

Reanalyze Fundamental Frequency Peak Delay in Mandarin

Lixia Hao, Wei Zhang, Yanlu Xie, Jinsong Zhang*

Beijing Language and Culture University, China

haolixiaxgl@sohu.com, wei_zhang_mail@126.com, {yanluxie, jinsong.zhang}@blcu.edu.cn

Abstract

In Mandarin, Fundamental Frequency (F0) peak delay has been reported to occur frequently in the rising (R) tone or high (H) tone succeeding by a low (L) tone. Its occurrence was ascribed to articulatory constraints within a conflicting tonal context: a high offset target followed by a low onset target. To further examine the underlying mechanism of the phenomenon, the current study tests the possibility that valley delay, as opposed to peak delay, may occur in an L+H tonal context; and peak or valley delay may also occur within a compatible tonal context where adjacent tonal values are identical or similar. An experiment was done on Annotated Speech Corpus of Chinese Discourse to investigate the frequency of occurrence and amount of peak and valley delay. The results indicated that: F0 peak and valley delay frequently occurred in both conflicting and compatible tonal contexts; the phenomenon was found extensively in R tone and F (falling) tone, but barely in H tone and L tone. The findings suggest that while peak or valley delay is partially due to articulatory constraints in certain tonal contexts, the speakers' active effort-distribution strategy based on economical principle is also behind the phenomenon.

Index Terms: peak delay, valley delay, tonal context, economical principle

1. Introduction

F0 Peak delay (henceforth peak delay) refers to the phenomenon that an F0 peak sometimes occurs after the syllable it is associated with [1]. It was highly concerned and has been widely reported for a number of languages. Steele (1986) found the F0 peak on the nuclear stress could be located to the post-nuclear syllables in English [2]. Silverman and Pierrehumbert (1990) reported that F0 peak of a prenuclear pitch accent also occurred after the syllable it is phonologically supposed to align with [3]. Similar phenomena were commonly found in Spanish [4, 5], Korean [6] and Greek [7].

While F0 peak is associated with pitch accent in some languages such as English, it is also an acoustic manifestation of lexical tones in Mandarin. F0 peak delay in the R tone has been widely reported [8, 9, and 10]. However, it has not been found in the H tone until Xu [1]. He proposed that the occurrence of peak delay in Mandarin, rather than due to phonological misalignment between segments and pitch units, is attributable to the articulatory constraints in certain tonal contexts. Specifically, when an R or H tone is succeeded by an L tone, just like the data collected in [1], a sharp turn has to be implemented at the syllable boundary. It takes time for the larynx to first produce a pitch rise, then stop, turn and start to lower the pitch. As a result, F0 peak associated with R tone or H tone appears in the following L-carrying syllable. The explanation revealed the underlying mechanism of peak delay in Mandarin, and at the same time raised new questions.

First, if the peak delay in Mandarin generally occur when there is a sharp turn caused by a high offset followed by a low onset, we would expect to see a valley delay, as opposed to peak delay, at a sharp turn caused by a low offset followed by a high onset. In both tonal contexts, the adjacent tonal values are different. These are defined as “conflicting” contexts by Xu [11]. In contrast, when adjacent tonal values are identical or similar, it is a “compatible” context. Among four lexical tones in Mandarin (see Table 1), H tone and R tone have offset high targets that may lead to a F0 peak in speech, while L tone and F tone have offset low targets that may lead to a F0 valley. The possibility of occurrence of valley delay needs to be tested.

Table 1: Four lexical tones in Mandarin.

Tonal category	Phonological description
Tone 1 or H tone	High-level
Tone 2 or R tone	Mid-rising
Tone 3 or L tone	Low-dipping
Tone 4 or F tone	High-falling

Second, if the occurrence of peak delay in Mandarin, and maybe also valley delay, is conditioned in the conflicting context, F0 peak and valley are supposed not to delay in a compatible tonal context where no sharp turn existed at syllable boundary, such as a tonal sequence of R+F or F+R (see Figure 1). However, unlike what is shown in Figure 1(a), the F0 peak has been observed within the F-carrying syllable regularly in a RF tonal combination [10]. Whether the F0 valley appears within the R-carrying syllable in a FR tonal combination? If so, there should be coherent links between these F0 patterns in compatible tonal contexts.

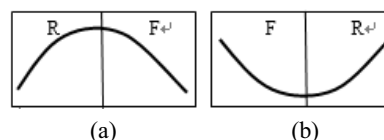


Figure 1: Schematic illustration of underlying pitch targets of tonal combinations: RF (a) and FR (b).

Third, assuming both the peak and valley delay occur in the compatible tonal contexts, the articulatory constraints proposed for conflicting contexts may be inadequate to fully account for the phenomena. Are there other mechanisms behind it? These unsolved issues need to be addressed.

2. Experiment

In earlier studies on peak delay, experimental materials were often specially designed. On the one hand, it is convenient for controlling irrelevant factors and manipulating independent variables; on the other hand, the simple designed materials might cover the complexity of the issue, and have only limited explanatory power due to small amount and less natural semantic structure. The present study therefore selects target tonal sequences from an annotated discourse corpus.

2.1. Corpus

We conduct the experiment on Annotated Speech Corpus of Chinese Discourse [12] in which each speaker read 18 passages and 8762 syllables altogether at normal speech rate. We use the data of F001 (female speaker, number 1) for this study.

2.2. Data processing

2.2.1. Syllable selection

From all the 16 tonal combinations in Mandarin, we select two tonal pairs to verify peak delay in conflicting contexts (H tone + L tone, and R tone + L tone); and two pairs to test the possibility of the occurrence of valley delay (L tone + H tone, and F tone + H tone), and other two pairs to examine whether peak and valley delay occur in the compatible contexts (R tone + F tone, and F tone + R tone). Then, some materials are taken out in order not to affect the judgments of the occurrence of peak or valley delay. These include tonal pairs with voiceless initial consonants in the second syllables, to avoid the F0-raising effect of the unvoiced initial segments [13, 14 and 15] and to reduce possible errors in pitch tracking; and those with punctuation marks in between; as well as some tonal pairs whose F0 contours have largely deviated from their underlying targets, such as a R tone has a falling contour or a F tone has a rising contour. The distorted F0 contours are recognized applying tone nuclei model [16]. The syllable selection step is done through a set of Perl programs written for current study. Table 2 presents the selected amount of the six tonal pairs.

Table 2: Tonal pairs processed in the experiment.

Tonal pair	HL	RL	RF	LH	FH	FR
Selected num.	75	66	83	13	35	135

2.2.2. Data extraction

The annotations of ASCCD include (while not limited to) segmentation labels between syllables and between the initial and the final within a syllable. For further analysis, we extracted the F0 and time of syllable onset and offset, peak or valley. The F0 peak or valley was positioned by finding the maximum or minimum value in the F0 contour. F0 contour used here starts from the onset vowel in syllable 1 (S1) to the offset of syllable 2 (S2). We complete the data extraction by writing another Perl program and then analyze the output in SPSS 21.0 and G*Power 3.1.9.2. Some missed F0 values in the output, due to pitch-period skipping errors or creaky voice of the L tone, are manually inserted with the aid of Praat.

2.2.3. Measurements

The measurements include: 1) syllable duration; 2) Peak/Valley-to-S2 (ms) which is the location of the F0 peak or valley relative to the onset of S2 (also the offset of S1). It refers to distance of peak or valley delay. The value is positive if the F0 peak or valley occurs later than S1, otherwise, it is negative. 3) Peak/Valley-to-S2 (Hz), which is the F0 value of peak or valley minus that of the onset of S2. It refers to magnitude of peak or valley delay. When peak or valley delay occurs, the value is positive for the former, and negative for the latter.

3. Results

The frequency of occurrence, distance and magnitude of peak and valley delay are shown as follows.

3.1. Frequency of occurrence of peak and valley delay in various tonal contexts

Figure 2 displays frequency distribution of occurrence of peak or valley delay. H, R, L and F represent H tone, R tone, L tone and F tone respectively. As can be seen, F0 peak or valley occurs after the syllable offset frequently in two conflicting tonal contexts (RL, FH), and regularly in two compatible tonal contexts (RF, FR); generally in R tone (RL, RF) and F tone (FH, FR) while barely in H tone (HL) and L tone (LH).

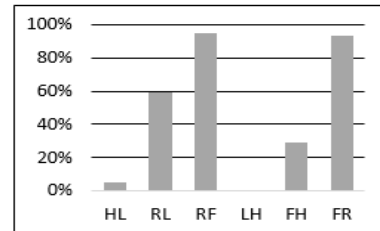


Figure 2: Frequency of occurrence of peak or valley delay in various tonal contexts.

3.2. Distance of peak or valley delay in R and F tone, and its correlation with syllable duration

Figure 3 shows Peak/Valley-to-S2 (ms), which indicates distance of peak or valley delay, in R and F tone. As earlier mentioned, the phenomena barely found in the other two tones.

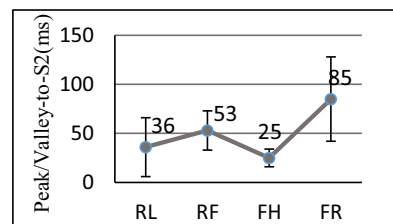


Figure 3: Distance of peak or valley delay in R and F tone.

The peak or valley delayed further in compatible contexts than in conflicting contexts in either R tone (RF > RL) or F tone (FR > FH). To examine the relationship between distance of F0 peak or valley delay and the duration of the syllable it is associated with, we plotted Peak/Valley-to-S2 (ms) against Duration of S1 (see Figure 4).

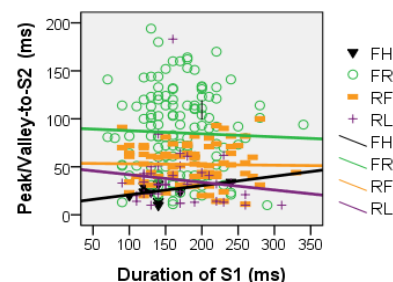


Figure 4: Peak/Valley-to-S2 (ms) plotted against Duration of S1. Different symbols represent various tonal contexts. Equations for all regression lines and r^2 values are given below.

FH: $y = 0.0969x + 11.156$; $r^2 = 0.1609$; FR: $y = -0.0325x + 90.924$; $r^2 = 0.0012$; RF: $y = -0.0073x + 53.867$; $r^2 = 0.0004$; RL: $y = -0.08x + 49.843$; $r^2 = 0.0176$.

As demonstrated in Figure 4, the coefficient of regression and the coefficient of determination (r^2) for all tonal pairs are quite small. This indicates that the location of F0 peak or valley

scarcely vary with syllable duration. Syllable duration does not play an important role in determining F0 peak or valley location.

3.3. Magnitude of peak or valley delay in R and F tone

Similar as what we found in 3.2, the magnitude of peak or valley delay is greater in compatible contexts than in conflicting contexts in either R tone (RF > RL) or F tone (FR > FH), as illustrated in Figure 5.

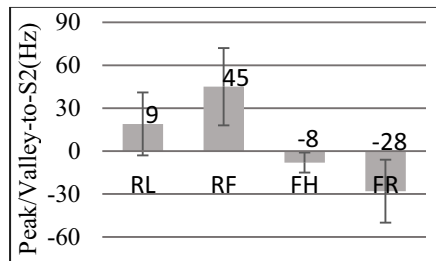


Figure 5: Magnitude of peak or valley delay in R and F tone.

3.4. Significance test of difference

Two-factor ANOVAs are performed on magnitude and distance of peak or valley delay with tone (R and F) and tonal contexts (conflicting and compatible) as the independent variables. The value of Valley-to-S2 (Hz) for F tone is converted to absolute value for the convenience of comparison. Beside F value and significance level, we also calculate effect size (η^2) and statistical power ($1-\beta$) for evaluating the strength of statistical claims and the possibility of the effect actually exists, as presented in Table 3. Dis and Mag are short for distance and magnitude respectively.

Table 3: The results of two-factor ANOVAs.

Factor	Variable	df	F	Sig.	η^2	$1-\beta$
Tone	Dis.	1	2.72	.100	.011	.174
	Mag.	1	10.26	.002	.039	.728
Tonal context	Dis.	1	33.28	.000	.117	.999
	Mag.	1	25.57	.000	.093	.993
Tone *	Dis	1	11.1	.001	.043	.768
Tonal context	Mag	1	.557	.456	.002	.034

It is revealed that the main effects of tone on magnitude, and tonal context on both distance and magnitude, as well as the interactive effect of tone*tonal context on distance are significant at 0.001 or 0.01 level (highlighted in light grey). Among them, however, only the main effects of tonal context on distance and magnitude are medium effect size and have large statistical power, according to Cohen's guideline [17]. To sum up, the statistical results demonstrate that both the distance and magnitude of peak or valley delay are significantly greater in compatible tonal contexts than in conflicting tonal contexts.

4. Discussion

4.1. The F0 peak delay and valley delay occur in both conflicting and compatible tonal contexts

Results of experiment support the hypotheses raised in Section 1 that both F0 peak and valley regularly occur after syllable offset, not only in conflicting tonal contexts, but also in compatible tonal contexts. Contrary to the peak delay, which has been discussed a lot, the valley delay seems has not been noticed before, except that Xu (2002) has mentioned its

possible occurrence [18]. When the F tone followed by the R tone, at syllable boundary, the larynx needs to terminate a sharp fall, turn, and then start to make a rise. The mechanism put forward for peak delay in Mandarin [1] can also account for valley delay. Actually, the phenomenon of valley delay in Mandarin were scattered in literature [10, 19, 20 and 21], but has been mostly ignored in these studies, and was not interpreted as valley delay.

The F0 peak and valley delay regularly occur in compatible tonal contexts (more than 90%, see Figure 2). One could argue that when the F0 peak appears in the F tone which is preceded by the R tone, it should be viewed as onset target of F tone [20] rather than peak delay. Then how can we explain the F0 valley appears in the R-carrying syllable within a FR tonal combination? Given the phonological description of R tone is mid-rising, as depicted in Table 1, its onset F0 is generally supposed to be higher than the offset F0 of F tone. Therefore, a F0 valley in R-carrying syllable preceded by an F tone should be considered as valley delay rather than the onset of R tone. Likewise, the F0 peak in the F tone preceded by an R tone is better to be treated as F0 peak delay since we view the phenomena in a coherent and systematic way.

4.1.1. Comparison of the results between current and earlier studies

The peak delay in H tone and R tone were investigated in both current study and [1]. At normal speech rate, the results are different in two aspects. First, peak delay in H tone occurs barely in this study (5%), while occasionally in [1] (14.3%). The occurrence of peak delay in the H tone has not been reported before [1]. His conclusions were based on mean values of four speakers while there were individual differences. In that study, the frequency of occurrence of peak delay in H tone was 0 for one male and one female speakers, and 28.6% for both the other two. The results might be considered as individual speaking styles rather than a general rule. The findings of the present study do not support the view that peak delay occurs in H tone. Second, the frequency of occurrence and the distance of peak delay in R tone succeeding by the L tone were smaller in this study. We will discuss the reasons in 4.3.

4.1.2. Different results between dynamic and static tones

The F0 peak or valley delay occurs extensively in R tone and F tone, but barely in H tone and L tone. As mentioned before, the former two tones are intrinsically dynamic and the latter two are static high or low. Dynamic tones in Mandarin do not have binary tonal contrasts such as high rise vs low rise, or high fall vs low fall. Accordingly, the peak or valley delay in dynamic tones leads to few perceptual confusion. Therefore, Mandarin may be a language that would allow for more frequent peak delay with greater amount in R tone, comparing to other tonal languages with more than one rising tones such as Vietnamese and Cantonese. Even though such comparisons have not been actually made, we try to create possibilities for the mechanisms proposed here to be verified, or falsified.

Two other reasons proposed before [21] may also account for the results. One is that the dynamic tones have two successive pitch targets to be realized. They are under bigger time pressure that results in offset target undershoot within syllable. Another is that beside the F0 height change, a change of F0 movement direction in dynamic tones also takes time.

The last question we raised in Section 1 is the underlying mechanism of the peak or valley delay.

4.2. The underlying mechanism of the occurrence of F0 peak and valley delay in Mandarin

According to [1], peak delay is ascribed to the articulatory constraints when a sharp turn appears at syllable boundary. This proposal has explanatory power for both the peak and valley delay in conflicting tonal contexts. However, it is inadequate to fully explain the data presented in this study.

The first is peak and valley delay in compatible tonal contexts such as RF and FR where no sharp turn to be made. The second is the issue of time pressure. Results of regression analysis (see Figure 4) demonstrate that location of F0 peak or valley barely vary with syllable duration. Moreover, if time pressure is the case, speakers are able to lengthen syllable duration as long as there is a need. The well-known pre-boundary lengthening [22, 23] is an example. The third is the significant effect of tonal context with large statistical power (see Table 3). Both the distance and magnitude of peak and valley delay in conflicting contexts (RL and FH) are smaller than that in compatible contexts (RF and FR). In short, the articulatory constraints within certain tonal contexts did explain some data of peak and valley delay while the reasons for other observed data remains unclear. Then what is it?

The economical strategy actively applied by the speakers, we assume, is a critical cause. In a compatible context, two identical tonal targets appeared at syllable boundary. An obviously economical way is to complete just one H or L target for two successive syllables to share it. For instance, the target-sequence of tonal pair RF could be simplified from LHHH to LHL, and the H target can be realized in either S1 or S2. In this case, it is effort-saving to utilize the articulatory inertia and reach the H target somewhere in S2 rather than realizing it in S1 and keep vocal cord vibrating fast for a while. Similarly, that could be what happened in the tonal pair FR. In other words, the peak or valley delay is the way of effort-saving, especially in the compatible tonal context. Because two successive identical or similar pitch targets make the peak or valley delay more comfortable. Likewise, at fast speech rate, the speaker's effort available for each tone is smaller than that at normal speech rate, hence the peak delay occurs more extensively, as reported in [1]. The speakers' active effort-distribution-strategy is behind the phenomena. In this way, we can consistently and systematically explain the difference found in current and earlier studies: peak and valley delay occur greater in compatible tonal contexts than in conflicting ones, and more extensive at fast speech than at normal speech.

The economical strategy and articulatory constraints in certain tonal contexts, albeit seems very similar, are essentially different from each other. Speakers produce a peak or valley delay not only due to the pressure of time and lagged movement of larynx, but also out of compromise between economical principle and communication needs. The peak or valley delay, as a natural results of effort-saving strategy, is the phonetic form of hypo-articulation for tonal combinations. The pattern will be broken as long as speakers recognize a communication need and spend extra effort for it. For instance, as which has been found in Spain, a pause for semantic need could serve as an inhibitor of peak delay [24]. We assume the pattern without peak or valley delay is the form of hyper-articulation for tonal combinations. Lindblom's H&H theory [25] and the compromise between economical principle and communication needs explain the F0 peak and valley delay in discourse powerfully. The mechanism also account for the different results that peak delay in R tone occurs less extensively in

current study than in earlier study, as we discussed in 4.2.1. Compared to the simple sentences designed especially for experiment, the materials from discourse in this study carry more communication and functional information which requires extra effort to break the F0 pattern of hypo-articulation, i.e., peak or valley delay.

4.3. The influence of F0 peak delay in Mandarin on tonal acquisition of CSL.

The "uncompleted form" of tones caused by peak or valley delay extensively existed in communication. It is considered not as errors but as a natural way [26] due to a compensation effect of adjacent tonal information [27] as well as the ability of categorical perception of native speakers [28, 29]. In contrast, learners who study Chinese as a second language (CSL) generally do not have the abilities [30, 31]. A natural implication is that the peak or valley delay in teachers' language in CSL occurs less frequently with less amount than that in communication among native speakers. The issue has not been addressed yet. By making such a specific prediction, we again attempt to enable the mechanisms to be proven right, or wrong.

For CSL learners, a widely reported error is the confusion of R tone and L tone in both speech perception and production, practically across all language background [30, 31, 32 and 33]. One of the causes, we assume, is the F0 peak delay in the R tone. As introduced before, R tone is the only rising tone in Mandarin, nevertheless, the L tone associated with a monosyllable or appeared at the end of an utterance could show a low rise [24]. The main difference between the mid-rise and low rise is the offset pitch height which has been found as a critical cue to distinguish R tone and L tone [30]. The F0 contour of R tone under the effect of peak delay becomes much similar with the L tone with a low rise, and then leads to confusion for CSL learners. Especially, they learn Chinese from monosyllabic tones at the very beginning. The L tone with a low rise has strong influence as the first impression.

5. Conclusions

The present study examines the F0 peak and valley delay in Mandarin and the underlying mechanisms. The results demonstrate that apart from peak delay in the R tone, valley delay frequently occur in the F tone, within either conflicting or compatible tonal contexts. But the phenomena are barely found in H tone and L tone. The findings suggest that in addition to articulatory constraints in certain tonal contexts, the speakers' effort-distribution strategy based on economical principle probably is also an important underlying mechanism. The accounts offer a new way to understand the peak delay phenomenon and mechanism in Mandarin and possibly in other languages as well. The conclusions of this study need to be evaluated with more data. Besides, how the F0 peak and valley delay affect the production and perception of tones in CSL awaits further study.

6. Acknowledgements

This work is supported by funds of Advanced Innovation Center for Language Resources and Intelligence, and the Special Program for Key Basic Research fund of Beijing Language and Culture University (the Fundamental Research Funds for the Central Universities)(16ZDJ03), as well as the Research Funds of Beijing Language and Culture University (16YCX219). The asterisked author is the corresponding author.

7. References

- [1] Y. Xu, "Fundamental frequency peak delay in Mandarin," *Phonetica*, vol. 58, pp. 26-52, 1-2, 2001.
- [2] Steele, S. A, "Nuclear accent F0 peak location: effects of rate, vowel, and number of following syllables," *J. acoust. Soc. Am*, vol. 80, S51, 1986.
- [3] K. E. A. Silverman and J. B. Pierrehumbert, "The timing of prenuclear high accents in English," in Kingston, Beckman. *Papers in Laboratory Phonology 1: Between the Grammar and Physics of Speech*, pp. 72-106, Cambridge: Cambridge University Press, 1990.
- [4] J. M. Garrido, J. Llisterri, C. de La Mota and A. Ríos, "Prosodic differences in reading style: isolated vs. contextualized sentences," *Proceedings of Eurospeech '93*, 21-23 Sep, Berlin, 1993, pp. 573-576.
- [5] P. Prieto, J. V. Santen and J. Hirschberg, "Tonal alignment patterns in Spanish," *J. Phonet*, vol. 23, pp. 429-451, 1995.
- [6] K. D. Jong, "Initial tones and prominence in Seoul Korean," Ohio State University, *Working Papers Ling*, vol. 43, pp. 1-14, 1994.
- [7] A. Arvaniti, D. R. Ladd and I. Mennen, "Stability of tonal alignment: the case of Greek prenuclear accents," *J. Phonet*, vol. 36, pp. 3-25, 1998.
- [8] C. Shih, "Tone and intonation in Mandarin," *Working Papers, Cornell Phonetics Laboratory*, vol. 3, pp. 83-109, 1988.
- [9] C. Shih, "Understanding phonology by phonetic implementation," in *INTERSPEECH 2005 – Eurospeech, European Conference on Speech Communication and Technology, September, Lisbon, Portugal*, DBLP, 2005, pp. 2469-2472.
- [10] Y. Wang, "The coarticulation and acquisition of Tone 2," *Language Teaching and Linguistic Studies*, no. 4, pp. 94-104, 1997.
- [11] Y. Xu, "Production and perception of coarticulated tones," *J.A.S.A.*, vol. 95, no. 4, pp. 2240-2253, 1994.
- [12] 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.
- [13] I. Lehiste and G.E. Peterson, "Some basic considerations in the analysis of intonation," *J. acoust. Soc. Am*, vol.33, pp. 419-425, 1961.
- [14] J.M. Howie, "On the domain of tone in Mandarin," *Phonetica*, vol. 30, pp. 129-148, 1974.
- [15] Hombert and J. M, "Consonant types, vowel quality, and tone," in Fromkin, *Tone: linguistic survey*, pp. 77-111, New York: Academic Press, 1978.
- [16] J. Zhang and K. Hirose, "Tone nucleus modeling for Chinese lexical tone recognition," *Speech Communication*, vol. 42, no. 1, pp. 447-466, 2004.
- [17] J. Cohen, *Statistical power analysis for the behavioral sciences* (2nd Ed.). Hillsdale, NJ: Erlbaum, 1988.
- [18] Y. Xu, "Articulatory constraints and tonal alignment," in *Proceedings of the 1st International Conference on Speech Prosody*, Aix-en-Provence, France, 2002, pp. 91-100.
- [19] 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.
- [20] Y. Zhao, Y. Li, J. Zhang and Y. Xie, "A study of production of Tone 3 in Mandarin Disyllables by Korean Students," *Proceedings of NCMMSC*, Tianjin, China, 2015.
- [21] Y. Xu, "Effects of tone and focus on the formation and alignment of F0 contours," *J. Phonet*, vol. 27, pp. 55-105, 1999.
- [22] M. Lin, "Break and prosodic phrase in Mandarin utterance," *Contemporary Linguistics*, vol. 2, no. 4, pp. 210-217, 2000.
- [23] B. Wang, Y. Yang and S. Lv, "Acoustic analysis on prosodic hierarchical boundaries of Chinese," *ACTA ACUSTICA*, vol. 29, no. 1, pp. 29-36, 2004.
- [24] J. Llisterri, R. Marín, C. de la Mota and A. Ríos, "Factors affecting F0 Peak Displacement in Spanish," in *Proceedings of the 4th European Conference on Speech Communication and Technology*, Eurospeech '95, Sep. 18-21, Madrid, Spain, vol. 3, pp. 2061-2064, 1995.
- [25] Lindblom, "Explaining Phonetic Variation: A Sketch of the H&H theory," *Speech Production and Speech Modelling*, Springer Netherlands, pp. 403-439, 1990.
- [26] 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.
- [27] T. Lin and S-Y. Wang, "On tone perception," *Chinese Linguistics. The Journal of Linguistic Society of China*, no. 2, pp. 59-69, 1985.
- [28] W. S-Y. Wang, "Language change," *Annals of N.Y. Academy of Science*, vol. 280, no. 1, pp. 61-72, 1976.
- [29] Y. Wang and M. Li, "The Effects of Tone Pattern and Register in Perceptions of Tone 2 and Tone 3 in Mandarin," *Acta Psychologica Sinica*, vol. 42, no. 9, pp. 899-908, 2010.
- [30] Y. Wang and M. Li. "The perception of Tone 2 and Tone 3 in Mandarin by Korean native speakers," *Language Teaching and Linguistic Studies*, no. 1, pp. 17-25, 2011.
- [31] T. Zou, J. Zhang, W. Cao, "A Comparative Study of Perception of Tone 2 and Tone 3 in Mandarin by Native Speakers and Japanese Learners," *Proceedings of ISCSLP2012*, Dec, Hong Kong, 2012, pp. 431-435.
- [32] He. P., "Talk about the elementary Chinese phonetic teaching the Japanese learner," *Language Teaching and Linguistic Studies*, vol. 3, 1997.
- [33] Y. Liu, "Analysis of Mandarin tones produced by Japanese and Korean students," *Chinese Teaching in the World*, no. 1, pp. 95-100, 1998.