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

Authors: Yinghao Li

Date: 2019-06-20T00:00:00+00:00

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 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.

Full Text

Preamble

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 boundary effects on domain-initial segments in Standard Chinese. Two speech sounds—the voiceless unaspirated alveolar stop /t/ and the high front vowel /i/—were examined to investigate domain-initial strengthening in both spatial and temporal dimensions. Articulatory and acoustic parameters were compared across five prosodic constituents: Syllable, Foot, Immediate Phrase, Intonational Phrase, and Utterance. Results show that: (1) The production of domain-initial consonantal gestures is prosodically encoded. Linguopalatal contact and seal duration vary as a function of prosodic boundary strength, with linguopalatal contact dependent on seal duration in a nonlinear fashion. Among the acoustic properties of domain-initial stops, the total voiceless interval and voicing during closure serve as reliable acoustic correlates marking the hierarchical structure of prosody. (2) At the release moment of domain-initial stops, no consistent pattern supports domain-initial strengthening. The linguopalatal contact of vowels immediately following domain-initial consonants shows no clear strengthening trend; however, phonatory features of vowels indicate pitch reset at major prosodic boundaries. These findings suggest that domain-initial strengthening is restricted to the segment immediately following the boundary. In conclusion, Standard Chinese strengthens the phonetic features of domain-initial segments as a function of boundary strength, which serves as an important mechanism for marking prosodic structure.

Subject Keywords: Prosodic boundaries, Segment articulation, Electropalatography, Standard Chinese

1. Introduction

The production of phonological units in utterances is conditioned by prosodic structure. One way this structural information is encoded is through how phonological units are produced at the edges of prosodic constituents. Recent studies have observed that segments in prosodic domain-initial positions are produced more strongly than in domain-medial positions. This boundary-induced articulatory variation for individual speech segments has become an important topic known as domain-initial strengthening, which refers to the greater magnitude of phonetic realization in articulatory or acoustic dimensions 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 across numerous languages, including 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).

Previous studies have demonstrated that prosodic shaping of segmental features exhibits gradient variation in articulatory and/or acoustic dimensions. In electropalatographic studies on consonants across languages, peak linguopalatal contact and articulatory seal duration have been shown to increase progressively 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 voice onset time (VOT) for Korean aspirated stops became progressively shorter as boundary strength decreased. Conversely, Kuzla and Ernestus (2011) found the opposite pattern for prosodic positional effects on VOT of fortis plosives in German. These gradient variations as a function of domain strength have been interpreted within the duration-dependent undershoot model proposed by Lindblom (1990), which posits that the time available for producing a segment determines how completely its phonetic target is realized. Segments are likely hypoarticulated in weaker positions due to insufficient production time.

The linguistic motivation for this articulatory adjustment is hypothesized to involve enhanced syntagmatic contrasts with neighboring segments through magnifying their associated features (Fougeron and Keating 1997; Hsu and Jun 1998; Cho 2005). The prosodic strengthening effect is also temporally constrained, tending to manifest only at the first post-boundary segment, while subsequent segments within the same syllable are seldom affected (Fougeron and Keating 1997; Cho 2005; Byrd, Krivokapic, and Lee 2006). In an electropalatographic study, Fougeron and Keating (1997) found no consistent or reliable pattern of linguopalatal contact for the domain-initial consonant /n/ compared to non-initial positions. Cho (2005) examined boundary-induced articulatory and acoustic variation of post-boundary vowels in domain-initial CV syllables in English, finding no consistent tongue fronting or raising for high front vowel /i/, nor tongue lowering and backing for low back vowel / /. Byrd, Krivokapic, and Lee (2006) investigated the temporal scope of prosodic boundary effects on segment articulation, finding that post-boundary consonant gestures were significantly strengthened with longer duration and larger articulatory displacement, while following consonants were temporally shortened with smaller displacement.

However, some studies have found evidence of domain-initial strengthening effects on non-initial vowels in post-boundary CV syllables (Farnetani and Vayra 1996; Cho and Keating 2009; Kim and Cho 2011). Farnetani and Vayra' s (1996) electropalatographic study on prosodic effects in CV syllable production found that initial syllables were strengthened overall, with vowels in domain-initial syllables showing more open vocal tracts regardless of lexical accent. Cho and Keating (2009) found mixed results for the locality hypothesis in their electropalatographic and acoustic study on English segment production: articulatory parameters for vowels in post-boundary CV syllables were insensitive to boundary strength, yet vowel amplitude showed strengthening at utterance-initial positions compared to lower boundaries. Kim and Cho (2011) found in their electromagnetic articulograph (EMA) study that vowels following /h, ph/ at prosodic boundary initial positions showed gradient tongue movement magnitude affected by boundary strength, similar to domain-initial vowels in syllables. These findings suggest that domain-initial strengthening may interact with other confounding factors affecting post-boundary initial syllables.

In a simulation study, Byrd and Saltzman (2003) demonstrated that the speech production mechanism controlling articulators receives increasingly more time when approaching a prosodic boundary, with gestural actualization time decreasing as the boundary recedes. This suggests that articulators can fully realize phonetic targets when provided sufficient gestural preparation time, with this timing mechanism accelerating immediately after the boundary.

If domain-initial strengthening affects only the first segment after prosodic boundaries, this implies that vowels in post-boundary syllables are shaped by another prosodic mechanism rather than domain-initial strengthening. In tonal languages like Chinese, syllables carry lexical tones that contrast meanings. Previous studies have shown that Mandarin syllable tones are hierarchically determined, with tone specification encoded according to boundary strength (Tseng et al. 2005). This hierarchically structured tonal specification manifests as f0 reset, another prosodic means of marking boundaries in Standard Chinese. Recent studies have also found that voice quality may be affected by prosodic structure, with electroglottographic (EGG) studies showing that post-boundary vowel quality becomes progressively breathier at higher prosodic boundaries. This gradient variation in voice quality measures has been assumed to relate to f0 reset at major boundaries (Garellek 2014).

Previous studies demonstrate that domain-initial strengthening effects depend on segmental identity and language. Regarding segmental identity, alveolar fricatives appear more resistant to articulatory variation across prosodic positions. For instance, Fougeron (2001) found that French alveolar fricatives showed less prosodically-conditioned articulatory variation compared to alveolar stops. Similar results were found in Korean, where domain-initial strengthening effects on articulatory magnitude were less salient for fricatives than stops (Kim 2001). This segment-specific response is explained by the aerodynamic requirements of fricative production, resulting in more language-specific manifestations of domain-initial strengthening.

Cho and Keating (2001, 2009) found that VOT for Korean lax and aspirated alveolar stops /t, th/ and English /t/ increased at higher prosodic boundaries. However, VOT variations showed the opposite pattern for Dutch and German

/t/, with shorter VOT at stronger boundaries (Cho and McQueen 2005; Kuzla and Ernestus 2011). These positional VOT differences involve different laryngeal gestures across languages, suggesting that prosodic signatures on segment articulation may be represented through different phonetic dimensions. In an articulatory study on Tamil, prosodic structure affected duration and timing relationships in consonant clusters but not consonant articulatory magnitude (Byrd et al. 2000). In short, segment- and language-specific prosodic signatures on articulation are attributed to fine-grained articulatory control and cross-language differences in segment production.

The current research investigates how prosodic structure affects individual speech segments at prosodic domain-initial positions in Standard Chinese. By prosodic structure, we mean that spoken utterances are hierarchically organized, with higher prosodic domains decomposing into immediately lower constituents. Following the Strict Layer Hypothesis (Selkirk 1984), utterances are shaped by hierarchical prosodic structure wherein higher domains directly dominate one or more immediately lower domains, and each prosodic domain must be contained within an immediately higher domain.

The prosodic model for Standard Chinese in this paper follows Li (2002) and Lin (2002) with minor modifications. Based on autosegmental-metrical theory, Li (2002) and Lin (2002) assert that Standard Chinese speech utterances are hierarchically organized, with larger prosodic units composed of several immediately lower constituents. These include syllable, foot, prosodic word, minor phrase, major phrase, and utterance, distinguished by pitch contour and break. For the two prosodic phrases, we use the terms "intermediate phrase" and "intonational phrase" to facilitate comparison with other languages. The foot domain is normally syntactically defined and constitutes the basis for the higher prosodic domain (prosodic word), which comprises a foot and/or a following unassigned monosyllable (Li 2002, 526). This paper focuses on the bi-syllabic foot domain as it represents the basic unit of metrical organization in Standard Chinese (Wang 2008).

The intermediate phrase domain comprises one or several prosodic words (here, foot serves as the immediately subordinate component) and is characterized by phrasal accent and noticeable pause (silent or filled). The intonational phrase domain is cued by longer domain-ending pauses and noticeable pitch contour resetting. The utterance, constituted by one or several major phrases, is marked by sentence accent, substantial durational compression of domain-ending syllables, and the longest pause (Wang 2008, 257). Figure 1 [Figure 1: see original paper] illustrates the prosodic hierarchy used in this study.

Extensive research on Standard Chinese prosodic structure has shown that f0 reset, syllable durational patterns, and pauses within utterances are the three main acoustic correlates cueing prosodic structure (Li 2002; Wang, Yang, and Chen 2004; Hu, Xu, and Huang 2002; Lin 2002). However, the role of domain-initial strengthening in marking Chinese prosodic structure has received little attention. In her acoustic study on segmental lengthening, Cao (2005, 165) suggested: "Domain-initial segmental lengthening likewise functions to mark boundary strength, and this function cannot be underestimated. It is more direct and reliable for indicating prosodic boundary strength···meanwhile, post-boundary consonantal duration is positively correlated with perceived boundary strength, progressively increasing as a function of boundary strength."

Keating et al. (2003, 161) hypothesized that lexical tone languages such as Taiwan Hokkien might show more salient domain-initial strengthening effects than English because they "should have less recourse to pitch to mark domain edges." Although their results did not support this hypothesis, the cumulative effect they observed demonstrated the universality of prosodic conditioning on domain-initial segments. Wang (2008) argued that consonant initials inside foot domains—the basic metrical template for rhythmic organization in Standard Chinese—were reduced compared to those heading foot domains, yet no gestural reduction was found at prosodic constituent initial positions above the foot level. This argument directly supports domain-initial strengthening principles.

Cao and Zheng (2006) found in their electropalatographic study that phrase-initial consonants showed greater linguopalatal contact than phrase-medial consonants in Standard Chinese. Li and Kong (2011) investigated articulatory strengthening for domain-initial alveolar stop /t/ in Standard Chinese, finding cumulative increases in linguopalatal contact and alveolar seal duration from syllable to utterance boundaries based on data from one female speaker. This paper extends that research by examining the same consonant and the tautosyllabic high front vowel /i/ from two speakers.

The central research question is how prosodic position affects the production of consonantal and vocalic gestures in CV syllables in Standard Chinese. Two aspects of domain-initial strengthening are examined: (1) Does prosody affect the articulatory and acoustic properties of domain-initial segments in a cumulative pattern? (2) Does domain-initial strengthening affect all segments in the post-boundary syllable?

The universality of domain-initial strengthening leads us to predict that domain-initial segments are strengthened with a cumulative scale indicative of linguistically intrinsic biomechanical encoding. Previous single-speaker studies indeed show strong tendencies for such prosodically-conditioned articulation, with linguopalatal contact for unaspirated /t/ positively correlating with prosodic hierarchy. Our second prediction concerns the temporal scope of domain-initial strengthening: we hypothesize that the effect is limited to the first segment immediately following the boundary, with the following vocalic segment not subject to domain-initial effects. The strengthening effect should quickly fade toward the end of initial segment production, as predicted by Byrd and Saltzman (2003).

2.1 Electropalatography (EPG)

Tongue-palate contact for lingual consonants, as captured by electropalatography, provides a reliable indicator of articulatory magnitude, where greater linguopalatal contact indicates stronger oral constriction and greater articulatory magnitude. This study used the WinEPG Electropalatography system produced by Articulate Instruments, which employs custom-made pseudo-palates with a thin acrylic base covering the palate from the upper teeth root to the anterior portion of the soft palate. The layout of 62 electrodes is designed relative to anatomical landmarks to enable inter-subject comparison. When the tongue contacts an electrode, the circuit completes 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 across five prosodic domains: syllable (SYL), foot (FT), intermediate phrase (ip), intonational phrase (IP), and utterance (U), within two symmetrical vocalic environments (/a/ and /i/). The high front vowel /i/ in syllable /ti/ served as the other test segment for investigating the temporal scope of domain-initial strengthening. This syllable was preceded by the five prosodic domains immediately before the initial stop and followed by syllables starting with alveolar stop /t/. To avoid confounding effects, the low tone (T3) was avoided as much as possible because it could affect linguopalatal contact of /i/ (Hoole and Hu 2004). The falling tone was used in most cases to examine vowel-initial f0 and voice quality under varying boundary conditions.

Table 1 shows sample stimuli from the sentence set for test consonant /t/. The utterance domain was elicited by a full period, with speakers instructed to make a long pause afterward. Commas parsed intonational phrase domains, with syllable counts in the two sentence components ranging from 6 to 12 syllables; speakers were instructed to ignore these commas. The intermediate phrase domain consisted of noun phrases comprising at least two feet. The foot domain was the second word inside an immediate phrase. The syllable domain was defined as boundaries occurring within a foot. The five-level break index based on C-ToBI (Li 2002) coded prosodic domains: break index 4 (B4) transcribed the Utterance domain, break index 0 (B0) the Syllable domain, and intermediate numbers transcribed intermediate domains.

2.3 Speakers and Procedures

Two university students (one male, one female) participated in the experiment. Both had lived in northern Chinese provinces or municipalities before university and worked as part-time announcers at the university TV station. At recording, both were 27 years old and received payment for participation.

Recording took place in a sound-attenuated booth at Peking University. Before recording, participants had 30-50 minutes to adapt to reading with a pseudo-palate installed and to familiarize themselves with the sentences. During recording, sentences were randomized in blocks, with each sentence repeated three times for the female speaker and five times for the male speaker. Sentence blocks appeared on a computer screen positioned about one meter away, and participants read at a normal speech rate. No specific prosodic phrasing instructions were given except to pause intentionally after orthographic periods. Electropalatographic, electroglottographic, and speech signals were recorded simultaneously (see Figure 2 [Figure 2: see original paper] for an example). The EPG signal sampling rate was 100 Hz, while speech and EGG signals were sampled at 22 kHz. Respiratory signals for chest and stomach breathing were also recorded for the male speaker but are not analyzed here. After recording, sentences with unclear test segment pronunciation or signal problems were eliminated, leaving 264 sentences for analysis.

Articulatory and acoustic analysis was performed using a Matlab program developed to process electropalatographic and speech signals. First, the electropalatographic signal was temporally aligned with the acoustic signal using the algorithm from Li and Pan (2012). Each utterance was then parsed from the recording and annotated in PRAAT (Boersma 2001) on two tiers: syllable and break index. Figure 3 [Figure 3: see original paper] demonstrates the break index tier used in this paper.

2.4 Articulatory Measurement

The articulatory gestures of /t/ and /i/ exhibit distinct tongue-palate contact patterns. In Figure 4 [Figure 4: see original paper], the first four rows are defined as the Front Region, closely related to alveolar stop closure formation. Linguopalatal contact in the back four rows relates to coarticulatory effects of the following vowel on the preceding consonant, which was not examined in this study.

To study boundary effects on post-boundary alveolar stops, two key time points for tongue-to-palate contact were defined: the point of maximum contact frame (PMC) and the release frame. The PMC was identified as the frame with maximal linguopalatal contact during the alveolar closure interval (when complete alveolar closure was observed) or acoustic closure interval (when no complete closure was found). The contact pattern at PMC directly reflects articulatory excursion magnitude for the segment (Cho and Keating 2009). The release frame was defined as the last frame of the alveolar stop closure interval; if no alveolar closure was observed in the acoustic closure interval (common in FT or SYL domains), the release frame was the frame immediately before the acoustic stop release. When alveolar stops were realized as voiced approximants (see Cho and Keating 2001), no release frame was available. As Cho and Keating (2009, 470) argue, the release frame "might reveal additional information about the prosodically-conditioned articulatory variation."

For both key frames, the percentage of contacted electrodes in the Front Region was computed, as this region relates to the tongue tip/blade gesture. Seal duration (SD) was measured as the interval between the first and last frames of alveolar closure; if no closure frames were found or the alveolar stop was realized as a voiced approximant, seal duration was zero.

To test the locality hypothesis of domain-initial strengthening, we investigated linguopalatal contact of the high front vowel /i/ in domain-initial syllable /ti/. Previous studies showed /i/ to be resistant to coarticulatory effects regardless of boundary strength (Cho 2004). If boundary effects are restricted to the initial position of post-boundary syllables, vowel /i/ in /ti/ should show no strengthening. To reduce confounding domain-final strengthening/lengthening effects, syllable /ti/ was designed to be followed by a Syllable boundary (though Syllable-initial tokens were followed by Foot boundaries due to the use of bi-syllabic feet in test sentence construction). The maximal linguopalatal contact frame for /i/ was selected between one-third and one-half of the vocalic interval to maximally reduce coarticulatory effects from adjacent segments. The percentage of contacted electrodes across the whole region was computed to measure /i/' s articulatory magnitude.

In addition to linguopalatal measures for vowel /i/, we examined vowel-initial f0, open quotient (OQ), and speed quotient (SQ) for vowels /i/ and /a/, as these tend to be conditioned by prosodic structure. As noted in the Introduction, f0 manifests f0 reset in Standard Chinese. OQ and SQ reflect vowel voice quality: OQ is the inverse of contact quotient, showing the portion of time vocal folds are open in each glottal period (higher OQ indicates larger glottal opening), while SQ computes the ratio between opening and closing phases in each glottal period from the EGG signal (higher SQ indicates faster vocal fold adduction). Recent research on voice quality strengthening in English and Spanish indicated that vowels after domain-initial glottal stops showed lower contact quotient (higher OQ) at higher prosodic boundaries (Garellek 2014). In this paper, f0 was computed from the derivative of the EGG signal, and OQ and SQ were obtained using the Hybrid method (Howard 1995). Figure 5 [Figure 5: see original paper] shows the definition of critical moments and intervals, along with three equations for obtaining these measures.

2.5 Acoustic Measurement

First, syllable durational properties were measured to verify appropriate utterance production for this study. A key acoustic parameter marking prosodic structure is pre-boundary vocalic duration (V1 duration), which demonstrates final lengthening—the durational expansion of rhymes in pre-boundary syllables (Wightman et al. 1992). Here, V1 duration was defined by observable F2 formant trajectory in the pre-boundary vowel. Another acoustic parameter indexing prosodic structure is acoustic closure interval, measured from the termination point of F2 contour in the previous vowel to the stop burst of the following consonant /t/. Acoustic closure duration was not measured at utterance-initial position because most of the silent interval was not due to alveolar closure gesture. These two acoustic measurements were compared with previous results on boundary-induced lengthening.

Acoustic properties of test segments were investigated because if domain-initial articulatory strengthening manifests acoustically, listeners might use those attributes for phrasing utterances, enabling future perceptual relevance studies. Acoustic measures included: (1) Voice onset time (VOT) of the voiceless unaspirated alveolar stop, measured from stop release to voice onset in the following vowel (signaled by F2 trajectory onset in the spectrogram). VOT was zero for approximant realizations or voiced stops. (2) Voicing during stop closure and total voiceless interval, computed to partially demonstrate vocal fold state during oral closure (Cho and Keating 2001). Voicing during closure was the percentage of acoustic closure duration showing voicing, represented by low-frequency voicing bars in the spectrogram or cyclical abduction-adduction in the EGG signal. Total voiceless interval was the duration of silent interval in acoustic closure plus VOT. These measures were not taken at utterance-initial position. (3) RMS burst energy, defined following Cho and Keating (2001) as acoustic energy at burst calculated from the RMS value of frequencies above 500 Hz in an FFT spectrum (low frequencies eliminated to reduce possible voicing effects over stop release).

3.1 The Prosodic Hierarchy Indicated by the Utterances

To ensure speech materials were appropriate for analysis, two durational measures were collected and compared with previous studies on acoustic correlates of Standard Chinese prosodic structure.

Although pre-pausal lengthening is well-accepted for indicating prosodic boundaries, its status in Standard Chinese remains unclear. Cao (2004, 2005) argued that pre-boundary rhyme lengthening conditioned by prosodic hierarchy mainly occurred at prosodic phrase levels but not at sentence and paragraph levels. Qian, Chu, and Pan (2001) similarly found pre-boundary syllable lengthening only at right edges of immediate and intonational phrases, with unstable effects at prosodic word level. Wang, Yang, and Chen (2004) found temporal expansion salient at lower prosodic boundaries (prosodic word and phrase), though its function in marking boundary strength became less important in higher domains where pause and f0 reset were more salient, even while cumulative final lengthening remained observable.

Figure 6 [Figure 6: see original paper] shows average pre-boundary V1 duration for both speakers. High agreement was achieved regarding syllable durational patterns domain-finally. For both speakers, V1 duration was significantly longer at domain-final positions in higher prosodic domains (ip, IP, and U) than in lower ones (SYL, FT). Meanwhile, duration was shorter at FT-final than SYL-final positions, with this difference reaching significance for the male speaker. For the female speaker, V1 duration was shorter at U-final than IP- or ip-final positions, though not significantly. These results generally confirm findings by Cao (2004, 2005) and Qian, Chu, and Pan (2000), showing that rhymes at ip- or IP-final positions tend to lengthen compared to SYL- or FT-final positions, without cumulative final lengthening.

Figure 7 [Figure 7: see original paper] shows acoustic closure interval across four prosodic boundaries. As predicted, acoustic closure interval varied with prosodic boundary strength. Both speakers produced longer closure intervals at higher prosodic boundaries, with the female speaker showing much longer intervals at IP- and ip-initial positions, attributed to her slower speech rate during stimulus production.

3.2.1 Linguopalatal Contact and Seal Duration

One-way ANOVA was conducted separately for articulatory measures from both speakers. Table 2 shows ANOVA results and post-hoc comparisons for articulatory measures in both vocalic contexts. Figure 8 shows means with standard error for linguopalatal contact at both frames.

Overall, linguopalatal contact at PMC and release frames, and alveolar seal duration, were larger and longer at higher prosodic constituents than lower ones. At the PMC frame, a clear cumulative increase in linguopalatal contact as a function of boundary strength was found in both vocalic contexts for the female speaker, who distinguished two boundary types in /a/ context (lower boundaries SYL and FT vs. higher boundaries ip, IP, and U) and four of five boundaries in /i/ context. For the male speaker, boundary strength indicated by PMC linguopalatal contact fell into two categories: lower boundaries (SYL and FT) and higher boundaries (ip, IP, and U). Figure 8(b) shows that linguopalatal contact at SYL-initial position was slightly larger than at FT-initial position in /i/ context, with peak contact at IP-initial being modestly largest among higher boundaries.

Linguopalatal contact patterns at the release frame were similar to PMC patterns except in /a/ context for the female speaker. For her, linguopalatal contact varied with boundary strength in /i/ context, but no such relation appeared in /a/ context, where significantly lower contact occurred at U-initial compared to ip-initial positions. For the male speaker, SYL and FT boundaries were distinguished from higher ip and IP domains in /i/ context, while /a/ context patterns matched PMC frame results. Linguopalatal contact at the release frame may also suggest boundary effects on segment articulation, though these effects might fade toward plosive release.

Figure 9 [Figure 9: see original paper] shows sample tokens of linguopalatal contact for /t/ at five domain-initial positions from the PMC frame. Both rows clearly show that Front Region linguopalatal contact, closely related to alveolar closure gesture, decreases as boundary strength weakens, regardless of vocalic environment. The rightmost frame in the third row shows incomplete alveolar closure at SYL-initial position. Careful examination reveals no such tokens at ip-, IP-, and U-initial positions. For the female speaker, 18% of SYL-initial tokens showed incomplete alveolar closure, while none appeared at FT-initial position. For the male speaker, incomplete alveolar closure occurred in 21% and 26% of tokens at SYL- and FT-initial positions, respectively.

When measuring alveolar seal duration, the starting point for alveolar closure could not be determined for most tokens at IP- or U-initial positions in /i/ context produced by the female speaker, because full contact frames without linguistic meaning preceded linguistically meaningful tongue-palate contact in the alveolar region. Therefore, seal duration for /t/ in /i/ context at U- or IP-initial positions was excluded from further analysis. Table 2 and Figure 10 [Figure 10: see original paper] show alveolar seal duration for domain-initial alveolar stops. For the female speaker, alveolar closure duration increased with boundary strength in both vocalic contexts, with three distinctions identified in /a/ context and all three boundaries distinguished in /i/ context (shortest at SYL-initial, longest at ip-initial). For the male speaker, only two boundary types were distinguished: alveolar closure duration at SYL- and FT-initial positions was significantly shorter than at higher prosodic domain-initial positions.

Articulatory magnitude of consonants at different domain-initial positions may be duration-dependent: given sufficient time, consonantal gestures can be fully realized. In their EPG study of Korean consonants, Cho and Keating (2001) found that the relationship between linguopalatal contact and seal duration was asymptotic rather than linear as contact increased. Figure 11 [Figure 11: see original paper] shows scatter plots of Front Region linguopalatal contact against seal duration with curve fitting functions accounting for large variance portions. As indicated in Figures 11(a, c, d), polynomial fits reveal nonlinear relationships between linguopalatal contact and alveolar closure duration. A special case occurred in /a/ context for the female speaker, showing an exponential relationship. Figure 11(b) reveals that linguopalatal contact increased nearly linearly when alveolar seal duration was below 0.15 seconds, but full Front Region contact was achieved above this threshold. This suggests a time threshold for full contact in the front four rows, below which alveolar closure began at the first row (alveolar ridge) and extended posteriorly as more closure time was provided at higher prosodic constituents.

3.2.2 Vocalic Linguopalatal Contact

Maximal linguopalatal contact of /i/ in domain-initial syllable /ti/ was investigated to test the locality hypothesis of domain-initial strengthening. As predicted, if domain-initial strengthening affects only the initial segment after boundaries, /i/’ s articulatory magnitude should not vary with boundary strength. Figure 12 [Figure 12: see original paper] shows maximal linguopalatal contact for /i/ from both speakers. No clear cumulative boundary effect emerged. Linguopalatal contact at IP- or U-boundary positions was consistently and significantly lower than at ip-boundary position for the female speaker (Figure 12a) and at FT-boundary position for the male speaker (Figure

12b).

3.2.3 F0 and Voice Quality Measures

Figure 13 [Figure 13: see original paper] shows vowel-initial f0, OQ, and SQ for vowels /a/ and /i/ in prosodic domain-initial syllables. Vowel-initial f0 was consistently higher at domain-initial positions of 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 emerged for OQ: the female speaker showed progressive OQ increase in higher prosodic domains (F(4,121)=13.72, p<0.0001, SYL,FT<IP,U, SYL<ip), supporting Garellek (2014), while the male speaker showed consistent OQ across positions except significantly lower OQ at U-initial position (F(4,276)=12.48, p<0.0001, U<IP, ip, FT, SYL). Decreased SQ in stronger prosodic domains appeared 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).

In summary, linguopalatal contact and seal duration measures generally show segment strengthening in higher prosodic constituents, though inter-personal and segmental positional factors are not negligible. Inter-personal variability, widely documented in previous studies, reflects different articulatory strategies adopted by speakers. Slightly different linguopalatal contact patterns at PMC and release frames suggest that strengthening effects may gradually fade away from the boundary. Domain-initial strengthening does not appear to extend to the vocalic interval immediately following domain-initial stops. The hierarchically structured vowel-initial f0 indicates that f0 reset functions to mark prosodic structure. Voice quality measures suggest that vocal folds tend to abduct for progressively longer portions of each abduction-adduction cycle at edges of higher prosodic constituents, though this gradient variation may be speaker-dependent. Additionally, the adduction gesture becomes relatively slower in higher prosodic constituents.

3.3 Acoustic Measures

Table 3 shows ANOVA and multiple comparison results for acoustic measures of the alveolar stop. Acoustic measures at SYL boundary in /i/ context from the male speaker were excluded because most tokens were realized as voiced stops or approximants.

Figure 14 [Figure 14: see original paper] shows mean VOT values at post-boundary positions in both vocalic contexts. In /a/ context, no boundary effect appeared, with VOT remaining around 0.015 seconds across boundary types. Boundary types influenced VOT in /i/ context: for the female speaker, VOT was longer at SYL than other domains, though no significant differences emerged; for the male speaker, boundary effects were significant, with VOT marginally longer at ip boundary than U boundary.

The other two acoustic measures reflecting vocal fold state show clearer cumulative boundary strength effects across speakers and vocalic contexts. Table 3 and Figure 15 [Figure 15: see original paper] indicate that voicing occupied larger portions of acoustic closure duration at SYL and FT boundaries than at higher boundaries. Voicing during closure was zero at IP and U boundaries across contexts and speakers, and at ip boundary in /a/ context for both speakers. Total voiceless interval variation showed the reverse pattern, increasing with boundary strength.

RMS burst energy showed no significant boundary strength effect in /a/ context. In /i/ context, no systematic pattern emerged despite some significant differences in multiple comparisons.

In summary, the four acoustic measures of the alveolar stop show different patterns under prosodic structural conditioning. VOT for unaspirated alveolar stops is not reliably conditioned by prosodic structure, though VOT tends to be slightly higher in lower prosodic domains in /i/ context. Voicing during closure interval and total voiceless interval are reliable acoustic measures showing cumulative post-boundary stop strengthening: as boundary strength increases, voicing ratio over acoustic closure duration decreases while total voiceless interval increases. No clear systematic pattern appears for RMS burst energy.

4. Discussion and Conclusion

This paper addresses two research questions. First, is domain-initial strengthening represented as gradient variation in articulatory and/or acoustic domains in Standard Chinese? Articulatory evidence shows that post-boundary unaspirated alveolar stop articulation varies with boundary strength, with increasingly larger linguopalatal contact and longer alveolar seal duration at higher prosodic domains. This cumulative effect on articulatory magnitude is salient at the PMC frame across speakers and vocalic contexts but fades toward release, as no robust effect appears at the release frame. Weakening of alveolar closure gesture at progressively lower prosodic constituents accompanies reduction of tongue blade contact on the post-alveolar area. The tongue tip gesture may also weaken at SYL or FT domain-initial positions, resulting in approximant or voiced stop realizations.

Vowel-initial f0 shows a clearly hierarchically-nested pattern reflecting global intonational conditioning. Voice quality measures tend to vary with boundary strength: vocal folds abduct more in higher prosodic domains for the female speaker, though no clear trend appears for the male speaker. However, both speakers show progressively lower SQ in higher prosodic domains, indicating slower vocal fold closure gestures when boundary strength increases. These voice quality results support Kong’ s finding that SQ is negatively correlated with f0, and the female speaker’ s OQ results align with Garellek’ s (2014) recent findings.

Regarding acoustic properties of /t/, VOT is not a reliable marker of boundary strength, though follow-up studies are warranted as VOT in /i/ context shows tendency to vary with boundary strength. Voicing during closure interval and total voiceless interval are reliable measures for marking boundary strength: as boundaries strengthen, voicing ratio decreases while total voiceless interval increases. These acoustic measures may be important perceptual correlates for boundary strength, deserving further perceptual investigation. RMS burst energy was expected to show boundary effects with smaller values at stronger boundaries, but no clear pattern emerged in either vocalic environment.

The second research question concerns the scope of domain-initial strengthening effects. The high front vowel /i/ in post-boundary syllable /ti/ was used to test the locality hypothesis. No clear trend appeared for /i/ linguopalatal contact, showing that domain-initial strengthening is restricted to the temporal domain of the first segment immediately following the boundary, with effects quickly fading toward that segment' s release. However, the vocalic interval is hierarchically organized by another prosodic device—intonation. For both speakers, starting f0 of the falling tone progressively decreases as boundary strength weakens. Cumulative variation in the two voice quality measures may be by-products of changing f0 (Kong 2001) or f0 reset at major boundaries (Garellek 2014).

While most results align with previous studies, one issue merits note. Articulatory properties of domain-initial /t/ from both speakers are rather similar at U and IP levels. Figure 11 suggests this may reflect a ceiling effect: when time thresholds are surpassed, tongue tip and blade gestures are fully realized for alveolar stops. At this point, total voiceless interval may serve as the most important acoustic cue distinguishing U- and IP-boundaries. Articulatory properties of /t/ at ip-initial position pattern with higher domains when time thresholds are met (see Table 2 and Figure 11), yet are distinguished from higher domains by total voiceless interval in both vocalic contexts. Articulatory properties of /t/ at SYL- and FT-initial positions appear similar, but closer examination reveals differences: for the female speaker, incomplete alveolar closure gestures occurred more frequently in SYL than FT domains, indicating SYL-initial alveolar gestures are more susceptible to gestural reduction than FT-initial positions.

In conclusion, this study confirms previous findings that domain-initial strengthening is a universally salient effect cueing prosodic hierarchy across languages. For Standard Chinese, the gradient nature of segment articulation can serve as an important cue marking prosodic structure.

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 insights. Part of these results was presented orally 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 Lexicon, 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.

________. 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 the 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 / , 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.

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 Papers 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.

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.

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* 中國語文 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.

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.

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-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 錢瑤, 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.

Note: Figure translations are in progress. See original paper for figures.