PROSODIC BOUNDARIES EFFECT ON SEGMENT ARTICULATION IN STANDARD CHINESE: AN ARTICULATORY AND ACOUSTIC STUDY

Yinghao Li*
Yanbian University, Yanji (China)

ABSTRACT

This paper presents an electropalatographic (EPG) and acoustic study of prosodic boundaries effect on the domain-initial segments in Standard Chinese.[1] Two speech sounds, namely, the voiceless unaspirated alveolar stop /t/ and the high front vowel /i/, were studied to examine the domain-initial strengthening in both spatial and temporal dimensions. The articulatory and acoustic parameters of the speech sounds were compared in initial positions of five prosodic constituents in Standard Chinese, namely, a Syllable, a Foot, an Immediate Phrase, an Intonational Phrase, and an Utterance. The results show that: (1) the production of the domain-initial consonantal gesture was prosodically encoded. The linguopalatal contact and the seal duration varied as a function of the prosodic boundary strength. The linguopalatal contact was dependent on the seal duration in a nonlinear fashion. Of the acoustic properties of the domain-initial stop, the total voiceless interval and voicing during closure were found to be reliable acoustic correlates that mark the hierarchical structure of the prosody. (2) At the release moment of the domain-initial stop, no consistent pattern was found to support the domain-initial strengthening. The linguopalatal contact of the vowel immediately following the domain-initial consonant did not show a clear trend of domain-initial strengthening; however, the phonatory features of vowels were indicative of pitch reset at major prosodic boundaries. These

---

* **Yinghao Li** 李英浩 is associate professor in the Foreign Languages College of Yanbian University. Address：Foreign Languages College, Yanbian University, Yanji, Jilin Province 133002, China; [youngho_lee@aliyun.com]

indicate that the domain-initial strengthening is restricted on the segment immediately following the boundary. In conclusion Standard Chinese strengthens the phonetic features of the domain-initial segments as a function of boundary strength, which serves as an important way to mark prosodic structure in Standard Chinese.

SUBJECT KEYWORDS
Prosodic boundaries **S**egment articulation **E**lectropalatography
Standard Chinese

1. INTRODUCTION

The production of phonological units in utterances is subject to the conditioning of the prosodic structure. One way this structure information is encoded is how a phonological unit is produced at edges of prosodic constituents. In recent years, an increasing number of studies have observed that the segments at prosodic domain-initial position are produced more strongly than at domain-medial position. The study of this boundary-induced articulatory variation for individual speech segments has become an important topic called the domain-initial strengthening. The domain-initial strengthening refers to the greater magnitude of phonetic realization in the articulatory or acoustic dimension of a phonological unit at the initial position of prosodic constituents. The well-attested assertion that "the stronger the position, the stronger the articulation" (Cho and Keating 2001, 156) has been tested in various languages, namely, English (Fougeron and Keating 1997; Byrd and Saltzman 1998; Byrd 2000; Cho 2001; Cho and Keating 2009; Keating et al 2003), French (Fougeron 2001), Dutch (Cho and McQueen 2005), Korean (Cho and Keating 2001), German (Bombien et al 2010; Kuzla and Ernestus 2011), Taiwan Hokkien (Hsu and Jun 1998; Hayashi, Hsu, and Keating 1999; Keating, et al 2003), and Standard Chinese (Cao and Zheng 2006; Li and Kong 2011).

In previous studies the prosodic shaping of features of a segment has demonstrated a gradient variation in the articulatory and/or acoustic dimension as a function of the boundary strength. In the electropalatographic (EPG) studies on consonants in various languages, the peak linguopalatal contact and the articulatory seal duration were progressively greater at stronger prosodic boundaries (Fougeron and Keating 1997; Fougeron 2001; Cho and Keating 2001, 2009; Keating et al 2003; Li and Kong 2011). In the acoustic domain Jun (1993) found that the voice onset time (VOT) for Korean aspirated stop was progressively shorter as the boundary strength becomes weaker. The reverse was the case for the prosodic positional effect on the VOT of fortis plosives in German (Kuzla and Ernestus 2011). The gradient variations as a function of domain strength has been considered to be related to the duration-dependent undershoot model as postulated by Lindblom (1990). That is to say, the time consumes the production of a segment determines how much the phonetic target is realized. The segment is probably hypoarticulated at weaker position because insufficient time is consumed. The linguistic motivation for this articulatory adjustment is hypothesized to be attributed to enhanced syntagmatic contrasts from the neighbouring segments through magnifying the associated features of the segments (Fougeron and Keating 1997; Hsu and Jun 1998; Cho 2005).

The prosodic strengthening effect is also temporally constrained. The scope for the domain-initial strengthening effect tends to be manifested at the first post-boundary segment, whereas the following segment in the same syllable is seldom affected (Fougeron and Keating 1997; Cho 2005; Byrd, Krivokapic, and Lee 2006). In an electropalatographic study Fougeron and Keating (1997) found that there was no consistent and reliable pattern of the linguopalatal contact for non-initial /o/ in the domain-initial syllables compared with domain-initial consonant /n/. Cho (2005) examined boundary-induced articulatory and acoustic variation of post-boundary vowels in domain-initial CV syllable in English. No consistent tongue fronting or raising was found for the high front vowel /i/, nor was tongue lowering and backing found for low back vowel /ɑ/. Byrd, Krivokapic, and Lee (2006) investigated the temporal scope of prosodic boundary effect on the segment articulation. They found that the articulatory gesture of the post-boundary consonant was significantly strengthened with a longer duration and larger articulatory displacement. The following consonants were temporally shortened, accompanied by smaller articulatory displacement. However, some studies have found evidence of the manifestation of the domain-initial strengthening effect on non-initial vowel in post-boundary CV syllable (Farnetani and Vayra 1996; Cho and Keating 2009; Kim and Cho 2011). In their electropalatographic study on prosodic effect on the production of CV syllables, Farnetani and Vayra (1996) found that the initial syllable was altogether strengthened and the vowel in the domain-initial syllable has a more open vocal tract, regardless of lexical accent condition. In the electropalatographic and acoustic study on the boundary effect on English segment production, Cho and Keating (2009) found mixed results for the locality hypothesis. They found that the articulatory parameters for V in post-boundary CV syllables were insensitive to the boundary strength; however, the vowel amplitude did show strengthening at utterance-initial position compared with that for lower boundaries. In the electromagnetic articulograph (EMA) study on the domain-initial strengthening effect, Kim and Cho (2011) found that the vowel following /h, p^h/ at the initial position of prosodic boundaries showed gradient tongue movement magnitude as affected by the boundary strength, which was rather similar with the case for the domain-initial vowel in the syllable. The above findings indicate that the domain-initial strengthening effect might interact with other confounding factors in affecting the post-boundary initial syllables. In a simulation study, Byrd and Saltzman (2003) showed that speech production mechanism controlling articulator(s) gets increasingly longer time when approaching a prosodic boundary, and this gestural actualization time becomes shorter when the boundary recedes. That is to say, the articulator(s) can be orchestrated to fully realize the phonetic target if provided with sufficient gestural preparatory time, and this timing mechanism speeds up immediately after the boundary.

If the domain-initial strengthening only affects the first segment after the prosodic boundaries, it indicates that the vowel in the post-boundary syllable is subject to the shaping of another prosodic mechanism instead of the domain-initial strengthening. In a tonal language such as Chinese, syllables are specified with lexical tones, which contrast meanings. In previous studies the syllable tones of Mandarin was proved to be hierarchically determined, which means that the tone specification is encoded according to the boundary strength (Tseng et al. 2005). Nevertheless, this hierarchically structured tonal specification is the manifestation of f0 reset, which is another prosodic means to mark the prosodic boundaries in Standard Chinese. In recent studies it was found that voice quality may also be affected by the prosodic structure. Electroglottographic (EGG) studies have shown that voice quality of the post-boundary vowel became progressively breathier at higher prosodic boundaries. For the gradient variation of voice quality measures, it was assumed to be related to f0 reset at major boundaries (Garellek 2014).

Previous studies have shown that the domain-initial strengthening effect was dependent on the segmental identity and language. Regarding the segmental identity, it was found that the alveolar fricative is more resistant to articulatory variation as a function of prosodic positions. For example, in the electropalatographic study on the domain-initial strengthening effect on the production of the French alveolar fricative, Fougeron (2001) found that the linguopalatal contact pattern for the alveolar fricative was less subject to the prosodically-conditioned articulatory variations as evidenced in the production of alveolar stops. A similar result was also found in Korean in that the domain-initial strengthening effect on the articulatory magnitude for the phonological units was not as salient for fricatives as for stops (Kim 2001). This segment-specific response to the prosodic conditioning is explained away as the aerodynamic requirements for producing fricatives which results in rather rigid tongue gestures. In terms of the language-specific manifestation for the domain-initial strengthening effect, Cho and Keating (2001, 2009) found that the VOT for Korean lax and aspirated alveolar stops /t, t^h/ and English /t/ was increased in higher prosodic boundaries. However, the VOT variations as incurred by the prosodic positions are in the inverse direction for /t/ in Dutch and German, with shorter VOT appearing at stronger boundaries (Cho and McQueen 2005; Kuzla and Ernestus 2011). These results concerning the positional-dependent VOT difference involve the different laryngeal gestures across languages. The prosodic signature on segment articulation may be represented by different phonetic dimensions. In the articulatory study on the Tamil language, the duration and timing relationship in consonant clusters was affected by prosodic structure, but no such effects were found for the consonant articulatory magnitude (Byrd et al 2000). In short, the segment- and language-specific prosodic signature on the segment articulation is attributed to the fine-grained articulatory control, on the one hand, and cross-language differences in segment articulation, on the other.

The current research investigated the effect of prosodic structure on the production of individual speech segments at prosodic domain-initial positions in Standard Chinese. By prosodic structure, we mean that a spoken utterance is hierarchically organized with the higher prosodic domain being decomposed into immediately lower constituents. Following the Strict Layer Hypothesis (Selkirk 1984), the utterance is shaped by the hierarchical prosodic structure in that the higher prosodic domain directly dominates one or more immediately lower prosodic domains, and a given prosodic domain must be contained by an immediately higher prosodic domain. The prosodic model for Standard Chinese used in the current paper follows Li 李 (2002) and Lin (2002) with some minor modifications. Based on the autosegmental-metrical theory, Li 李 (2002) and Lin (2002) assert that the speech utterance of Standard Chinese is hierarchically organized in that the larger prosodic units are composed of several immediately lower prosodic constituents (domains). These constituents include syllable, foot, prosodic word, minor phrase, major phrase, and utterance, which were claimed to be distinguished by the pitch contour and break. For the two prosodic phrases, we use the terms intermediate phrase and intonational phrase for the comparison with the results in other languages. The domain of foot is normally syntactically defined, which constitutes the basis for the higher prosodic domain-prosodic word, and the latter comprises a foot and/or a following unassigned monosyllable (Li 李 2002, 526). In the current paper, the bi-syllabic foot domain is to be studied because it is the basic unit of the metrical organization in Standard Chinese (Wang 2008). The domain of intermediate phrase comprises one or several prosodic words (in this paper we use foot as the immediately subordinate component for intermediate phrase), and is characterized by a phrasal accent and noticeable pause (either silent or filled pause). For intonational phrase, the prosodic cues include longer domain-ending pause and noticeable pitch contour resetting. The utterance, constituted by one or several major phrases, is cued by a sentence accent, substantial durational compression of domain-ending syllable and longest pause (Wang 2008, 257).

Figure 1 shows the prosodic hierarchy used in the current paper.

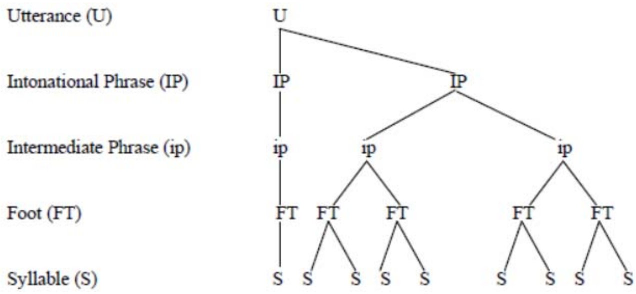

Figure 1 The hierarchy of prosodic structure of Standard Chinese

Extensive research on the prosodic structure in Standard Chinese has shown that the f0 reset, the syllable durational pattern, as well as the pause inside utterances are the three main acoustic correlates cueing the prosodic structure (Li 李 2002; Wang, Yang, and Chen 2004; Hu, Xu, and Huang 2002; Lin 2002). However, the function of the domain-initial strengthening in marking the Chinese prosodic structure has received little attention. In her acoustic study on the segmental lengthening, Cao (2005, 165) suggested:

> 'The domain-initial segmental lengthening, likewise, has the function to mark the boundary strength, and this function can not be underestimated. It is more direct and reliable to indicate the prosodic boundary strength. …meanwhile, the post-boundary consonantal duration is positively correlated with the perceived boundary strength, and it is progressively increased as a function of the boundary strength.'

Keating et al (2003, 161) hypothesized that a lexical tone language such as Taiwan Hokkien might have a more salient domain-initial strengthening effect than English because "it should have less recourse to pitch to mark domain edges". Although their result did not support the hypothesis, the cumulative effect that was found did show the universality of the prosodic conditioning of domain-initial segments. Wang (2008) argued that the consonant initial inside a foot domain, the basic metrical template for rhythmic organization in Standard Chinese, was reduced somehow compared with one that heads a foot domain; however, no gestural reduction was found on prosodic constituent initial position above the foot domain. This argument directly supports the tenets of the domain-initial strengthening in that the consonant gesture tends to be undershot in prosodically weak position. However, it is worth investigating whether gradient articulatory gesture exists at the initial position of prosodic boundaries of different strengths.

In the electropalatographic study, Cao and Zheng (2006) found that the linguopalatal contact for the phrase-initial consonant was greater than that for the phrase-medial consonant in Standard Chinese. Li and Kong (2011) investigated the articulatory strengthening phenomenon for domain-initial alveolar stop /t/ in Standard Chinese, and found a cumulative increase of linguopalatal contact as well as alveolar seal duration from syllable to utterance boundaries based on one female speaker's electropalatographic data. In this paper we will extend our previous study by investigating the same consonant, and the tautosyllabic high front vowel /i/ uttered by two speakers. The sole research question to be addressed in the current paper is how the prosodic position affects the production of consonantal and vocalic gestures in CV syllables in Standard Chinese. Two aspects of the domain-initial strengthening are to be covered: (1) Does the prosody affect the articulatory and acoustic properties for the domain-initial segment in the cumulative pattern? (2) Does the domain-initial strengthening affect all segments in the post-boundary syllable?

The universality of the domain-initial strengthening effect leads us to predict that the segment at domain-initial position is strengthened and that a cumulative scale exists, which is indicative of the linguistic encoding of the segmental production rather than an intrinsic biomechanical process. As a matter of fact, previous studies with one speaker showed a strong tendency of this prosodically: conditioned segment articulation in that the linguopalatal contact for unaspirated /t/ was positively correlated with the prosodic hierarchy. The second prediction is concerned with the temporal scope of the domain-initial strengthening effect. It is hypothesized that the domain-initial strengthening effect scope is limited in the domain-initial segment, and the following vocalic segment is not subject to the domain-initial effect.

The domain-initial strengthening quickly fades away toward the end of the initial segment production, as predicted by Byrd and Saltzman (2003).

2. METHOD

2.1 Electropalatography (EPG)

The tongue-palate contact for lingual consonants captured by electropalatography is a reliable indicator of articulatory magnitude. Thus, more linguopalatal contact indicates more oral constriction and greater articulatory magnitude. WinEPG Electropalatography produced by Articulate Instruments was used to capture the tongue-palate contact signal. WinEPG uses custom-made pseudo-palates of thin acrylic base. The pseudo-palate covers the palate from root of upper teeth to the anterior portion of the soft palate. The layout of the 62 electrodes is designed to be relative to the anatomical landmarks in order to carry out inter-subject comparison. When the tongue contacts an electrode, the circuit is completed and the signal is recorded by the WinEPG system.

2.2 Speech Stimuli

The test segments were the voiceless unaspirated alveolar stop /t/ and the high front vowel /i/. The stop was placed at domain-initial positions of five prosodic domains, namely, syllable (SYL), foot (FT), intermediate phrase (ip), intonational phrase (IP), and utterance (U); and the stop was placed in two symmetrical vocalic environments, low vowel /a/ and high front vowel /i/. The high front vowel /i/ in syllable /ti/ was the other test segment for investigating the temporal scope of the domain-initial strengthening. The syllable was preceded by the five prosodic domains immediately before the initial stop, and was followed by syllables starting with the alveolar stop /t/. For the tonal specification of the test syllables, the low tone (T3) was avoided as much as possible because it would possibly affect the linguopalatal contact of /i/ (Hoole and Hu 2004). The falling tone was used in most cases for examining the vowel-initial f0 and voice quality of the vowels with varying boundary conditions.

Table 1 shows the samples of the stimuli of the sentence set for test consonant /t/. The utterance domain was elicited by a full period, and speakers were instructed to make a long pause after it. The comma was used to parse the intonational phrase domain, with the syllable number in the two sentence components ranging from 6 to 12 syllables. The speakers were instructed to ignore the comma. The intermediate phrase domain was a noun phrase that comprised at least two feet. The foot domain was the second word inside an immediate phrase. The syllable domain was defined as the boundary occurring inside a foot. The five-level break index based on C-ToBI (Li 2002) was used to code the prosodic domains: the break index number 4 (B4) was used to transcribe the Utterance domain, and break index 0 (B0) the Syllable domain. The rest break numbers were used to transcribe the intermediate domains.

Table 1. Sample sentences for the test consonant /t/. (The underscored characters and syllables were the locations where the test consonant /t/ headed the prosodic domains. In the right column of each cell, the stimuli sentence was given first, and the phonetic transcription between slash brackets and with the literal translation in parenthesis following.)

U-initial	国家对高等教育的投入加大。< B4>大学教师的工资不断提高。 /kuo35 tɕia55 tuei51 kɑu55 təŋ214 tɕiɑu51 y51 də0 t^{h}ou35 ʐu51 pu35 tuan51 tɕia55 ta51。ta51 ɕyɛ35 tɕiɑu51 ʂʅ55 də0 kuŋ55 tsɿ55 pu35 tuan51 t^{h}i35 kɑu55/ (The state increases its input in higher education. The salary for university teachers is increasing.)
IP-initial	这一地带昼夜温差大，<B3>大麦颗粒饱满。 /tʂɤ51 i35 t^{h}i 51 tai51 tʂou51 iɛ51 uən55 tʂ^{h}a55 ta51, ta51 mai51 k^{h}ɤ55 li51 pɑu35 man214/ (The temperature difference is dramatic between day and night in this area, thus the barley seed appear plump.)
ip-initial	负债重重的加大<B2>大力削减学生的奖学金计划。 /fu51 tʂai51 tʂ^{h}uŋ35 tʂ^{h}uŋ35 də0 tɕia55 ta51 ta51 li51 ɕy55 tɕiɛn214 ɕyɛ35 ʂəŋ55 də0 tɕiɑŋ214 ɕyɛ35 tɕin55 tɕi51 xua51/ (The heavily-indebted UC cut down the student scholarship program.)
FT-initial	华夏<B1>大地充满了勃勃的生机。 /xua35 ɕia 51 ta51 ti51 tʂ^{h}uŋ55 man214 lə0 po35 po35 də0 ʂəŋ55 tɕi55/ (The China land is full of vigor.)
SYL-initial	负债重重的加<B0>大大力削减学生的奖学金计划。 / fu51 tʂai51 tʂ^{h}uŋ35 tʂ^{h}uŋ35 də0 tɕia55 ta51 ta51 li51 ɕy55 tɕiɛn214 ɕyɛ35 ʂəŋ55 də0 tɕiɑŋ214 ɕyɛ35 tɕin55 tɕi51 xua51/ (The heavily-indebted UC cut down the student scholarship program.)

2.3 Speakers and Procedures

Two university students (one male and one female) participated in the experiment. They had lived in a northern province or municipality in China before entering into university. After enrolling in the university they had worked as part-time announcers at the university TV station. At the time of the recording session they were 27 years old and were paid for their participation.

The recording was conducted in a sound-attenuated booth at Peking University. Before the recording session the participants were given 30~50 minutes to adapt to the reading task with a pseudo-palate installed in their mouths and to familiarize themselves with the sentences. In the recording session, the sentences were randomized blocked and each sentence was repeated three times for the female speaker and five times for the male speaker. The sentence blocks were presented on the computer screen which was positioned about one meter in front of them, and they were instructed to read the sentence list at a normal speech rate. No specific instructions on prosodic phrasing were given, except that they were required to make an intentional pause after an orthographic period in the sentence. The electropalatographic, electroglottographic, and the speech signal were simultaneously recorded into a computer (see Figure 2 for an example). The sampling rate for EPG signal was 100 Hz, and that for speech and EGG signal was 22 kHz. For the male speaker the respiratory signals for chest and stomach breathing were also recorded, which will not be analyzed in the present paper. After the recording the sentences that had unclear pronunciation of the test segments or signal problem were eliminated. In the end a total of 264 sentences were submitted for further analysis.

The articulatory and acoustic analysis was carried out on the Matlab program developed to process the electropalatographic and speech signals. First, the electropalatographic signal was temporally aligned with the acoustic signal by using the algorithm in Li and Pan (2012). Then each utterance was parsed from the recording item and annotated at PRAAT (Boersma 2001) for two tiers, the syllable tier and the break index tier. Figure 3 demonstrates the break index tier in this paper.

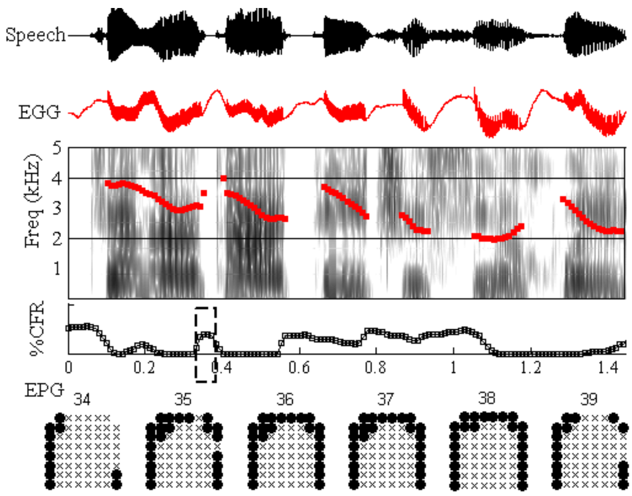

Figure 2 The simultaneously-recorded speech signal, EGG and EPG signals. (From top to bottom: speech signal, EGG signal, the spectrogram with the f0 contour plotted on it, one EPG contact measures, namely, % contact of the Front Region (defined later), and six consecutive EPG frames that show the alveolar closure gesture between 0.33 and 0.38 second. The filled dots indicate tongue-palate contact, and the "x" no contact.)

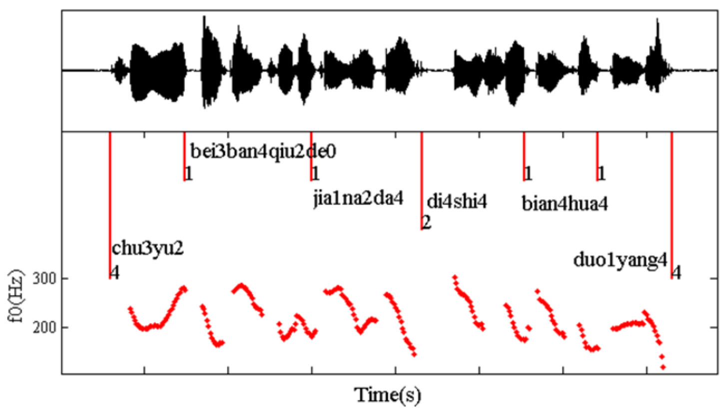

Figure 3 The break index tier of the utterance "处于北半球的加拿大地势变化多样 (Canada, located in the northern hemisphere, has diversified geographic features.)。". (The numerals on the right of the line indicate break index number. The phonetic annotation is in Chinese *pinyin*, and the numerals following each syllable stand for the tones, 1 for high level tone, 2 for rising tone, 3 for low tone, 4 for falling tone, and 0 for neutral tone.)

2.4 Articulatory Measurement

The articulatory gestures of /t/ and /i/ have distinct tongue-palate contact pattern. In Figure 4 we defined the first four rows as the Front Region, which was closely related with the closure-formation of the alveolar stops. The linguopalatal contact at the back four rows was closely related with the coarticulatory effect of the following vowel on the preceding consonant, which was not studied in this paper.

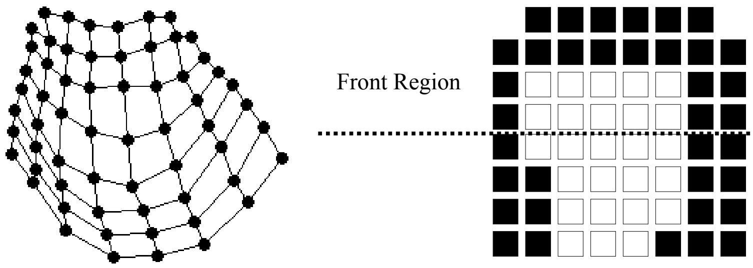

Figure 4 The three-dimensional illustration of the placement for 62 electrodes on pseudo-palate (left) and the division of the two regions of the electrodes (right). The linguopalatal contact on the right chart is the point of maximum contact frame for unaspirated /t/ in syllable /ti/. The shaded squares refer to the contact captured by the WinEPG system.

To study the boundary effect on the post-boundary alveolar stop, two key time points for tongue-to-palate contact were defined: one was the point of maximum contact frame (PMC) and the other the release frame. The PMC was defined as the frame that had the maximal linguopalatal contact during the alveolar closure interval (when the complete alveolar closure was observed), or acoustic closure interval (when no complete alveolar closure was found). The contact pattern at the PMC is considered to directly reflect the magnitude of articulatory excursion for the segment in question (Cho and Keating 2009). The release frame was defined as the last frame of the closure interval of the alveolar stop. If no alveolar closure was observed in the acoustic closure interval (normally in FT or SYL domain), the release frame was defined as the frame immediately before the acoustic release of the stop. In case the alveolar stop was realized as a voiced approximant (see Cho and Keating 2001), no release frame was available. As argued by Cho and Keating (2009, 470), the release frame "might reveal additional information about the prosodically-conditioned articulatory variation". For the two key frames, the percent of contacted electrodes in the Front Region was computed, for this region was related with the tongue tip/blade gesture. The seal duration (SD) was taken as the interval between the first and last frame of the alveolar closure. In case no closure frames were found or the alveolar stop was realized as voiced approximant, the seal duration was zero.

To test the locality hypothesis of the domain-initial strengthening, we investigated the linguopalatal contact of the high front vowel /i/ in the domain-initial syllable /ti/. In the previous studies, the high front vowel /i/ was shown to be resistant to the coarticulatory effect, regardless of the boundary strength (Cho 2004). Provided that the boundary effect is restricted at the initial position of the post-boundary syllable, the vowel /i/ in the syllable /ti/ would not show strengthening effect. To reduce the confounding domain-final strengthening/lengthening effect, the syllable /ti/ was designed to be followed by a Syllable boundary. However, the Syllable-initial tokens were followed by a Foot boundary because the bi-syllabic foot was used to construct the test sentences as much as possible. In this case, the syllable was assigned to head a following foot domain. The maximal linguopalatal contact frame for /i/ was selected between the one third and half toward the vocalic interval. By doing so, the coarticulatory effect of the preceding and following segments of the vowel was maximally reduced. The percent of contacted electrodes in the whole region was computed to measure the articulatory magnitude of /i/.

Besides the linguopalatal measure for vowel /i/, we also investigated the vowel-initial f0, open quotient (OQ), and speed quotient (SQ) of the vowels /i/ and /a/, for they tended to be conditioned by the prosodic structure. As shown in Introduction, the f0 manifests the f0 reset in Standard Chinese. The OQ and SQ reflect the voice quality of vowels. The OQ is the inverse of the contact quotient (or closed quotient), which shows the portion of time when the vocal folds are opened in each glottal period. Higher OQ is related with larger glottis opening. The SQ computes the time portion ratio between the opening phase and closing phase as defined in EGG signal in each glottal period. The SQ is associated with how fast the vocal folds adduct. The higher the SQ is, the faster the vocal folds adduct. A recent study on the voice quality strengthening in English and Spanish had indicated that vowels after the domain-initial glottal stops showed lower contact quotient, or higher OQ, at higher prosodic boundaries (Garellek 2014). In the current paper the f0 was computed based on the derivative of the EGG signal, and the OQ and SQ were obtained by Hybrid method (Howard 1995). Figure 5 shows the definition of critical moments and intervals, and three equations to obtain the measures.

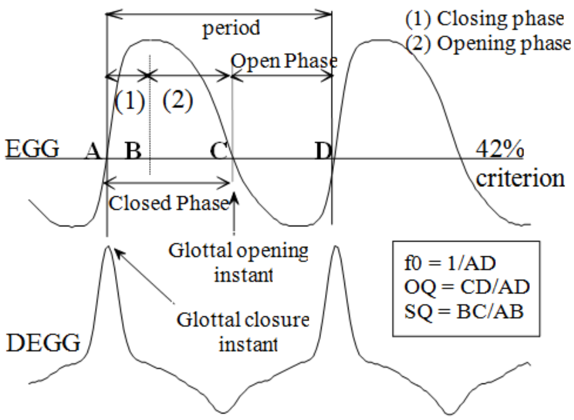

Figure 5 The definition of critical moments and intervals, and three equations to obtain f0, OQ and SQ.

2.5 Acoustic Measurement

As the first step, the durational properties of syllables were measured to check whether the utterances were produced appropriately for the current study. One important acoustic parameter that marks the prosodic structure is the pre-boundary vocalic duration (or V1 duration). This measure is used to demonstrate the final lengthening, which refers to the durational variations of the rhyme in pre-boundary syllable (Wightman et al 1992). In this paper the V1 duration was defined by the observable F2 formant trajectory in the pre-boundary vowel. Another acoustic parameter that indexes the prosodic structure is the acoustic closure interval. In this paper it was taken as the interval started from the termination point for F2 contour in the previous vowel to the stop burst for the following consonant /t/. The acoustic closure duration would not be measured at the utterance-initial position because the better part of the silent interval was not the result of the alveolar closure gesture. These two acoustic measurements were to be compared with the previous results of boundary-induced lengthening.

The acoustic properties of the test segments were investigated because if the domain-initial articulatory strengthening was demonstrated in the acoustic domain, the listeners would possibly pick up those acoustic attributes for phrasing the utterances, and the research of the perceptual relevance of the domain-initial strengthening can be carried out in future. The acoustic measures included: (1) Voice onset time (VOT) of the voiceless unaspirated alveolar stop. The VOT was measured from the point of the stop release to the voice start of the following vowel, signalled by the onset of F2 trajectory in the spectrogram. In case of approximant realization or voice stop, the VOT was zero. (2) Voicing during stop closure and total voiceless interval. These two measures were computed to partially demonstrate the vocal folds state during the oral closure interval (Cho and Keating 2001). The vocal folds state during the oral occlusion of the alveolar stop might as well reflect the conditioning of boundary strength. The voicing during stop closure was the percent of the voicing of the acoustic closure duration. The voicing was represented by the voicing bar at low-frequency of the spectrogram, or cyclical abduction-adduction of vocal folds in EGG signal. The total voiceless interval was the duration of the silent interval in the acoustic closure duration plus VOT. No measurement was conducted at the Utterance-initial position for the two measures. (3) RMS burst energy. Following Cho and Keating (2001) the RMS burst energy for /t/ was defined as the acoustic energy at the burst through calculating the RMS value of frequencies above 500Hz of an FFT spectrum. The low frequency was eliminated because the possible effect of the voicing over the stop release.

3. RESULT

3.1 The Prosodic Hierarchy Indicated by the Utterances

To ensure that the speech material was appropriate for the further analysis, two durational measures were collected and compared with the previous study regarding the acoustic correlates of the prosodic structure in Standard Chinese.

Although the function of the pre-pausal lengthening in indicating the prosodic boundaries is a well-accepted notion, the case for Standard Chinese is far from clear. Cao (2004, 2005) argued that the pre-boundary rhyme lengthening as conditioned by prosodic hierarchy was mainly found at the prosodic phrase level but not at the sentence and paragraph level. Qian, Chu, and Pan (2001) also found that pre-boundary syllable lengthening only existed at the right edges of immediate and intonational phrases, and this effect was unstable at prosodic word level. In their acoustic investigation on the durational pattern of the pre-pausal rhyme, Wang, Yang, and Chen (2004) found the temporal expansion was salient at the lower prosodic boundaries, such as prosodic word and prosodic phrase, whereas its function in marking the prosodic boundary strength became less important in higher prosodic domains when the pause and f0 reset were the two salient acoustic correlates for marking prosodic boundaries, though at this time the cumulative final lengthening was still observable.

Figure 6 shows the average pre-boundary V1 duration produced by the two speakers. High agreement was achieved regarding the syllable durational pattern domain-finally. For both speakers, the V1 duration was significantly longer at domain-final positions in higher prosodic domains (ip, IP, and U) than in lower ones (SYL, FT). Meanwhile, it was shorter at FT-final position than at SYL-final position, and the difference between the two for the male speaker reached significant level. For the female speaker, the V1 duration was shorter at U-final position than at IP- or ip-final positions, though no significant difference was found. These results basically confirm the findings by Cao (2004, 2005) and Qian, Chu, and Pan (2000). On the one hand, the rhyme at the ip- or IP-final positions tends to be lengthened compared with that at SYL- or FT-final positions, and no cumulative final lengthening is found, on the other.

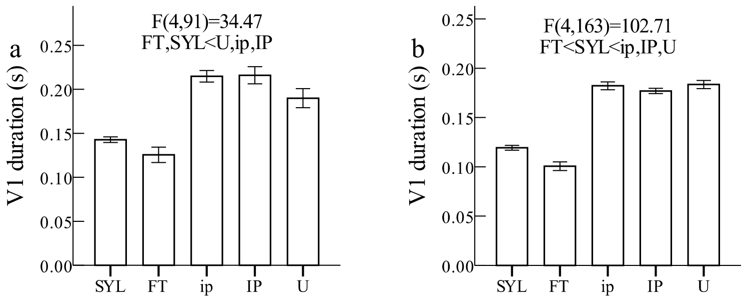

Figure 6 Pre-boundary average V1 duration for female speaker (a) and male speaker (b). The error bar stands for one standard error.

Figure 7 shows the acoustic closure interval at four prosodic boundaries. As predicted, the acoustic closure interval varied depending on the strength of the prosodic boundaries. For both speakers, the longer acoustic closure interval appeared at higher prosodic boundaries. The female speaker produced much longer closure interval at IP- and ip-initial positions, which was attributed to the slower speech rate when the female speaker produced the speech stimuli.

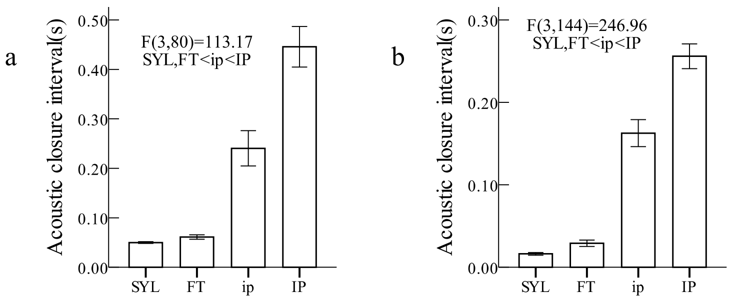

Figure 7 The acoustic closure interval of /t/ at domain-initial positions for (a) female and (b) male speaker. (The measurement at the U-initial position was excluded because it was not the result of alveolar closure gesture on the whole.)

3.2. Articulatory Measures

3.2.1 Linguopalatal contact and seal duration

The one-way analysis of variance (ANOVA) was separately conducted for the articulatory measures produced by the two speakers.

Table 2 shows the results of the ANOVA and the post-hoc comparisons of the articulatory measures in two vocalic contexts. Figures 8 shows the means with one standard error for the linguopalatal contact at two frames.

Table 2 The ANOVA and multiple comparison (Bonferroni method, $p<0.05$) results for the articulatory measures for /t/. Direction of the difference in the measures is indicated by "<" for less percentage or length. The symbol ※ indicates trend effect ($p<0.1$).

Measure	Female speaker	Male speaker
	Context a	
% contact(PMC)	$F(4,67)=29.91$, $p<0.0001$	$F(4,118)=65.71$, $p<0.0001$
	SYL, FT<ip, IP, U	SYL, FT<ip, IP, U
% contact(Release)	$F(4,67)=5.48$; $p<0.0001$	$F(4,118)=12.25$, $p<0.0001$
	SYL, FT<ip, U<ip	SYL, FT<ip, IP, U
Seal duration	$F(4,67)=44.51$, $p<0.0001$	$F(4,118)=107.81$, $p<0.0001$
	SYL, FT<ip<IP, U	SYL, FT<ip, IP, U
	Context i	
% contact(PMC)	$F(4,19)=10.24$, $p<0.0001$	$F(4,40)=7.64$, $p<0.0001$
	SYL<ip, IP, U	SYL, FT<ip, IP, U
% contact(Release)	$F(4,19)=10.04$, $p<0.0001$	$F(4,40)=4.29$, $p<0.01$
	SYL<ip, IP, U	SYL<ip, IP; FT<ip, IP※
Seal duration	$F(2,15)=75.34$, $p<0.0001$	$F(4,40)=16.08$, $p<0.0001$
	SYL<FT<ip	SYL, FT<ip, IP, U

Generally speaking, the variation of the linguopalatal contact at the PMC and the release frame, and the alveolar seal duration, was larger and longer at higher prosodic constituents than at lower ones. Figure 8 presents the linguopalatal contact of /t/ at the PMC and the release frames. At the PMC frame, a clear tendency of cumulative increase of linguopalatal contact as a function of the boundary strength was found in both vocalic contexts for the female speaker, who distinguished two types of boundaries in /a/ context, namely, lower boundaries of SYL and FT, and higher boundaries of ip, IP, and U. In /i/ context four out of five boundaries were distinguished. For the male speaker, the boundary strength as indicated by the linguopalatal contact at the PMC frame falls into two categories, lower boundaries (SYL and FT), and higher boundaries (ip, IP, and U). A close look at Figure 8 (b) shows that the linguopalatal contact at SYL-initial position is slightly larger than that at FT-initial position in /i/ context, and the peak contact at IP-initial is the modestly largest among the three higher boundaries.

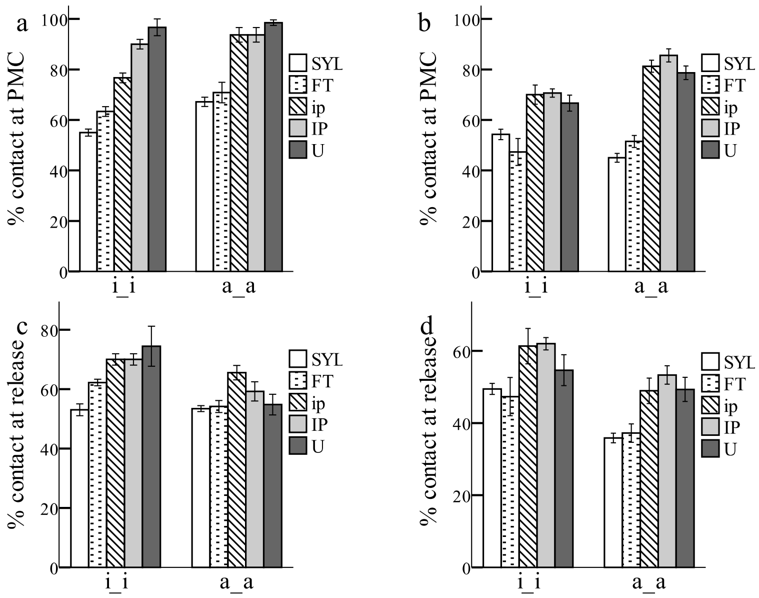

Figure 8 The linguopalatal contact (%) in the Front Region at the PMC and release frame. (The left column is for the female speaker, and the right one male speaker.)

The linguopalatal contact pattern at the release frame was similar to that at the PMC frame except in the /a/ context produced by the female speaker. For the female speaker, the linguopalatal contact varied as a function of the boundary strength in the /i/ context, but no such relation was found in the /a/ context in which condition the significantly lower linguopalatal contact was found at the U-initial position compared with that at the ip-initial position. For the male speaker, the SYL and FT boundaries were distinguished from the higher domains of the ip and IP in the /i/ context, whereas the case in /a/ context was identical to the result obtained at the PMC frame. The linguopalatal contact obtained at the release frame might as well be suggestive of the boundary effect on the segment articulation; however, this effect might fade away toward the plosive release.

Figure 9 shows the sample tokens of linguopalatal contact of /t/ at five domain-initial positions, taken from the PMC frame. As clearly seen in the two rows, the linguopalatal contact in the Front Region, which is closely related with the alveolar closure gesture, is decreasing as the boundary strength becomes progressively weaker, regardless of the vocalic environment. The rightmost frame in the third row shows an incomplete alveolar closure that happened at the SYL-initial position. A careful examination of occurrence ratio of this incomplete alveolar shows that no such tokens existed in the ip-, IP- and U-initial position. For the female speaker 18% tokens have incomplete alveolar closure at SYL-initial position, but no such token is found at FT-initial position. The occurrences of the incomplete alveolar closure produced by the male speaker are 21% and 26%, respectively, at SYL- or FT-initial positions.

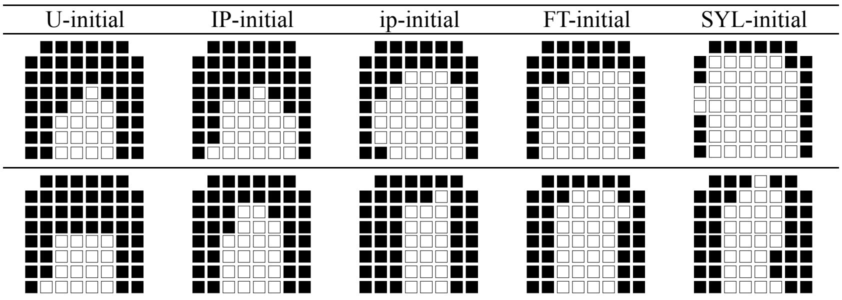

Figure 9 Sample tokens of linguopalatal contact for /t/ produced by the male speaker. (The first row refers to the boundary type. The second and third rows respectively refer to the peak contact frame in /a/ context and /i/ context.)

When measuring the alveolar seal duration, it was not possible to determine the starting point for the alveolar closure in the majority of tokens at IP- or U-initial positions produced by the female speaker in /i/ context. This was because full contact frames that had no linguistic meaning preceded the linguistically meaningful tongue-palate contact in the alveolar region. Thus, the seal duration for /t/ in /i/ context at U- or IP-initial positions was excluded from further analysis. Table 2 and Figure 10 show the alveolar seal duration for the domain-initial alveolar stop. For the female speaker the alveolar closure duration increases as a function of boundary strength in both vocalic contexts. In the /a/ context three distinctions were identified by the alveolar closure duration. In the /i/ context all three boundaries were distinguished with the alveolar closure duration being shortest at SYL-initial position and longest at ip-initial position. For the male speaker, only two types of boundaries are distinguished: the alveolar closure duration at SYL- and FT-initial positions was significantly shorter than at other higher prosodic domain-initial positions.

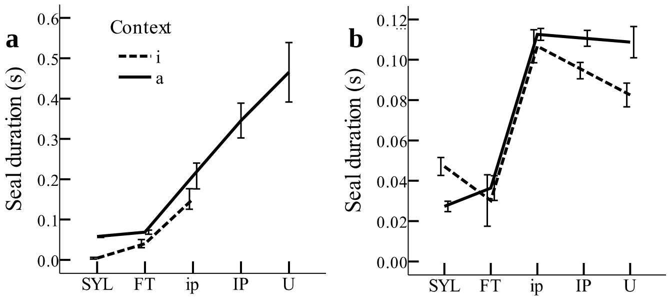

Figure 10 The alveolar seal duration after five prosodic boundaries in two vocalic contexts produced by (a) female and (b) male speakers.

The articulatory magnitude of a consonant at different domain-initial positions might be duration-dependent. Provided sufficient time is given, the consonantal gesture is to be fully realized. In their EPG study on the Korean consonants at different domain-initial positions, Cho and Keating (2001) found that the relationship between the linguopalatal contact for consonants and the seal duration was asymptoti, instead of linear when the linguopalatal contact became larger. Figure 11 shows the scatter plot of the linguopalatal contact in the Front Region against the seal duration, with the curve fitting function that can account for a large portion of variance. As indicated in Figure 11(a, c, d), a polynomial fit was obtained to show that the nonlinear relationship between the linguopalatal contact and the alveolar closure duration. The special case occurred in the /a/ context produced by the female speaker, which shows an exponential relationship between the two measures. A close look at Figure 11b shows the linguopalatal contact increases nearly linearly when the alveolar seal duration was below 0.15 second; however full contact in the Front Region was achieved above 0.15 second. This is an interesting finding in that a time threshold for a full contact in the front four rows tends to exist. It was observed that below this threshold the alveolar closure started from the first row, or the alveolar ridge, and extended to posterior rows, or the post-alveolar/palatal area, as more alveolar closure time was provided in higher prosodic constituents.

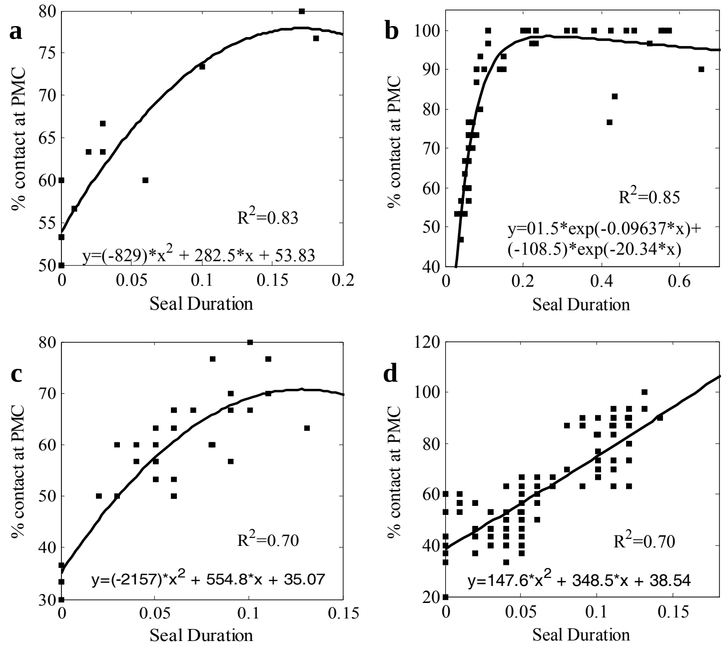

Figure 11 Scatter plot of the linguopalatal contact percent in the front region against the seal duration. (Female speaker, see (a) for /i/ context and (b) for /a/ context. Male speaker, see (c) and (d)).

3.2.2 Vocalic linguopalatal contact

The maximal linguopalatal contact of /i/ in the domain-initial syllable /ti/ was investigated to test the locality hypothesis of the domain-initial strengthening. As predicted in the Introduction, if the domain-initial strengthening only affects the initial segment immediately after the boundary, the articulatory magnitude of /i/ will not vary as a function of the boundary strength. Figure 12 shows the maximal linguopalatal contact for /i/ uttered by the two speakers.

No clear trend of cumulative boundary effect was found for both speakers. The linguopalatal contact at IP- or U-boundary position was consistently and significantly lower than that at ip-boundary position for the female speaker (Figure 12a) and at FT-boundary position for the male speaker (Figure 12b).

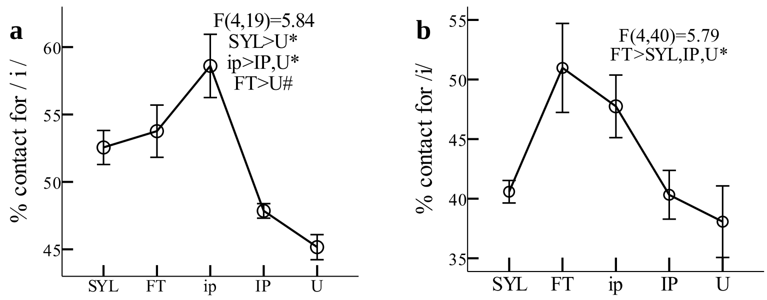

Figure 12 The maximal linguopalatal contact for /i/ in syllable /ti/ after five prosodic boundaries. (female speaker (a) and male speaker (b), * $p<0.05$, # $p<0.1$)

3.2.3 F0 and voice quality measures

Figure 13 shows the vowel-initial f0, OQ and SQ for vowels /a/ and /i/ in the prosodic domain-initial syllables. The vowel-initial f0 was consistently higher at the domain-initial position of the higher prosodic boundaries for both speakers (Female: $F(4,121)=27.95$, $p<0.0001$, SYL<FT, ip<U, SYL<IP; Male: $F(4,276)=30.24$, $p<0.0001$,SYL<FT, ip<IP, U). Different patterns were observed for the OQ of the two speakers. For the female speaker the OQ progressively increased in higher prosodic domains ($F(4,121)=13.72$, $p<0.0001$, SYL,FT<IP,U, SYL<ip), which supports the finding by Garellek (2014). But for the male speaker, the OQ tends to be consistent at different positions, except that significantly lower OQ was found at the U-initial position ($F(4,276)=12.48$, $p<0.0001$, U<IP, ip, FT, SYL). Decreased SQ in stronger prosodic domains was found for both speakers (Female: $F(4,121)=11.87$, $p<0.0001$, SYL, FT>IP, U, SYL>ip; Male: $F(4,276)=2.93$, $p<0.05$, SYL>U).

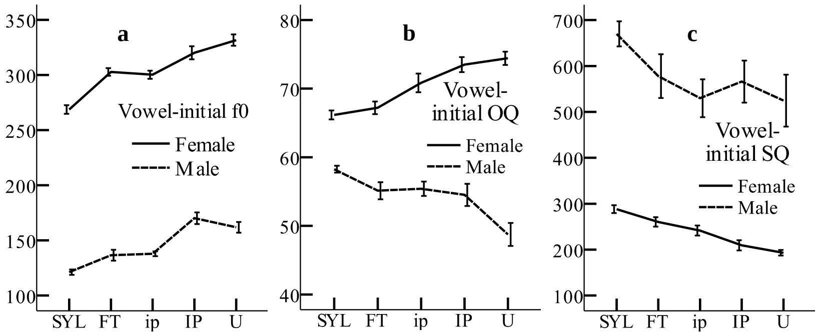

Figure 13 The vowel-initial f0 (a), OQ (b) and SQ (c) for the two speakers.

In summary, the linguopalatal contact and seal duration measures show a general trend in that the segment is strengthened in higher prosodic constituents. However, inter-personal and segmental positional factors are not negligible. As a matter of fact, inter-personal variability has been widely found in previous studies, which show different articulatory strategies adopted by speakers. The slightly different linguopalatal contact at the PMC and the release frames tend to indicate that the strengthening effect might gradually fade away when it is far away from the boundary. The domain-initial strengthening tends not to extend to the vocalic interval that immediately follows the domain-initial stop. The hierarchically structured vowel-initial f0 indicates that the f0 rest is functioning in marking the prosodic structure. The voice quality measures indicate that the vocal folds tend to abduct for a progressively longer portion of time in each abduction-adduction cycle at the edge of higher prosodic constituents. But this gradient variation as a function of boundary strength might be speaker-dependent. In addition, the adduction gesture becomes relatively slower in higher prosodic constituents.

3.3 Acoustic Measures

Table 3 shows the results of the ANOVA and the multiple comparisons for the acoustic measures of the alveolar stop. The acoustic measures at the SYL boundary in the /i/ context produced by the male speaker were excluded because the majority of tokens was either realized as voiced stop or approximant.

Table 3 ANOVA and multiple comparison (Bonferroni method, $p<0.05$) results for acoustic measures. Direction of the difference in the measures is indicated by "<" for less percentage or length.

Measure	Female speaker	Male speaker
Context a		
VOT	n.s.	n.s.
Voicing during stop closure	$F(4,67)=30.59$; $p<0.0001$ SYL, FT>ip, IP, U	$F(4,118)= 42.25$, $p<0.0001$ SYL, FT>ip, IP, U
Total voiceless interval	$F(4,67)=156.42$; $p<0.0001$ SYL, FT<ip<IP<U	$F(4,118)= 259.10$, $p<0.0001$ SYL, FT<ip<IP<U
RMS burst energy	n.s.	n.s.
Context i		
VOT	$F(4,19)=2.35$, $p=0.09$	$F(3,16)=4.28$, $p<0.05$ ip>U # (p=0.07)
Voicing during stop closure	$F(4,19)=21.85$, $p<0.0001$ SYL, FT>ip, IP, U	$F(3,16)= 65.39$, $p<0.0001$ FT>ip>IP, U
Total voiceless interval	$F(4,19)=108.96$, $p<0.0001$ SYL<ip<IP<U FT<IP	$F(3.16)=177.47$, $p<0.0001$ FT<ip<IP, U
RMS burst energy	$F(4,19)=4.19$, $p=0.01$ SYL, U<IP	$F(3,16)=10.63$, $p<0.001$ FT>U

Figure 14 shows the mean values of the VOT at post-boundary position in the two vocalic contexts. In the /a/ context no boundary effect was found, and the VOT stood around 0.015 second, regardless of the boundary types. The boundary types tended to influence the VOT in the /i/ context. For the female speaker the VOT appeared longer at SYL than other domains, but no significant difference was found across boundary types. For the male speaker the boundary effect was significant, and the VOT was marginally longer at the ip boundary than at the U boundary.

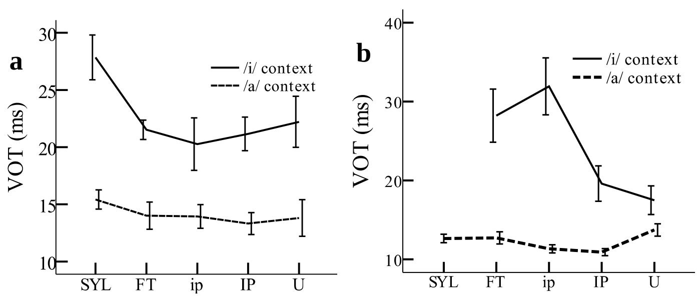

Figure 14 The VOT at post-boundary position for (a) female and (b) male speaker.

The other two acoustic measures that reflect the state of the vocal folds show a clearer cumulative effect of boundary strength across speakers and vocalic contexts. Table 3 and Figure 15 indicate that the voicing takes up the larger portion of the acoustic closure duration at the SYL and FT boundaries than at the three higher boundaries. Voicing during the closure interval was zero at the IP and U boundaries across vocalic contexts and speakers. And at the ip boundary in /a/ context the vocal folds did not vibrate in the acoustic closure duration for both speakers. The variation of the total voiceless interval shows a reverse pattern. It increased as the boundary strength became stronger.

The RMS burst energy shows no significant effect of the boundary strength in /a/ context. In /i/ context no systematic pattern emerges though some significant differences were found in the multiple comparisons.

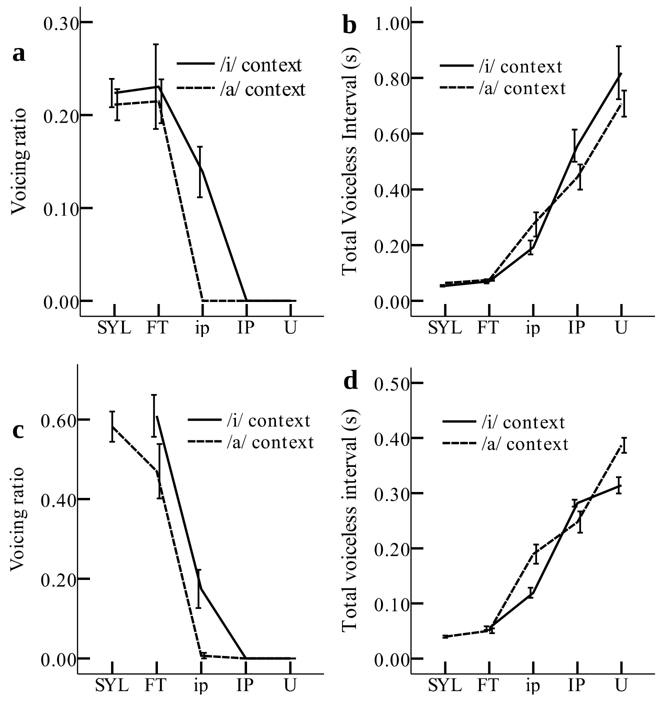

Figure 15 Voicing during the stop closure (a,c) and the total voiceless interval (b,d) produced by the two speakers. (female speaker: a,b; male speaker: c,d)

To summarize, the variation of the four acoustic measures of the alveolar stop shows different pictures under the conditioning of the prosodic structure. The VOT of the unaspirated alveolar stop tends not to be conditioned by the prosodic structure, though the VOT is slightly higher in lower prosodic domains in the /i/ context. The voicing during closure interval and the total voiceless interval are two reliable acoustic measures that show a cumulative strengthening of the post-boundary stop. With the boundary strength becoming stronger, the voicing ratio over the acoustic closure duration decreases, and the total voiceless interval increases. No clear and systematic pattern is found for the RMS burst energy of the stop.

4. DISCUSSION AND CONCLUSION

Two research questions are addressed in this paper. The first concerns whether the domain-initial strengthening is represented as a gradient variation in the articulatory and/or acoustic domains in Standard Chinese? The articulatory evidence shows that articulation for the post-boundary unaspirated alveolar stop varies as a function of boundary strength with increasingly larger linguopalatal contact and longer alveolar seal duration at higher prosodic domains. This cumulative effect on the articulatory magnitude is salient at PMC frame across speakers and vocalic environment. But it fades away toward the release moment, for no robust effect is found at the release frame. The weakening of alveolar closure gesture at the boundary of progressively lower prosodic constituents is accompanied by the reduction of the tongue blade contact on the post-alveolar area. The tongue tip gesture can also be weakened at the initial position of the SYL or FT domains, resulting in the approximant or voiced stop realization.

The vowel-initial f0 shows a clearly hierarchically-nested pattern, which reflects the global intonational conditioning. The voice quality measures tend to vary with the boundary strength. The vocal folds tend to be abducted in higher prosodic domains for the female speaker, but no clear trend is found for the male speaker. However, both speakers have progressively lower SQ in higher prosodic domains, which indicates that the vocal folds closure gesture becomes slower when the boundary strength is stronger. The results for voice quality supports Kong's finding that the

SQ is negatively correlated with the f0, and the result of the OQ produced by the female speaker complies with the recent finding of Garellek (2014).

In terms of acoustic properties of /t/, VOT is not a reliable candidate for marking the boundary strength. But a follow-up acoustic study on stops is worth carrying out because the VOT in the /i/ context shows a tendency to vary as a function of boundary strength. The voicing over the acoustic closure interval and the total voiceless interval are two reliable measures to mark the boundary strength. As the boundary becomes stronger, the former was progressively smaller and the latter progressively longer. These two acoustic measures may be important acoustic correlates for the perception of the boundary strength, which deserves a further perceptual study. The RMS burst energy is expected to show boundary effect with the value becoming smaller at a stronger boundary. However, no clear pattern emerges in both vocalic environments.

The second research question is what is the scope of the domain-initial strengthening effect? In this regard, the high front vowel /i/ in post-boundary syllable /ti/ is used to test the locality hypothesis. No clear trend is found for the linguopalatal contact for /i/. This result shows that the domain-initial strengthening effect is restricted on the temporal domain of the first segment immediately following the boundary, and the strengthening effect quickly fades away toward the release of the first segment. However, the vocalic interval is hierarchically organized by another prosodic device, the intonation. For both speakers, the starting f0 of the falling tone progressively decreases as the boundary strength becomes weaker. As for the cumulative variation of the two voice quality measures, they might be the by-product of changing f0 (Kong 2001) or f0 reset at major boundaries (Garellek 2014).

Though most results in the current paper complied with the previous study, there is one issue worth noting. The articulatory properties of the domain-initial /t/ produced by the two speakers are rather similar in most cases of U and IP. From Figure 11, we can conclude that this might be attributed to the ceiling effect, namely, tongue tip together with the tongue blade gesture are fully realized for the alveolar stop, provided the time threshold is surpassed. At this point, the total voiceless interval could serve as the most important acoustic cue to mark between U- and IP-boundary. The articulatory properties of /t/ at the ip-initial position tend to pattern with the higher domain in articulatory domains when the time threshold is guaranteed (see Table 2 and Figure 11); however, it is distinguished from the higher domains by the total voiceless interval in both vocalic environments. The articulatory and acoustic properties of /t/ at SYL and FT-initial positions are rather similar; however, a close look at the linguopalatal contact pattern shows a different picture. For the female speaker, the occurrence of incomplete alveolar closure gesture is more frequent in SYL domain compared with the FT domain. This indicates that SYL-initial alveolar gesture is more liable to undergo gestural reduction compared to FT-initial position.

To conclude, the results in the current paper confirm the previous findings in that the domain-initial strengthening is a universally salient effect that cues the prosodic hierarchy in languages. Regarding Standard Chinese, the gradient nature of segment articulation could serve as the important cue that marks the prosodic structure in this language.

NOTES

1. This work was supported in part by the National Natural Science Foundation of China (61073085), Humanity and Social Science Research Fund of the Ministry of Education of China (11YJAZH055) and Yanbian University Social Sciences Research Startup Fund. Professor Kong Jiangping read the draft and offered his insights. Part of the results of this paper was orally presented at the 17th Phonetic Science Congress, 2011, Hong Kong.

REFERENCES

BOERSMA, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9/10):341-345.

BOMBIEN, L., C. Mooshammer, P. Hoole, and B. Kuhnert. 2010. Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics* 38:388-403.

BYRD, D. 2000. Articulatory vowel lengthening and coordination at phrasal juncture. *Phonetica* 57:3-16.

BYRD, D. and E. Saltzman. 1998. Intra-gestural dynamics of multiple prosodic boundaries. *Journal of Phonetics* 26:173-199.

_____. 2003. The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics* 31:149-180.

BYRD, D., A. Kaun, S. Narayanan, and E. Saltzman. 2000. Phrasal signatures in articulation. In *Acquisition and the Lexico,* Papers in Laboratory Phonology V, ed. M. B. Broe and J. B. Pierrehumbert, 70-87. Cambridge: Cambridge University Press.

BYRD, D., J. Krivokapic, and S. Lee. 2006. How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America* 120:1589-1599.

CAO, Jianfen 曹劍芬. 2004. Restudy of segmental lengthening in Mandarin Chinese. Paper presented at the Second International Conference on Speech Prosody, Nara, Japan, in its proceedings ed. B. Bel and I. Marlien, 231-234. http://isca-speech.org/archive/sp2004/papers/sp04_231.pdf, (accessed 2/10/2015).

_____. 2005. Yinduan yanchang de butong leixing jiqi yunlu jiazhi 音段延長的不同類型及其韻律價值 (Different types of segmental lengthening and their prosodic value). *Nanjing Shifan Daxue Wenxueyuan Xuebao* 南京師範大學文學院學報 4:160-167.

CAO, Jianfen and ZHENG, Yuling. 2006. Articulatory strengthening and prosodic hierarchy. In *Proceedings of the Third International Conference on Speech Prosody*, ed. R. Hoffmann and H. Mixdorff, 289-292. Dresden: Technische Universität Darmstadt Press.

CHO, T. 2001. Effects of prosody on articulation in English. PhD diss., University of California, Los Angeles.

_____. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32:141-176.

_____. 2005. Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /ɑ, i/ in English. *Journal of the Acoustical Society of America* 117:3867-3878.

CHO, T. and P. Keating. 2001. Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics* 29:155-190.

_____. 2009. Effects of initial position versus prominence in English. *Journal of Phonetics* 37:466-485.

CHO, T. and J. M. McQueen. 2005. Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics* 33:121-157.

FARNETANI, E. and M. Vayra. 1996. The role of prosody in the shaping of articulation in Italian CV syllables. Paper presented at the First ETRW (ESCA Tutorial and Research Workshop) on Speech Production Modelling, Autrans, France, in its proceedings, ed. ESCA, 9-12. http://www.isca-speech.org/archive/spm_96/papers/sps6_009.pdf, (accessed 2/10/2015).

FOUGERON, C. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29:109-135.

FOUGERON, C. and P. Keating. 1997. Articulatory strengthening at the edge of prosodic domains. *Journal of the Acoustical Society of America* 101:3728-3740.

GARELLEK, M. 2014. Voice quality strengthening and glottalization. *Journal of Phonetics* 45:106-113.

HAYASHI, W., C. Hsu, and P. Keating. 1999. Domain-initial strengthening in Taiwanese: A follow-up study. UCLA Working Paper in Phonetics 97:152-156.

HOOLE, P. and F. Hu. 2004. Tone-vowel interaction in Standard Chinese. Paper presented at the International Symposium on Tonal Aspects of Languages, Beijing, China, in its proceedings ed. B. Bel and I. Marlien, 89-92. http://www.isca-speech.org/archive/tal2004/papers/tal4_089.pdf, (accessed 2/10/2015).

HOWARD, D. M. 1995. Variation of electrolaryngographically derived closed quotient for trained and untrained adult female singers. *Journal of Voice* 9:163-172.

HSU, C. and S. Jun. 1998. Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic? UCLA Working Papers in Phonetics 96:69-89.

HU, W. X., B. Xu, and T.Y. Huang. 2002. An experimental study on prosodic boundary in Chinese Mandarin. *Journal of Chinese Information Processing* 16(1):43-48.

JUN, S. 1993. The phonetics and phonology of Korean prosody. PhD diss., Ohio State University.

KEATING, P., T. Cho, C. Fougeron, and C. Hsu. 2003. Domain-initial articulatory strengthening in four languages. In *Phonetic Interpretation. Papers in Laboratory Phonology VI*, ed. J. Local, R. Ogden, R. Temple, 145-163. Cambridge: Cambridge University Press.

KIM, S. 2001. The interaction between prosodic domain and segmental properties: Domain initial strengthening of fricatives and post obstruent tensing rule in Korean. M.A. thesis, University of California, Los Angeles.

KIM, S. and T. Cho. 2011. Effects of prosodic boundary versus accent in the articulation of English /æ/ in #VC and #CVC. Paper presented at the 17th International Congress of Phonetic Sciences, Hong Kong, China, in its proceedings ed. W. S. Lee and E. Zee, 1082-1085.

KONG, Jiangping 孔江平. 2001. Lun yuyan fasheng 論語言發聲 (On language phonation). Beijing: Press of Minzu University of China.

KUZLA, C. and M. Ernestus. 2011. Prosodic conditioning of phonetic details in German plosives. *Journal of Phonetics* 39:143-155.

LI, Aijun 李愛軍. 2002. Putonghua duihua zhong yunlu tezheng de shengxue biaoxian 普通話對話中韻律特徵的聲學表現 (Prosodic analysis on conversations in Standard Chinese). *Zhongguo Yuwen* 中国语文 6:525-535.

_____. 2002. Chinese prosody and prosodic labelling of spontaneous speech. Paper presented at the First International Conference on Speech Prosody, Aix-en-Provence, France, in its proceedings ed. B. Bel and I. Marlien, 39-46. http://www.isca-speech.org/archive/sp2002/papers/sp02_039.pdf, (accessed 2/10/2015).

LI, Yinghao and Jiangping Kong. 2011. Prosodic boundary effects on segment articulation and V-to-V coarticulation in Standard Chinese. Paper presented at the 17th International Congress of Phonetic Sciences, Hong Kong, China, in its proceedings ed. W. S. Lee and E. Zee, 1218-1221.

LI, Yinghao and Xiaosheng Pan. 2012. The temporal alignment algorithm for electropalatographic and speech signals in long utterance. In *Proceedings of International Conference on Audio, Language and*

Image Processing, Volume 2, ed. W. G. Wan, F. L. Luo, and X. Q. Yu, 957-960. Piscataway: IEEE.

LIN, Maocan 林茂灿. 2002. Putonghua yuju de yunlu jiegou he jipin(F0) gaodixian goujian 普通話語句的韻律結構和基頻(F0)高低線構建 (The prosodic structure in Putonghua and the construction of high- and low-end F0 contour). *Dangdai Yuyanxue* 當代語言學 4:254-265.

LINDBLOM, B. 1990. Explaining phonetic variation: A sketch of the H & H theory. In *Speech Production and Speech Modeling*, ed. W. J. Hardcastle and A. Marchal, 403-440. Dordrecht: Kluwer Academic Publishers.

QIAN, Yao 錢瑶, Min Chu 初敏, and Wuyun Pan 潘悟雲. 2001. Putonghua yunlv danyuan bianjie de shengxue fenxi 普通話韻律單元邊界的聲學分析(The acoustic correlates of the prosodic boundaries in Standard Chinese). In *Xinshiji de xiandai yuyinxue*: *diwujie quanguo xiandai yuyinxue xueshu huiyi* 新世紀的現代語音學: 第五屆全國現代語音學學術會議 (The Phonetic Research in the New Century: Proceedings of the 5th National Conference on Modern Phonetics), ed. Lianhong Cai 蔡蓮紅, Tongchun Zhou 周同春,and Jianhua Tao 陶建華, 70-75. Beijing: Qinghua daxue chubanshe.

SELKIRK, E.O. 1984. *Phonology and Syntax: the Relationship between Sound and Structure*. Cambridge: MIT Press.

TSENG, C., S. Pin, Y. Lee, H. Wang, and Y. Chen. 2005. Fluent speech prosody: Framework and modeling. *Speech Communication* 46(3-4):284-309.

WANG, B., Y. F. Yang, and S. N. Lu. 2004. Acoustic analysis on prosodic hierarchical boundaries of Chinese. *Acta Acoustica* 29(1):29-36.

WANG, Hongjun 王洪君. 2008. Hanyu feixianxing yinxixue 漢語非線性音系學 (*Non-linear Phonology of Mandarin Chinese*). Beijing: Beijing Daxue Chubanshe.

WIGHTMAN, C. W., S. Shattuck-Hufnagel, M. Ostendorf, and P. J. Price. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91(3):1707-1717.

韻律邊界對漢語普通話音段發音的影響—基於發音生理和聲學的研究

李英浩*
延邊大學，延吉（中國）

提要
本文使用動態電子齶位元(EPG)和聲學分析的方法，考察漢語普通話韻律邊界對韻律單元域首音段的發音生理和聲學特徵的影響。我們選取普通話的清不送氣齒齦塞音/t/和前高母音/i/，從音段產生的空間域和時間域分析域首發音增強現象。普通話的韻律層級包括音節、音步、小韻律短語、大韻律短語和話語。通過比較不同韻律層次域首音段的發音生理和聲學參數，我們發現：（1）單元域首輔音的發音動作受到普通話韻律結構的制約。輔音的舌齶接觸和生理持阻時長與韻律邊界的強度密切相關；輔音的舌齶接觸與生理持阻時長之間呈現出非線性關係；輔音聲學時段的清聲段時長和濁聲時長比能夠有效地標記韻律邊界的強度。（2）輔音除阻時刻的舌齶接觸以及輔音後接母音的最大舌齶接觸受邊界強度的影響較小，且後接母音的嗓音特徵與較大韻律邊界的基頻重設有關。這說明域首發音增強的作用域限於韻律邊界後面的音段。研究結果表明，普通話韻律單元域首音段的發音特徵得到增強，且增強程度與邊界強度密切相關，這是標示普通話韻律結構的一個重要方式。

主題詞
韻律邊界　**音**段發音　**動**態電子齶位元　**漢**語普通話

* **李英浩**為延邊大學副教授，執教于延邊大學外國語學院英語系；通訊：中國吉林省延吉市，延邊大學外國語學院英語系，郵編 133002; [youngho_lee@aliyun.com]