., Peng, H., & McBride-Chang, C. (2008). Phonological awareness in young Chinese children. *Developmental Science*, *11*(1), 171–181. https://doi.org/10.1111/j.1467-7687.2007.00654.x

Singh, L., & Fu, C. S. L. (2016). A New View of Language Development: The Acquisition of Lexical Tone. *Child Development*, *87*(3), 834–854. https://doi.org/10.1111/cdev.12512

Siok, W.T., & Fletcher, P. (2001). The role of phonological awareness and visual-orthographic skills in Chinese reading acquisition. *Development Psychology, 37*, 886–899.

Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. In I. D. I. S. (ed.), *The crosslinguistic study of language acquisition.* (pp. 1157–1256).

Stephen D. Goldinger. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(5), 1166–1183.

Steven Pinker. (1998). Words and Rules. *Lungua*, *106*, 219–242.

Steven Pinker. (1999). Kids Say the Darnedest Things. In *words and rules* (pp. 189–210).

Steven Pinker. (1998). Words and Rules. *Lungua*, *106*, 219–242.

Sumner, M., Kim, S. K., King, E., & McGowan, K. B. (2014). The socially weighted encoding of spoken words: A dual-route approach to speech perception. *Frontiers in Psychology*, *4*(JAN), 1–13. https://doi.org/10.3389/fpsyg.2013.01015

Tang, P., Yuen, I., Rattanasone, N. X., Gao, L., & Demuth, K. (2019). The acquisition of phonological alternations: The case of the Mandarin tone sandhi process. *Applied Psycholinguistics*, *40*(6), 1495–1526. https://doi.org/10.1017/S0142716419000353

Tomas, E., Van De Vijver, R., Demuth, K., & Petocz, P. (2017). Acquisition of nominal morphophonological alternations in Russian. *First Language*, *37*(5), 453–474. https://doi.org/10.1177/0142723717698839

Tong, X., Tong, X., & McBride-Chang, C. (2015). Tune in to the Tone: Lexical Tone Identification is Associated with Vocabulary and Word Recognition Abilities in Young Chinese Children. *Language and Speech*, *58*(4), 441–458. https://doi.org/10.1177/0023830914562988

Treiman, R. (1987). Levels of Phonological Awareness. In *the Annual Meeting of the American Educational Research Association*.

Ullman, M., & Pinker, S. (2002). The past and future of the past tense. *Trends in*

Cognitive Sciences, *6*(11), 456–463.

Ullman, M. T. (1999). Acceptability ratings of regular and irregular past-tense forms: Evidence for a dual-system model of language from word frequency and phonological neighbourhood effects. *Language and Cognitive Processes*, *14*(1), 47–67. https://doi.org/10.1080/016909699386374

Ullman, M. T. (2001). The declarative/procedural model of lexicon and grammar. *Journal of Psycholinguistic Research*, *30*(1), 37–69. https://doi.org/10.1023/A:1005204207369

Ullman, M. T., & Pierpont, E. I. (2005). Specific language impairment is not specific to language: The procedural deficit hypothesis. *Cortex*, *41*(3), 399–433. https://doi.org/10.1016/S0010-9452(08)70276-4

Van Der Lely, H. K. J., & Christian, V. (2000). Lexical word formation in children with grammatical SLI: A grammar-specific versus an input-processing deficit? *Cognition*, *75*(1), 33–63. https://doi.org/10.1016/S0010-0277(99)00079-7

Vijver, R. Van De, & Baer-henney, D. (2006). *On the development of the productivity of plural suffixes in German*. 1–3.

Wan, I. P., & Jaeger, J. (1998). Speech errors and the representation of tone in Mandarin Chinese. *Phonology*, *15*(3), 417–461. https://doi.org/10.1017/S0952675799003668

Wang, C. (2011). *Children's Acquisition of Tone 3 sandhi in Mandarin*.

Wang, H. S. (1993). Taiyu biandiao de xinli texing. [On the psychological status ofT Taiwanese tone sandhi.]. *Tsinghua Xuebao [Tsinghua Journal of Chinese Studies]*, *23*, 175–192.

Wang, W. S.-Y., & Li, K.-P. (n.d.). Tone 3 in pekinese. *Journal of Speech & Hearing Research*, *10*(3), 629.

Wang, W. S. (1967). *Phonological Features of Tone Linked references are available on JSTOR for this article : International Journal of American Linguistics*. *33*(2), 93–105.

Wewalaarachchi, T. D., & Singh, L. (2016). Effects of suprasegmental phonological alternations on early word recognition: Evidence from tone Sandhi. *Frontiers in Psychology*, *7*(MAY), 1–14. https://doi.org/10.3389/fpsyg.2016.00627

Wong, PS; Schwartz, RG; Jenkins, J. (2005). *Perception and production of lexical tones by 3-year-old, Mandarin-speaking children*. *48*(5), 1065–1079.

Wong, P. (2012). Acoustic characteristics of three-year-olds' correct and incorrect monosyllabic Mandarin lexical tone productions. *Journal of Phonetics*, *40*(1), 141–151. https://doi.org/10.1016/j.wocn.2011.10.005

Wong, P., & Strange, W. (2017). Phonetic complexity affects children's Mandarin tone production accuracy in disyllabic words: A perceptual study. *PLoS ONE*, *12*(8), 1–21. https://doi.org/10.1371/journal.pone.0182337

Xinjie, W., Jin, Y. U., Hui, L. I., Shanshan, Z., & Yiming, Y. (2015). *A Study of Tone Sandhi Acquisition of Speaking Children*. https://doi.org/10.16499/j.cnki.1003-5397.2015.04.018

Xu, B.-H., Tang, Z.-Z., & Qian, N.-R. (1981). Xinpai Shanghai fangyan de liandu biandiao (Tone sandhi in New Shanghai). *Fangyan (Dialects)*, 145–155.

Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, *25*(1), 61–83. https://doi.org/10.1006/jpho.1996.0034

Yan, H., & Zhang, J. (2016). Pattern substitution in Wuxi tone sandhi and its implication for phonological learning. *International Journal of Chinese Linguistics*, *3*(1), 1–44. https://doi.org/10.1075/ijchl.3.1.01yan

Yang, C. (2016). *The Price of Linguistic Productivity: How Children Learn to Break the Rules of Language*. 280. https://doi.org/10.7551/mitpress/9780262035323.001.0001

Yang, C. (2018). *A User's Guide to the Tolerance Principle*. *July*, 1–12.

Yin, L., Li, W., Chen, X., Anderson, R. C., Zhang, J., Shu, H., & Jiang, W. (2011). The role of tone awareness and pinyin knowledge in Chinese reading. *Writing Systems Research*, *3*(1), 59–68. https://doi.org/10.1093/wsr/wsr010

Yip, M. J. W. (1980). *The tonal phonology of Chinese*. 1–18.

Yuan, J., & Chen, Y. (2014). 3rd tone sandhi in Standard Chinese: A corpus approach. *Journal of Chinese Linguistics*, *42*(1), 218–236.

Zamuner, T. S., Kerkhoff, A., & Fikkert, P. (2006). Acquisition of Voicing Neutralization and Alternations in Dutch. *Proceedings of the 30th Annual Boston University Conference on Language Development*, 701–712.

Zee, E., and Maddieson, I. (1980). Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis. *Glossa*, *14*, 45–88.

Zhang and Lai, Yuwen, J. (2006). Testing the role of phonetic naturalness in Mandarin tone sandhi. *Kansas Working Papers in Linguistics*, *28*, 65–126. https://doi.org/10.1017/S0952675710000060

Zhang, C. C., & Gang, P. (2013). Productivity of Mandarin Third Tone sandhi: A wug test. *Eastward Flows the Great River: Festschrift in Honor of Prof. William S-Y. Wang on His 80th Birthday*, *August*, 256–282.

Zhang, C., Xia, Q., & Peng, G. (2015). Mandarin third tone sandhi requires more effortful phonological encoding in speech production: Evidence from an ERP study. *Journal of Neurolinguistics*, *33*, 149–162. https://doi.org/10.1016/j.jneuroling.2014.07.002

Zhang, H., & Lin, C. (1979). *Revised Chinese version of Wechsler Intelligence Scale* *for 6-16 year-old Children (WISC-CR)*.

Zhang, J. (2014). Tones, Tonal Phonology, and Tone Sandhi. *The Handbook of Chinese Linguistics*, 443–464. https://doi.org/10.1002/9781118584552.ch17

Zhang, J. (2016). Using Nonce-probe Tests and Auditory Priming to Investigate Speakers' Phonological Knowledge of Tone Sandhi. *Tonal Aspects of Languages 2016*, *2016*(May), 12–18. https://doi.org/10.21437/tal.2016-3

Zhang, J. (2018). *Speakers Treat Transparent and Opaque Alternation Patterns Differently — Evidence from Chinese Tone*.

Zhang, J., & Lai, Y. (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology*, *27*(1), 153–201. https://doi.org/10.1017/S0952675710000060

Zhang, J., & Lai, Y. (2008). Phonological Knowledge beyond the Lexicon in Taiwanese Double Reduplication. *Proceedings of the 27th West Coast Conference on Formal Linguistics*, 458–466.

Zhang, J., Lai, Y., & Sailor, C. (2009). Opacity, phonetics, and frequency in Taiwanese tone sandhi. In *Current issues in unity and diversity of languages: Collection of papers selected from the 18th International Congress of Linguists* (pp. 3019–3038). Linguistic Society of Korea.

Zhang, J., Lai, Y., & Sailor, C. (2011). Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua*, *121*(2), 181–206. https://doi.org/10.1016/j.lingua.2010.06.010

Zhang, J., Lai, Y., & Turnbull-Sailor, C. (2006). Wug-testing the “tone circle” in Taiwanese. *Proceedings of the 25th West Coast Conference on Formal Linguistics*, *25*, 453–461.

Zhang, J., & Meng, Y. (2016). Structure-dependent tone sandhi in real and nonce disyllables in Shanghai Wu. *Journal of Phonetics*, *54*, 169–201. https://doi.org/10.1016/j.wocn.2015.10.004

Yip, M. (2002). *Tone . (Cambridge Textbooks in Linguistics.) Cambridge : Cambridge University Press.*

Zhou, X., & Marslen-Wilson, W. D. (1997). The abstractness of phonological representation in the Chinese mental lexicon. In *Cognitive Processing of Chinese and Related Asian Languages* (pp. 3–26).

Zhu, H. (2002). Phonological development in specific contexts: Studies of Chinese-speaking children. In *Clevedon, England: Multilingual Matters*.

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

Zhu, X.-N. (1999). Shanghai tonetics. *München: Lincom Europa*.

Zhu, X. (2005). *Shanghai sheng diao shi yan lu [An experimental study in Shanghai tones]. Shanghai: Shanghai jiao yu chu ban she.*