The International Journal Of Engineering And Science (Ijes)
||Volume||2 ||Issue|| 11||Pages|| 17-22||2013||
Issn(E): 2319 – 1813 Issn(P): 2319 – 1805

Effect of Vowel-To-Vowel Coarticulation of Disyllabic Sequences in Chinese

Maolin Wang, Wei Xiong
College of Chinese Language and Culture, 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. The stimuli are in the form of $C_1V_1.C_2V_2$, which is designed for V_1 to affect V_2. The subjects are twelve native speakers of standard Chinese, and the F_2 onset value, as well as the F_2 delta value, is analyzed. Results show that, due to long temporal separation between vowels for the aspirated stops relative to the unaspirated ones, the effect of aspiration is obvious, and in the contexts of aspirated stops, the effect of place of articulation is restrained.

KEYWORDS : *Coarticulation, formant, speech, stop, vowel*

Date of Submission:. 6 November, 2013 Date of Acceptance: 20 November 2013

I. INTRODUCTION

The realization of a segment may be influenced by the neighboring segments, which is called 'coarticulation'. 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. Phonetic segments possess inherent properties, referred to as 'coarticulatory resistance', that limit the extent to which they can be influenced by neighboring segments [3]. 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. There have been a number of studies on the coarticulatory effect of segments in Chinese, including the analysis of the acoustic coarticulatory patterns of voicelese 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 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 carry-over coarticulation.

II. METHODOLOGY

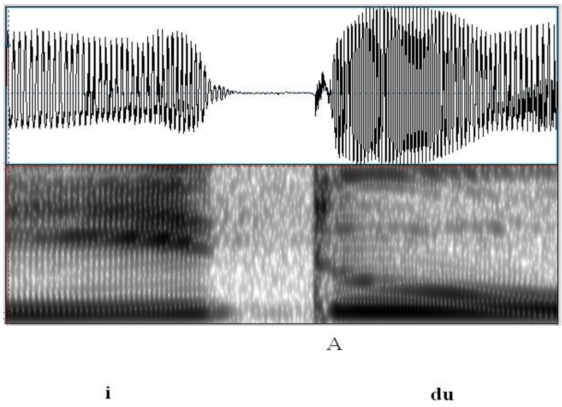
2.1 Speakers, Stimuli and Recording

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_1 providing the 'changing' vowel context, V_2 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/ is used to influence the onset 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 is 'jidu' and 'sudu', which means 'extremely' and 'speed' 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.

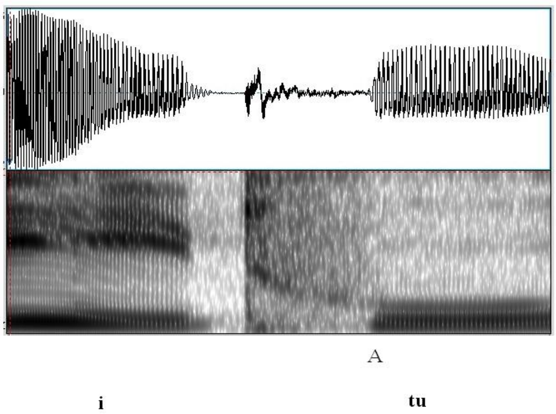
2.2 Procedure and Measurements

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).

1) F_2 onset 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 onset value of the fixed vowel. F_2 onset frequency is taken at the onset of the fixed vowel V_2, both for unaspirated and aspirated stop contexts. Fig. 1 displays the waveform and spectrogram of 'idu' (Fig. 1a) and 'itu' (Fig. 1b), as in the sequences of 'jidu' (extremely) and 'ditu' (map), with the former an example of unaspirated stop context, and the latter of aspirated context. For both of the two cases, F_2 onset values are taken at the onset points of the vowels, that is, point 'A' on both graphs.



(a) 'idu'



(b) 'itu'

Fig. 1. Waveform and spectrogram of (a) 'idu' and (b) 'itu'

2) F_2 delta: In order to compare the extent of coarticulatory effects under various consonant contexts, besides the F_2 onset values, their differences caused by the changing V_1 contexts are also calculated. Coarticulation effects due to changing V_1 contexts are indexed by F_2 delta values obtained at the V_2 F_2 onset, and F_2 delta (Hz) is derived by calculating the difference in onset frequencies of the fixed vowel in each sequence pair. Fig. 2 displays the F_2 contours of the sequence pair 'jidu' (extremely) and 'sudu' (speed), with the contour of 'jidu' in solid line, and that of 'sudu' 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_2 F_2 onsets differ in this pair, then it is reasonable to attribute the frequency difference to the high vs. low F_2 contexts in V_1. The greater the F_2 delta value is, the greater the coarticulatory effect of V_1 is on V_2.

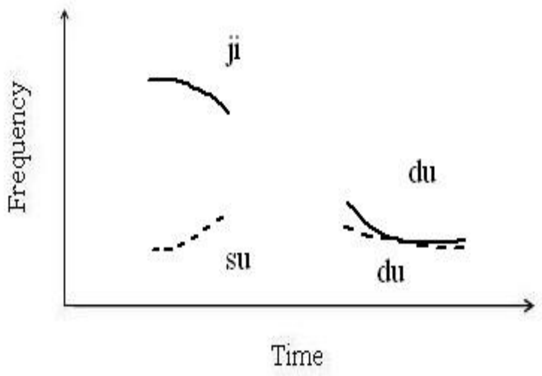


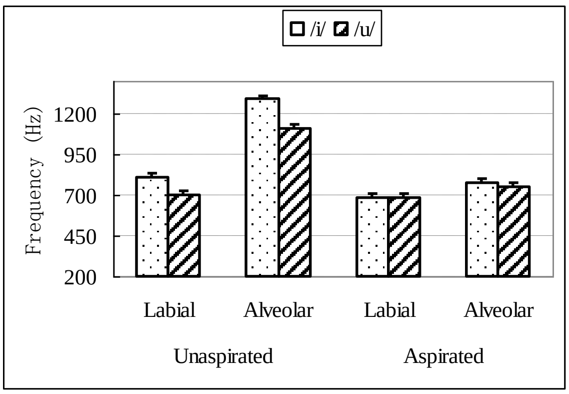
Fig. 2. F_2 contours of the sequence pair 'jidu' and 'sudu', with the former in solid line, and the latter in dashed line

A repeated measures ANOVA was performed with two within-subjects factors— aspiration (unaspirated, aspirated) and place of articulation (labial, alveolar). F_2 delta was the dependent variable in all ANOVA analyses.

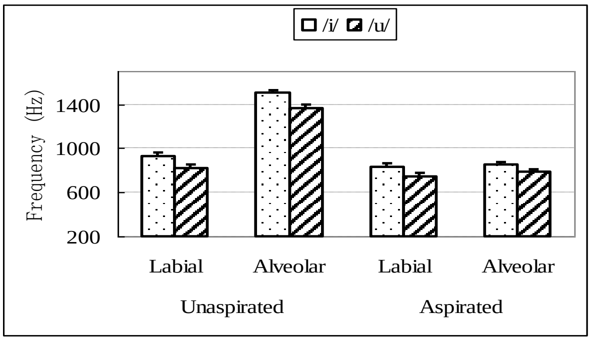
III. RESULTS

3.1 F_2 VALUE

Fig. 3 graphs the F_2 onset 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, the effects of aspiration, place of articulation, and changing vowel are all significant, aspiration: $F(1, 71) = 446.9$, $p < 0.001$; place of articulation $F(1, 71) = 661.8$, $p < 0.001$; changing vowel $F(1, 71) = 127.4$, $p < 0.001$, with F_2 onset values in the unaspirated context greater than the aspirated context, in the alveolar context greater than the labial context, in the /i/ context greater than in the /u/ context.



(a) F2 values of male speakers



(b) F2 values of female speakers

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

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

3.2 F_2 Delta Value

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 unaspirated stop contexts much greater than that of the aspirated ones. When interactive effect is examined, it is shown that the place of articulation × aspiration interaction is significant: $F(1, 71) = 6.02$, $p = 0.017$, which is attributable to the disproportional effect of place of articulation and aspiration. The effect of place of articulation is not consistant under different aspirational condition. In the next subsection the effects of the factors will be described in detail to help inform and elaborate on the main effects.

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

		Mean	Statistical result
Place of articulation	Labial	71.7	$F(1, 71) = 7.47$, $p = 0.008$
	Alveolar	111.9	
Aspiration	Unaspirated	138.7	$F(1, 71) = 43.0$, $p < 0.001$
	Aspirated	44.9	

3.2.1 The effect of place of articulation

a) The unaspirated stop contexts: Fig. 4 shows the F_2 delta under the effects of place of articulation and aspiration. Result from repeated measures ANOVA shows that, in the unaspirated stop contexts, the effect of place of articulation is significant: $F(1, 71) = 9.03$, $p = 0.004$, with the extent of alveolar context exceeding that of labial context.

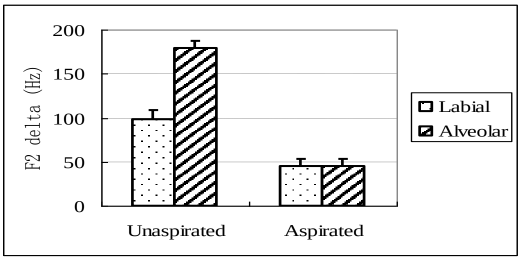


Fig. 4. F_2 delta under the effects of place of articulation and aspiration

b) The aspirated stop contexts: When the intervocalic stops are aspirated, there is no significant difference: F(1, 71) = 0.000, p = 0.985. No significant difference exists in the coarticulatory effect of the contexts of place of articulation in the case of aspirated stop context.

3.2.2 The effect of aspiration

a) The contexts of labial: It is shown from repeated measures ANOVA result that the difference between unaspirated and aspirated stop contexts is significant: F(1, 71) = 5.49, p = 0.022, with greater extent of coarticulation for unaspirated stop context than the aspirated one.

b) The contexts of alveolar: In the context of alveolar stops, the difference between unaspirated and aspirated stop contexts is also significant: F(1, 71) = 43.8, p < 0.001. The unaspirated stop context exceeds the aspirated one in the extent of coarticulation.

IV. DISCUSSION

Analysis in the previous section shows that, when F_2 onset value is investigated, the effect of the changing vowel context is significant, with F_2 onset value following vowel /i/ higher than that following /u/. This implies that trans-consonantal carry-over vowel to vowel effect exists in Chinese. In regard to the F_2 delta values, when main effects are examined, 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 unaspirated stop contexts greater than that of the aspirated ones. The DAC model [13] predicts that obstruents, particularly alveolopalatals, that maximally engage the tongue dorsum for the occlusion gesture would reduce V effects, that is, stops like /d/ and /t/ exhibite 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 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.

As for the effect of aspiration, analysis in the previous section shows a general tendency for coarticulation effect to be greater for unaspirated stops than the aspirated ones, regardless of the context of place of articulation. From this a conclusion can be drawn that aspiration can be taken as one of the important factors in blocking coarticulatory effect, and an explanation for the great effect of aspiration might be due to the long temporal separation of the aspirated stops relative to their unaspirated counterparts. Table 2 displays the mean temporal intervals between V_1 and V_2 for different stop contexts. From Table 2 it is shown that the intervals for aspirated stops are much longer than the unaspirated ones. Result from a one-way ANOVA shows that there is significant difference between them (F(1, 575) = 581.9, p<0.001). Therefore, it is shown that long temporal separation between V_1 and V_2 reduces the coarticulation effect. Hence, it is reasonable to suggest that the long time interval between measurement points for aspirated stops account for the low coarticulatory effects.

Table 2 Mean intervals (in ms) between V_1 and V_2 for various stop contexts

		Place of articulation		Mean
		Labial	Alveolar	
Aspiration	Unaspirated	118.6	116.3	117.5
	Aspirated	186.4	187.9	187.2

Detailed examination on the coarticulatory effect under various consonant conditions shows that, the effect of place of articulation exists in the unaspirated stop contexts, with the extent of alveolar exceeding that of labial. As is mentioned above, this is due to the greater degree of involvement of lips in articulating labial than that of alveolar. However, when the intervocalic stops are aspirated, the effect of place of articulation disappears. We speculate that this is also because of the long temporal separation between V_1 and V_2 in the aspirated stop context. In this context, the time intervals between the vowels are longer than those of the unaspirated ones, as is shown in Table 2. The long intervals reduce and interfere with the coarticulation effect, as a result, the effect of place of articulation is restrained in these contexts.

V. CONCLUSION

In this study, the V-to-V coarticulation in the VCV sequences is investigated, and it is found that, due to long temporal separation between vowels for the aspirated stops relative to the unaspirated ones, the effect of aspiration is obvious, with the extent in the unaspirated stop contexts exceeding the aspirated ones. Long temporal separation between V_1 and V_2 reduces and interferes with the coarticulation effect, and aspiration is a prominent factor in blocking coarticulation effect. Regarding the effect of place of articulation, the result is not in accordance with the DAC prediction, which implies that the pattern of 'degrees of articulatory constraint' is not consistent among languages. In the contexts of aspirated stops, due to the long intervals between the vowels, the effect of place of articulation is restrained.

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

ACKNOWLEDGEMENTS

This work was supported in part by the Social Science Research Project of Guangdong Province, Grant No. GD11CWW04, as well as the Humanity and Social Science Research Project for Colleges in Guangdong Province, Grand No. 11WYXM012.

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.