International Journal of Engineering Research and Development
e-ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 10, Issue 6 (June 2014), PP.09-15

Study on the Anticipatory Coariticulatory Effect of Chinese Disyllabic Sequences

Maolin Wang
College of Chinese Language and Culture / Overseas Chinese Education Research Institute, Jinan University, Guangzhou, China

Abstract:- In this study, the Vowel-to-Vowel (V-to-V) coarticulatory effect in the Vowel-Consonant-Vowel (VCV) sequences is investigated, and the F2 offset value of the first vowel is analyzed. Results show that, in the trans-segment context, anticipatory coarticulation exists in Chinese. Due to high articulatory strength of aspirated obstruents, in the context of subsequent vowel /i/, the V1 F2 offset value is higher in the context of aspirated stops than the unaspirated ones. Results from this study are not in accordance with the DAC prediction, which implies that the pattern of degrees of articulatory constraint is not consistent among languages.

Keywords:- Aspiration, coarticulation, formant, place of articulation, vowel

I. INTRODUCTION

Coarticulation refers to a situation in which a conceptually isolated speech sound is influenced by, and becomes more like, a preceding or following speech sound. For example, all else being equal, the back vowel [u] in 'two' is produced farther forward than the same vowel in 'who' due to the influence of the adjacent coronal consonant. Coarticulation effect may vary with its specific context or the phonological system of a language. In his classic spectrographic study of VCV sequences in three languages, Öhman [1] found that F_2 values of target vowels varied more due to vowel context in English and Swedish than in Russian. He attributed the coarticulatory differences to the languages' consonant systems, arguing that the requirements on the tongue body imposed by contrastive palatalization in Russian, but not in English or Swedish, restricted transconsonantal coarticulation in Russian. Consonant restrictions on V-to-V coarticulation have also been reported by Recasens [2], who found less V-to-V coarticulation across the velarized lateral of Catalan than across the 'clear' lateral of Spanish and German. He and his colleagues ascribed the coarticulatory differences to different lingual constraints for these laterals.

Regarding to the extent of coarticulation, Bladon and Al-Bamerni [3] originated the concept of 'coarticulatory resistance' that claimed phonetic segments possess inherent properties that limit the extent to which they can be influenced by neighboring segments. Using this concept within a coarticulatory approach to speech production, Recasens [4] developed the 'degrees of articulatory constraint' (DAC) model to account for coarticulatory effects of both vowels and consonants. Recasens' model predicts that the more a specific region of the tongue is involved in the occlusion for the C, the more the C affects V, but the less it can be shaped by the vowel, and the less the transconsonantal V-to-V coarticulation.

In Chinese, there have been a number of studies on the coarticulatory effect of segments, including the analysis of the acoustic coarticulatory patterns of voiceless fricatives in CVCV [5], the study of vowel formant pattern and the coarticulation in the voiceless stop initial monosyllables [6], the acoustic study of intersyllabic anticipatory coarticulation of three places of ariculation of C_2 in CVCV [7], vowel segmental coarticulation in read speech in Standard Chinese [8], and anticipatory coarticulation in V1#C2V2 sequences [9]. It is found that coarticulation exists in segment adjacent and trans-segmental contexts in Chinese.

Coarticulation is a common phenomenon in languages, and it is believed that coarticulation affects the smoothness and naturalness of the synthesized speech in Text-to-Speech. Therefore, the naturalness of synthesized speech will be greatly improved if speech coarticulation is properly solved [10]. The research presented in this paper aims to investigate the V-to-V coarticulation in VCV sequences in Chinese. Coarticulation may be generally classified as carry-over (left-to-right) or anticipatory (right-to-left) ones [11], and the present study will focus on anticipatory coarticulation.

II. METHODOLOGY

A. Speakers, stimuli and recording

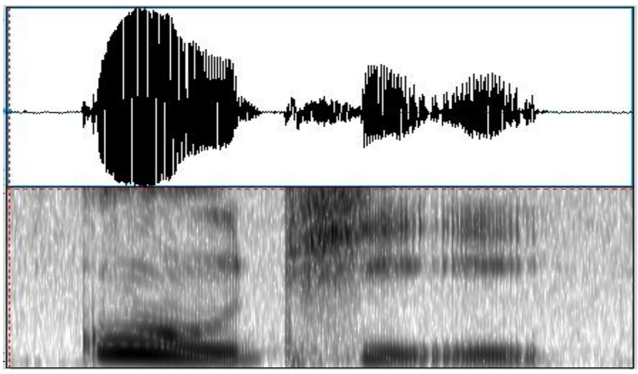
Twelve native speakers of Standard Chinese, six male and six female, participated in the recording. Regarding the stimuli, disyllabic words, in the form of $C_1V_1.C_2V_2$, are used, with V_2 providing the 'changing' vowel context, V_1 the 'fixed' vowel, which is designed for the changing vowels to affect the fixed vowel. The fixed vowel is /u/, and for the changing vowel context, vowels /i/ vs. /u/ are used to influence the offset of the F_2 frequencies of the fixed vowel. The intervocalic consonant C_2 includes /b, p, d, t/, two unaspirate stops /b/, /d/ and two aspirated ones /p/, /t/. All the words used are in normal stress, without neutral tone syllables. An example of a pair of words used are 'guti' and 'gutu', which mean 'solid' and 'homeland' respectively. Two sets of words of identical combinations are used, so there are 16 words in the word list (4 stops × 2 changing vowel contexts × 2 sets).

Recording was done in a sound-treated room, and the acoustic data were recorded directly into the computer at a sampling rate of 16 kHz using the recording software of Cool Edit Pro. The speakers were asked to read the word list three times, in random order for each repetition, in normal pace, so each speaker produced 48 tokens: 16 words × 3 repetitions. In total, 576 tokens were acoustically analyzed (48 tokens × 12 speakers).

B. Procedure and measurements

1) F_2 offset value: This study aims at investigating the extent of V-V coarticulation in VCV sequences, and vowel formant is examined. Formant values are extracted using Praat [12], and the effect of trans-consonantal coarticulation is analyzed by measuring the F_2 offset value of the fixed vowel. F_2 offset frequency is taken at the offset of the fixed vowel V_1.

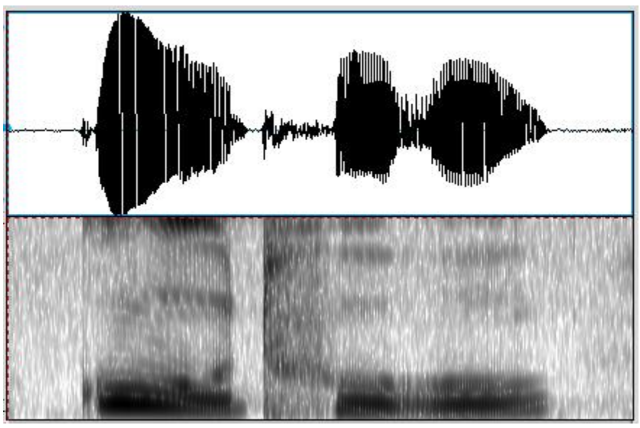
Fig. 1 displays the waveform and spectrogram of 'uti' (Fig. 1a) and 'utu' (Fig. 1b), as in the sequences of 'guti' (solid) and 'gutu' (homeland), with the former an example of unaspirated stop context, and the latter of aspirated context. For both of the two cases, F_2 offet values are taken at the offset points of the first vowel, that is, point 'A' on both graphs.



A

gu ti

(a) 'uti'



A

gu tu

(b) 'utu'

Fig. 1. Waveform and spectrogram of (a) 'uti' and (b) 'utu'

2) F_2 delta: In order to compare the extent of coarticulatory effects under various consonant contexts, besides the F_2 offset values, their differences caused by the changing V_2 contexts are also calculated. Coarticulation effects due to changing V_2 contexts are indexed by F_2 delta values obtained at the V_1 F_2 offset, and F_2 delta (Hz) is derived by computing the difference in offset frequencies of the fixed vowel in each sequence pair, as is shown in formula (1)

$$\Delta F_2 = F_{2i} - F_{2u} \qquad (1)$$

In (1), F_{2i} and F_{2u} refer to the V_1 F_2 offset values preceding vowel /i/ and /u/ respectively, and ΔF_2 is the F_2 delta at the offset of V_1.

Fig. 2 displays the F_2 contours of the sequence pair ‘guti’ (solid) and ‘gutu’ (homeland), with the contour of ‘guti’ in solid line, and that of ‘gutu’ in dashed line. In this sequence pair, for the changing vowel context, F_2 of /i/ is high and that of /u/ is low. If V_1 F_2 offset values differ in this pair, then it is reasonable to attribute the frequency difference to the high vs. low F_2 contexts in V_2. The greater the F_2 delta value is, the greater the coarticulatory effect of V_2 is on V_1.

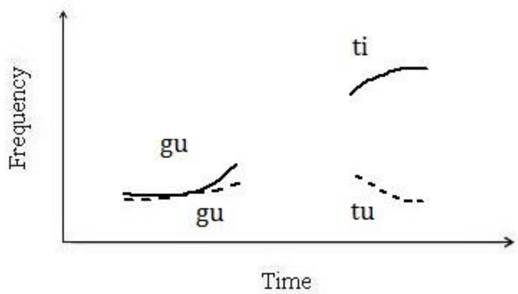


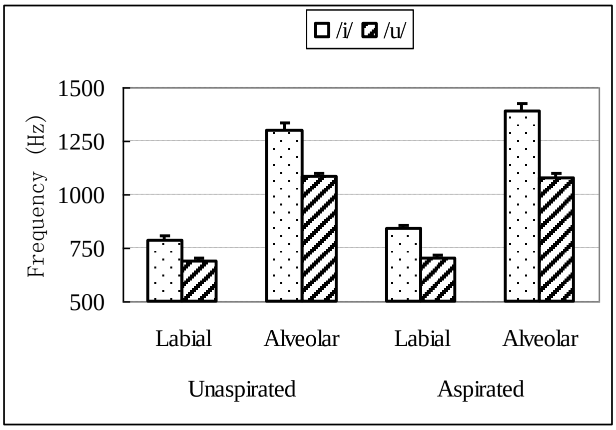
Fig. 2. F_2 contours of the sequence pair ‘guti’ and ‘gutu’, with the former in solid line, and the latter in dashed line

A repeated measures ANOVA was performed with two within-subjects factors— aspiration (unseparated, aspirated) and place of articulation (labial, alveolar).

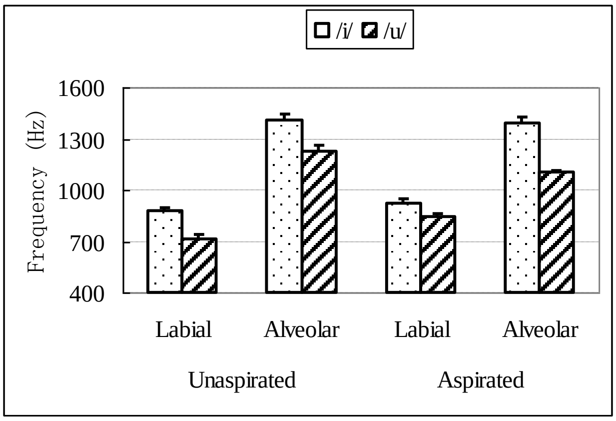
III. RESULTS

A. F_2 value

Fig. 3 graphs the F_2 offset values for male speakers (Fig. 3a) and female speakers (Fig. 3b), broken down by the contexts of aspiration, place of articulation and changing vowels. The changing vowel contexts are indicated by /i/ and /u/, which refer to changing vowel context of /i/ and /u/ respectively. Repeated measures ANOVA results show that, as far as main effect is concerned, there are significant effects for place of articulation and the changing vowel context, place of articulation: $F(1, 71) = 751.4$, $p < 0.001$; changing vowel context: $F(1, 71) = 373.2$, $p < 0.001$, with V_1 F_2 offset values greater in the alveolar context than in the labial context, and greater in the vowel context of /i/ than of /u/. However, there are no significant effect of aspiration: $F(1, 71) = 3.52$, $p = 0.065$.



(a) F_2 values of male speakers



(b) F_2 values of female speakers

Fig. 3. F_2 offset values for male speakers (a) and female speakers (b), broken down by the contexts of aspiration, place of articulation and changing vowels

Further analysis shows that there is significant aspiration × place of articulation interactive effect: $F(1, 71) = 6.58$, $p = 0.012$, aspiration × changing vowel interactive: $F(1, 71) = 5.62$, $p = 0.02$, place of articulation × changing vowel interactive effect: $F(1, 71) = 39.1$, $p < 0.001$, and aspiration × place of articulation × changing vowel interactive effect: $F(1, 71) = 9.25$, $p = 0.003$. In the next subsections, the effects will be elaborated in detail.

1) The effect of aspiration

a) In the context of /i/: When the subsequent vowel is /i/, it is shown from ANOVA results that there is significant effect for it: $F(1, 143) = 5.34$, $p = 0.022$, with V_1 F_2 offset values higher in the context of aspirated consonants than that in the unaspirated context.

b) In the context of /u/: In the context of vowel /u/, ANOVA results show that there is no significant effect of aspiration on the V_1 F_2 offset values: $F(1, 143) = 0.015$, $p = 0.903$, with no significant difference of V_1 F_2 offset values between the unaspirated and the aspirated contexts.

2) The effect of place of articulation

a) In the context of /i/: As for the effect of place of articulation on the V_1 F_2 offset values, When the subsequent vowel is /i/, repeated measures ANOVA results show that there is significant effect of place of articulation on the V_1 F_2 offset values: $F(1, 143) = 608$, $p < 0.001$, with V_1 F_2 offset values in the alveolar context higher than that in the labial context.

b) In the context of /u/: Similar to the context of /i/, in the context of changing vowel /u/, it is shown from ANOVA results that the effect of place of articulation on the V_1 F_2 offset values is also significant: $F(1, 143) = 517.4$, $p < 0.001$, with V_1 F_2 offset values in the alveolar context higher than that in the labial context.

For the purpose of elaborating the coarticulatory effect under various consonant contexts in detail, F_2 delta value will be analyzed in the next subsection. To be specific, the extent under the contexts of place of articulation and aspiration will be presented.

B. F_2 delta value

Fig. 4 shows the F_2 delta under the effects of place of articulation and aspiration, and Table 1 presents the F_2 delta means and significance results for the main effects. From Table 1 it can be seen that, in terms of overall main effect, there is significant effect for both place of articulation and aspiration, with the effect in the alveolar contexts greater than that of labial, and the effect of aspirated stop contexts greater than that of the unaspirated ones.

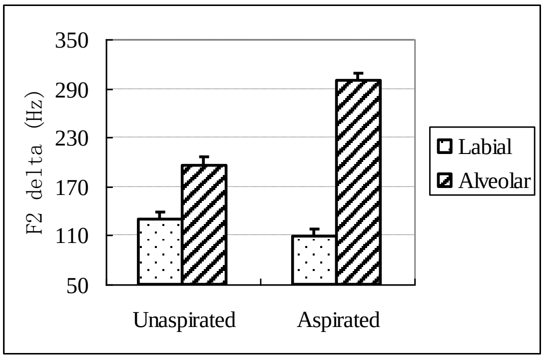


Fig. 4 F_2 delta under the effects of place of articulation and aspiration

Table 1 F_2 delta means (in Hz) and statistical results for the main effects

		Mean	Statistical result
Place of articulation	Labial	119.2	**F(1, 71) = 39.1, p < 0.001**
	Alveolar	248.1	
Aspiration	Unaspirated	162.6	**F(1, 71) = 5.62, p = 0.02**
	Aspirated	**204.7**	

When interactive effects are examined, it is shown that the place of articulation × aspiration interaction is significant: $F(1, 71) = 9.25$, $p = 0.003$, which is due to the asymmetrical effect of place of articulation and aspiration. Detailed effect will be elaborated in the following sections.

1) The effect of place of articulation

a) The unaspirated stop contexts: Result from repeated measures ANOVA shows that, in the unaspirated stop contexts, the effect of place of articulation is significant: $F(1, 71) = 5.3$, $p = 0.024$, with the extent of alveolar context exceeding that of labial context.

b) The aspirated stop contexts: When the intervocalic stops are aspirated, the coarticulatory extent in the alveolar context is also significantly greater than that in the labial context: $F(1, 71) = 44.1$, $p < 0.001$.

2) The effect of aspiration

a) The contexts of labial: It is shown from repeated measures ANOVA result that, in the context of labial, there is no significant difference between the unaspirated and aspirated stop contexts: $F(1, 71) = 0.78$, $p = 0.379$. The extents of coarticulation between the unaspirated and aspirated stop contexts are roughly the same.

b) The contexts of alveolar: In the context of alveolar stops, the difference between unaspirated and aspirated stop contexts is significant: $F(1, 71) = 11.2$, $p = 0.001$. The aspirated stop context exceeds the unaspirated one in the extent of coarticulation.

IV. DISCUSSION

Analysis in the previous section shows that, when V_1 F_2 offset value is analyzed, the effect of the changing vowel context on it is significant, with F_2 offset value preceding vowel /i/ higher than that preceding /u/. This study aims at investigating the anticipatory coarticulatory effect of V_2 on V_1, with the changing vowels of V_2 as /i/ and /u/. The F_2 value of /i/ is high, while that of /u/ is low. The V_1 F_2 offset values are significantly different when preceding /i/ and /u/. This implies that trans-consonantal anticipatory vowel to vowel effect exists in Chinese.

Results in the previous section shows that, in the context of subsequent vowel /i/, the V_1 F_2 offset value is higher in the context of aspirated stops than the unaspirated ones, which is due to the high articulatory strength of aspirated obstruents. Generally speaking, phonetic segments can be classified into two groups: ‘fortis’ and ‘lenis’, which refer to consonants that are produced with greater and lesser energy respectively, such as in energy applied, articulation, etc. Fortis and lenis were coined as less misleading terms to refer to consonantal contrasts in languages that do not employ actual vocal fold vibration in their voiced consonants, but instead involved amounts of articulatory strength. For example, in English there are fortis consonants, as in ‘come’ and ‘put’ that exhibit a longer stop closure and shorter preceding vowels than their lenis counterparts, as in ‘grass’ and ‘bed’. In Chinese, aspirated consonants are fortis, while unaspirated ones are lenis.

Since the aspirated consonants in Chinese are fortis, their articulatory strength is great. Generally speaking, consonants with high articulatory strength tend to exert great effect on the preceding vowels. In Chinese, consonants and vowels are combined into one unit: syllable. Syllables with consonants of high articulatory strength may exert great effect on the preceding vowel. As a result, the anticipatory coarticulatory effect is great when C_2 is aspirated consonant. The F_2 value of vowel /i/ is high, so in this case, the V_1 F_2 offset value will be comparative high. As in the context of aspirated consonants, the anticipatory coarticulatory effect is comparatively great, the V_1 F_2 offset value will be higher in the context of aspirated consonants than the in unaspirated context.

However, in the context of vowel /u/, it is not the case, i.e., there is no significant difference between the V_1 F_2 offset values in the unaspirated and aspirated contexts. In this experiment, the pre-consonant vowel is /u/, whose second formant is low. When post-consonant vowel is also /u/, vowels preceding and following the consonants are the same, and their effect on each other will diminish. Even if the aspiration condition of the intervocalic consonant changes, the effect of the post-consonant vowel on the pre-consonant one remains the same. Therefore, when the post-consonant vowel is /u/, there is no significant difference between the V_1 F_2 offset values in the unaspirated and aspirated contexts.

As for the effect of place of articulation, there is significant effect of place of articulation on the V_1 F_2 offset value, with V_1 F_2 offset value preceding alveolars higher than preceding labials, which indicates that the effect of place of articulation is significant. Generally speaking, the F_2 value of a vowel may be reduce in the context of a labial, while increased in the context of an alveolar [9]. Chen [7] investigated the effect of coarticulation between neighboring as well as trans-segmental vowels, and her results proved these features of labials and alveolars. Since the F_2 value may be reduced in the context of labials, the V_1 F_2 offset value will be low in this context. Therefore, the V_1 F_2 offset value is higher preceding alveolars than preceding labials.

In regard to the F_2 delta values, when main effects are examined, there is significant effect for place of articulation, with the effect in the alveolar contexts greater than that of labial. The DAC model [13] predicts that in VCV sequences, an increase in the degree of constraint for the consonant should yield an increase in the prominence of the C-to-V effects and a decrease in the strength of the V-to-C and V-to-V effect. According to the DAC model, the diversity in the involvement of the articulators in the production of different obstruents results in the variation of the degree of articulatory constraint.

With respect to coarticulation, coarticulatory sensitivity, which is the magnitude and temporal extent of the coarticulatory effect at a given articulator, is shown to be inversely related to the degree of articulatory constrait: highly constrained phonetic segments are generally more resistant to coarticulation than those specified for a lower degree of articulatory constraint, and thus less sensitive to coarticulatory influence from neighboring segments. The model also predicts that coarticulatory dominance is positively related to the degree of articulatory constraint: phonetic segments with high DAC value and coarticulation resistant usually have prominent coarticulatory effects on neighboring phonetic segments. Obstruents, particularly alveolopalatals, which maximally engage the tongue dorsum for the occlusion gesture would reduce V effects, that is, stops like /d/ and /t/ exhibit reduced extents of V-V coarticulation.

In this study, regarding the main effect of place of articulation, the result is not in accordance with the DAC prediction. This implies that the pattern of ‘degrees of articulatory constraint’ is not consistent among languages. For consonants of similar place of articulation, it is possible that their degrees of articulatory constraint are diverse among languages. The articulators for labials are the lips, and that for the alveolar is the tongue tip. The tongue tip is small and flexible, so in the production of alveolars, its contacting area with the gingiva might not be quite large. On the other hand, when producing the labials, the contacting area of the lips is larger than that of the tongue tip, so the degree of articulatory constraint is high for labial, and low for alveolar in Chinese.

With regard to the aspiration of C_2, it is shown from the previous section that, when the intervocalic consonant is alveolar, coarticlation effect is greater in the context of aspirated consonant than that of unaspirated one, which is due to the high articulatory strength of aspirated consonants. The articulatory strength of the aspirated consonant is high, and its effect on the preceding vowel is great. In Chinese, consonants and vowels are combined into syllable. Syllables with consonants of high articulatory strength will have high effect on the preceding vowel. Therefore, in the context of alveolar, coarticulatory effect is greater in the aspirated context that in the unaspirated context. However, in the context of labial, there is no significant difference of the coarticulatory extent between the aspirated and unaspirated conditions. This is due to the low extent of coarticulation in the labial context. When the inter-vocalic consonant is labial, the degree of articulation constraint is high, and the vowel to vowel coariticulatory effect is low. It is so low that the difference of aspiration condition has no effect on the extent of coarticulation. Therefore, in the context of labial, there is no effect of aspiration on the coarticulatory extent.

V. CONCLUSION

In this study, the V-to-V coarticulation in the VCV sequences is investigated, and it is found that there is significant difference between the V_1 F_2 offset values in the contexts of subsequent vowels /i/ and /u/, which means that trans-segmental anticipatory coarticulation exists in Chinese. Due to high articulatory strength of aspirated obstruents, in the context of subsequent vowel /i/, the V_1 F_2 offset value is higher in the context of aspirated stops than the unaspirated ones. As F_2 value of a vowel may be reduce in the context of a labial, while increased in the context of an alveolar, there is significant effect of place of articulation on the V_1 F_2 offset value, with V_1 F_2 offset value preceding alveolars higher than preceding labials. Results from this study are not in accordance with the DAC prediction, which implies that the pattern of degrees of articulatory constraint is not consistent among languages.

This study is significant in speech engineering. In speech synthesis, the effect of trans-consonantal coarticulation must be taken into consideration. The extent of coarticulation in the context of labial exceeds that of alveolar, and the extent of aspirated consonants exceeds that of unaspirated ones, so much attention should be paid in these contexts. However, in some case, the difference of coarticulatory effects can be neglected, as it disappears. Therefore, this study is helpful in speech engineering technology.

REFERENCES

[1] S. E. G. Öhman, “Coarticulation in VCV utterances: spectrographic measurements,” Journal of the Acoustical Society of America, vol. 39, pp. 151–168, 1966.

[2] D. Recasens, “An acoustic analysis of V-to-C and V-to-V coarticulatory effets in Catalan and Spanish VCV sequences,” Journal of Phonetics, vol. 15, pp. 299–321,1987.

[3] R. A. W. Bladon and A. Al-Bamerni, “Coarticulation resistance in English /l/,” Journal of Phonetics, vol. 4, pp. 137–150,1976.

[4] D. Recasens, “V-to-C coarticulation in Catalan VCV sequences,” Journal of Phonetics, vol. 12, pp.61–73, 1984.

[5] Z. Wu and G. Sun, “Acoustic coarticulatory patterns of voicelese fricatives in CVCV in Standard Chinese”, Report on Phonetic Research, Institute of Linguistics, CASS, 1990, pp. 1–19.

[6] J. Yan, “A study of the vowel formant pattern and the coarticulation in the voiceless stop initial monosyllable of Standard Chinese”, Report on Phonetic Research, Institute of Linguistics, CASS, 1990, pp. 30–54.

[7] X. Chen, “A study on the coarticulation of segments in Chinese, Chinese Linguistics”, vol. 5, pp. 345–350, 1997.

[8] G. Sun, “Vowel segmental coarticulation in read speech in Standard Chinese”, Report on Phonetic Research, Institute of Linguistics, Chinese Academy of Social Sciences, 1998.

[9] Y. Li and J. Kong, “Anticipatory coarticulation in V1#C2V2 sequences in standard Chinese,” *Journal of Tsinghua University*, Vol. 51, No. 9, pp 1220-1225, 2011.

[10] M. Wang, P. Chai and M. Xie, “Research on the Coarticulation of Two-syllable Word for Mandarin Speech Synthesis,” Computer Engineering, vol. 28, no. 5, pp. 182-184, 2002.

[11] G. Modarresi, H. Sussman, B. Lindblom and E. Burlingame, “An acoustic analysis of the bidirectionality of coarticulation in VCV utterances,” Journal of Phonetics, vol. 32, pp. 291–312, 2004.

[12] P. Boersma, “Praat, a system for doing phonetics by computer”, Glot International, 5:9/10, pp. 341–345, 2001.

[13] D. Recasens, “An EMA study of VCV coarticulatory direction,” Journal of the Acoustical Society of America, vol. 111, pp. 2828–2841, 2002.