TAIWANESE TONE SANDHI VIEWED FROM AN INTENSITY PERSPECTIVE

Fran H-L Jian
Dept. Linguistic Science, University of Reading, U.K

ABSTRACT

Tone sandhi is a phenomenon that occurs in tone languages such as Taiwanese, Mandarin and Cantonese. It is a set of rules that defines how words placed at specific positions in sentences carry particular linguistic tones or fundamental frequency contours. Most investigations into Taiwanese are of an impressionistic nature and the few acoustic studies are limited to the study of fundamental frequency. In this work we study the effect of tone sandhi on the intensity contour in Taiwanese speech materials. The results show that short tone words subject to tone sandhi are not only significantly modified in the fundamental frequency domain but changes in the intensity contour can be measured.

1. INTRODUCTION

Tone sandhi is a phenomenon that occurs in tone languages such as Taiwanese, Mandarin and Cantonese. It is a set of rules that defines how words placed at specific positions in sentences carry particular linguistic tones or fundamental frequency contours. In a tone language the fundamental frequency contour of a word defines its lexical meaning. Unlike Mandarin, which has a relatively simple tone structure and has been subject to much research, Taiwanese has a more complex and subtle tonal structure that has received little attention. Most investigations into Taiwanese are of an impressionistic nature and the few acoustic studies are limited to the study of prosodic features such as fundamental frequency and duration.

The documented sandhi rules for Taiwanese state how the long tones change into other long tones, and how the short tones change into the opposite short tones. In this work we focus on the sandhi rules that apply to the short tones. It is traditionally accepted that a tone belongs to the short tone category if and only if it ends with unreleased and unvoiced obstruents [p], [t], [k] or a glottal stop. This results in morphemes with such codas to demonstrate steeper energy decay than a long tone word [1, 4]. Further, the short tone sandhi rules state that a short tone word changes into the other short tone and as a result of tone sandhi the glottal stop is lost [2, 21, 22, 24]. Obviously, this rule contradicts the definition of a short tone.

In this work we clarify these contradicting impressionistic views through a quantitative acoustic study. Taiwanese utterances at two stages of sandhi change are produced by a native Taiwanese subject and digitally recorded. The recordings are analysed using Xwaves that provides fundamental frequency and energy contour plots. The time-varying distributions for the energy-contour slopes of the words that undergo sandhi-tone changes are compared.

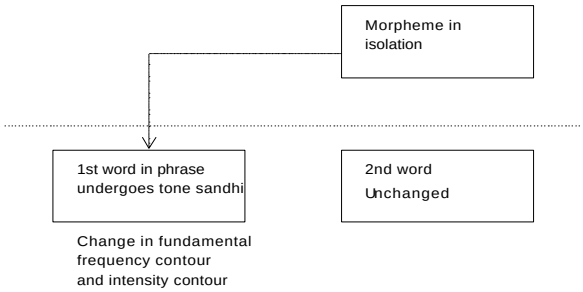

Figure 1. Example of tone sandhi

The preliminary results show that in addition to noticeable changes in the fundamental frequency domain there are significant changes in intensity. However, morphemes undergoing tone-sandhi get even steeper intensity decays than morphemes in isolation, indicating that the obstruent coda is preserved and even reinforced. Consequently, we found no evidence that a short tone word loses its glottal stop during tone sandhi. For disyllable phrases the intensity contour is steepest for the 1st morpheme and normal for the second morpheme. For phrases of triple adjective reconstruction there does not appear to be a distinctive trend regarding the change in the intensity contour.

The reminder of this paper is organized as follows: First we give a brief introduction to the Taiwanese language, followed by a tone sandhi primer. This is followed by a description of the data acquisition, followed by results and discussion. The paper is closed by a set of concluding remarks.

2. TAIWANESE

Taiwanese is also known as Amoy, Hokkien or Fukienese. It belongs to the Chinese dialect group of southern Min, within the larger Sino-Tibetan language family. The language originates from the Fujian province of southeast China. Although Mandarin Chinese is the official language in Taiwan, Taiwanese is widely used in conversation. It is the main language used by older generations and it is mutually unintelligible from Mandarin and other Chinese dialects.

2.1. Tone languages

In common with many Asian tongues such as Thai, Cantonese and Mandarin, Taiwanese is a tone language, where the interpretation of an utterance depends on the tone or fundamental frequency contours applied to the utterance [20]. It is widely accepted that Taiwanese has seven citation tones [2, 5, 6, 7, 10, 13], also known as inherent tones [15] or underlying tones [7], where there are two level tones, four falling tones and one falling-rising or rising tone [19].

2.2. The Taiwanese tone inventory

Besides fundamental frequency movement, Chinese dialects including Cantonese and Taiwanese have the notion of “free” and "checked" syllables. “Checked’’ syllables are those with voiceless obstruent codas [p, t, k, ?] and “free” syllables refer to those with sonorant or vocalic codas [20]. The seven Taiwanese tones have been classified into two main groups based on this concept. In traditional Chinese terminology, syllables with obstruent codas are called “entering” tones since they end in a stop. Likewise, syllables with sonorant or vocalic endings are categorised as “non-entering” tones. Alternatively, the former is classified as “short” tones and the latter as “long” tones respectively [9]. In this paper we adopt the alternative terms to refer to the Taiwanese tonal inventories, namely the “long” and “short” tones.

2.3. The suprasegmental attributes of Taiwanese speech

Studies by Jian [1, 4] demonstrated that ``short’’ tones have a steeper intensity contour decay than ``long’’ tones. The explanation for the differences in intensity contour is that short tone morphemes have obstruent codas consisting of [p, t, k, ?]. The simultaneous closure of the glottis results in steep intensity decays when plotting intensity against time. The sudden closure ``removes’’ the energy from the vibrating vocal folds, so instead of a naturally decaying energy contour, the result is a sudden and violent energy decay. Here we investigate how intensity is affected by tone sandhi for ``short’’ tones. Extracting energy features is more complex than extracting fundamental frequency features, because the energy location depends on the intensity of the utterance and distance from the speaker. Further, the energy movement is the sum of many factors. In this paper we employ a strategy first proposed by Jian [4] that allows the envelope of a complex speech sound to be represented using only a small set of parameters.

3. TONE SANDHI

The term ``sandhi" originates from a Sanskrit word, meaning ``joining" or ``combining". Sandhi forms refer to the cases where specific modifications occur during specific circumstances. The tones in Taiwanese undergo changes when they appear in different positions of the utterances. Tones under these changes are called sandhi tones or derived tones. Sandhi tones in Taiwanese are generally conditioned by syntactic structures. The tone of a syllable not occurring at the final position of the domain of a tonal group (such as a noun phrase, a verb phrase, etc.) undergoes the sandhi rules. There are various types of sandhi rules. Apart from the general type, there are other types of sandhi rules in which the domains of sandhi are defined syntactically, while the sandhi is conditioned morphologically [16, 17, 21, 22, 23]. The non-general types of tone sandhi occur regardless of the syllable type and the tonal inventory. These different tone sandhi apply to the syllable before the diminutive suffix [-a] with tonal vaule ``53", triple adjective construction and the neutral tone.

3.1. The General type of tone sandhi

The general type of tone sandhi is called ``South Min or Min-Nan tone circle" [18, 23]. According to this rule of tone circle, the long tones (i.e., the free syllables) and short tones (i.e., the checked syllables) have their own system in tonal changes. For the five long tones, they form a circle of change pattern within themselves. That is, one long tone changes into another long tone, and that specific one only. Tone 1 always turns into tone 7, tone 7 changes into tone 3, tone 3 becomes tone 2, and tone 2 transforms into tone 1, while tone 5 always becomes tone 7. The short tones form their own subsystem. The principal tone changes for long and short tones are illustrated in Figure 2.

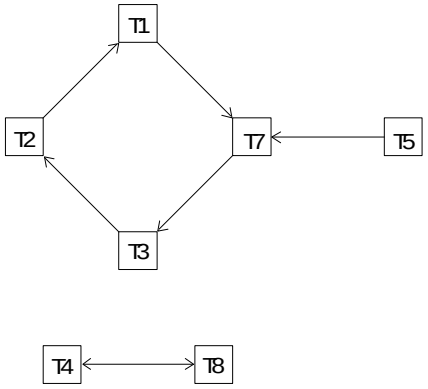

Figure 2 Example tone sandhi rules

3.2. Syllable before diminutive suffix /a/

Diminutive suffix ``a" shows a tonal value of high falling ``53’’. The tone of the syllable preceding the suffix is subjected to a different rule of tone change from that of the general type of tone sandhi. According to Tung [21, 22] and Zhang [12], the syllable first undergoes a general sandhi rule, and if the result of that is a high falling tone, it will further change to high level ``55’’ tone, or it will be a rising ``24’’ tone (considered as phonologically high) if the outcome is a non-high level general sandhi ``33’’ or ``11’’; all others remain unchanged. Chow [11] and Cheng [13, 14, 15] find level tones (high level or mid level) only in the syllables directly preceding the diminutive /a/ unless the syllables preceding it act as kinship terms or proper names then they keep unchanged.

3.3. Triple adjective construction

The triple adjective construction is a special type of adjective phrase in Taiwanese speech. When a monosyllabic adjective morpheme repeats itself three times, the first syllable undergoes a different sandhi rule but the second syllable follows the

general sandhi rule. The final syllable tone remains the same if it is at the end of a tonal group. Various rules have been reported for the first syllable by different researchers. However, a general trend is noted. The first syllable initially undergoes the general sandhi rule. If the resulting tone is not considered phonologically high, this non-high sandhi value turns into a rising tone ``15'' or ``13'' [5, 13, 14], or ``24'' [21, 25, 26]. Tonal values such as long tone ``21'', long tone ``33'' and short tone ``21'' are not considered phonologically high [5, 13, 14]. Hence, the high versus non-high tonal value of the general sandhi forms is treated as an important feature in conditioning the tone sandhi of this morphological structure. If the first syllable already has an underlying phonologically high rising tone ``24'', then it does not undergo alteration. If the resulting sandhi tone of the first syllable is already a high tone other than a rising tone, such as ``53'', this sandhi value will be kept without further change. In essence, the first syllable of the adjective triples has a high tone, the second syllable has either a high or a low tone according to the general tone sandhi, and the third syllable keeps its citation tone at the final position of a tonal group.

4. METHOD

Speech material spoken by one male Taiwanese subject was recorded and digitized. The data was subsequently post-processed using Xwaves, a software application that can extract fundamental frequency contours and intensity traces. The corpus consisted of disyllabe sandhi phrases and triple-adjective reconstruction phrases where all the morphemes were ``short'' tones. The intensity traces were further post-processed as outlined in the procedure below.

4.1. Intensity movements

As illustrated, Taiwanese short tones demonstrate a characteristic intensity contour that is different from that of long tones [1, 4]. To put it in another way, the short tone words can be viewed as "bursts" of energy, ending abruptly and stridently, while the long tone words have more "soft" or smooth contours. Previous attempts have been made at extracting intensity features from spoken utterances. For example, Lee [27] introduced an intensity factor that is the slope from the maximum to the 10% threshold of the maximum. The problem with such simple factors is that the drop in intensity may occur just at the end of the word and will in such cases provide a similar result as for an ordinary long tone word.

4.2. Intensity extraction algorithm

The first thing to note is that the start segments of the words are similar, thus we ignore the attack (the attack is the segment from the start of the word to the maximum RMS amplitude). The real question is how the slope of the energy contour changes with time. The following approach was adopted. The segment of the word starting at the maximum energy peak to the end of the word is split into three parts, the start part, middle part and end part. For each part a histogram of slopes is generated. The histograms summarize the distribution of the energy contour, i.e. the percentage of the word with level slope, the percentage of the word with medium slope and the percentage of the word with a steep slope. Only negative slopes are considered since we are interested in the nature of the word intensity decay. Figure 3 illustrates the steps involved.

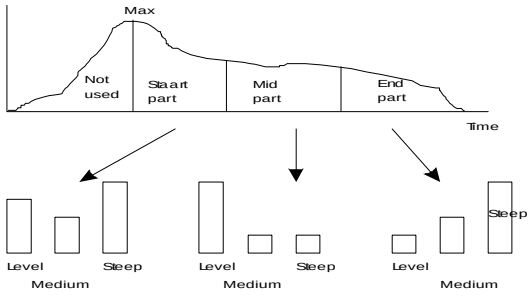

Figure 3: Extraction of intensity contour features.

One advantage of this representation is that a complex energy contour is modeled using a vector of 9 scalars. Further, it encompasses the change in slope over time and it is independent of absolute word length and intensity.

5. RESULTS

The results from the experiment are presented in Table 1. This table lists the nine intensity scalars for the 1st and 2nd words of disyllable phrases and the 1st, 2nd and 3rd intensity vectors for the triple adjective construction. An intensity contour with a steep intensity contour is identified by a high percentage of medium and steep slopes, especially towards the end segment of the morphemes. From the table it is therefore clear that the 1st word from disyllabic phrases has a steeper energy contour than the 2nd word.

Category		Start			Middle			End		
		Level	Medium	Steep	level	Medium	Steep	Level	Medium	Steep
Disyllable	1st	10.4	5.5	2.9	15.8	7.4	6.9	21.0	10.2	10.2
	2nd	9.0	5.7	9.3	14.8	11.2	2.3	27.2	9.1	1.3
Tri-syllables	1st	18.1	2.3	1.9	14.9	11.1	3.4	31.1	3.6	6.5
	2nd	9.3	2.7	1.7	18.4	4.9	6.0	8.8	31.3	2.1
	3rd	11.4	5.7	1.0	14.1	13.1	5.8	18.7	17.7	7.8

Table 1 The intensity features for words in disyllabic and tri-syllabic sandhi.

In particular, the mean of the 1st word in the disyllabic phrases has 6.8 % of steep decay for the middle segment, and 10.2 % of

medium decay and 10.2 % of steep decay for the end segment. This is significantly larger than the same result for the 2nd word which yields 2.3 % of steep decay for the middle segment and 9.1 % medium decay and 1.3 % steep decay for the end segment.

The results are more complex for the triple adjective reconstruction. For the 1st word there is 11.1 % of medium decay and 3.4 % of steep decay for the middle segment and 3.6 % of medium decay and 6.5 % of steep decay for the end segment. This is similar to the mean for the 3rd word, where there is 13.1 % of medium decay and 5.8 % of steep decay for the middle segment and 17.7 % of medium decay and 7.8 % of steep decay for the end segment. The 2nd word is slightly different, where there is 4.9 % of medium decay and 6.0 % of steep decay for the middle segment, and a massive 31.3 % of medium decay and only 2.1 % steep decay for the end segment.

6. DISCUSSION

The results show that the 1st word in a disyllable phrase is ``deepened'' due to the tone sandhi. By "deepened" we mean that the intensity contour is steeper reinforcing the impression of a short tone. The corpus in this pilot study did not provide a clear indication whether there is a distinctive change in energy for the triple adjective construction.

Since the intensity contours of the morphemes that have undergone tone sandhi is ``deeper'' or similar to the unchanged morpheme, we can conclude that the tone sandhi has not resulted in the ``removal'' of the glottal stop as claimed in the literature. This is because the absence of the obstruent codas would result in a more ``shallow'' energy trait, or a word with a softer intensity contour-decay.

7. CONCLUSIONS

The results analysed in the paper show that during tone sandhi a morpheme undergoes intensity changes as well as changes to the fundamental frequency contour. Morphemes in the 1st position of a disyllable phrase are reinforced, but there is no apparent change for triple adjective morphemes. Since there is no reduction in intensity contour decays, there is no evidence of loss of glottal stop. Therefore, the claim that a ``short'' tone word subject to tone sandhi loses its glottal stop does not appear to stand up. This preliminary study only investigated disyllabic and tri-syllabic phrases. However, future studies will include other tone sandhi structures such as the diminutive–suffix /a/. Further, with a larger corpus it should be possible to study the various tones in detail.

ACKNOWLEDGMENT

The author thanks Frode Eika Sandnes for his kind assistance with the data acquisition.

REFERENCES

[1] F.H.Jian, Perception of long and short tones in Taiwanese Speech, *The Journal of the Acoustical Society of America*, vol. 102, No.5, Pt.2, p. 3095 (Abstract), 1997.

[2] Robert L Chen, Tone sandhi in Taiwanese, *Linguistics*, vol. 41, pp. 19-42, 1968.

[3] Matthew Y. Chen, The syntax of Xiamen tone sandhi, *Phonology yearbook* 4, 1987.

[4] F.H.L.Jian, Classification of Taiwanese tones based on pitch and energy movements, *The 5th International Conference on Spoken Language Processing*, 30 Nov-4 Dec, Vol.2, pp. 329-332, 1998.

[5] R.L.Cheng, Taiwanese and Mandarin structures and their developmental trends in Taiwan, *Yuan-Liou Publishing Co., Ltd.*, 1997.

[6] Chin-Chu Yang, Bilingual dictionary of Mandarin and Taiwanese, *Dun Li Publishing Co. Inc*., 1996.

[7] K.T.Hsu, A study of the Taiwanese sound pattern representation and orthography, *Wen-He Ltd*, 1988.

[8] Eric Zee, Duration and intensity as correlates of F0, *Journal of Phonetics*, 6, 213-220, 1978.

[9] Tsai-Chwun Du, Tone and stress in Taiwanese, *University of Illinois at Urbana-Champaign, Ph.D. Thesis*, 1988.

[10] Shu-Hui Peng, Production and perception of Taiwanese tones in different tonal and prosodic contexts, *Journal of Phonetics*, 25, 371-400, 1997.

[11] James Chow, A study of /a/ in Amoy Chinese, ms. East Asian Language Department, University of Hawaii, 1980.

[12] Z.S.Zhang, Tone and Tone Sandhi in Chinese, *PhD Thesis Ohio State University* ,1988.

[13] Robert L Cheng, Some notes on tone sandhi in Taiwanese, *Linguistics*, vol.100, pp.5-25, 1973.

[14] Robert L. Cheng and Susie S. Cheng, Phonological structure and romanization of Taiwanese Hokkian, *Student Bookstore Co., Ltd*, 1977

[15] Robert L.Cheng, A comparison of four types of sound pattern representations of Southern Min, *Wen-He Publishing Ltd*, 1988.

[16] W. S-Y. Wang, Phonological features of tone, *International Journal of American Linguistics*, vol.33, pp. 93-105, 1967.

[17] Chi-Lin Shih, The prosodic domain of tone sandhi in Chinese, *Ann Arbor, Mi.: University Microfilms International*, 1986.

[18] Ji-Shiow Wu, Tone sandhi in the Min-Nan dialects, *University of Texas at Austin MA Thesis*, 1986.

[19] H-C. Li, Preface to Hokkien Grammar, *Wen-He Publishing Ltd*., 1988.

[20] K.L.Pike, Tone languages, *Ann Arbor:University of Michigan Press*, 1948.

[21] Tung-Ho Tung, A South Min Dialect of Taiwan, *Monographs Series* A.24. Taipei, Taiwan, 1967.

[22] Tung-Ho Tung, Phonology of the Amoy dialect, *Academia Sinica, Bulletin of the Institute of History and Philology*, vol. XXIX, 1957.

[23] Chinese Language Research Centre, Xiaman University, The Dictionary of Mandarin Chinese and Min-Nan language, *Tai-Lee Publisher*, 1993.

[24] Dah-An Ho, Two diachronic implications of phonemic variation, with discussion on original values of Chin-Chiang tones, *Academia Sinica, Bulletin of the Institute of History and Philology*, 55 (1), pp.101-152, 1984.

[25] Pang-Xing Ting, Taiwan Yuyang Yuanliou, *Student Book Co., Ltd.* Taipei, 1980.

[26] Jia-Hua Yuan et al, Hanyu Fangyan Gaiyao, *Wenzi Gaige chuban She*, 1960.

[27] Tan Lee et. al., Tone recognition of isolated Cantonese syllables, *IEEE Trans. on Speech and Audio Processing*, 1995.