

# The Effect of Carry-over Coarticulation of Bi-Syllabic Words in Chinese

Maolin Wang and Guangling Zi, *Member, IACSIT*

**Abstract**—In this study, trans-consonantal vowel-to-vowel coarticulation in Chinese is investigated. The target words are in the form of 'bV<sub>1</sub>ba', and the subjects are eight native speakers of standard Chinese. Vowel formants are examined at the onset, middle and offset points of the target vowel. Results show that trans-segmental coarticulation exists in Chinese, especially at the onset point, as well as for the second formant of the target vowel. Coarticulation is more likely to occur on F<sub>2</sub>, and in Chinese, coarticulatory effect does not extend to the offset point of the vowel.

**Index Terms**—Coarticulation, consonant, formant, speech, vowel.

## I. INTRODUCTION

In connected speech, segments are not isolated, and they will be affected by the neighboring segments. Coarticulation refers to the fact that speech sounds are not produced as isolated gestures, but are superimposed in a complex, context-dependent fashion. While it is clear that all natural speech is coarticulated, the magnitude and temporal extent of coarticulation in different contexts has been difficult to explain. The need to adequately describe the varieties of coarticulation has been a key factor driving the development of competing models of speech production and also perception [1]. Researches on speech production reveal systematic differences between languages in the spatiotemporal characteristics of coarticulation [2].

Öhman [3] compared the coarticulatory effects in three languages, and found that the F<sub>2</sub> values of target vowels varied more in English and Swedish than in Russian, due to vowel context. He attributed the coarticulatory differences to the languages' consonant systems, arguing that the requirements on the tongue body imposed by contrastive palatalization in Russian restricted trans-consonantal coarticulation. Following Öhman's [3] ground breaking work on transconsonantal vowel-to-vowel coarticulation, researchers have sought to understand how different factors influence this effect, and it is shown that a language's system of vowel contrasts may influence V-to-V effect.

Beddor *et al.* [4] conducted three experiments to test the hypothesis that V-to-V coarticulatory organization differs in Shona and English. An acoustic study of Shona and English

trisyllables shows that the two languages differ in the coarticulatory effect of stressed and unstressed vowels on each other, and the relation between the production and perception data suggests that listeners are attuned to native-language coarticulatory patterns. It is shown from research results that languages with larger vowel systems tend to exhibit weaker V-to-V coarticulatory effects than those with smaller systems. Weaker effects have been shown for English compared to the much smaller five-vowel systems of Shona, Swahili [5].

Cho [6] examined how the degree of vowel-to-vowel coarticulation varies as a function of prosodic factors, and results show that vowels in prosodically stronger locations are coarticulated less with neighboring vowels, but do not exert a stronger influence on the articulation of neighboring vowels. An examination of the relationship between coarticulation and duration reveals that accent-induced coarticulatory variation cannot be attributed to a duration factor, and some of the data with respect to boundary effects may be accounted for by the duration factor. He proposed that prosodically conditioned V-to-V coarticulatory reduction is another type of strengthening that occurs in prosodically strong locations. The prosodically driven coarticulatory pattern can be taken as part of the phonetic signaling of the hierarchically nested structure of prosody.

It is shown from studies on various languages that vowel-to-vowel coarticulatory effects not only in transitions, but extending into the steady-state period of the transconsonantal vowel both in palatographic data [7], [8] and in acoustic data [9], [10]. While there is ample evidence of the existence of vowel-to-vowel coarticulatory effects, factors have been cited which affect the extent of these effects. For instance, these effects may be constrained by intervocalic palatals and velars, whose production requires use of the tongue body in conflict with the production of vowels, thereby restricting vowel-to-vowel coarticulation [7], [11].

Early researches indicate that V-to-V coarticulation might be a relatively local phenomenon [12], but subsequent work has shown that this is not always the case. Instances of long-distance coarticulation, involving effects crossing two or more intervening segments have been found [13], [14]. Magen [15] analyzed [bVb**ə**bVb] sequences produced by four English speakers and found evidence of coarticulatory effects between the first and final vowel, meaning that such effects can cross foot boundaries and multiple syllable boundaries. More recently, Grosvald [16] studied long-distance vowel-to-vowel coarticulation in English, and found that anticipatory V-to-V coarticulation can occur over at least three vowels' distance in natural discourse, and that even such long-distance effects can be perceived by some

Manuscript received September 9, 2013; revised November 9, 2013. 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.

Maolin Wang and Guangling Zi are with the College of Chinese Language and Culture, Jinan University, Guangzhou, China (e-mail: wmljd@126.com).

listeners.

There has been some research work on the coarticulation of segments in Chinese. Wu and Sun [17] analyzed the acoustic coarticulatory patterns of voiceless fricatives. The fricatives are designed in CVCV contexts in which three peripheral vowels are combined with the fricatives in  $C_1$  and  $C_2$ . Acoustic data, including frequencies and durations of the lower margins, concentration bars, vowel formants and the onset and offset transitions, are measured. It is found that there are three types of coarticulation effects: homorganic, heterorganic and contiguous.

Yan [18] studied the vowel formant pattern and the coarticulation effect in the voiceless stop onset monosyllables, and it is found that there is no effect of tones on formant values of the vowel. However, there is significant effect of aspiration on the formant values at the onset point of the vowel. Chen [19] investigated the intersyllabic anticipatory coarticulation in CVCV sequences, and found that there is no effect of articulation manner on the formant values. There is coarticulatory effect of  $C_2$  on  $V_1$ , as well as  $V_2$  on  $V_1$ . Sun [20] analyzed the coarticulation effect of vowels in read speech, and found that, as the speaking rate increases, the deviation of the vowel formant is also magnified, as is shown in the apparent centralization of the vowel. In fast speech, the formant values of /i/ and /u/ are significantly different, due to the frontness and the backness of the following consonants. Besides, there is also the study on the anticipatory coarticulation in  $V_1\#C_2V_2$  sequences [21], and coarticulatory effect in VCV sequences [22]. However, as far as we know, there have no studies on coarticulatory effect at different vowel points in Chinese.

The present study will investigate the effect of trans-segmental coarticulation in Chinese. In particular, it will try to answer the following questions. Does trans-segmental coarticulation occur in Chinese? What is the extent of coarticulation in Chinese?

## II. EXPERIMENT 1

### A. Methodology

#### 1) Speakers and stimuli

The speakers for this experiment were eight native speakers of Standard Chinese, four males and four females. The stimulus list is comprised of 6 stimuli, embedded in three carrier sentences of Chinese, each containing an item from a set of two target words. The target words are 'Biba' and 'Baba', which are supposed to be two persons' names. They are in the form of 'bV<sub>1</sub>ba', with /a/ as the target vowel, and V<sub>1</sub> providing the 'changing' vowel contexts, namely, the contexts of /a/ and /i/.

Labial /b/ is chosen to minimize the effects of consonant articulation on lingual articulations [3]. Within the carrier sentences, the target item is located at the sentence initial position. One example is shown as follow,

Biba de jiejie lai le.

Biba's sister has come.

#### 2) Procedure and measurements

The speakers were asked to read the sentences three times,

in random order for each time, in normal speed, so each speaker produced 18 tokens: 6 sentences  $\times$  3 repetitions. In total, 144 tokens were acoustically analyzed (18 tokens  $\times$  8 speakers).

This study aims at investigating the extent of V-to-V coarticulation in VCV sequences, and vowel formants were examined. Formant values were extracted using Praat [23], and the extent of trans-consonantal coarticulation was analyzed by comparing the formant values of the target vowel at three points: the onset point, the middle point and the offset point. That is, formant values at the onset, middle and offset points of the target vowel /a/ were extracted, and the values at different vowel contexts were compared. As is mentioned in the previous subsection, there are two contexts for the target vowel, /i/ and /a/. Coarticulatory effect exists if there is significant difference between the formant values in the two vowel contexts. On the contrary, there is no coarticulatory effect if there is no significant difference. A repeated measures ANOVA was performed for the comparison, and statistic analysis was done in SPSS.

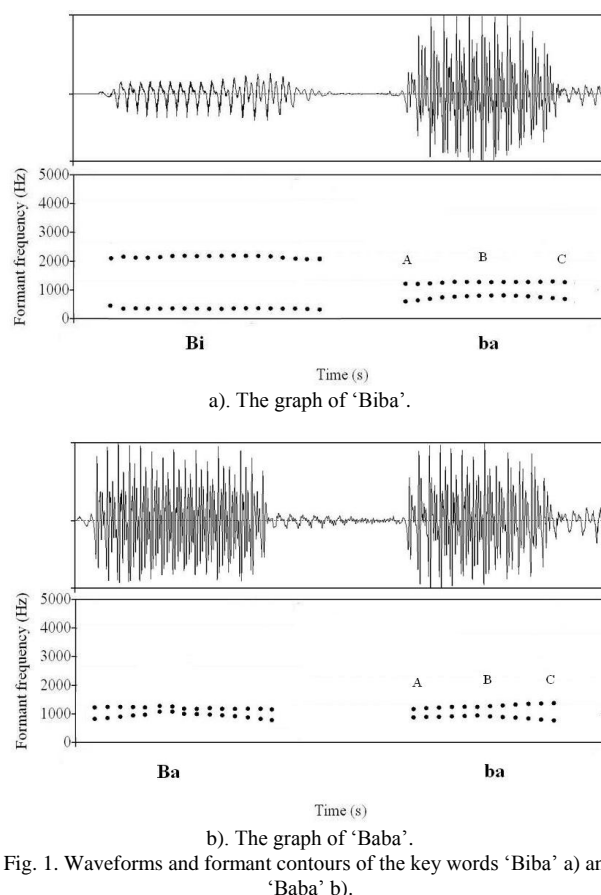
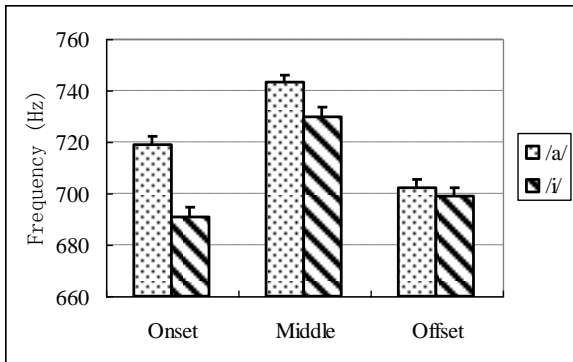


Fig. 1. Waveforms and formant contours of the key words 'Biba' a) and 'Baba' b).

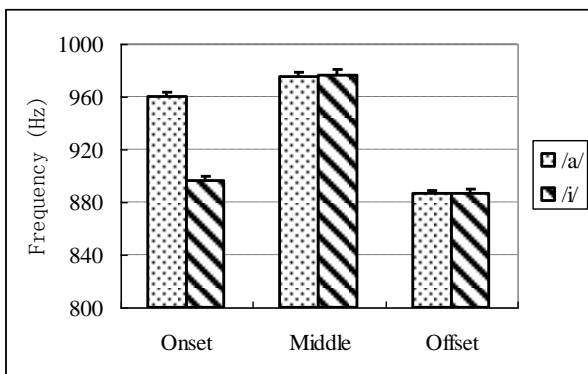
Fig. 1 displays the waveforms and formant contours of the two key words 'Biba' (a) and 'Baba' (b). In the graphs, the second syllable 'ba' is the key syllables, and the preceding syllables 'bi' and 'ba' provide the changing vowel contexts. Formant values are extracted from the onset, middle and offset points of the second syllable 'ba', which correspond to point A, B and C in the graphs. Comparison is made for formant values of the target values in the two contexts at each of the three points. Taken the onset point, point A, as an example, since the contexts shown in graph (a) and graph (b) are different, if the formant values at point A in the two contexts are significantly different, it can be concluded that

there is an effect of coarticulation at that point. In this study, comparison will be made for both  $F_1$  and  $F_2$  values at each of the three points, the onset, middle and the offset point.

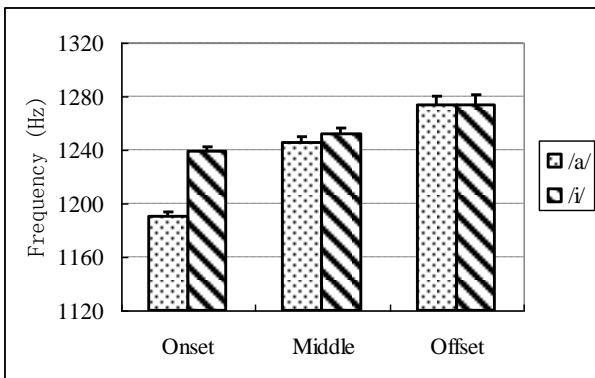
**B. Result**



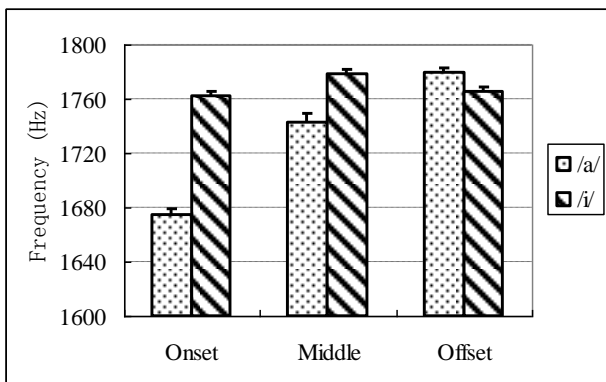
a). F1 for male speakers



b). F1 for female speakers



c). F2 for male speakers.



d). F2 for female speakers.

Fig. 2.  $F_1$  (a and b) and  $F_2$  (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points.

Fig. 2 presents the mean  $F_1$  (a and b) and  $F_2$  (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points. One of the aims of this study is to investigate the extent of the coarticulation, so formant is measured at three points, the onset, middle and the offset points of the target vowel. Analysis is given in the following subsections.

*1) Onset point*

Results from a repeated measures ANOVA show that, at the onset point of the target vowel, with data of both male and female speakers pooled together, the effect of changing vowel is significant for both  $F_1$  and  $F_2$ ,  $F_1$ :  $F(1, 71) = 40.09, p < 0.001$ ;  $F_2$ :  $F(1, 71) = 78.31, p < 0.001$ . That is, coarticulatory effect exists at the onset point of the target vowel.

*2) Middle point*

Analysis shows that, at the middle point of the target vowel, the effect of changing vowel is significant for  $F_2$ , but not for  $F_1$ ,  $F_1$ :  $F(1, 71) = 0.63, p = 0.431$ ;  $F_2$ :  $F(1, 71) = 4.65, p = 0.034$ . Coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

*3) Offset point*

Coming to the offset point of the target vowel, it is shown that the effect of changing vowel is not significant for either  $F_1$  or  $F_2$ ,  $F_1$ :  $F(1, 71) = 0.01, p = 0.912$ ;  $F_2$ :  $F(1, 71) = 0.58, p = 0.45$ . Coarticulatory effect does not exist at that point.

III. EXPERIMENT 2

*A. Methodology*

In experiment 2, the speakers and the stimuli are the same as in experiment 1, and the procedure and measurements are also the same as that of experiment 1. However, the target item is located at the sentence medial position. One example is shown as follow,

Zhe shi Biba de jiejie.

This is Biba's sister.

*B. Result*

Fig. 3 presents the mean  $F_1$  (a and b) and  $F_2$  (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points. Analysis is given in the following subsections.

*1) Onset point*

Results from a repeated measures ANOVA show that, when key words occur at sentence medial position, at the onset point of the target vowel, with data of both male and female speakers pooled together, the effect of changing vowel is significant for  $F_2$ , but not for  $F_1$ .  $F_1$ :  $F(1, 71) = 0.04, p = 0.847$ ;  $F_2$ :  $F(1, 71) = 47.72, p < 0.001$ . That is, coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

*2) Middle point*

At the middle point of the target vowel, it is shown that, similar to that at the onset point, the effect of changing vowel is significant for  $F_2$ , but not for  $F_1$ .  $F_1$ :  $F(1, 71) = 0.39, p = 0.537$ ;  $F_2$ :  $F(1, 71) = 7.28, p = 0.009$ . Coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

3) Offset point

Coming to the offset point of the target vowel, it is shown that the effect of changing vowel is not significant for either  $F_1$  or  $F_2$ ,  $F_1$ :  $F(1, 71) = 3.82, p = 0.55$ ;  $F_2$ :  $F(1, 71) = 0.25, p = 0.621$ . Coarticulatory effect does not exist at that point.

IV. GENERAL DISCUSSION

In reviewing the results reported in the previous section we note that, for Experiment 1, trans-segmental coarticulation does exist in Chinese, especially at the onset point of the target vowel. When formant values are examined at the onset point of the target vowel, it is found that coarticulatory effect exists for both  $F_1$  and  $F_2$ .

In this experiment, coarticulatory effect is examined at the onset, middle and offset points of the target vowel respectively. It is found that, as far as carry-over coarticulation is concerned, the coarticulatory effect is strong at the onset point, reduced at the middle point, and disappeared at the offset point. To be specific, at the onset point of the target vowel, coarticulatory effect exists for both  $F_1$  and  $F_2$ . However, at the middle point of the target vowel, coarticulatory effect exists for  $F_2$ , but not for  $F_1$ .

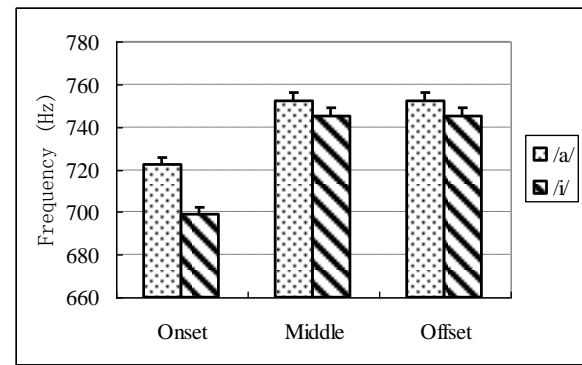
The effect on the first and the second formants are not consistent with each other at the middle point of the target vowel, with  $F_2$  affected, while  $F_1$  unaffected. We speculate that the reason for this is that the difference between /a/ and /i/ for  $F_2$  is larger than that for  $F_1$ . According to the report of Bao [24], the mean formant values by 8 male speakers are as follow, /a/:  $F_1 = 984$  Hz,  $F_2 = 1157$  Hz; /i/:  $F_1 = 283$  Hz,  $F_2 = 2350$  Hz. The differences between /a/ and /i/ for  $F_1$  and  $F_2$  are 701 Hz and 1193 Hz respectively. The difference of  $F_2$  is much larger than that of  $F_1$ . If the formant difference is large, the force for changing the course of formant contour is also large. Vowel coarticulatory effect is more likely to occur on cases with great formant difference, therefore, it occurs on  $F_2$ , not on  $F_1$ .

When the offset point of the target vowel is investigated, it is found that coarticulatory effect does not exist. That is, for either  $F_1$  or  $F_2$ , there is no coarticulatory effect. This result is caused by the ‘distance effect’: at the onset point, the distance from the measured point to the changing vowels is close, and the effect is great; at the middle point, the distance gets farther, and the effect is reduced; while at the offset point, the distance gets even farther, and the effect disappears.

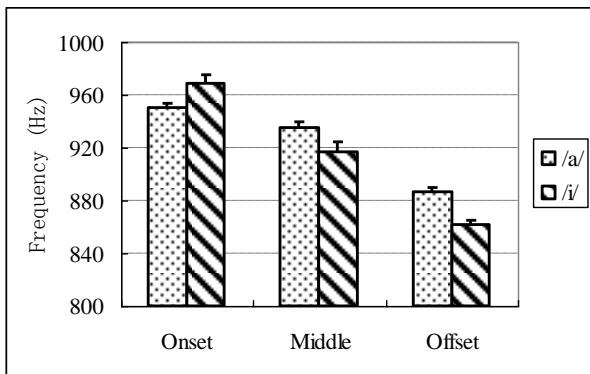
As for Experiment 2, it is noted that, trans-segmental coarticulation also exists, especially for the second formant value of the vowel. It is found that, coarticulatory effect exists for the second formant values at the onset and the middle points of the vowel. To be specific, at the onset and the middle points of the vowel, there are significant differences between the two changing vowel contexts for the second formant, but not the first formant.

Magen [15] investigated the extent of vowel-to-vowel coarticulation in English trisyllabic utterances, and it was found that coarticulatory effects can, in some instances, extend beyond the bounds that previous research had assumed; coarticulatory effect can extend from one full vowel, through the medial schwa, and into the midpoint of the next full vowel. He proposed that foot does not define the domain over which coarticulatory effects operate. However, in the present study, it is found that coarticulatory effect does not extend to the end of the vowel. We speculate this is because that Chinese is of different language typology from English.

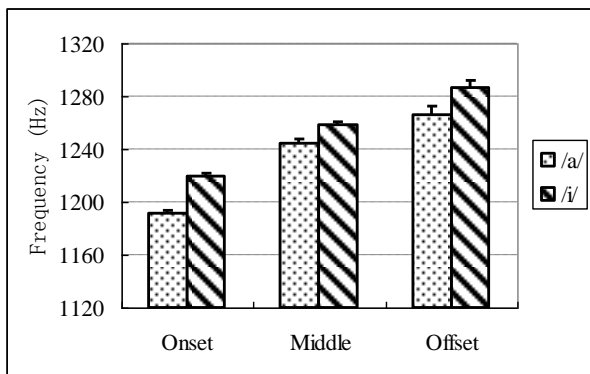
One of the well-known properties of language typology is the rhythm unit of a language. Languages have been



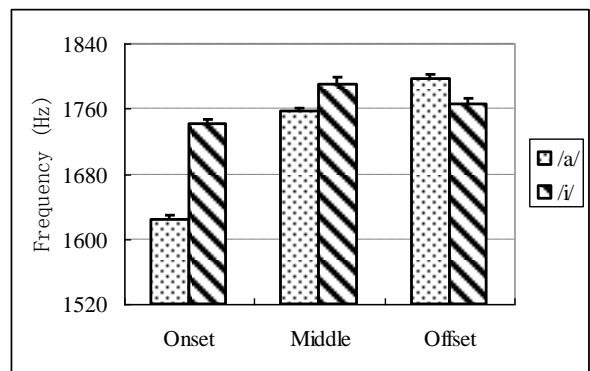
a). F1 for male speakers.



b). F1 for female speakers.



c). F2 for male speakers.



d). F2 for female speakers.

Fig. 3. F1 (a and b) and F2 (c and d) values of the target vowel /a/ for male (a and c) and female (b and d) speakers, in the contexts of /a/ and /i/, measured at the onset, middle and offset points.

categorized as mora-timed, such as Japanese, stress-timed, like English and German, and syllable-timed, such as Chinese and French [25], [26]. English is a stress-timed language, and the unstressed syllables are quite weak, so it is possible for coarticulatory effect to extend to the third syllable in English. Chinese is a syllable-timed language, in which syllables are rarely as weak as the unstressed ones in English. Compared to English, the degree of articulatory constraint (DAC) in Chinese is high. Therefore, the coarticulatory effect in Chinese is not as great as that in English.

## V. CONCLUSION

In this study, the vowel-to-vowel coarticulation effect in bi-syllabic words in Chinese is analyzed, and it is found that trans-segmental coarticulation exists in Chinese. When the key word occurs at the initial position of an utterance, the effect is great at the onset point of the target vowel, while when the key word occurs at the middle position, the effect is great for the second formant value of the vowel. Because of the 'distance effect', coarticulatory effect is great at the onset point, reduced at the middle point, and disappeared at the offset point of the target vowel. Coarticulation is more likely to occur on F<sub>2</sub> because the difference between /a/ and /i/ of F<sub>2</sub> is larger than that of F<sub>1</sub>. In Chinese, coarticulatory effect is not as great as that in English, as the two languages are of different prosodic typology.

This study is significant for speech engineering. In speech synthesis, the effect of trans-segmental coarticulation must be taken into consideration, especially at the onset point, as well as for the second formant of the target vowel. It is not necessary to consider the trans-syllabic coarticulatory effect at the end of the vowel, as in Chinese, coarticulatory effect does not extend to the end of the syllable. Therefore, this study is helpful for speech engineering technology.

## ACKNOWLEDGMENT

The research reported here was carried out in the Laboratory of Phonetics and Speech Signal Processing of the Department of Applied Linguistics, Jinan University.

## REFERENCES

- [1] W. F. Katz, "Anticipatory coarticulation and aphasia implications for phonetic theories," *Journal of Phonetics*, vol. 28, pp. 313–334, 2000.
- [2] S. Y. Manuel, "Cross-language studies: relating language-particular coarticulation patterns to other language-particular facts," in W. J. Hardcastle and N. Hewlett (eds.), *Coarticulation: theory, data and techniques*, Cambridge, U.K.: Cambridge University Press, pp. 179–198, 1999.
- [3] S. E. G. Öhman, "Coarticulation in VCV utterances: spectrographic measurements," *Journal of the Acoustical Society of America*, vol. 39, pp. 151–168, 1966.
- [4] P. S. Beddor, J. D. Harnsberger, and S. Lindemann, "Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates," *Journal of Phonetics*, vol. 30, pp. 591–627, 2002.
- [5] S. Y. Manuel and R. A. Krakow, "Universal and language particular aspects of vowel-to-vowel coarticulation," *Haskins Laboratories Status Reports on Speech Research*, SR-77/78, pp. 69–78, 1984.
- [6] T. Cho, "Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English," *Journal of Phonetics*, vol. 32, pp. 141–176, 2004.
- [7] D. Recasens, "Vowel-to-vowel coarticulation in Catalan VCV sequences," *Journal of the Acoustical Society of America*, vol. 76, pp. 1624–1635, 1984.
- [8] D. Recasens, "An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences," *Journal of Phonetics*, vol. 15, pp. 299–312, 1987.
- [9] S. Y. Manuel, "The role of contrast in limiting vowel-to-vowel coarticulation in different languages," *Journal of the Acoustical Society of America*, vol. 88, pp. 1286–1298, 1990.
- [10] D. H. Whalen, "Coarticulation is largely planned," *Journal of Phonetics*, vol. 18, pp. 3–35, 1990.
- [11] E. T. Purcell, "Formant frequency patterns in Russian VCV utterances," *Journal of the Acoustical Society of America*, vol. 66, pp. 1691–1702, 1979.
- [12] T. Gay, "Articulatory movements in VCV sequences," *Journal of the Acoustical Society of America*, vol. 62, pp. 183–193, 1977.
- [13] P. West, "The extent of coarticulation of English liquids: An acoustic and articulatory study," in *Proc. the international conference of phonetic sciences*, San Francisco, 1999, pp. 1901–1904.
- [14] S. Heid and S. Hawkins, "An acoustical study of long-domain /t/ and /l/ coarticulation," in *Proc. the fifth seminar on speech production: Models and data*, Bavaria, Germany: Kloster Seeon, 2000, pp. 77–80.
- [15] H. S. Magen, "The extent of vowel-to-vowel coarticulation in English," *Journal of Phonetics*, vol. 25, pp. 187–205, 1997.
- [16] M. Grosvald, "Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation," *Journal of Phonetics*, vol. 37, pp. 173–188, 2009.
- [17] Z. Wu and G. Sun, "Acoustic coarticulatory patterns of voiceless fricatives in CVCV in Standard Chinese," *Report on Phonetic Research*, Institute of Linguistics, CASS, pp. 1–19, 1990.
- [18] 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, pp. 30–54, 1990.
- [19] X. Chen, "A study on the coarticulation of segments in Chinese," *Chinese Linguistics*, *Zhongguo Yuwen*, vol. 5, pp. 345–350, 1997.
- [20] G. Sun, "Vowel segmental coarticulation in read speech in Standard Chinese," *Report on Phonetic Research*, Institute of Linguistics, Chinese Academy of Social Sciences, 1998.
- [21] 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.
- [22] M. Wang, W. Yan, and Z. Xiong, "Coarticulatory effect in VCV sequences in Chinese," *Journal of Tsinghua University*, vol. 51, no. 9, pp. 1244–1248, 2011.
- [23] P. Boersma, "Praat, a system for doing phonetics by computer," *Glott International*, vol. 5, no. 9/10, pp. 341–345, 2001.
- [24] H. Bao, "The physiological explanation of monophthongs in Chinese," *Zhongguo Yuwen*, vol. 2, pp. 117–127, 1984.
- [25] K. L. Pike, *The intonation of American English*, Ann Arbor: University of Michigan Press, 1945.
- [26] I. Lehiste, *Suprasegmentals*, Cambridge, MA: MIT Press, 1970.



**Maolin Wang** was born in Hebei, China, in 1965. He received PhD from the Chinese Academy of Social Sciences in 2003. He is an associate professor in Jinan University, Guangzhou, China. His research interest is phonetics and speech signal processing.