Interspeech 2018
2-6 September 2018, Hyderabad

A Preliminary Study on Tonal Coarticulation in Continuous Speech

Lixia Hao[1, 2], *Wei Zhang*[1], *Yanlu Xie*[1], *Jinsong Zhang*[1]*

[1] Beijing Advanced Innovation Center for Language Resources
Beijing Language and Culture University, Beijing, China
[2] Dongguan University of Technology, Dongguan, China
haolixiaxgl@sohu.com, wei_zhang_mail@126.com, {yanluxie, jinsong.zhang}@blcu.edu.cn

Abstract

Tonal variations in continuous speech are complicated in nature and it is a challenge to identify the effect of tonal coarticulation given several influencing factors. To address the issue, the present study proposes a scheme for labeling tonal coarticulation in Mandarin continuous speech by applying Hypo- and Hyper-articulation theory. We assume that the bidirectional tonal coarticulation (both carryover and anticipatory effects) as patterns of Hypo-articulation results from the economical articulatory rule. The effects may partially disappear under the influence of specific stress patterns and become unidirectional (carryover or anticipatory). At a prosodic boundary, the effects of tonal coarticulation may completely disappear and lead to the occurrence of patterns of Hyper-articulation. Based on the scheme, we have labeled the data in the Annotated Speech Corpus of Chinese Discourse. It is shown that: three annotators are consistent at a fairly high level (86.2%) on average, and the acoustic parameters of four kinds of tonal coarticulation are significantly different. Therefore, we conclude that the proposal is feasible for investigating tonal coarticulation in Mandarin continuous speech. Though the labeling scheme is language dependent, it may well have cross-linguistic implications.

Index Terms: tonal coarticulation, continuous speech, Hypo- and Hyper-articulation, bidirectional and unidirectional effect

1. Introduction

Tonal coarticulation refers to any influence of one tone upon its neighbors. Previous studies in various tonal languages focused on the directionality and magnitude of tonal coarticulation within words and phrases [1-15]. It has been found that the direction of tonal coarticulation could be both progressive (i.e., the carryover effect) and regressive (i.e., the anticipatory effect). The magnitude of the carryover effect is greater than the anticipatory effect for Vietnamese [1, 2], Thai [3-5], Mandarin [6], Cantonese [7] and Tianjin Chinese [8, 9]; while for Nanjing Chinese [10], Malaysian Hokkein (i. e. Southern Min) [11] and Mizo [12], these two effects appear to be symmetric.

It is interesting to note that there are some conflicts regarding the directionality of tonal coarticulation in Mandarin. Shen [13] and Xu [6] investigated nonsense tri-tonal and bi-tonal sequences and concluded that the carryover and anticipatory effect occur simultaneously, i. e., the effect of tonal coarticulation in Mandarin is bidirectional in nature. However, Lin & Yan studied meaningful quadrisyllabic words and phrases with normal stress, and discovered that tonal coarticulation was unidirectional rather than bidirectional [14]. Each tone was affected by either a carry-over effect or an anticipatory effect. They argued that the manifestation of tonal coarticulation in natural words and phrases with normal stress was different from that in nonsense tonal sequences with even or unknown stress. The conflicts suggest that tonal coarticulation under the influence of natural stress patterns awaits further study.

The patterns of stress are relatively natural and diverse in the meaningful continuous speech. However, different from tonal coarticulation within words and phrases that has been investigated in a number of languages, the phenomenon in meaningful continuous speech has barely been examined. This can be attributable to the fact that tones exhibit considerable variations in continuous speech due to local tonal coarticulation [6, 13, and 15] and multiple global events, such as stress, focus, prosodic phrasing, speed and speaking style, etc. For example, the stressed syllable often lead to expansion of pitch range of the tone by making high target higher and low target lower [16, 17]; and the pitch range of its succeeding syllable compresses [18-20]. Besides, a prosodic boundary between two neighboring tones may set F0 contour of the tones free from coarticulatory effects [21]. The carryover and anticipatory effects of tonal coarticulation can even be absent at certain boundaries. With the influence of such factors, identification of the effect of tonal coarticulation in continuous speech becomes much complicated. Therefore the method successfully applied within words or phrases, where tonal coarticulation exists between any two adjacent tones, is not suitable in continuous speech task.

The present study aims to explore the method of examining tonal coarticulation in meaningful continuous speech by taking Mandarin as an example. Then we further examine how the effect of tonal coarticulation in Mandarin continuous speech is similar to or different from that within words and phrases. Finally, the distribution of tonal coarticulation in Mandarin continuous speech and its influencing factors will be discussed.

2. Method

Four basic lexical tones in Mandarin (referred to as T1, T2, T3 and T4) are characterized by different forms of F0 contours (see Figure 1).

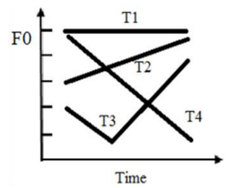

Figure 1: *Canonical F0 forms of 4 lexical tones in Mandarin.*

The phonological description for the four tones are high, rising, low and falling respectively. The rising tail of T3 barely

 10.21437/Interspeech.2018-1849

appears in the continuous speech. As mentioned above, tonal variations in Mandarin continuous speech, is complicated under the influence of multiple factors, including tonal coarticulation. For this reason, it is not feasible to identify and compare the effects of tonal coarticulation by directly analyzing the acoustic parameters of tones, as done in the previous studies on words and phrases. The speech data needs to be labeled to identify the presence of the bidirectional or unidirectional effect of tonal coarticulation. Therefore, the current study proposes a scheme for labeling tonal coarticulation in Mandarin continuous speech. Based on the scheme, we label a dataset of a speech corpus of Chinese discourse.

2.1. Theoretical basis of the labeling scheme

In order to explain the complex F0 variations in continuous speech, Zhang [22] introduced the Hypo- and Hyper-articulation theory (H&H theory) [23] into Mandarin for the recognition of tones. It is understood that there is a regular and relatively stable acoustic pattern for any disyllabic words in Mandarin [24-26]. Zhang attributed the basic pattern to the economical rule of articulation, i.e. the results of Hypo-articulation. When the specific pattern is affected by additional articulatory efforts due to functional needs, new patterns appear as the results of Hyper-articulation [22]. The form of Hypo-articulation is caused by the tonal context, while the form of Hyper-articulation is caused mainly by prosodic factors.

Applying the H&H theory, the F0 variations in continuous speech can be explained in a systematic way. The theory can build a connection between local and global factors which may affect the realization of tones in continuous speech. We therefore have reason to assume that the H&H theory is potentially applicable to the labeling of tonal coarticulation in continuous speech. de Jong & Beckman [27] have suggested that H&H theory can be used to characterize many differences in supralaryngeal features between less and more accented syllables. A natural implication is that it may also be used in characterizing different laryngeal features between less and more coarticulated tones. Based on the H&H theory, we draft a labeling scheme for tonal coarticulation in continuous speech.

2.2. Labeling scheme for tonal coarticulation in Mandarin continuous speech

It is revealed that tonal coarticulation affects the adjacent tones and not across syllables [13, 14]. Thus the labeling scheme is designed for each of 16 bi-tonal combinations in Mandarin continuous speech. According to the previous discussions, Hypo- and Hyper articulation may result in the presence and absence of tonal coarticulation. Such as a prosodic boundary could make the effect of tonal coarticulation disappear. In the cases of the presence of the effect, we assume it to be either bidirectional or unidirectional, depending on specific stress patterns. As shown in Table 1, we use tag 1, 0B, 0L and 0R to label different effects.

Table 1: *Tags for different effects of tonal coarticulation.*

H&H	Tonal coarticulation	Tag
Hyper-articulation	Nonexistent	1
	Bidirectional effects	0B
Hypo-articulation	Carryover effect only	0L
	Anticipatory effect only	0R

2.2.1. *Acoustic manifestation of Hypo- and Hyper-articulation*

The Hypo- and Hyper-articulation cost less to greater articulatory efforts [23]. According to the Fujisaki model, small articulatory effort leads to a smooth trajectory in the process of F0 contour generation [28]. The smoothness of F0 contours requires the onset of a tone assimilated to the offset of a previous tone, i. e. the assimilatory effect of carryover and/or anticipatory tonal coarticulation [6]. Extra efforts increase distinctiveness of articulation while decrease coarticulation of neighboring segments or tones [21, 27]. Therefore, the smoothness of F0 contours serves as an important reference to distinguish Hypo- and Hyper- articulation, i.e. the presence and absence of tonal coarticulation.

The next factor is the extent of realization of tonal targets. Greater effort makes high tone higher, low tone lower, and the pitch range of dynamic tones expanded [16, 17]. Greater to less amount of efforts leads a tone to realize or deviate from its underlying target [29, 30]. The better the tonal target is realized, the less it is influenced and coarticulated with neighboring tones. Beside the pitch-related parameters, the duration of syllables is also important. Generally it takes some time to fully realize tonal targets. What needs to be taken into account for T2 (rising tone) and T4 (falling tone) is that a speaker needs more time to raise the pitch than to lower it [31-33]. Last but not the least factor is the duration of silence between syllables that frequently break the specific pattern of tonal coarticulation.

We present the acoustic patterns of Hypo- and Hyper-articulation by taking T4T4 as an example. T4T4 is a pair of falling tones (see dashed lines in Figure 2). Under the carryover and anticipatory assimilatory effects, the offset of first T4 and the onset of the following T4 will get close to each other, as illustrated by the solid lines in Figure 2(a). As a result, the former will be raised while the latter will be lowered. This pattern was originally highlighted by Chao [16] and then has been recognized as the default economical form of T4T4. Both the carryover and anticipatory effects are observed in Figure 2(a). We therefore label it as “0B”: bidirectional coarticulatory effects (Table 1).

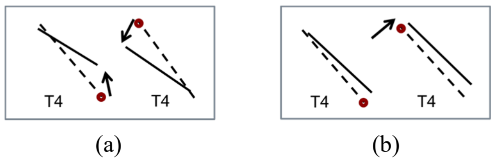

Figure 2: *A schematic depiction of F0 patterns of Hypo- (a) and Hyper- articulation (b) for T4T4.*

When there exists a functional need that results in extra efforts in articulation, the coarticulation and trade-off relationship between the successive tones will reduce or even disappear and new patterns occur. As can be seen in Figure 2(b), both T4s almost fully realize the falling target and the transition is sharp. We therefore label it as “1”: representing that the effect of tonal coarticulation is absent.

2.2.2. *Sub-classes of Hypo-articulation (0B, 0L and 0R)*

The contextual coarticulation under the influence of certain patterns of stress or prosodic strength may cause the deviation of a tone from its underlying form. For example, T4 may be produced with a level, or even rising contour in continuous speech. The deformation of tones with concomitant short duration is not rare in both spontaneous and read speech [34]. It can be commonly seen in metrically weak positions [35], such as in the medial positions of a polysyllabic word [36, 37]. When the pitch target of the first tone is fully realized, and a deviation or distortion of F0 contour is observed in the second tone, it indicates that the preceding tone has an effect on the succeeding one, but not vice versa. These patterns with only carryover effect is labeled as "0L". On the contrary, those patterns with only anticipatory effect is labeled as "0R". 0L, 0R and 0B, are the subclasses of Hypo-articulation.

We again consider T4T4 as an example. Figure 3 is a part of a sentence consisting of five successive T4s, read by a female Beijing speaker in a formal manner. Five syllables with tones are transcribed as follows (number 4 here refers to T4).

Ju4 zhu4 "Zheng4 yi4 lun4"
The great work "A Theory of Justice"

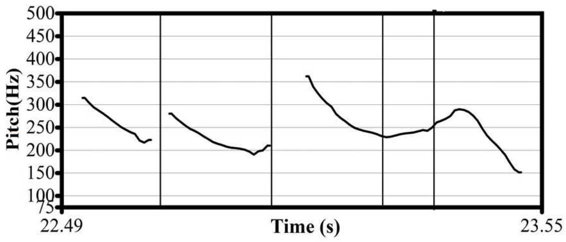

Figure 3: *Illustration of different acoustic patterns of T4T4.*

In the figure, the vertical solid lines mark boundaries of syllables. The tags (0B, 1, 0L and 0R) at the boundaries represent the specific coarticulatory effects between successive T4s. The acoustic pattern of the first T4T4 shows bidirectional effects of tonal coarticulation as depicted in Figure 2(a). At the onset of the third T4, a big raise indicates the presence of a pitch reset and therefore a prosodic boundary between the second and the third T4s. The boundary releases the coarticulatory relationship and generates a pattern of "1". The pitch contour of the fourth T4 is heavily distorted over a short duration that seems like a transitional tone between the third and the fifth T4s. It is under the influence of contextual coarticulation, but too weak to exert any effects on the other tones. This leads to the presence of the unidirectional effects (0L and 0R).

According to the labeling scheme driven by H&H theory, the annotation is done based on visual speech signals rather than listening to the audio.

3. Annotation

The annotation is conducted on Annotated Speech Corpus of Chinese Discourse (ASCCD) [38]. It contains 18 passages and altogether 8762 syllables read in a formal speaking style at normal speech rate. The corpus has been transcribed based on C-ToBI. Data of F001 (female speaker, number 1) is labeled by three transcribers based on the following steps.

First, the annotators discuss the labeling scheme. We revise the scheme repeatedly until the members reach an agreement. Then, three transcribers label half of the data independently for consistency check that is focused on the standards practiced by each of them. After the discussion on confused tags, lastly, another half of data is annotated. The tag chosen by the majority is considered as the final one.

For statistical analysis of acoustic parameters, we identify tone nucleus of syllables through a script, and then one of the transcribers manually checks and adjusts the onset and offset of nucleus (see Figure 4) which is represented by the F0 contour between H1 and L1 or H2 and L2, t and t' indicate the onsets and offsets of nuclei.

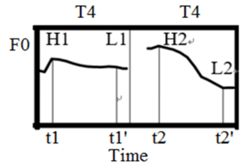

Figure 4: *Illustration of labeling of the tone nuclei.*

4. Results and Discussion

4.1. Consistency check

Table 2 shows the averages of results of a pairwise consistency check based on three transcribers. Four tonal pairs are listed as examples of 16 bi-tonal sequences. We observe 80.3% to 93.1% agreement on the suggested labels for effects of tonal coarticulation. The average of consistency of 16 tonal pairs is 86.2% which indicates that the results of different transcribers are highly consistent.

Table 2: *Results of consistency check.*

Tonal pair	**T4T1**	**T4T2**	**T4T3**	**T4T4**
Num. of tags	530	600	497	873
Consistency	80.3%	84%	89.4%	93.1%

4.2. Statistical analyses

We perform the statistical analysis to quantify the difference of acoustic manifestation of Hypo- and Hyper-articulation.

Firstly we select and extract relevant parameters. According to the analysis in 2.2, two basic sets of parameters involved are, (a) those which reflect the relationship of two contiguous tones; (b) those related with the extent of realization of tonal targets. The first group includes: pitch reset at the syllable boundary (P_R), difference of F0 value between peaks of two successive tones (ΔH), and two valleys (ΔL), as well as the duration of silence between adjacent syllables (Sil). The second group of parameters vary with different tonal targets. T1, T2, T3 and T4 in Mandarin are known as high-level, mid-rising, low-falling and high-falling tones respectively. Based on this, the pitch-related parameters include: pitch range (PR) and average pitch (P_{Avr}) for T1; rising range (RR) and slope (K) for T2; falling range (FR) and the lowest pitch (P_L) for T3; FR and K for T4; along with the duration of two syllables (Dur_{S1} and Dur_{S2}). Among them, Dur_{S1} and Dur_{S2} as well as "Sil" are time-domain parameters which are common for each tonal pair, while others are pitch-related and vary with the specific category of tones in pair. For example, ΔL is not relevant in T1T1. For T4T4, the pitch-related parameters include: FR_{S1}, FR_{S2}, K_{S1}, K_{S2}, P_R, ΔH and ΔL. These parameters are computed as follows (refer to Figure 4).

$$\mathrm{FR} = F0_H - F0_L \quad (1)$$

$$\mathrm{K} = FR/(t' - t) \quad (2)$$

$$P_R = F0_{H2} - F0_{L1} \quad (3)$$

$$\Delta\mathrm{H} = F0_{H2} - F0_{H1} \quad (4)$$

$$\Delta\mathrm{L} = F0_{L2} - F0_{L1} \quad (5)$$

Based on the labeling of tone nuclei (Figure 4), we extract physical parameters through a Perl script. Then we conduct One-way ANOVAs in SPSS 21.0 with coarticulatory effect (1, 0B, 0L and 0R) as the independent variable. The results presented in Table 3, indicate that the main effect of tonal coarticulation are significant ($p<.01$) on all the acoustic parameters; and the effect sizes (η^2) are medium or large according to Cohen's conventions.[39] The results demonstrate that the acoustic parameters of four types of coarticulatory effects could be statistically distinguished.

Table 3: *Results of One-way ANOVAs for T4T4.*

Parameters	F	p	η^2
FR1	36.08	0.000	0.35
FR2	33.03	0.000	0.34
K1	32.73	0.003	0.33
K2	26.72	0.000	0.34
Pr	262.37	0.000	0.98
ΔH	218.91	0.001	0.36
ΔL	102.06	0.000	0.34
DurS1	47.33	0.000	0.45
DurS2	24.64	0.005	0.38
Sil	25.4	0.000	0.41

4.3. The distribution of different effects of tonal coarticulation in Mandarin continuous speech

Table 4 displays the distribution of four kinds of coarticulatory effects in Mandarin continuous speech.

Table 4: *Distribution of coarticulatory effects of 16 tonal pairs.*

	1	0B	0L	0R	Total
Num.	3331	1890	991	802	7014
%	48.9%	27.6%	13.5%	10.1%	100%

As can be seen, the effect of tonal coarticulation is not observed in 48.9% (tag 1) of 16 bi-tonal combinations. This may be due to the influence of prosodic boundaries. As shown in Figure 3, the prosodic boundary between the second and the third syllable set the F0 contours of both syllables free from the influence of each other. This is in accordance with the results of the tone recognition model in [21]. It also agrees with the findings on segmental coarticulation in Mandarin, that the higher the prosodic boundary level is, the less is the cross-boundary segmental coarticulation [40].

Among the coarticulated tonal pairs, 27.6% of the coarticulatory effects is bidirectional, and about 24% is unidirectional (Table 4). The coexistence of bidirectional and unidirectional tonal coarticulation in continuous speech is attributable to the different patterns of stress or prosodic strength. For example, a syllable with weak prosodic strength, i.e. less articulatory efforts, will be affected by rather than exert influence on neighboring syllables [29, 30]. This leads to the presence of unidirectional coarticulatory effect.

The results in Table 4 demonstrate that while tonal coarticulation in words or phrases is generally found to be bidirectional, its effect in continuous speech may partially or completely be absent. The prosodic boundaries and different patterns of stress in continuous speech can account for the phenomena which are not found within words and phrases.

By comparing the results of each of 16 bi-tonal combinations, we find that the distribution of coarticulatory effects in some tonal pairs are apparently different from the others. For instance, in Table 5, the rate of 0L (highlighted) is much higher than that of 0B and 0R in T2T2; while 0R occurs more frequently than both 0B and 0L in T2T4. We therefore assume the distribution of coarticulatory effects is influenced by the specific tonal context. In order to test the possibility, we conduct Chi-square test of independence. The result shows that the effect of tonal context is highly significant, $\chi2\ (45) = 768.92$, $p< .001$. It indicates that the distribution of coarticulatory effects is significantly affected by the specific tonal context. This result provides further support for the findings of Gandour et al. [4, 5] and Wang [41] that the extent of tonal variations and the direction of tonal coarticulation can partly be attributable to the tonal context.

Table 5: *Distribution of coarticulatory effects of T2Tx.*

Tonal pair	**1**	**0B**	**0L**	**0R**
T2T1	47.7%	24.9%	6.8%	20.5%
T2T2	46.4%	16.9%	29.0%	7.7%
T2T3	60.8%	19.7%	8.0%	11.5%
T2T4	45.5%	17.8%	9.7%	26.9%

5. Conclusion

The present study proposes a labeling scheme for examining tonal coarticulation in Mandarin continuous speech. The proposal contributes in three ways. First, by applying H&H theory, we explain the systematic nature of complex acoustic variations of tones in Mandarin continuous speech. Second, the scheme unites the facts and findings in previous studies on both local tonal coarticulation and global prosodic events. Finally, the labeling scheme establishes acoustic rather than perceptual standards to determine the labels. These standards are explicit and therefore easy for transcribers to apply.

It was shown that the results of transcribers are highly consistent, and the different types of tonal coarticulation are statistically distinct. The results also prove our assumption that the effect of tonal coarticulation in continuous speech can be bidirectional or unidirectional, or even absent. We conclude that the proposed labeling scheme is feasible for investigating tonal coarticulation in Mandarin continuous speech. It also sheds some light on the future studies of other tonal languages.

6. Acknowledgements

This work is supported by Advanced Innovation Center for Language Resources and Intelligence (KYR17005), Research funds of State Language Commission (ZDI135-51), Wutong Innovation Platform of Beijing Language and Culture University (16PT05) and the Fundamental Research Funds for the Central Universities (18YJ030006). The asterisked author is the corresponding author.

7. References

[1] M. S. Han, and K. Kim, "Phonetic variation of Vietnamese tones in disyllabic utterances," *Journal of Phonetics*, 2, pp. 223 – 232, 1974.

[2] M. Brunelle, "Northern and Southern Vietnamese tone coarticulation: A comparative case study," *Journal of Southeast Asian Linguistics,* vol. 1, pp. 49–62, 2009.

[3] A. S. Abramson, The coarticulation of tones: An acoustic study of Thai. *Bangkok: Chulalongkorn University Press,* 1979: 1-9.
[4] J. Gandour, S. Potisuk, S. Dechongkit, and S. Ponglorpisit, "Tonal coarticulation in Thai disyllabic utterances: a preliminary study," L*inguistics of the Tibeto - Burman Area*, 1992, 15, 93 – 110.
[5] J. Gandour, S. Potisuk, and S. Dechongkit, "Tonal coarticulation in Thai," *Journal of Phonetics*, 1994, 22, 477 – 492.
[6] Y. Xu, "Contextual tonal variations in Mandarin," *Journal of Phonetics*, vol. 25, pp. 61–83, 1997.
[7] Y. Wong, "Contextual tonal variations and pitch targets in Cantonese," *Proceedings of Speech Prosody, May 2-5, Dresden, Germany*, 2006, pp. 317-320.
[8] J. Zhang and J. Liu, "Tone Sandhi and Tonal Coarticulation in Tianjin Chinese," *Phonetica*, 2011, 68(3):161.
[9] Q. Li and Y. Chen, "An acoustic study of contextual tonal variation in Tianjin Mandarin," *Journal of Phonetics,* 2016, 54(8):123-150.
[10] S. Chen, C. Wiltshire, and B. Li, "An updated typology of tonal coarticulation properties," *Taiwan Journal of Linguistics.* 16(2), accepted.
[11] Y. Chang, and F. Hsieh, "Tonal coarticulation in Malaysian Hokkien: a typological anomaly?" *Linguistic Review,* 2012, *29*(1), 37-73.
[12] P. Sarmah, L. Dihingia, and W. Lalhminghlui, "Contextual variation of tones in Mizo," in *INTERSPEECH 2015 – 16th Annual Conference of the International Speech Communication Association, Dresden, Germany, September 6-10,* 2015, pp. 983–986
[13] X. S. Shen, "Tonal articulation in Mandarin," *Journal of Phonetics*, vol. 18, pp. 281–295, 1990.
[14] M. Lin and J. Yan, "Tonal coarticulation patterns in quadrisyllabic words and phrases with normal stress of Mandarin," *Chinese Journal of Acoustics,* 1993(3):245-257.
[15] Y. Xu, "Production and perception of coarticulated tones," *J.A.S.A*, vol. 95, no. 4, pp. 2240–2253, 1994.
[16] Y. R. Chao, *A Grammar of Spoken Chinese*. Berkeley: University of California Press, 1968.
[17] Y. Quan, "An experimental study on stress in Mandarin sentence," *Ph.D. dissertation, Peking University*, 2001.
[18] Y. Jia, "Phonetic realization and phonological analysis of focus in standard Chinese," *Ph.D. dissertation, Nankai University*, 2009.
[19] W. Cao, *The Prosodic Realization of Focus and Stress in Chinese.* Peking: Beijing language and Culture University Press, 2010.
[20] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *Journal of Phonetics*, vol. 27, pp. 55–105, 1999.
[21] J. Zhang and H. Kawanami, "Modeling carryover and anticipation effects for Chinese tone recognition," In *Proceeding of Eurospeech'99, September, Budapest, Hungary*, 1999, pp. 747–750.
[22] J. Zhang and K. Hirose, "Tone nucleus modeling for Chinese lexical tone recognition," *Speech Communication*, vol. 42, no. 1, pp. 447–466. 2004.
[23] Lindblom, "Explaining Phonetic Variation: A Sketch of the H&H theory," *Speech Production and Speech Modelling*, Springer Netherlands, pp. 403–439, 1990.
[24] M. Lin, "An experimental study on tone sandhi of disyllabic words in Mandarin," *Studies of the Chinese Language*, no. 1, 1980.
[25] Z. Wu, "Tonal variations in Mandarin Speech," *Studies of the Chinese Language*, no. 6, 1982.
[26] J. Kong and S. Lv, "A VQ Study on Pitch Models of Disyllable in Mandarin," *ACTA ACUSTICA*, vol. 25, no. 2, pp. 166–174, 2000.
[27] K. de Jong, M. E. Beckman, and J. Edwards, "The interplay between prosodic structure and articulation," *Language and Speech*, vol. 36, no. 2, 3, pp. 197–212, 1993.
[28] H. Fujisaki and K. Hirose, "Analysis of voice fundamental frequency contours for declarative sentences of Japanese," *J. Acoust. Soc. Jpn*, vol. 5, no. 4, pp. 233–242, 1984.
[29] C. Shih, "Understanding phonology by phonetic implementation." in *INTERSPEECH 2005 – Eurospeech, European Conference on Speech Communication and Technology, Lisbon, Portugal, September*, 2005, DBLP, 2469-2472.
[30] C. Shih, "Linking phonology and phonetics: an implementation of model of tones," *Interfaces in Chinese Phonology: Festschrift in Honor of Matthew Y. Chen on His 70th Birthday*, pp. 99–120, 2008.
[31] Ohala, J. John, and W. G. Ewan, "Speed of Pitch Change," *Journal of the Acoustical Society of America* 53. no.1. 1973.
[32] J. Sundberg, "Maximum speed of pitch changes in singers and untrained subjects,'' *Journal of Phonetics*, vol. 7, pp. 71–79. 1979.
[33] Y. Xu and X. Sun, "Maximum speed of pitch change and how it may relate to speech," *Journal of the Acoustical Society of America*, vol. 111, no. 3, pp. 1399-1413, 2002.
[34] G. Kochanski, C. Shih, and H. Jing, "Quantitative measurement of prosodic strength in Mandarin," *Speech Communication*, vol. 41, no.4, pp.625–645, 2003.
[35] G. N. Clements and K. Ford, "Kikuyu tone shift and its synchronic consequences," *Linguistic Inquiry*, vol. 10, pp. 179–210, 1979.
[36] M. Y. Liberman and P. Alan, "On stress and linguistic rhythm," *Linguistic Inquiry*, vol. 8, no. 2, pp. 249–336, 1977.
[37] M. Y. Chen, "Metrical structure: evidence from Chinese poetry," *Linguistic Inquiry*, vol. 10, no. 3, pp. 371–420, 1979.
[38] A. Li, M. Lin, X. Chen, et.al, "Speech corpus of Chinese discourse and the phonetic research," in *6th ICSLP 2000, October 16–20, Beijing, China, Proceedings*, 2000, pp. 42–47.
[39] J. Cohen *Statistical power analysis for the behavioral sciences*. (2nd Ed.) New Jersey: Lawrence Erlbaum Associates, 1988.
[40] Y. Zheng, J. Cao, and H, Bao "Coarticulation and prosodic hierarchy," *Proceeding of Tal, April, Laruchel, France,* 2006, pp. 27-29.
[41] Y. Wang, "The coarticulation and acquisition of Tone 2," *Language Teaching and Linguistic Studies*, no. 4, pp. 94-104, 1997.