Acquisition of Tone Sandhis by English Speaking Learners of Chinese

Wenhua Jin
Kennesaw State University

Abstract

This study examines the acquisition of five Chinese tone sandhis, including Yi, Bu, T2, Half-T3, and T3 Sandhis by English speaking learners of Chinese. Fifty-three college students from different levels of Chinese classes participated in this study. Investigation results show different acquisition patterns for the various sandhi processes and learner levels, revealing thereby multiple factors at play in L2 learners' acquisition of the tone sandhis: the instruction factor, phonetic motivation, learning effect, inherent intra-lingual factor, and interference from English intonation. This study provides valuable implications for L2 Chinese pedagogy as well as Second Language Acquisition in general and also presents new evidence for the understanding of the long debated nature of Chinese Tone 3 from the perspective of L2 Chinese acquisition.

Keywords: tone sandhi, tone acquisition, L2 Chinese, Chinese pronunciation

1. Introduction[1]

Mandarin Chinese (Chinese hereafter) is long believed to be one of the most difficult languages to learn as an adult, especially due to its complicated tonal system (Bluhme and Burr, 1971; Chen, 2000; Shen, 1989; Shih, 1986; Sun, 1998; Wang et al., 1999; Ross, 2001; H. Zhang, 2013). As evidenced in previous studies, problems with suprasegmental features contribute more considerably to the L2 foreign accent than those from segments (Shen, 1988, 1989). At the suprasegmental tonal level, Chinese has four basic tones as well as a short and weak neutral tone. Traditionally, the four basic tones are called Yinping 阴平, Yangping 阳平, Shangsheng 上声, and Qusheng 去声 (Chen 1975). Chao (1930) designed a five-level numerical scale to represent the pitch height, ranging from 1 through 5 with the latter being the highest pitch level. In this representation system, the four basic tones are labeled as "55", "35", "214", and "51" respectively. They are also described as high-level, mid-rising, low-dipping, and high-falling tones (Chao, 1968). In language instruction, these four basic tones are usually referred to as Tone 1 (T1), Tone 2 (T2), Tone 3 (T3), and Tone 4 (T4), and are marked with iconic diacritics above the nucleus vowel in the official Pinyin Romanization system (as illustrated in Table 1).

[1] This paper is made possible with the support from the CHSS Faculty Summer Research Grant at Kennesaw State University. The author is also grateful for the arrangement of a data transcription assistant, Ms. Liyuan Yan, from CIKSU.

Tone Category	Traditional Name	Tone Description	Pitch Value	Examples in Pinyin	English Meaning
T1	Yinping	High-level	55	Fāng	'Square'
T2	Yangping	Mid-rising	35	Fáng	'House'
T3	Shangsheng	Low-dipping	214	Fǎng	'Visit'
T4	Qusheng	High-falling	51	Fàng	'Put'

Table 1. The four basic tones in Mandarin Chinese

Unlike the four basic tones, the neutral tone (T0) does not have its own default pitch value and must be preceded by at least one syllable that carries one of the four basic tones (Lin, 2007). Thus, depending on the preceding tone, T0 exhibits different pitch levels (Chao, 1933; Dreher and Lee, 1966; Shen, 1986). After T1 and T2, the neutral tone has a pitch value of 2; after T3, the pitch value is 4; after T4, the pitch value is 1 (Chao, 1933).

In addition to the tones, there is another extensive suprasegmental feature in Chinese called tone sandhi---a process that alters the phonetic shape of adjacent tones when they come into contact with each other in a sequence of syllables (Chen, 2000), producing thereby surface sandhi tones that are different from the underlying tones. There are several sandhi rules in Chinese that involve T2, T3, T4, the morphemes yī (一, 'one') and bù (不, 'no'), and the numerals qī and bā (七 and 八, 'seven' and 'eight') (Chao, 1968; Norman, 1988; Lin, 2007; Sun, 2006). The details of the sandhi rules are addressed in the following section 2.

The intriguing nature and importance of Chinese tones and tone sandhis have triggered great interest in the study of tonal acquisition of L2 Chinese in the past decades (Chen G-T., 1974;

Chen Q-H., 2000; Elliot, 1991; McGinnis, 1996; Miracle, 1989; Shen, 1989; Sun, 1998; Yang, 2011, 2016; H. Zhang, 2013, 2018; among others). While great efforts have been made in the previous studies on the perception and production of the four categorical lexical tones (Bent, 2005; Chen, 1997, 2000; Miracle, 1989; Shen, 1989; White, 1981), research efforts on the L2 acquisition of Chinese tone sandhis remain limited (H. Zhang 2007; Yang, 2016). For example, there is barely any empirical study on the acquisition of tone sandhis involving Yi and Bu, which are among the most frequently used words in daily speech, and the limited available studies on Tone 3 sandhis often led to contradicting conclusions (Yang, 2016; H. Zhang, 2013, 2018).

The importance of Chinese tonal system and its current research status in the field of L2 Chinese acquisition call for more empirical studies. In addition, the complexity of the tonal system of Chinese constitutes a considerable learning hurdle for L2 Chinese learners and a challenging pedagogical task for instructors. Thus, this study aims to fill the gap in the literature by addressing the acquisition patterns of five major tone sandhis that involve Yi, Bu, T2, and T3 by English speaking learners of Chinese, thereby providing pedagogical guidance for the L2 Chinese field and new empirical evidence for ongoing debates on relevant issues in Chinese linguistics field as well.

2. Chinese Tone Sandhis and Relevant Previous Studies

Tone sandhi refers to a tone changing process that is based on the context at the synchronic level. Chao (1968) and Norman (1988) distinguish between phonetic tone sandhi and phonemic tone sandhi, with the former resulting in a phonetic tonal variant or allophone of a tone, while the latter leading to a categorically different surface tone from the underlying one. Other scholars (e.g. Chen, 2000; Xu, 1994, 2004; Shih and Sproat, 1992) limit the concept of tone sandhi to those processes where the underlying tone is changed into a categorically different tone at the surface level and consider the rest of the cases as tonal coarticulation, where "the tonal targets remain the same, but the acoustic realization of the targets is varied due to their implementation in different tonal contexts and/or with different amounts of articulatory effort" (Xu, 2004:784). In this study, we

follow the former group of scholars (Chao, 1968; Norman, 1988; Yang, 2016; H. Zhang, 2013) and use the term tone sandhi to cover both the phonetic and phonemic synchronic tone changes in Chinese.

The two most important tone sandhi processes in Chinese are the ones that involve T3, the low dipping tone with pitch value of [214]. A T3 preceding another T3 changes into T2, i.e. the rising tone with pitch value of [35]: [214]➐[35] /___T3. This T3 sandhi process is a phonemic change, as it changes the underlying T3 into a different surface tone T2. However, before any other tones including T1, T2, and T4, a T3 becomes what is commonly known as a "half third" low tone [21] by losing its final rise part [4]: [214]➐ [21] /___T (T≠ [214]), hence the name Half-T3 Sandhi. In contrast to T3 Sandhi, Half-T3 Sandhi is a phonetic change with stronger phonetic motivation as it reduces the articulatory efforts on the speakers without losing the lexical distinction.

Another common Yi sandhi process is specifically relevant to the morpheme yī ([55], 'one'): before T1, T2, and T3, it changes into a T4, i.e. yī [55]➐ yì [51] /___T (T≠ [51]), but before another T4, it is pronounced as T2, i.e. yī [55]➐ yí [35] /__T4. The Bu sandhi is similar to the Yi sandhi in that bù ([51] 'no') remains as T4 before T1, T2, and T3, i.e. bù [51]➐ bù [51] / ___T (T≠ [51]), but changes into T2 before a following T4: bù [51]➐ bú [35] / ___T4. These two sandhi rules and the T3 sandhi are often understood as a dissimilation process motivated by the Obligatory Contour Principle (OCP) (Leben, 1973; Goldsmith, 1976; McCarthy, 1986). As analyzed in H. Zhang (2013) and Yip (2002), adjacent identical tones are prohibited in the above-mentioned cases in Chinese.

T2 sandhi is a process that changes a T2 into T1 in conversational speech (Chao, 1968) when it is the second tone in a trisyllabic word or phrase in which the first syllable is either T1 or T2: [35]➐ [55] / [55, 35] __ T. Another less commonly discussed tone sandhi is half-T4 sandhi: T4 loses its final portion before another T4 to become a half fall, i.e. [51]➐ [53] / ___T4. In addition, the numerals qī 'seven' and bā 'eight' can optionally change to a T2 before a following T4 syllable: [55]➐[35] / ____T4. Of all these tone

sandhi rules, we choose to examine five sandhis that are most relevant to L2 Chinese pronunciation. Table 2 lists the sandhi rules examined in this study.

Sandhi Name	Sandhi Process	Example	English
T3 Sandhi	[214]➐[35] /___T3	mǎi bǐ ➐ mái bǐ	'buy pens'
Half-T3 Sandhi	[214]➐ [21] /___T (T≠ 214)	mǎi shū ➐ mǎi [21] shū	'buy books'
Yi Sandhi	[55]➐ [51] /___T (T≠ [51])	yī běn➐ yì běn	'a copy'
	[55]➐ [35] /__T4	yī biàn ➐ yí biàn	'one time'
Bu Sandhi	[51]➐ [51] /___T (T≠ [51])	bù gāo➐ bù gāo	'not tall'
	[51]➐ [35] / ___T4	bù shì ➐bú shì	' is not'
T2 Sandhi	[35]➐ [55] / [55, 35] __ T	shēn lán sè➐ shēn lān sè	'dark blue'

Table 2. Five tone sandhi processes examined in this study

Although no acoustic studies are available on Yi and Bu sandhis, all the rest of sandhi patterns in table 2, i.e. T3 Sandhi, Half-T3 Sandhi, and T2 Sandhi were found to manifest themselves acoustically in a number of instrumental studies in the literature (e.g.,

Lin, 1985; Lin et al., 1980; Lin & Yan, 1991; Shen X., 1990; Shih, 1988; Wang and Li, 1967; Wu, 1984; Xu, 1994, 1997, 1999). However, acquisitional studies on these tone sandhis in the field of L2 Chinese are very limited. Even though they are most frequently used in daily speech, there is barely any research on the acquisition of Yi and Bu sandhis. There are hardly any T2 sandhi acquisition studies, though Xu (1994) and Shih and Sproat (1992) revealed interesting acoustic patterns of T2 sandhi. The few available studies on the acquisition of T3 and Half-T3 Sandhis often led to contradicting results. For example, in the novel words experiment studies on native Mandarin speakers, J. Zhang and Lai (2010) showed that native Mandarin speakers applied greater accuracy with the phonetically more motivated Half-T3 Sandhi than with the phonetically less motivated T3 Sandhi. This was later confirmed by Yang (2016) in his study of L2 Chinese T3 acquisition by English-speaking learners with the finding of a significantly higher accuracy rate for Half-T3 Sandhi than for T3 Sandhi. However, in her study of English, Korean, and Japanese-speaking learners' sandhi patterns of T3, H. Zhang (2018) found that learners processed or acquired the T3 Sandhi (with less phonetic motivation) better than Half-T3 Sandhi, thereby contradicting the findings of the previous studies.

In connection with these Tone 3 sandhi acquisition studies, there is a related debate about the nature of Tone 3, that is, should the basic form of Tone 3 be a dipping tone with a pitch value of [214] as originally proposed by Chao (1930), or should it simply be described as a low tone of [21]? As early as the 1970s and 80s, scholars started to advocate for a change from the traditional description of [214] to the primary status to [21] for Tone 3 (Chen, 1973; Chin, 1987; Lin, 1985). Their arguments are: 1). [21] is the most frequently occurring variant of tone 3. 2). [21] is acquired earlier than [214] in L1 Chinese. 3). The Chinese tone system would be much more simplified with the resulting symmetrical two-pair contrast of high vs. low and rising vs. falling. 4). The observed confusion between tone 2 and tone 3 would thus be eliminated. 5). It is easier for learners to comprehend and retain. J. Zhang and Lai (2010), however, found it difficult to defend the encoding of [21] into the underlying representation of tone 3 because Mandarin Chinese has "right-dominant" sandhis where non-final tones are changed and

domain-final tones are protected. In addition, while simplification processes such as [214]➐[21] as in Half-T3 Sandhi are very common cross linguistically, contour complications in final position such as [21]➐[214] for T3 words at the phrase or sentence final positions are universally marked and unusual. In more recent L2 Chinese acquisition studies, H. Zhang (2013) and Yang (2016) echoed the early advocates of 1970s and 80s by suggesting that a switch from the traditional norm of [214] to [21] would be most beneficial to the L2 Chinese learners. This way, there would be just one sandhi rule related to tone 3 instead of two, i.e. the T3 Sandhi, and it would substantially reduce the cognitive load in the acquisition process. As mentioned above, however, Yang's (2016) findings about the acquisition of T3 and Half-T3 sandhis are contradictory to those of H. Zhang's (2013) study. In the most recent survey of the L2 Chinese pronunciation teaching in higher education in USA, Yang and Jin (2018) reported that 54% of the respondents taught Tone 3 as a low tone and 20% of them taught Tone 3 as a low-dipping tone. As for the two Tone 3 sandhis, 86% of the respondents taught the T3 Sandhi, while only 53% of them lectured on the Half-T3 Sandhi. Given the general overview from the survey on current pedagogical practices in L2 Chinese field, it would be desirable to conduct further research on students' learning outcomes of the current instructional practices to help us better understand where the problems are, if any, and what needs to be adjusted in terms of pedagogy and how.

The importance of Chinese tonal system and its current research and pedagogy status in the field of L2 Chinese acquisition call for more empirical studies. Therefore, this study investigates one of the understudied issues in L2 Chinese field, i.e. the acquisition patterns of major tone sandhi processes by English speaking learners of Chinese from different levels, with an aim to provide practical guidance for instructors, but also for L2 learners that would achieve better tonal teaching/learning results. It is our hope that the L2 Chinese acquisition data in this current study may also provide new empirical evidence for the ongoing debate about the nature of Tone 3.

3. Methodology

3.1 Participants

Fifty-six L2 learners of Chinese from four different levels of classes in a public university in the Southeast of the United States were recruited to participate in this study. Three participants were excluded from the analysis because their native language was not English, and the remaining fifty-three participants were all native English speakers. Table 3 lists all the participants for this study. The terms Level 1 through Level 4 were defined in terms of which level of Chinese language courses they were registered in at the time of data recording. For example, Level 1 included students who were taking first year beginning 1000 level Chinese language courses; Level 2 included those who were taking second year 2000 level Chinese language courses, and so on.

Levels	Female	Male	Total
Level 1	5	10	15
Level 2	10	5	15
Level 3	7	8	15
Level 4	4	4	8
Total	26	27	53

Table 3. Participants of this study

3.2 **Speech Stimuli**

In L2 Chinese acquisition, students are found to be able to produce isolated tones correctly, but often have difficulty producing correct tones in connected speech (Guo and Tao, 2008; White, 1981; P. Zhang and Xu, 1981). Therefore, connected speech in the formats of passage reading or spontaneous speech, which are also found to be scarce in previous studies (Chen, 2000; Yang, 2016), would constitute a better setting than isolated words to investigate the tone sandhis. Spontaneous speech, however, may be problematic for our purpose because the samples may contain morphological and syntactic errors that affect participants' pronunciation and participants may also be able to avoid difficult L2 sounds or sound sequences if they are asked to talk about something freely (Piske, MacKay, and Flege, 2001). There would be no guarantee that all the to-be-examined target tone sequences will be present in spontaneous speech samples. For these reasons, the data elicitation method decided on for this study was passage reading, where all the target words are embedded into the sentences of the passage. The presence of carrier sentences would increase the magnitude of coarticulatory tonal variation (Xu 1994), thereby providing an ideal environment to examine the learners' production of tone sandhis.

For this study, a total of 36 words/phrases (see Appendix A) were used and embedded into a coherent paragraph (see Appendix B). Disyllabic words that present all the sandhi environments were selected for the examination of the sandhis of T3, Half-T3, Yi, and Bu because it is widely recognized that disyllabic units are the dominant metrical structure and the basic prosodic domain in Chinese where these tone sandhis are obligatory (Chen, 1975; Speer, Shih, and Slowiaczek, 1989 ; Shih, 2013; H. Zhang, 2013). For the examination of T2 sandhi, three-syllable words/phrases were used, depicting all the environments where T2 sandhi may occur, i.e. the first syllable is T1 or T2, and the third syllable is any one of the four basic tones (Lin, 2007). While the presence of Pinyin Romanization and tonal diacritics might ease the task of reading, it may also lead the participants to produce the tones accordingly, as Yang (2016) mentioned in his explanation of the occurrence of full tone 3 in its "canonical" form. Thus, to aid in the naturalness of the speech (Miracle, 1989), the passage was presented only in Chinese characters.

Since the investigation focused on the production of the tone sandhis, the passage was made morphologically and syntactically as simple as possible, with no new words or grammar to the participants of the lowest level, therefore variables which might influence the participants' phonetic production were excluded.

3.3 **Procedure**

The passage was sent to the classes one week in advance, and students were asked to practice reading aloud before coming to the actual recording. The recording was conducted by the researcher in a quiet room on the university campus with LogicPro X software installed in a MacBook Pro laptop at the sampling rate of 44.1kHz, using a unidirectional Shure MS58 microphone attached to the computer. The participants were aware that the research was about L2 Chinese acquisition, but they were unaware of the specific focus on tone sandhis. They were instructed to read the passage as naturally as possible, and if they made mistakes for an embedded target word/phrase or read certain sentences too slowly, they were asked to reread the relevant part. Participants were recorded only one time for their speech. After the recording, participants were asked to fill out a brief questionnaire about their language learning background. A total of 1908 (36*53) tokens were collected for further analysis.

3.4 **Data Transcription and Analysis**

Following the practices in previous studies in L2 Chinese acquisition, this current research relied on human ears for the judgement of the data. While PRAAT is a great tool to measure the values of Fundamental Frequency (F0), the acoustic correlates of tones, it cannot judge if the L2 tonal productions are acceptable or not to native Chinese speakers, and the human ears remain the final judge of any speech performance (Chen, 2000; H. Zhang, 2013). If T3 for example, with the pitch value of 214 on the five point scale, is pronounced as 213 or 313 or 314, these productions are still perceived as a T3 (Tseng 1981) despite the phonetic differences between them.

For this study, the researcher and another native Chinese speaker, who was a graduate student in Teaching Chinese as Foreign

Language with a couple of years' Chinese teaching experience, listened to all the data and transcribed target words/phrases independently. The tones for all the produced Yi, Bu, T3 words, and T2 words in all the target words/phrases were transcribed into three categories: original tone, sandhi tone, and other[2]. The researcher transcribed the data twice with three weeks' interval and found an intra-judge consistency rate of 97. 3%. For cases of discrepancy, the researcher listened to the words again and made a semi-final judgement, which was then checked with the transcription results from the other judge. The inter-judge consistency rate was 93.8%. For any discrepancy found between the two judges, they met and listened together and agreed on a final judgement. The final transcribed tone data were then submitted to SPSS v.25 for statistical analysis.

4. Results

In terms of the overall accuracy, participants produced the highest accuracy rate of 56.6% for Half-T3 Sandhi, followed by an accuracy rate of 47.8% for T3 Sandhi, and then by 45% for Bu Sandhi and 42% for Yi Sandhi. The overall ranking of the accuracy rate of these tone sandhis then is Half-T3 > T3 > Bu > Yi as shown in Figure 1. ANOVA analysis indicated that the differences between them, however, were not statistically significant.

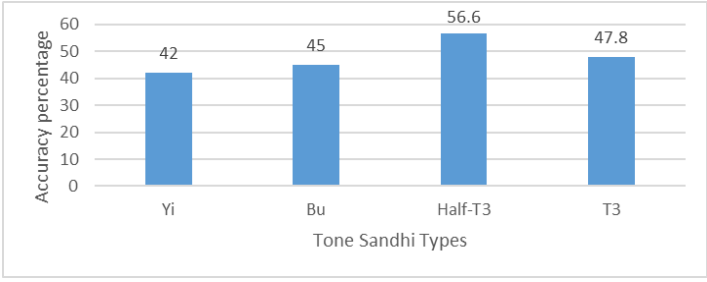

Figure 1. Overall accuracy rates for the different tone sandhi processes

[2] For Bu Sandhi, the underlying tone for Bu is the same as the sandhi tone before T1, T2, and T3. The occurrence of Bu as T4 in these environments was transcribed as sandhi tone.

Across the different levels, Level 3 students exhibited the highest accuracy rate of 60% for these tone sandhis, followed by Level 4 students with 51.3%, and then by Level 2 students with 45% and Level 1 students with 34.5%. Figure 2 shows the accuracy ranking among the different levels as Level 3 > Level 4 > Level 2 > Level 1. The statistically significant difference was only detected between Level 1 and Level 3 ($p=0.02$ at 95% confidence interval). Post Hoc Scheffe test put Levels 2, 3, and 4 as a homogeneous subset and Level 1 as the other subset. No significant overall gender effect was detected.

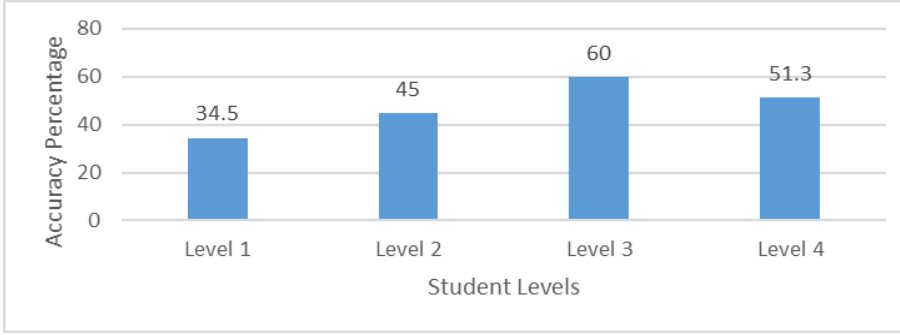

Figure 2. Overall accuracy rates for the different student levels

T2 Sandhi is not included in the above accuracy analysis, because it is optional and even if T2 surfaces faithfully as T2 in the T2 Sandhi environment, it is still considered as correct pronunciation. We did find, however, an overall occurrence rate of 32.8% for T2 Sandhi in this passage reading experiment, which was also the lowest rate among all the sandhi processes examined in this study. In the following sections, we explore the results for each individual sandhi.

4.1 **Yi Sandhi**

In the examination of the Yi Sandhi data, we found that the correct sandhi tones of Yi (T4 before a T1, T2, or T3; T2 before a T4) and the original tone of Yi (T1) exhibited almost equal amount of distribution (see Figure 3).

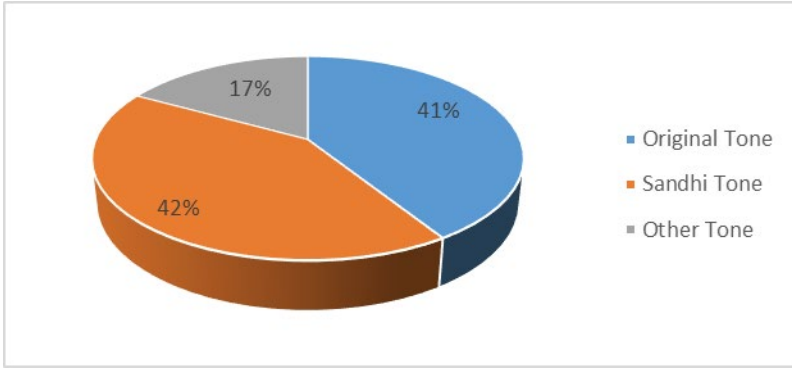

Figure 3. Distribution of different pronunciations of Yi

When we looked into the specific different environments of Yi Sandhi, i.e. Yi followed by T1, T2, T3 and T4 respectively, we found an accuracy ranking of Yi+T2 > Yi+T4 > Yi+T1 > Yi+T3 (see Figure 4). Although the differences between them were not statistically significant, a clear pattern emerged from this ranking, that is, Yi followed by T2 or T4 showed higher accuracy rates than the cases where Yi was followed by T1 or T3.

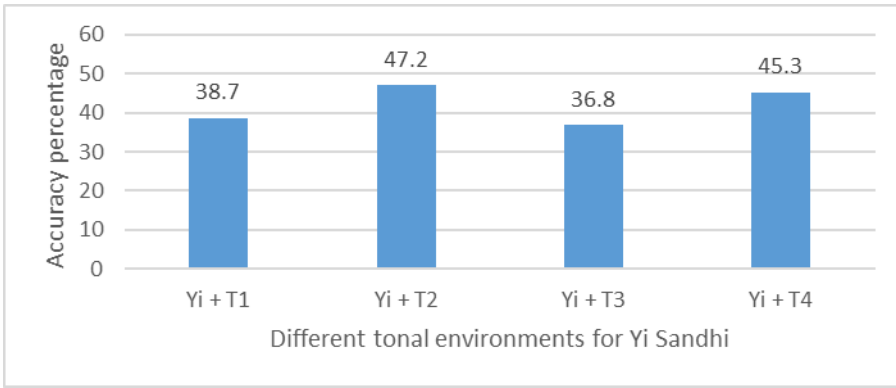

Figure 4. Accuracy rates for different tonal environments for Yi Sandhi

Across different student levels, ANOVA analysis revealed that Level 3 students performed significantly better than those of Level 1 and Level 2 ($p=0.000$ and $p=0.003$ respectively at 95% confidence interval), but no significant difference was found between any other two levels. Post Hoc Scheffe test grouped Levels 1 & 2 as

one homogeneous subset and Levels 3 & 4 as another subset. As in the overall analysis above, no significant gender effect was detected.

4.2 **Bu Sandhi**

In the production of Bu Sandhi words and phrases, we found 45% of the cases were pronounced as the correct Sandhi tones (i.e. T4 before a T1, T2, and T3; T2 before a T4) as shown in Figure 5. Unlike the Yi Sandhi production data where "Other Tone" occupied a small percentage, here it occurred more than 50% of the time. A further examination of these "Other Tone" cases showed a majority preference of T1 (55) for Bu.

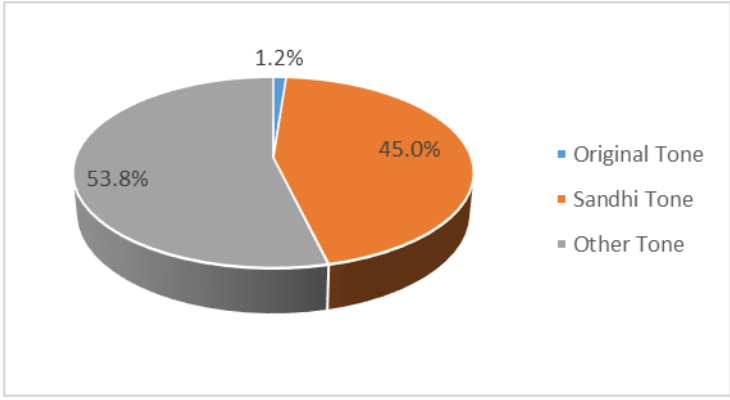

Figure 5. Distribution of different pronunciations of Bu[3]

The examination of the different Bu Sandhi environments (i.e. Bu+T1, Bu+T2, Bu+T3, and Bu+T4) revealed an accuracy ranking of Bu+T4 > Bu+T2 > Bu+T3 > Bu+T1 (see Figure 6). Although the differences in accuracy between these different environments were not statistically significant, we could easily see that students performed better for Bu followed by T4 or T2 than for Bu followed by T1 or T3. Grouping Bu+T4 and Bu+T2 into one

[3] Since the occurrence of Bu as T4 before a T1, T2 and T3 was transcribed as Sandhi Tone, the "Original Tone" portion in this figure refers to the cases where Bu was still pronounced as T4 before another T4.

subset, and Bu+T1 and Bu+T3 into another subset, we found the former group performed significantly better than the latter (t=2.399, p=0.017 at 95% confidence interval). We will address this interesting pattern in the discussion section. We did not find, however, any significant differences in Bu Sandhi across student levels or gender.

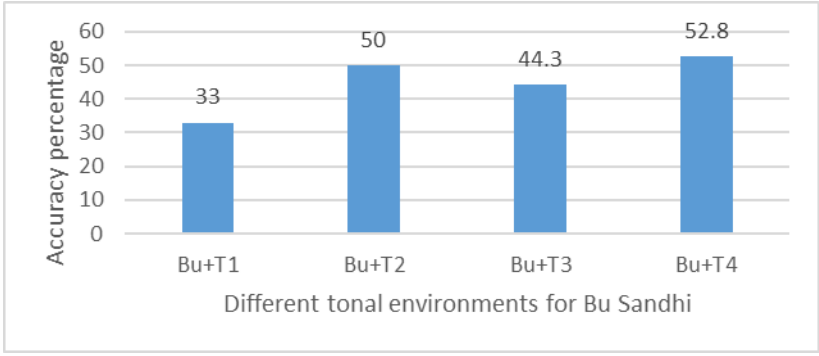

Figure 6. Accuracy rates for different tonal environments for Bu Sandhi

4.3 **T2 Sandhi**

As mentioned above, T2 Sandhi is typically expected to occur in fast speech, thus we were surprised to find 32.8% of T2 Sandhi occurrence in the students' passage reading speech (see Figure 7). Of all the cases of T2 Sandhi occurrence, 63.3% happened in T1+T2+OtherT environment, while 36.7% occurred in T2+T2+OtherT environment. Independent Samples T test showed that the difference in T2 Sandhi occurrence between the two environments was statistically significant (t=2.567, p=0.012 at 95% confidence interval).

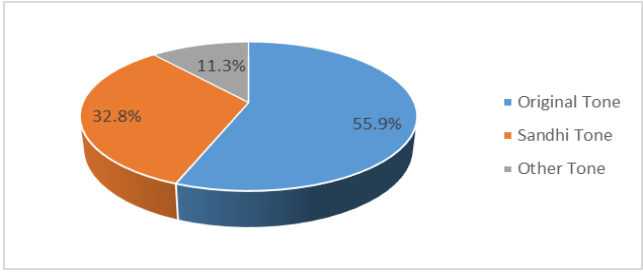

Figure 7. Distribution of different pronunciations of T2 in T2 Sandhi environment

In addition, we found that unlike other sandhi processes where the upper level of students (Levels 3 and 4) tended to produce higher percentage of correct tone sandhis than the lower level of students (Levels 1 and 2), for T2 Sandhi, it was the lower level students who produced more tone sandhi cases (see Figure 8), resulting in a ranking pattern of Level 1 > Level 2 > Level 4 > Level 3. Although ANOVA test showed no significant differences between the levels or different genders, further group comparison of lower level (levels 1 and 2) and upper level (levels 3 and 4) did reveal a significant difference in T2 Sandhi occurrence between the two larger groups (t=2.714, p= 0.008 at 95% confidence interval). It is worth exploring why T2 Sandhi would exhibit a pattern opposite to what has been observed of other tone sandhi processes.

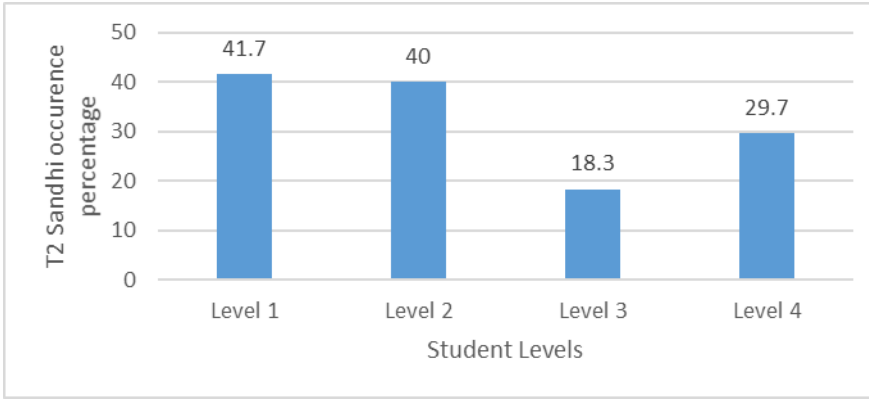

Figure 8. T2 Sandhi occurrence rate among different levels of students

In an effort to understand the possible effect of the tone of the third syllable on the preceding T2 in the T2 Sandhi environment, we further examined the T2 Sandhi cases with each of the four different tones for the third syllable. However, we found no significantly different effect from the tone of the third syllable.

4.4 **Half-T3 Sandhi**

For the half-T3 Sandhi process, we expected that the underlying Tone 3 would surface as one of the three different variants, i.e. the traditional full tone [214], the sandhi tone [21], or mistakenly as any other tone. In the production data set, however, we found the occurrences of the latter two, but no surfacing of traditional full tone was detected: correct sandhi tone occurred in 56.6% of the data, and other mistaken tones happened in the remaining 43.3% of the data (see Figure 9). In the correct Half-T3 Sandhi cases, T3+T4 environment showed the highest accuracy rate of 61.3%, followed by T3+T2 environment of 58.5%, which was followed in turn by T3+T1 environment with 50% accuracy. The accuracy differences between these tonal environments, however, were not statistically significant.

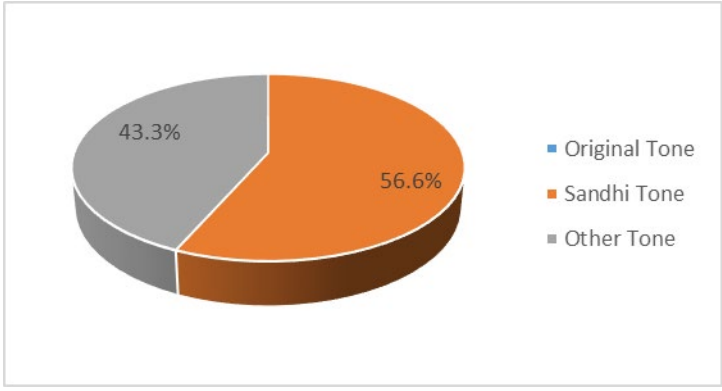

Figure 9. Distribution of different pronunciations of T3 in Half-T3 Sandhi environment

In the examination of the performance differences across different student levels, we found that Level 3 students showed the highest accuracy rate of Half-T3 Sandhi, and Level 1 students, not surprisingly, exhibited the lowest accuracy rate (see Figure 10).

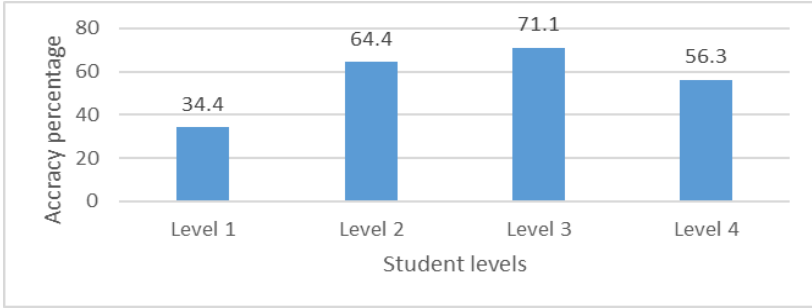

Figure 10. Accuracy rates of different student levels for Half-T3 Sandhi

ANOVA analyses revealed a significant difference between Level 1 and Level 2 ($p=0.001$); a significant difference between Level 1 and Level 3 ($p<0.001$); and a tendency towards a significant difference between Level 1 and Level 4 ($p=0.09$), resulting in a homogeneous subgroup of Levels 2, 3 and 4 and a separate group of Level 1. As in other sandhi processes, no significant gender effect was detected.

4.5 **T3 Sandhi**

In the examination of T3 Sandhi production data, we found a very small percentage of T3 surfaced as the traditional full tone [214], and the rest is split in half between sandhi tone [35] and other tones (see Figure 11).

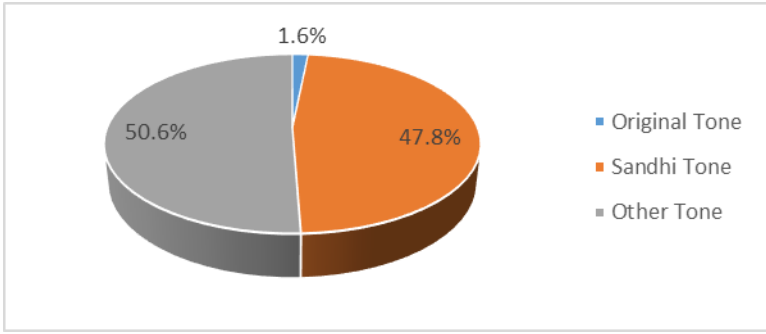

Figure 11. Distribution of different pronunciations of T3 in T3 Sandhi environment

Unlike other sandhi processes (except for T2 Sandhi), where Level 1 students consistently showed lower accuracy rates, T3 Sandhi was the only process where Level 1 students performed as proficiently as other levels (see Figure 12). Although statistical significance was not detected between different student levels, this unusual pattern exhibited by Level 1 in this specific Sandhi process deserves our further attention and we will address this in the discussion section. As with other sandhi processes, no significant gender effect was found.

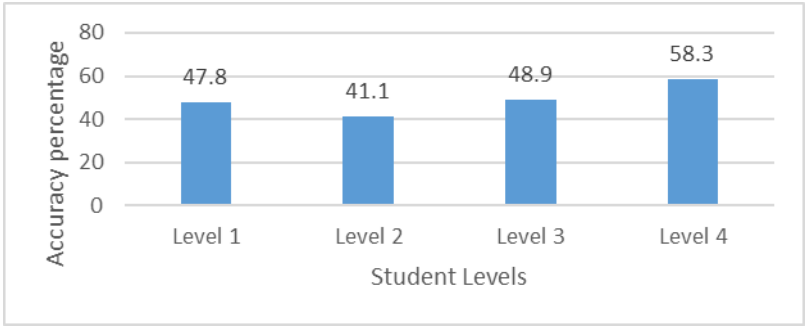

Figure 12. Accuracy rates of different student levels for T3 Sandhi

As Half-T3 Sandhi and T3 Sandhi are both relevant to the discussion on the nature of Tone 3, we further compared Half-T3 Sandhi and T3 Sandhi production data. The overall accuracy rates for Half-T3 Sandhi and T3 Sandhi were 56.6% and 47.8% respectively. Although the accuracy difference between the two sandhi processes was not statistically significant (t=1.58, p=0.12), we could easily tell that students performed better for the Half-T3 Sandhi than for the T3 Sandhi.

5. Discussion

Compared with the accuracy percentages of the individual lexical tones in L2 Chinese acquisition studies (e.g. Yang, 2016; H. Zhang, 2013), the relatively low accuracy rates of the tone sandhi processes found in this current study suggest two things: acquisition of tone sandhis is different from the acquisition of individual lexical tones; the students have not completely acquired the application of

tone sandhis in appropriate tone environments, often resulting in an easily perceivable foreign accent in their speech.

For the four tone sandhi processes that we discussed the accuracy rates, we found the accuracy hierarchy of Half-T3 > T3 > Bu > Yi. This finding is surprising in the sense that Bu and Yi are among the most frequently used words and are members of the T4 and T1 categories, which were reported to be of the lowest error rates (41.5% and 27.1% respectively) among all the lexical tones in H. Zhang (2013)'s experiment 1 on English speaking learners, but in the current study they exhibited lowest accuracy rates in sandhi processes. Thus, tone and tone sandhi acquisitions appear to be different, and a better acquisition of the former does not necessarily mean an equivalent acquisition of the latter.

The accuracy ranking reported here, however, is explicable from the perspective of the current pedagogical practices in both the local and national settings. In the questionnaire filled out by the participants in this study, 83% of the participants reported their instructors explicitly taught T3 Sandhi in their Chinese classes, while 54.7% and 50.9% of them reported an explicit classroom instruction on Yi and Bu Sandhis respectively. In the larger national context, the most recent survey of the instructors on the L2 Chinese pronunciation teaching in higher education in USA reported that 86% of the instructors taught the T3 Sandhi (Yang and Jin, 2018). Obviously, the substantial differences in the extent of emphasis and explicit lecture in the classroom resulted in different accuracy rates between the tone sandhis, i.e. T3 Sandhi with higher level of emphasis in classroom led to its higher accuracy rate than Yi / Bu Sandhis.

Another factor that has also played a role in the acquisition of the tone sandhis is the level of phonetic motivation involved in the sandhi processes. Only 18.9% of the participants in this study reported that their instructors explicitly taught Half-T3 Sandhi in class and 53% of the instructors from Yang and Jin's national survey reported that they lectured on Half-T3 Sandhi in class, both of which were much lower than those for T3 Sandhi. In spite of these relatively lower percentages, Half-T3 Sandhi exhibited a higher

accuracy rate than T3 Sandhi in this study, contrary to what we would have expected based on the classroom practices. Recall that Half-T3 Sandhi is the process where the second rising part of [214] is not phonetically realized and the tone is simplified into [21] in front of a T1, T2, or T4. This reduction process is then more natural and phonetically more motivated with less articulatory efforts than T3 Sandhi where [214] is changed into [35]. We believe this higher phonetic motivation has in effect overridden the pedagogical effect, resulting in the higher accuracy rate of Half-T3 Sandhi.

The phonetic motivation effect in tone sandhi acquisition was also evidenced in the participants' production of T2 Sandhi ([35]❼ [55] / [55, 35] __ T), which was expected to only optionally occur in fast speech. Results from the questionnaire suggested that T2 Sandhi was barely lectured on in class as only five participants reported that their instructors mentioned T2 Sandhi in class. Still, application of T2 Sandhi occurred in 32.8% of all the applicable T2 Sandhi cases, which we believe was a result mainly of the phonetic coarticulation effect from the preceding tone in the first syllable and the following tone in the third syllable. Figure 13 provides an illustration of T2 Sandhi application from a male participant. As we can see from the pitch contour (blue line) in Figure 13, the rising contour of the underlying T2 of the second syllable (Lán) was flattened into a surface T1 in between the preceding T1 of the first syllable (Shēn) and the T4 of the third syllable (Sè). As we reported above, our further examination of the T2 Sandhi cases with each of the four different tones for the third syllable revealed no significantly different effect from the tone of the third syllable. Thus, the flattening effect of the T2 in the second syllable was mainly due to the coarticulatory effect from the preceding tone in the first syllable rather than the following tone. Recall that of all the cases of T2 Sandhi application, 63.3% happened in T1+T2+OtherT environment, while 36.7% occurred in T2+T2+OtherT environment. The statistically significant difference in T2 Sandhi application between the two environments suggests that T1 of the first syllable in trisyllabic words and phrases in Chinese exerts greater tonal carryover effect on the second syllable than T2 of the first syllable. The strong phonetic motivation of tonal assimilation once again naturally overrides the pedagogical effect in the field.

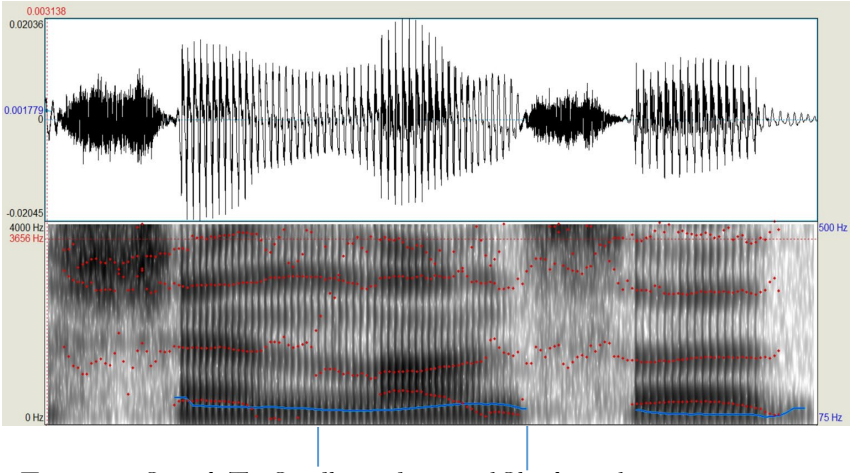

Figure 13. Sample T2 Sandhi production of Shēnlánsè from participant 4

As mentioned previously, the overall relatively low accuracy rates indicated that the students have not completely acquired tone sandhis in all applicable tonal environments. Different levels of students, however, exhibited different extents of accuracy as shown in the ranking of Level 3 > Level 4 > Level 2 > Level 1, suggesting a learning effect involved in the acquisition process. Generally speaking, a higher student level correlated with a higher accuracy rate, and there existed a significant difference between the beginning Level 1 and the advanced Level 3. The fact that Level 4 fell behind Level 3 posed an interesting question. For one thing, there were only eight students in Level 4, a much smaller pool than all other levels, and individual learner differences might have come into play. For another thing, this might be a result of students' reaching a tonal learning plateau at the fourth level. A follow-up longitudinal study on the current Level 3 students into their fourth level will possibly help illustrate the issue.

Of the four tone sandhi processes that we compared accuracy rates, T3 Sandhi is the only process where Level 1 students, who consistently showed lower accuracy rates than other levels, performed almost as well as Level 3 students (47.8% for Level 1 and 48.9% for Level 3). This again was related to the pedagogical practices. As mentioned previously, 83% of the participants reported

that their instructors explicitly lectured T3 Sandhi in class, and of all the participants who reported explicit T3 Sandhi lecture, 77.3% mentioned that the practicing of T3 Sandhi in class was at Level 1. The great pedagogical emphasis on T3 Sandhi at the beginning level resulted in better performance of Level 1 students on this sandhi process.

A word is in order here for T2 Sandhi that we did not include in the accuracy comparison with other sandhi processes. As mentioned in the results section, for T2 Sandhi, it was the lower level of students who applied T2 Sandhi significantly more often than the upper level students, resulting in an application ranking of Level 1 > Level 2 > Level 4 > Level 3, which was opposite to what has been observed in other sandhi processes. As indicated in the responses from the questionnaire, T2 Sandhi was barely mentioned in class lectures. In the absence of formal lecturing, the natural articulatory assimilation then came into full play. However, it appears that as the students' proficiency level increases, they become more and more aware of the norm of standard pronunciation in regular speech, which in this case is the faithful pronunciation of T2 as [35] rather than [55], and exert greater efforts to restrict the occurrences of the nonstandard variant, i.e. the sandhi tone [55]. For other sandhi processes, however, their respective sandhi tones are the standard norms. Thus, the higher the student level is, the more accurately they apply other sandhi processes, resulting in contradictory patterns between T2 Sandhi and other sandhi processes in terms of student levels.

Another important finding of this study is relevant to the accuracy rates of Yi and Bu Sandhis. For Yi Sandhi, we found an accuracy ranking of Yi+T2 > Yi+T4 > Yi+T1 > Yi+T3 and for Bu Sandhi, the accuracy ranking was Bu+T4 > Bu+T2 > Bu+T3 > Bu+T1. A clear pattern that emerged from both of these rankings was that Yi/Bu followed by T2 or T4 showed higher accuracy rates than the cases where Yi/Bu were followed by T1 or T3. Recall that Yi and Bu share the similar sandhi patterns, i.e. Yi/Bu appear as T4 [51] before a T1 [55], T2 [35], or T3 [214] but surface as T2 [35] before a T4 [51]. If we write the emerged pattern in terms of the five point pitch values, we derive the following:

Yi Sandhi: [51] + [35] or [35] + [51] > [51] + [55] or [51] + [214]

Bu Sandhi: [51] + [35] or [35] + [51] > [51] + [214] or [51] + [55]

In the [51] + [35] sequence, we see the ending pitch of the first syllable is low, and the beginning pitch of the second syllable is similar; in the [35] + [51] sequence, the ending pitch of the first syllable is high, and the beginning pitch of the second syllable is the same. Thus, we have on the left hand side of the equations perfect examples of "compatible" (Xu, 1994) phonetic context where adjacent phonetic units have identical or similar values along the pitch dimension. In the [51] + [55] sequence on the right hand side of the equations, however, the ending pitch of the first syllable is low, and the beginning pitch of the second syllable is high, producing thereby a "conflicting" (Xu, 1994) phonetic context where the adjacent phonetic units have very different values along the pitch dimension. Thus, while it is easy to produce the consecutive tones in a "compatible" phonetic environment, producing the ones in a "conflicting" phonetic environment in connected speech would be difficult as it requires the articulators responsible for pitch control to change very quickly from one state to another at the syllable boundaries if the underlying tonal values were to be fully realized within the second syllable. Therefore, producing the sandhi tone of Yi/Bu before a T2 or T4 (a "compatible" phonetic context) is easier than producing the sandhi tone before a T1 (a "conflicting" phonetic context), hence the higher accuracy rates of the former than the latter. For the [51] + [214] sequence, although the ending pitch of the first syllable and the beginning pitch of the second syllable are similar, the immediate opposite pitch trends from rising to falling, and then to rising again lead to complicated pitch contours that are more difficult to produce than the previously mentioned sequences in simple "compatible" contexts, hence the lower accuracy rate of the Yi/Bu before a T3. Again, we can see here that universal phonetic factors plays an important role in students' acquisition of tone sandhis.

Another phenomenon that deserves our attention is that "Other Tone" for Bu sandhi occurred 53.8% of the time and a further examination of these "Other Tone" cases showed a majority

preference of T1 (55) for Bu. This can be explained in terms of the relative markedness or easiness of the four Chinese tones. According to the universal, phonetically grounded, tonal markedness scale *R >> *F >> *L (Hyman & VanBik, 2004; Ohala, 1978), the rising tones are more complex than falling tones, which in turn are more complex than level tones. In Chinese then, this translates into a markedness scale of T2 >> T4 >> T1. In other words, T2 is the hardest to produce, T4 is the second hardest, and T1 is the easiest. As a complex contour tone, Chinese T3 is more marked than simple contour tones such as T2 and T4 (J. Zhang 2004; H. Zhang 2013), therefore it is also harder than T1, which means that of all the Chinese four tones, T1 is the easiest. Both Chinese L1 and L2 acquisition studies have confirmed that T2 and T3 are harder to acquire than T1 (e.g. Li and Thompson 1977; Chen, 1997). Thus, when Bu (T4) is pronounced in other tones, students would be more inclined to employ T1, the easiest one as the substitute, than to use the harder ones such as T2 or T3. T1 being the least marked and easiest tone is an inherent property of the Chinese language, and the substitution of T1 in the Bu Sandhi is an effect of this intra-lingual factor playing a role in the students' acquisition of tone sandhis.

However, this T1 substitution phenomenon in Bu Sandhi may also be a result of interference from English intonation. Chiang (1979) reported a case where the tones of non-terminal syllables in a Chinese sentence are subject to interference from the English enumeration intonation. The English enumeration intonation has a rising pitch similar to that of the Chinese T2 in the non-terminal syllables but a falling pitch in the terminal syllable: for example, Péars, plúms, prúnes, lìmes. When transferred, however, the pitch form may vary according to the condition in which it is produced. Thus, when the Chinese sentence is said quickly, the terminal syllable will probably still be produced like T4, but the non-terminal syllables will probably be pronounced as T1. For example, "wǒ yào chī yú," (I want to eat fish) would become "wō yāo chī yù," i.e. all the non-terminal syllables are pronounced as T1 regardless of their underlying tones, and the terminal syllable is pronounced as T4. This sweeping pattern of T1 is exactly what we see in our data. Many participants were very fluent in reading the passage, which provided a possible scenario for this pattern to occur. More interestingly, what was

observed was that this pattern was still possible, even when participants were reading at a regular normal speed. Figure 14 provides a sample where the participant reads, "tā bù shì zhōngguórén, zhōngwén bù hǎo," (He is not Chinese, and his Chinese is not good) at normal speed and produced all the non-terminal syllables in T1 and the terminal syllable in T4 (note the different underlying tones). This suggests that interference from English intonation is an inevitable factor that affects students' tone sandhi acquisition.

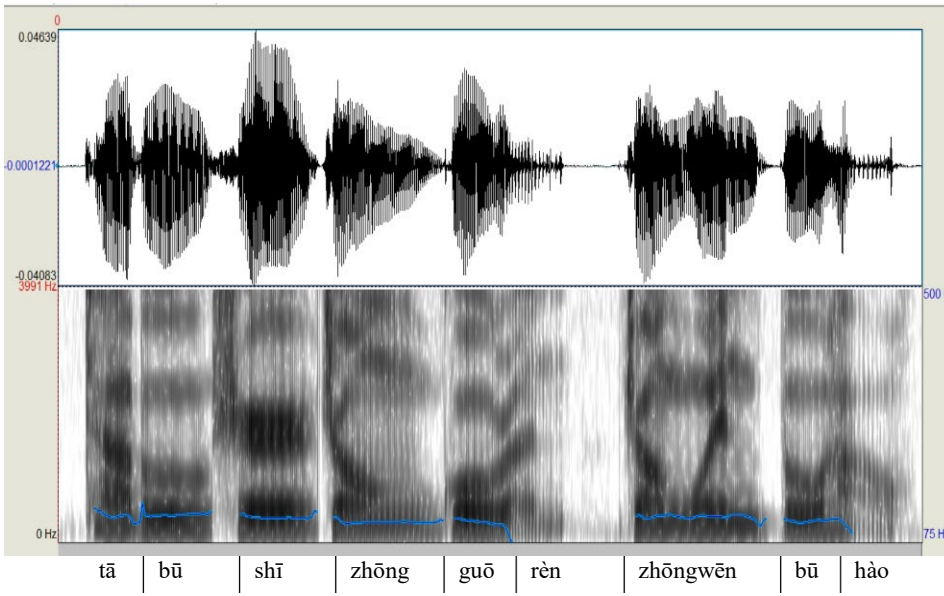

Figure 14. English intonation transfer in tā bù shì zhōngguórén, zhōngwén bù hǎo by participant 1 (the blue line indicates the pitch contour)

Finally, this study provides new evidence from the perspective of L2 Chinese acquisition for the understanding of the nature of Tone 3. The occurrence rate of the traditional full tone variant [214] of Tone 3 was extremely low in T3 Sandhi data (1.6%) and no surfacing of [214] was detected in half-T3 Sandhi data. The accuracy rate for Half-T3 Sandhi, which has stronger phonetic motivation, was 56.6% overall (and 71.1% for the advanced Level 3 students), which is better than that of T3 Sandhi with weaker phonetic motivation

(47.8%). These results are contradictory to H. Zhang's (2013) L2 Chinese findings that most Tone 3 before a non-Tone 3 surfaced as a full tone [214] and that T3 Sandhi exhibited a higher accuracy rate than Half-T3 Sandhi. However, they are consistent with the findings of J. Zhang and Lai's (2010) L1 Chinese study and Yang's (2016) L2 Chinese study, confirming that a phonetically motivated sandhi process is easier to acquire than a less motivated one.

The extremely low occurrence rate of the traditional full tone variant [214] may suggest a necessity to reconsider the default form of Tone 3 in Chinese pedagogy. Based on the above-mentioned tone markedness scales, Complex Contour >> Simple Contour and R >> F >> L, the three variants of Tone 3 exhibit a markedness hierarchy of [214] >> [35] >> [21], i. e. the traditional full tone variant is the most marked/difficult and the low [21] variant is the least marked/easiest. Since sandhi rules never output a tone which is more marked than its input (Hyman and VanBik, 2004) and contour complication in final position such as [21]❼[214] at the phrase or sentence final positions is universally marked and unusual (J. Zhang and Lai, 2010), it appears to make better sense, from the phonological perspective, to have [214] as the underlying form for Tone 3. However, in the L2 Chinese field, the goal of L2 learners is to be able to pronounce correctly at the phonetic level and communicate successfully (Yang and Jin 2018), and the abstract phonological level does not really matter to the majority of L2 Chinese learners. That said, even at the phonological level, there have been proposals of Tone 3 as an underlying low tone (Yip, 1980, 2002), as all that matters for Tone 3 is the low feature. Furthermore, important empirical evidence needs to be taken into consideration: of the three variants of Tone 3, [21] variant is the most widely distributed, including before non-Tone 3 and even at phrase or sentence final positions (Hu, 1979; Tsung, 1987; H. Zhang, 2013); it occurs most frequently (Chen, 1983; Lin, 1985); it is the least marked and easiest as analyzed above; [21] variant is acquired before [214] in L1 Chinese (Li and Thompson, 1977) ; and [21] Half-T3 Sandhi with stronger phonetic motivation is easier to acquire than T3 Sandhi. Thus, [21] variant as the default form makes better practical phonetic sense to L2 learners and instructors as well. Therefore, we propose a separation approach of the phonetic level from the phonological level

and a focus on the phonetic level in L2 Chinese language pedagogy by teaching [21] variant as the default form for Tone 3, which we believe will substantially reduce the cognitive processing load for the L2 learners, resulting in better performances in communicative reality. All that learners need to remember is to change it into a rising tone before another Tone 3 and simply keep it as [21] elsewhere. They have the option to produce it as [214] at phrase or sentence final positions, though it is not necessary given real life speech patterns (Hu, 1979).

6. Conclusion

Problems with suprasegmental features contribute more considerably to the L2 learners' typical foreign accent (Shen, 1989), which may reduce intelligibility and may serve as a basis for negative social evaluation and discrimination (Lippi-Green, 1997; Munro, 2003). Thus, this study makes an effort to tackle the understudied suprasegmental process of Chinese tone sandhi acquisition. The analysis of the sandhi application data showed that the accuracy rates of the tone sandhi application are relatively low, suggesting that students have not completely acquired the tone sandhis in all applicable tonal environments. The accuracy hierarchies found in this study, including Half-T3 > T3 > Bu > Yi, Level 3 > Level 4 > Level 2 > Level 1, Yi+T2 > Yi+T4 > Yi+T1 > Yi+T3, Bu+T4 > Bu+T2 > Bu+T3 > Bu+T1, Half-T3 Sandhi > T3 Sandhi, etc., combined with the participant response data from the questionnaire, suggest that there are several factors that play important roles in the tone sandhi acquisition process of L2 learners.

First, the instruction factor: the difference in the extent of emphasis and explicit lecture in classroom results in different accuracy rates between different tone sandhis, i.e. T3 Sandhi with higher level of emphasis in classroom practices led to its higher accuracy rate than Yi / Bu Sandhis. The great pedagogical emphasis on T3 Sandhi at the beginning level results in the better performance of Level 1 students on this sandhi process.

Phonetic motivation is a second factor. Sandhi processes with higher phonetic motivation result in the higher accuracy rate of application, e.g. Half-T3 Sandhi > T3 Sandhi, despite its low

pedagogical emphasis in class. The strong tonal assimilation effect from a preceding tone may lead to a substantial application rate of T2 Sandhi, again despite its lack of emphasis in the pedagogical practices. In Yi/ Bu cases, consecutive tones in "compatible" phonetic environment are easier to produce than the tones in "conflicting" phonetic environment, resulting in higher accuracy rate of the former.

We also found the learning effect. As students advance to higher level of proficiency, their tone sandhi application shows a higher rate of accuracy, though they may reach a tonal learning plateau at the fourth level. Additionally, there is an inherent intra-lingual factor at play. T1 is the least marked and easiest tone is Chinese, and it is used as the most frequent substitute when Bu is pronounced as "Other tones" in the sandhi process. Finally, interference from English intonation is inevitable. The sweeping pattern of T1 in non-terminal syllables in Bu Sandhi environments is an inadvertent interference originated from the English enumeration intonation.

This study also provides new evidence for the understanding of the nature of Tone 3 from the perspective of L2 Chinese acquisition. It is suggested, based on multiple empirical evidence, that [21] variant as the default form for T3 makes better practical phonetic sense.

For the field of L2 Chinese, this study provides important pedagogical implications. First, it highlights the importance of the pedagogy of tone sandhi. L2 Chinese tone sandhi acquisition is mediated by pedagogy, and explicit instruction and practices in class result in better accuracy rates of tone sandhi applications. Tone sandhi problems need to be addressed carefully right in the beginning stage of learning and the long-term training process can never be over-emphasized, even at the advanced level. More specifically, we suggest that Yi, Bu, and T3 Sandhi rules be explicitly lectured on and practiced at the beginning semester, together with the introduction of Tone 3 as a default [21], which will make the Half-T3 Sandhi rule redundant and reduce the processing load for L2 learners. Before- and after-teaching experiments that focus on these sandhis can be

conducted to monitor the teaching/learning effects. As students need help noticing what they are doing, the teachers should analyze students' speech and help them select areas for practice based on empirical findings. More importantly, these efforts need to be made continuously and constantly through all levels into the advanced level. Secondly, a normal talking speed rather than the slow "teacher talk" is recommend in classroom as slow "teacher talk" often deprives the speech of all the natural tone sandhi applications by emphasizing the citation tones of the target words. Regular speaking speed may help Chinese L2 learners acquire important co-articulations in speech, and eventually assist the L2 learners in approximation of native-like speech production. Thirdly, at the second level, that is, after learners' initial exposure to the Chinese tone sandhis, a brief introduction of the contrastive analysis between English intonation and Chinese tones will be helpful so that learners will be aware of and pay special attention to areas of potential transfer from English. Fourthly, at the advanced level, introduction of some key Chinese phonetic phenomena in terms of universal grammar relevant to tone sandhis might help learners better understand the nature of the sandhi processes and solidify their perceptual knowledge. Finally, we suggest that L2 Chinese textbooks make adjustments and reflect the default form of Tone 3 as [21] at the phonetic level rather than the traditional form of [214].

While this study is specifically on Chinese, its implications are extendable to the teaching of other less commonly taught languages (LCTLs). Just as Chinese tone sandhi acquisition is mediated by pedagogy, we believe explicit instruction and practices in class on suprasegmental features of other LCTLs will result in better acquisition of these features, thereby reducing the foreign accent in L2 learners. The fact that Chinese L2 learners at the upper level may encounter a plateau in the acquisition of tonal sandhis suggests that other LCTLs might also need to pay attention to a potential bottleneck effect when their learners reach an upper level and take special measures to address the problem. The suggestion of regular speaking speed in class may assist learners of other LCTLs acquire important co-articulations in speech and eventually approximate native-like pronunciation. In addition, regardless of learners' different target languages, a contrastive approach between English and the

target language is recommended, which will successfully raise learners' awareness of potential areas of transfer from English. Finally, informing the learners with some relevant phonetics knowledge of the target language may help solidify their perceptual knowledge and facilitate their acquisition of the target language. For example, L2 Korean learners are aware of the three-way contrast between Korean tense, lax, and aspirated stop consonants, but have difficulty in producing the contrast. If the instructor is knowledgeable of Korean phonetics and explains to the learners about how the contrast is exhibited in Voice Onset Time of the consonant and the Fundamental Frequency (F0) of the immediately following vowel, students will have a profound understanding of the differences and an easier time producing them.

Future studies on Chinese tone sandhi acquisition may shed new light on our current understanding by conducting a longitudinal study on a same cohort of learners from beginning to advanced levels, and also by including participants from diversified L1 backgrounds.

References

Bent, Tessa. (2005). *Perception and Production of Non-native Prosodic Categories*. Ph.D dissertation. Northwestern University.

Bluhme, H., and Burr, R. (1971). An audio-visual display of pitch for teaching Chinese tones. *Linguistics,* 22: 51–57.

Chao, Yuen-Ren. (1930). "A system of tone-letters". *Le maitre phonetique* 45: 24-27.

Chao, Yuen-Ren. (1933). Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology*, 4 (3), 121-134.

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

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

Chen, Gwang-tsai (1974). Pitch Range of English and Chinese Speakers. *Journal of Chinese Linguistics*, 2(2), 159-171.

Chen, Gwang-tsai. (1975). Tone pair. *Journal of the Chinese Language Teachers' Association*, 10 (1), 25-27.

Chen, Matthew Y. (2000). *Tone sandhi: patterns across Chinese dialects*. Cambridge: Cambridge University Press.

Chen, Ming-yuan. (1983). Hanyu zuowei di er yuyan jiaoxue zhong de yuyin wenti (Phonetic problems in the teaching of Chinese as a second language). *Language Learning and Communication.* 2(3), 249-265.

Chen, Qinghai. 1997. Toward a sequential approach for tonal error analysis. *Journal of Chinese Language Teachers Association*, 32 (1), 21–39.

Chen, Qinghai. (2000). *Analysis of Mandarin Tonal Errors in Connected Speech by English–speaking American Adult Learners*. Brigham Young University: Dissertation manuscript.

Chen, Shirley. (1973). The third tone and see-saw pairs. *Journal of the Chinese Language Teachers' Association*, 8 (3), 145-149.

Chiang, Thomas. (1979). Some interferences of English intonations with Chinese tones. *International Review of Applied Linguistics in Language Teaching.* 17 (3), 245-250.

Dreher J. and P-C. Lee. (1966). *Instrumental investigation of single and paired Mandarin tonemes.* Douglas Advanced Research Laboratory.

Elliot, C.E. (1991). The relationship between the perception and production of Mandarin tones: An exploratory study. *University of Hawai'i Working Papers in ESL*, 10 (2), 177-204.

Guo, Lijuan and Liang Tao. (2008). Tone production in Mandarin Chinese by American students: a case study. *Proceedings of the 20th North American Conference on Chinese Linguistics*. Edited by Marjorie K.M. Chan and Hana Kang. Columbus, Ohio: The Ohio State University. 1:123-138.

Hu, Bingzhong. (1979). Sansheng biandiao ji qi jiaoxue (Phonetic variation and teaching of the third tone). *Yuyan Jiaoxue yu yanjiu*. 2: 55-59.

Hyman, L. M. & VanBik, K. (2004). Directional Rule Application and Output Problems in Hakha Lai Tone. *Language and Linguistics*. 5(4):821-861.

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

Lin, M., L. Lin, G. Xia, and Y. Cao. (1980). Putonghua erzici biandiao de shiyan yanjiu [An experimental study of tonal variation in disyllabic words in Standard Chinese]. *Zhongguo Yuwen* [*Chinese Linguistics*] 1: 74-79.

Lin, M., and J. Yan. (1991). Tonal coarticulation patterns in quadrisyllabic words and phrases of Mandarin. *Proceedings of the 12th International Congress of Phonetic Sciences*, 242-245.

Lin, T. (1985). Preliminary experiments on the nature of Mandarin neutral tone. *Working Papers in Experimental Phonetics*, ed. by T. Lin and L. Wang, 1-26. Beijing: Beijing University Press.

Lin, William C. J. (1985). Teaching Mandarin tones to adult English speakers: analysis of difficulties and suggested remedies. *RELC Journal*, 16 (2), 31-47.

Lin, Yen-Hwei. (2007). *The Sounds of Chinese*. Cambridge, UK: Cambridge University Press.

Lippi-Green, R. (1997). *English with an accent: language ideology and discrimination in the United States*. New York: Routledge.

McGinnis, S. (1996). Tonal distinction errors by beginning Chinese language students: A comparative study of American English and Japanese native speakers. In S. McGinnis (Ed.), *Chinese pedagogy: An emerging field*. Columbus, OH: Foreign Language Publications. 81-91.

Miracle, C. (1989). Tone Production of American Students of Chinese: A Preliminary Acoustic Study. *Journal of the Chinese Language Teachers' Association*, 24 (3), 49-65.

Munro, M. J. (2003). A primer on accent discrimination in the Canadian context. *TESL Canada Journal.* 20(2), 38-51.

Norman, Jerry. (1988). *Chinese*. Cambridge: Cambridge University Press.

Ohala, J. J. (1978). Production of Tone. *Tone: A Linguistic Survey*. Ed. By Victoria A. Fromkin. 3-39. New York: Academic Press.

Piske, Thorsten, Ian R. A. MacKay, James E. Flege. (2001). Factors affecting degree of foreign accent in an L2: a review. *Journal of Phonetics,* 29, 191-215.

Ross, Claudia. (2001). Evaluating intermediate Chinese textbooks. *Journal of the Chinese Language Teachers Association*, 36: 1-22.

Shen, X. (1986). Phonology of the prosody of Mandarin Chinese. *Cahiers de linguistique Asie Orientale*, 15 (1), 171-178.

Shen, X. (1988). Ability of learning the prosody of an intonational language-French by speakers of a tonal language-Chinese. *Texas Linguistics Forum*, 30: 326-335.

Shen, X. (1989). Toward a Register Approach in Teaching Mandarin Tones. *Journal of Chinese Language Teachers Association*, 24: 27-47.

Shen, X. (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics,* 18: 281-295.

Shih, Chilin. (1986). *The prosodic domain of tone sandhi in Chinese*. Doctoral dissertation, University of California, San Diego.

Shih, Chilin. (1988). Tone and intonation in Mandarin. *Working Papers, Cornell Phonetics Laboratory*, 3: 83-109.

Shih, Chilin. (2013). Mandarin third tone sandhi and prosodic structure. In *Studies in Chinese Phonology*, edited by Jialing Wang and Norval Smith. De Gruyter, Inc. 81-123.

Shih, Chilin. and R. Sproat. (1992). Variations of the Mandarin rising tone. *Proceedings of the IRCS Workshop on Prosody in Natural Speech* No. 92-37, 193-200. Philadelphia.

Speer, S. R., C-L. Shih, and M.L. Slowiaczek. (1989). Prosodic structure in language understanding: evidence from tone sandhi in Mandarin. *Language and Speech*, 32(4), 337 -354.

Sun, Chaofen. (2006). *Chinese: a linguistic introduction*. Cambridge: Cambridge University Press.

Sun, S. (1998). *The Development of a Lexical Tone Phonology in American Adult Learners of Standard Mandarin Chinese*. Honolulu: University of Hawai'i Press.

Tseng, C. Y. (1981). *An acoustic phonetic study on tones in Mandarin Chinese*. Ph. D dissertation. Brown University.

Tsung, Chin. (1987). Half-third first: on the nature of the third tone. *Journal of Chinese Language Teachers Association.* 22(1), 87-101.

Wang, W. S-Y., and K.-P. Li. 1967. Tone 3 in Pekinese. *Journal of Speech and Hearing Research*, 10: 629-636.

Wang, Y., Spence, M., Jongman, A., and Sereno, J. (1999) Training American listeners to perceive Mandarin tones. *1999 Acoustical Society of America*. 106: 3649–3658.

White, Caryn. 1981. Tonal pronunciation errors and interference from English intonation. *Journal of Chinese Language Teachers Association*, 16 (2), 27–56.

Wu, Z. 1984. Putonghua sanzizu biandiao guilü [Rules of tone sandhi in trisyllabic words in Standard Chinese]. *Zhongguo Yuyan Xuebao [Bulletin of Chinese Linguistics]*, 2: 70-92.

Xu, Yi. (1994). Production and perception of co-articulated tones. *Journal of the Acoustical Society of America*, 95: 2240–2253.

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

Xu, Yi. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, 27: 55-105.

Xu, Yi. (2004). Understanding tone from the perspective of production and perception. *Language and Linguistics*, 5(4), 757-797.

Yang, Chunsheng. (2011). *The Acquisition of Mandarin Prosody by American Learners of Chinese as a Foreign Language (CFL).* Dissertation. Ohio State University.

Yang, Chunsheng. (2016). *The acquisition of L2 mandarin prosody*. John Benjamins Publishing Company.

Yang, Chunsheng and Wenhua Jin. (2018). Chinese as a Second Language Pronunciation Teaching Survey. *Journal of the National Council of Less Commonly Taught Languages*, 23:153-189.

Yip, Moira. (1980). *The tonal phonology of Chinese*. Unpublished Ph.D Dissertation. Massachusetts Institute of Technology.

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

Zhang, Hang. (2007). *A Phonological study of second language acquisition of Mandarin Chinese tones*. MA thesis The University of North Carolina at Chapel Hill.

Zhang, Hang. (2013). *The Second Language Acquisition of Mandarin Chinese tones by English, Japanese and Korean speakers*. Unpublished Ph.D dissertation. The University of North Carolina at Chapel Hill.

Zhang, Hang. (2018). *Second Language Acquisition of Chinese Tones—Beyond First-Language Transfer*. Publisher: Brill

Zhang, J. (2004). The role of contrast-specific and language-specific phonetics in contour tone distribution. In Bruce Hayes, Robert Kirchner, and Donca Steriade (eds.), *Phonetically based phonology*. Cambridge University Press, Cambridge, UK. 157-190.

Zhang, J. and Lai. Y-W. (2010). Testing the role of phonetic knowledge in Mandarin tone sandhi. *Phonology* 27(1), 153-201.

Zhang, Pengpeng and Lumin Xu. (1981). Shi xi "yangqiang yangdiao" de wenti [A trial analysis of "foreign accent" problems]. *Yuyan Jiaoxue he Yanjiu [Language Teaching and Research]*, 3: 65-70.

Appendix A

Target words/phrases examined in this study

(All tones are in the underlying forms.)

Yi Sandhi

yīzhāng yīzhī yīpíng yītiáo yīběn yībǎ
yīgè yījiàn

一张 一支 一瓶 一条 一本 一把
一个 一件

Bu Sandhi

bùgāo bùduō bùxíng bùnéng bùhǎo bùxiǎo
bùdà bùshì

不高 不多 不行 不能 不好 不小
不大 不是

T2 Sandhi

Shēnhóngsè shēnlánsè chīwánfàn hēwánshuǐ

深红色 深蓝色 吃完饭 喝完水

háiméichī háiméiwán háiméizǒu háiméikàn

还没吃 还没完 还没走 还没看

Half-T3 Sandhi

hěnduō mǎishū qǐchuáng mǎiyú xiězì mǎiròu

很多 买书 起床 买鱼 写字 买肉

T3 Sandhi

xiǎolǐ mǎibǐ hěnzǎo mǎishuǐ suǒyǐ mǎijiǔ

小李 买笔 很早 买水 所以 买酒

Appendix B

Passage

小李的床上有 一本书、 一瓶水、一把小刀、一件深红色的毛衣、和一条深蓝色的裤子。她用一支笔在一张不大的纸上写字 ，写下要买的东西，因为她要跟Mike去商店。要买的东西很多， 所以她很早就起床了。吃完饭， 喝完水， 她还没走， Mike就来了。 Mike不高， 他不是中国人， 中文不好，不能一个人去商店。Mike 说他太忙了， 书还没看， 功课还没完， 早饭还没吃。 小李说不行，要他吃早饭。 到了商店，人不多， 他们先买书、买笔 ，然后去买鱼、 买肉、买水，最后去买酒。那鱼真不小，他们很高兴！