PROSODIC BOUNDARY EFFECTS ON SEGMENT ARTICULATION AND V-TO-V COARTICULATION IN STANDARD CHINESE

Yinghao Li[a] *& Jiangping Kong*[a,b]

[a]Department of Chinese Language and Literature; [b]Center for Chinese Linguistics, Peking University, Beijing, China

leeyoungho@pku.edu.cn; kongjp@gmail.com

ABSTRACT

This paper investigates the prosodic conditioning of the segment production and vowel-to-vowel coarticulation in the Standard Chinese through electropalatographic and acoustic analysis. The articulatory and acoustic measures were obtained for un-aspirated alveolar consonant /t/ and vowels /i/ and /a/. Results show that the domain-initial consonant is strengthened in a hierarchical manner in higher prosodic domains. The vocalic gesture after the strengthened consonant tends to be reduced, and the domain-final vocalic gesture shows more linguapalatal contact and longer duration. The vocalic anticipatory effect is shown to appear up to intermediate phrase boundary, but is constrained by the articulatory constraint for vocalic gesture. The vocalic carryover effect is likely constrained by foot domain.

Keywords: electropalatography, prosodic structure, Standard Chinese, vowel-to-vowel coarticulation

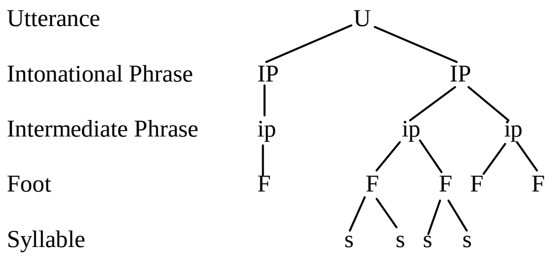
1. INTRODUCTION

Previous articulatory and acoustic studies have shown that segment articulation is modulated by prosodic structure in a hierarchical fashion in languages [1, 2, 3, 4, 8]. Segment articulation at the prosodic boundary edges is different from that at the medial position of prosodic domain, and the boundary effect on segment articulation is cumulative along the prosodic hierarchy. The strength of prosodic boundary is also shown to harnesses the vowel-to-vowel coarticulation in that the vocalic effect is more salient across lower vs. higher prosodic boundaries [2, 9].

In the Standard Chinese (SC), the pitch contour and pause are the main apparatus for marking the prosodic boundaries, but the foot boundary rely less on the two acoustic cues. In [13] it is hypothesized that the bisyllabic foot is phonetically characterized by the closer gestural overlap and tonal coarticulation within the foot than across foot domain. But the tone sandhi does not strictly comply with prosodic domain in the SC [10]. It is thus hypothesized that the articulatory and acoustic correlates for foot boundary might be related to the consonantality of initial consonant in that it is more consonant-like at the foot-initial than at syllable-initial position within foot domain [13].

The paper aims at finding out the prosodic signatures in articulatory and acoustic properties of segmental articulation in the SC with a special interest in comparing the foot-internal and foot-boundary segment articulation. The boundary constraint on vocalic anticipatory and carryover effects is also discussed.

Figure 1: A five-level prosodic hierarchy for the SC.



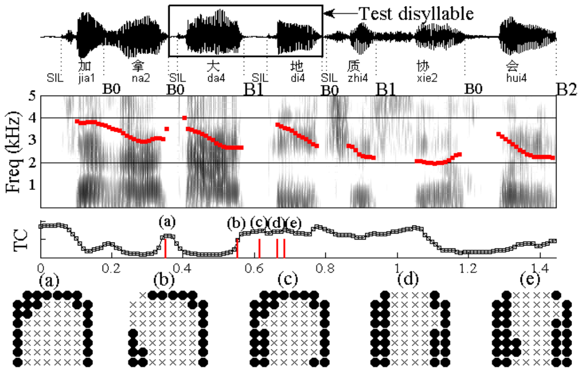
2. METHOD

2.1. Speech material

A five-level prosodic hierarchy in [11] with the prosodic word level being replaced by foot was tentatively adopted and shown in Figure 1. The test segments for articulatory strengthening experiment were unaspirated alveolar stop /t/ and the vowel /i/. In the experiment of boundary constraint for vocalic coarticulation, additional low vowel /a/ was used to construct four bisyllable sequences, /ta#ti/, /ta#ta/, /ti#ta/, and /ti#ti/ (# stands for morpheme boundary).

Five boundary conditions were placed either before the first or the second syllable, and a total of 32 utterances were designed. The syllable was always ended with /a/ preceding the test disyllables while the segment following the disyllables was not controlled. The tonal condition for the disyllable was not controlled either. A part of example utterance was shown in Figure 2.

Figure 2: A part of an annotated utterance (From top to bottom are speech signal with three tiers of prosodic annotation, spectrogram superimposed with pitch contour derived from EGG signal, total contact profile, and five EPG frames).



2.2. Recording

The recording was taken in a sound-attenuated booth. Each sentence was repeated for three times, and the 96 sentences were divided into six blocks in random with two dummy sentences placed in the first and last position in each block. One female speaker participated in the experiment and was instructed to read the sentences at normal speed. The electropalatographic signal was recorded by 62-electron electropalatography (100Hz), and the speech signal was recorded at a sampling rate of 22 kHz.

2.3. Measurements

The maximum linguapalatal contact (MaxC) of the alveolar closure was measured for /t/ (Frames (a) and (c) in Figure 2). The alveolar seal duration (ASD) was defined as the interval from the first to last frame that showed alveolar closure. The acoustic silent duration (AD) was the interval from the last pitch pulse to the energy burst for stop. The maximum linguapalatal contact (MaxV) for /i/ was taken between the one third and one half intervals into the vowel (Frame (e) in Figure 2). The center of gravity (CoG) of the linguapalatal contact was defined following [7], and the centrality index (CC) was defined following [5]. The vocalic duration (VD) and F1/F2 near the MaxV-matched time point were measured for both vowels.

For the second experiment, the linguapalatal contact of the posterior four rows was measured respectively in the final frame over V1 interval (POS_E) and the first frame over V2 interval (POS_S)(Frame (b) and (d) in Figure 2). The V1-end F2 (F2_E) and V2-start F2 (F2_S) were first derived from 20-order covariance LPC in PRAAT, and manually adjusted in the EPG analysis platform developed on Matlab.

3. RESULTS

3.1. Domain-initial consonant strengthening

A series of one-way ANOVA were conducted for three measures of the test consonant, and significant differences were yielded at 0.0001 level. Table 1 showed the summary of the LSD post hoc multiple comparisons for the three measures for /t/.

Table 1: Summary of LSD post hoc comparison results for /t/ ($p<0.05$).

Measures	/i/ context	/a/ context
MaxC	S<F<ip, IP<U	S, F<ip, IP<U
ASD	S, F<ip<IP<U	S, F<ip<IP<U
AD	S, F<ip<IP<U	S, F<ip<IP<U

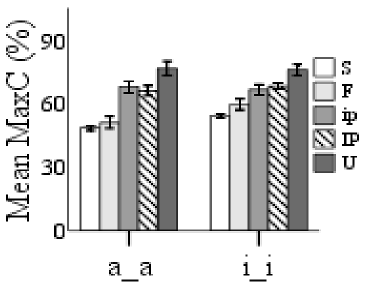
Figure 3 shows the MaxC for /t/ at five domain-initial positions in two symmetrical flanking vowel contexts. In both contexts the speaker successfully distinguished the utterance from the immediate lower prosodic domain, namely intonational phrase. However, the difference between the IP and the ip was not significant across vocalic contexts, with ip-initial position appears to have more contact than IP-initial position in the /a/ context. This might be attributed to the punctuation effect, for the speaker made purposeful pauses when encountering comma. The MaxC for lowest two prosodic domains, syllable and foot, were significantly lower than that for higher domains, but it only differed in the symmetrical /i/ vs. /a/ context. A careful examination looking into the frame tokens for maximum linguapalatal contact frames in the /i/ context found that 3 out of 12 tokens lacking alveolar closure at syllable-initial position, which was reflective of increased vocalic effect on the production of tongue tip closure gesture in a short interval of time.

The post hoc results indicated that the alveolar closure duration (ASD) and acoustic silent interval (AD) were progressively longer above foot domains across both vocalic contexts, with those measures not distinguished below foot domains. However, the result of additional independent *t*-tests showed that the syllable-initial ASD was significant less than foot-initial ASD in both vocalic contexts ($p<0.05$), showing a genuine cumulative effect of prosodic boundaries.

The contextual effect on two articulatory measures was also compared on domain-matched basis through independent *t*-tests. The results showed that syllable-initial /t/ had more contact in the /i/ vs. /a/ context, and the seal duration was shorter in the /i/ vs. /a/ context at both syllable and foot boundaries. This may explain the distinction between syllable and foot domains in the /i/ context in that the consonant gesture is reduced in lower domain because of the vocalic gesture influence as the articulatory interval for consonant gesture becomes shorter.

The post hoc results for the AD patterned with those for the ASD, but no difference was found between syllable and foot domains across vocalic contexts through independent *t*-tests. On the one hand, this result indicates that acoustic silent interval is an effective device to mark higher prosodic domain, on the other hand it fails to distinguish lower domains.

Figure 3: Mean MaxC of /t/ at five domain-initial positions. Error bars refer to one standard error.



The above results indicate that the prosodic structure modulates the domain-initial consonant production in a hierarchically cumulative manner. The prosodic domains above foot are well distinguished acoustically, but the distinction between syllable and foot boundaries relies more on the articulatory measures. The articulatory reduction for consonant at the syllable-initial position tends to show a combined effect of flanking vowel and time interval.

3.2. Boundary effect on high front vowel

The articulatory and acoustic measures for /i/ in the syllable /ti/ marked with both left and right boundaries were submitted to the one-way ANOVA analysis, and the LSD post hoc results were shown in Table 2.

A systematic articulatory variation for the production of /i/ was found depending on the boundary strength for either left or right boundary. When the domain-initial consonant was progressively strengthened, the gestural magnitude for /i/ tended to be reduced accordingly, with reduced vocalic duration and centralized tendency in F1/F2 space. However, when the boundary condition immediately after /i/ ascended along the prosodic hierarchy, the production of /i/ tended to be more target-like. But the result for vocalic duration only partially supported the possible strengthening of the vocalic gesture, for U-final vocalic duration was significant shorter than IP- or ip-final one. This might the result of the segmentation procedure when trailing portion of /i/ at U-final position was truncated because of irregularity of vocalic pulses. No significant difference was found for F1/F2, but the average F1 was lower and average F2 higher in higher prosodic domain.

Table 2: Summary of LSD post hoc comparison results for /i/ ($p<0.05$).

Measures	Right boundary	Left boundary
MaxV	S<F<IP,U / S<ip	S,F,ip>IP,U
CoG	S<F,ip	S,F,ip>IP,U
CC	S,F<IP,U	S,ip>IP,U / F>U
Duration	S<F,U<IP,ip	S>IP,ip,U / F,IP>U
F1	n.s.	F<U, S<U ※
F2	n.s.	S,F,ip,IP>U

※ p = 0.08

The above results show an interesting interaction of domain-initial consonant strengthening and domain-final vocalic strengthening in a syllable. On the one hand the strengthened consonantal articulation suppresses the vocalic influence in higher domain, on the other hand vocalic gesture immediately after the consonant tends to be reduced as the preceding consonant is strengthened. The vocalic strengthening is conditioned by the boundary condition immediately after the vowel.

3.3. Vowel-to-vowel coarticulation

The hypothesized closer gestural overlap in lower prosodic domain implies that vocalic coarticulation is more salient than in higher domains. The vocalic anticipatory effect was investigated by comparing the boundary-matched POS_E and F2_E when V1 was held constant while V2 varied. The vocalic carryover effect was investigated by comparing the boundary-matched POS_S and F2_S when V2 was held constant while V1 varied. Table 3 showed the results of the independent *t*-tests. The V1-end linguapalatal contact and F2 showed significant vocalic anticipatory effect on domains below ip for V1 /a/. The same effect was found for V1 /i/ at Syllable domain, but a significant difference also propped up at IP domain. The latter may be attributed to the tone condition, for the third tone (T3) was applied for the first syllable in /ti (B3) ta/ whereas the first tone (T1) for that in /ti (B3) ti/, with the longer duration under T3 vs. T1 condition facilitating the larger linguapalatal contact.

Table 3: Independent *t*-test results for vocalic anticipatory and carryover effects.

Anticipatory effect		
Measures	POS_E	F2_E
/a/	S*, F*, ip**	S***, F***, ip**
/i/	S**, IP*	S***
Carryover effect		
Measures	POS_S	F2_S
/a/	n.s.	S***
/i/	S※	S**

* $p < 0.05$; ** $p<0.01$; *** $p<0.001$; ※ $p<0.1$

The V2-start linguapalatal contact and F2 showed significant vocalic carryover effect for V2 /i/, but significant difference was only found for F2 for V2 /a/. Considering that carryover effect is generally not salient in the SC, it is safe to say here that carryover effect is constrained by the foot domain.

The above results indicate that the vocalic anticipatory effect is conditioned by the boundary strength on the one hand, and the articulatory constraint for vowels on the other. By careful examination of all tokens, it is found that timing of the initiation of the V2 gesture relative of the V1 gesture is responsible for the all-or-none vocalic anticipatory effect. A second factor involves the specific articulatory constraint for vowels which is reminiscent of the model proposed in [12].

4. CONCLUSION

The results of the first experiment indicate that the prosodic structure shows a hierarchical effect on the consonant articulation in the SC with cumulative increase in linguapalatal contact and seal duration for stop consonant in higher domain-initial positions. It also supports the hypothesis in [13] that the foot and syllable boundaries can be distinguished by the articulatory strength for consonant production instead of the acoustic silent duration. The vocalic gesture is shown to be affected by the preceding consonant production and the boundary condition immediately on its right. When the initial consonant is progressively strengthened the following vowel tends to be progressively reduced, but this is only a tentative conclusion because only one speaker's data is used in the current study. The final-lengthening for vowel is accompanied by the strengthening of the vocalic gesture, which also supports the previous studies. The vocalic anticipatory effect is prominent up to the ip boundary but is constrained by the articulatory constraints for the vocalic gesture. However, the vocalic carryover effect is likely to be constrained in foot domain.

5. ACKNOWLEDGEMENTS

This research was supported by China Social Sciences Funds of the Ministry of Education (Grant 10JJD740007) and China National Natural Science Funds (Grant 61073085). We thank two anonymous reviewers for the comments on the earlier version of the paper.

6. REFERENCES

[1] Byrd, D., Kaun, A., Narayanan, S., Salzman, E. 2000. Phrasal signatures in articulation. In Broe, J.B., Pierrehumbert, J.B. (eds.), *Papers in Laboratory Phonology, Vol. V: Acquisition and the Lexicon*. Cambridge: Cambridge University Press, 70-87.

[2] Cho, T. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *J. Phon.* 32, 141-176.

[3] Cho, T., Keating, P. 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. *J. Phon.* 29, 155-190.

[4] Cho, T., Keating, P. 2009. Effects of initial position versus prominence in English. *J. Phon.* 37, 466-485.

[5] Fontdevila, J., Pallares, M.D., Recasens, D. 1994. The contact index method of electropalatographic data reduction. *J. Phon.* 22, 141-154.

[6] Fougeron, C. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *J. Phon.* 29, 109-135.

[7] Hardcastle, W.J., Gibbon, F.E., Nicolaidis, K. 1991. EPG data reduction methods and their implications for studies of lingual coarticulation. *J. Phon.* 19, 251-266.

[8] Keating, P., Cho, T., Fougeron, C., Hsu, C. 2003. Domain-initial strengthening in four languages. In Local, J., Ogden, R., Temple, R. (eds.), *Papers in Laboratory Phonology, Vol. 6: Phonetic Interpretation*. Cambridge: Cambridge University Press, 145-163.

[9] Kondo, Y. 2006. Within-word prosodic constraint on coarticulation in Japanese. *Language and Speech* 49, 393-416.

[10] Kuang, J.J., Wang, H.J. 2006. An acoustic study on the third tone sandhi across various prosodic boundaries. *Proc. 7th Phonetic Conference of China* Beijing.

[11] Li, A.J. 2002. Prosodic analysis on conversations in Standard Chinese. *Zhong Guo Yu Wen* 291, 525-535.

[12] Recasens, D., Pallares, M.D., Fontdevila, J. 1997. A model of lingual coarticulation based on articulatory constraints. *J. Acoust. Soc. Am.* 102, 544-561.

[13] Wang, H.J. 2008. *Non-linear Phonology of the Standard Chinese*. Beijing: Beijing University Press.