Journal of Phonetics 71 (2018) 65–80

Contents lists available at ScienceDirect

Journal of Phonetics

journal homepage: www.elsevier.com/locate/Phonetics

Special Issue: Marking 50 Years of Research on Voice Onset Time, eds. Cho, Docherty & Whalen

Prosodic-structural modulation of stop voicing contrast along the VOT continuum in trochaic and iambic words in American English

Sahyang Kim [a], Jiseung Kim [b], Taehong Cho [c,*]

[a] Department of English Education, Hongik University, Seoul, Republic of Korea
[b] Department of Linguistics, University of Michigan, United States
[c] The Institute for Phonetics and Cognitive Sciences of Language, Department of English Language and Literature, Hanyang University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 13 December 2017
Received in revised form 19 July 2018
Accepted 19 July 2018
Available online 18 August 2018

Keywords:
American English
Stop voicing contrast
VOT
Prosodic enhancement
Domain-initial strengthening
Focus
Prominence
Stress
Iambic
Trochaic

ABSTRACT

This study explores the phonetic nature of phonological stop voicing contrast in American English by investigating how phonetic implementation of the voicing contrast is modulated by the prosodic structure along the continuum of phonetic voicing. In particular, the present study examines (1) the effects of two kinds of prosodic strengthening that can arise with prosodic structuring, a boundary-related domain-initial strengthening (DIS) and a prominence-induced strengthening, and (2) the possible enhancement types of linguistic contrasts that can underlie prosodic strengthening. The phonetic voicing was estimated using the *Integrated Voicing Index* (IVI), taking into account both the voicing lag (positive VOT) and the voiced interval during the closure. Results obtained with initial stops in both trochaic and iambic words are encapsulated as follows. Under the influence of DIS, both voiced and voiceless stops were produced with an increase in voicelessness, showing an enhancement of structurally motivated syntagmatic (CV) contrast. The effect size was larger for voiced stops, yielding a boundary-induced phonetic *reduction* of voicing contrast. Under the influence of prominence (focus), both voiced and voiceless stops showed an increase in voicelessness only in trochaic words, but this time, it was voiceless stops that showed a far greater effect, resulting in a *maximization* of voicing contrast—i.e., an enhancement of paradigmatic contrast. Moreover, the *reduced* voicing for voiced stops under prominence even in the medial position indicates that voiced stops are realized in reference to the phonetic feature {vl. unaspirated} rather than {voiced}. These findings imply that seemingly non-contrastive low-level variation is indeed systematically modulated by the prosodic structure in reference to phonetic representations that regulate the phonetic implementation of phonological contrast in a given language. An alternative account is also discussed in the framework of Articulatory Phonology.

1. Introduction

It has been well established in the field of phonetics and phonology that when an utterance is produced, phonological constituents of various levels (such as syllables, words, and phrases) must be put together in a hierarchically organized way according to the prosodic structure stipulated by the grammatical system of a given language (e.g., Beckman, 1996; Shattuck-Hufnagel & Turk, 1996). A growing body of studies on the phonetics–prosody interface has further suggested that the phonetic realization of individual segments is fine-tuned systematically depending on where in a prosodic structure they occur (e.g., Cho, 2016; Fletcher, 2010). An important assumption that underlies the phonetics–prosody interface is that prosodically conditioned phonetic granularity operates systematically at the subphonemic (phonetic) level, such that phonological units are fleshed out with fine-grained phonetic content in a way that serves the linguistic functions assumed by the prosodic structure (Cho, 2011; Fletcher, 2010; Keating & Shattuck-Hufnagel, 2002), often modulating phonetic implementation of phonological contrast (e.g., de Jong, 1995, 2004; Cho & McQueen, 2005; Cho, Lee, & Kim, 2014). In the present study, we build on that premise by exploring how the phonetic implementation of phonological *voicing* contrast of stops in American English can be modulated by prosodic structure and how the prosodically conditioned fine-tuning of

* Corresponding author at: Hanyang Institute for Phonetics and Cognitive Sciences of Language, Department of English Language and Literature, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 133-791, Republic of Korea.

E-mail addresses: sahyang@hongik.ac.kr (S. Kim), jiseungk@umich.edu (J. Kim), tcho@hanyang.ac.kr (T. Cho).

https://doi.org/10.1016/j.wocn.2018.07.004
0095-4470/

voicing contrast illuminates the phonetic nature of phonological stop voicing contrast.

1.1. Background

The modulation of phonetic implementation according to the prosodic structure has been discussed in terms of prosodic strengthening, which arises with boundary and prominence marking (see Fletcher, 2010, or Cho, 2016, for a review). Boundary-induced and prominence-induced strengthening refer to a spatiotemporal expansion of segmental realization at the edges of a prosodic constituent (e.g., phrase-initial/final positions) and in stressed/accented syllables, respectively (e.g., Beckman & Edwards, 1994; Cho, 2005, 2006; de Jong, 1995; Fougeron & Keating, 1997; Turk & White, 1999, *inter alia*). The two kinds of prosodic strengthening can be linked to the dual functions of prosodic structure (the delimitative function for boundary marking and the culminative function for prominence marking) and are often construed to enhance different kinds of linguistic contrast, such as syntagmatic or paradigmatic contrast (see Fougeron, 1999; and Cho, 2011, 2016, for a review). The term *syntagmatic* pertains to the structural relationships between neighboring linguistic elements that form a sequence in speech. The boundary-marking function of a prosodic structure can be syntagmatically, or structurally, motivated, to enhance the contrast between neighboring segments (or the syntagmatic contrast) localized at prosodic junctures. The term *paradigmatic*, on the other hand, pertains to the relationship among linguistic units such as phonemes (or words) that can substitute for one another in a given context. The paradigmatic contrast enhancement used here generally describes the maximization of phonemic distinction of contrastive sounds, which is often considered to be associated with prominence. Given the potentially different functions of prosodic structures and their relevance to linguistic contrast with different locality conditions (edges vs. stressed syllables), we specifically address the relationship between English stop voicing contrast and enhancement associated with different prosodic strengthening effects. In what follows, we elaborate on specific issues, along with our research questions and hypotheses.

1.2. Issues and research questions about boundary-related stop voicing contrast

Research on domain-initial strengthening (DIS), which arises with boundary marking, has indicated that the DIS effect is closely linked to phonetic feature enhancement. For example, in an acoustic-aerodynamic study of the DIS effect on three-way contrastive stops in Korean (lenis, fortis, aspirated; e.g., Cho, Jun & Ladefoged, 2002), Cho and Jun (2000) reported that voice onset time (VOT) was more lengthened in domain-initial than in domain-medial positions for aspirated stops, and it was shortened for fortis stops. These results were interpreted as indicating enhancements of different laryngeal features: [spread glottis] for the former and [constricted glottis] for the latter. In a similar vein, Cho and McQueen (2005) showed that the DIS effect in Dutch induced a shortening of VOT for phonologically voiceless stops, the opposite of the DIS effect found in English (Pierrehumbert & Talkin, 1992; Cho & Keating, 2009), despite the fact that the voiceless stop in both languages can be specified with the same phonological feature [-voice] (e.g., Keating, 1984, 1990; Kingston & Diehl, 1994). The asymmetrical boundary-induced modulation of VOT between the two languages was attributed to language-specific constraints on what phonetic features can be involved in the phonetic implementation of the phonological feature [-voice]—i.e., {vl. unaspirated} ({-spread glottis}) vs. {vl. aspirated} ({+spread glottis}) for voiceless stops in Dutch vs. English. In other words, it is not the phonological feature but the language-specific phonetic feature with phonetic content that operates in fine-tuning phonetic implementation under prosodic strengthening. This is in line with Keating's (1984; cf. 1990) view that stops in world languages can be further distinguished in terms of three phonetic categories, {vl. aspirated}, {vl. unaspirated}, and {voiced}, based on which actual phonetic content is determined (but see Cho & Ladefoged, 1999 for linguistic arbitrariness in choosing a modal VOT value in a given language; cf. Chodroff & Wilson, 2017).

Under the assumption that English voiceless stops are phonetically implemented on the basis of the phonetic feature {vl. aspirated}, the boundary-related enhancement of {vl. aspirated} for English voiceless stops might be evident in an increase in the amount of glottal opening (e.g., Cooper, 1991) and longer VOT (Cho & Keating, 2009; Cho et al., 2014; Pierrehumbert & Talkin, 1992), which can be interpreted as a case of paradigmatic enhancement. The increased glottal width and longer VOT, however, could also be interpreted as evidence for a syntagmatic (CV) enhancement because the augmented voicelessness (as reflected in the larger glottal width and longer VOT) would make the consonant more consonant-like, enhancing its structural distinction from neighboring vowels.

One way of testing these possible explanations of enhancement would be to examine how *voiced* stops are phonetically realized compared to voiceless stops under the influence of DIS. If the DIS effect is driven by an enhancement of paradigmatic (phonemic) contrast, voiced stops in the domain-initial position would be produced with an increase in voicing in the direction of phonological contrast between voiced and voiceless stops. The expected polarization effect is schematized as Type 1 in Fig. 1, which shows a leftward polarization of voiced stops (with voicing lead) along the phonetic voicing continuum to be maximally contrastive with the voiceless counterpart. Type 2 in Fig. 1, in which the phonetic voicing for voiced stops is assumed to remain more or less stable, could also be acceptable evidence of paradigmatic enhancement, given that the polarization is still achieved by an increase in voicelessness for the voiceless counterpart. (See below for further discussion on this possibility under prominence-induced strengthening.) Alternatively, however, if the DIS effect is driven by a syntagmatic enhancement of CV contrast, voiced stops are expected to be produced with an increase in voicelessness, just as voiceless stops are, to enhance their consonantality, as schematized as Type 3 in Fig. 1.

However, our understanding of how voiced stops are actually realized along the phonetic voicing continuum under the influence of DIS has been extremely limited, making it difficult to test these possibilities. DIS effects have been explored on some voiced segments in English (e.g., /b, n/) but only in the

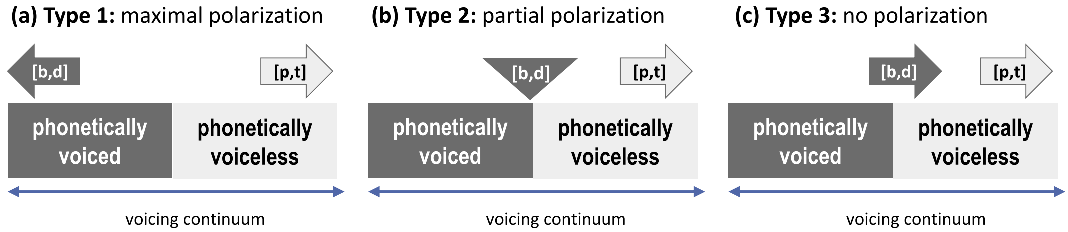

Fig. 1. Three types of phonetic implementation of voicing contrast.

supralaryngeal articulatory dimension (e.g., Byrd, Krivokapić, & Lee, 2006; Cho, 2005; Fougeron & Keating, 1997). Cole, Kim, Choi, and Hasegawa-Johnson (2007) indeed sampled the acoustic realization of English voiced and voiceless stops in varying prosodic positions, including measurements of VOT and closure duration (CD) in the 'lab news' speech, and reported no DIS effect on VOT or CD for /t, d/. However, they did not measure 'voicing' during closure for voiced stops, nor did they control lexical stress or syllable contexts. Thus they provide insufficient data with which to draw firm conclusions about DIS effects on voicing contrast in English. Most recently, Davidson (2016) explored how voiced stops in English are phonetically realized, especially with respect to phonetic voicing (or phonation) during closure. It was reported that a majority of the voiced stops were produced without voicing (phonation) before the release in utterance-initial positions (about 75%), but the proportion of tokens with some degree of voicing during closure increased to 65% in the phrase-medial but word-initial position and to 85–90% in the phrase-medial and word-medial position. Based on the discussion in the literature on DIS, Davidson interpreted the reduction of voicing at higher prosodic domains as stemming from either domain-initial articulatory strengthening that would create aerodynamic constraints that impede phonation or from an enhancement of phonetic features such as {-voiced, -spread glottis}. Although that interpretation illuminates the nature of DIS effects on voiced stops in English, the study did not directly compare voiced and voiceless stops (but see Davidson (2017) for results and related discussion on voiceless stops in English), nor did it systematically examine how the DIS effect interacts with the prominence that can arise with both lexical stress and phrasal accent.

In the present study, we therefore examine the phonetic realization of phonological voicing contrast for stops in English (e.g., /p, t/ vs. /b, d/ as in ***p****ánel* vs. ***b****ánner* and ***t****ánner* vs. ***D****ániel*) as a function of boundary, stress, and accent and their interactions, with a view to testing which of the three voicing contrast types schematized in Fig. 1 best maps the actual phonetic realization of voicing contrast along the phonetic voicing continuum under the influence of DIS.

Another important question that arises when testing DIS effects on stop voicing contrast is the extent to which DIS effects could be further conditioned by prominence distribution over an initial word. It has been reported in the literature that DIS tends to interact with phrase-level accents (nuclear pitch accent). So, for example, an increase in VOT for English voiceless stops due to DIS is more likely to be observed in a relatively low prominence condition (i.e., when the initial syllable does not receive an accent), whereas the effect often disappears when the accent falls on the word (Cho & Keating, 2009; Cho, Lee, & Kim, 2011, 2014; Cole et al., 2007). Such a *prominence-dependent* DIS effect on VOT can be understood as a ceiling effect of prominence on VOT. Because VOT is lengthened under prominence as well, it might leave no room for further temporal expansion due to DIS, as discussed in Cho (2016). To the best of our knowledge, however, DIS has only been studied in an environment in which the initial syllable is lexically stressed (e.g., Byrd et al., 2006; Cho, 2005, 2006; Cho & Keating, 2009; Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003). It is therefore not entirely clear how DIS effects might be modulated by a differential degree of prominence. Given that a DIS effect is more robust in a less prominent condition (stressed but unaccented) than in a more prominent condition (stressed *and* accented), one might expect the effect to still be clearly or even more clearly observable when the initial syllable is *unstressed*—i.e., in an even less prominent condition. We therefore test this possibility by including test words with initial stress (trochaic words, e.g., *pánel*) and non-initial stress (iambic words, e.g., *panáche*). Because the degree of prominence associated with the unstressed syllable can be further modulated by the presence or absence of an accent on the stressed (second) syllable (in the case of iambic words) through the possible leftward spreading of accentual lengthening (e.g., Turk & White, 1999), we control the phrasal accent factor in our investigation to observe how the higher-order prominence factor (i.e., accentuation) influences the manifestation of DIS on trochaic and iambic words.

1.3. Issues and research questions about prominence-related stop voicing contrast

Along with those questions about the effects of DIS on stop voicing contrast in English, we also explore the directions in which the phonological voicing contrast can be phonetically polarized under prominence. The three possible polarization types, as schematized in Fig. 1, will also be considered in connection with prominence-induced strengthening. As briefly mentioned above, prominence-induced strengthening is often assumed to be associated with an enhancement of paradigmatic contrast, which de Jong (1995) describes as localized hyperarticulation—i.e., localized to the stressed syllable as opposed to a communicatively driven hyperarticulation that is extended globally to the whole utterance, as in hyper- & hypo-articulation theory (Lindblom, 1990). We expect that this type of prosodic strengthening will enhance distinctive features to maximize phonemic (and lexical) contrasts. For example, de Jong (1995) showed that /ʊ/ in English is produced with a more retracted tongue position when it is in the accented (focused) condition than when it is in the unaccented condition, showing an enhancement of the [+back] feature of the vowel. De Jong (2004) also demonstrated that vowel duration in English is used as an important phonetic feature for marking phonological voicing contrast of the following stops, and its effect is enhanced (with the durational contrast being polarized) when the syllable is focused. The same durational phonetic feature, however, did not show an enhancement pattern in Arabic, in which duration is used to mark vowel quantity contrast (de Jong & Zawaydeh, 2002). The prominence-induced strengthening effect, therefore, especially when it is realized with focus, has been used to diagnose which phonetic content or phonetic feature is regulating the phonemic contrast in a given language (de Jong & Zawaydeh, 2002; de Jong, 2004). In the present study, we also use focus, expressed by a nuclear pitch accent, to test how stop voicing contrast is phonetically implemented under prominence and which phonetic feature operates in association with polarizing voicing contrast.

If a voiced stop is produced operating on {vl. unaspirated}, it is likely to show a partial polarization effect, roughly similar to Type 2 in Fig. 1. Because of the voicelessness associated with {vl. unaspirated}, voicing of a voiced stop is expected to be implemented in the voiceless territory along the phonetic voicing dimension. A relevant finding that lends support to this prediction is in Smiljanic and Bradlow's study (2008). They showed that VOT is lengthened for word-initial (but phrase-medial) voiceless stops in clear speech (vs. casual speech), but the VOT of voiced stops in the same position remained relatively stable regardless of speech style modification. Cho et al. (2014), however, demonstrated that VOT for a voiceless stop in the /s/-stop cluster that is already shortened by an allophonic rule is shortened even more under focus-induced accent, showing an enhancement of the allophonically derived phonetic feature {vl. unaspirated}. Thus, a voiced stop, if phonetically implemented on {vl. unaspirated}, can still be produced with a shortening of VOT, showing some degree of polarization insofar as the effect remains in the positive (voicing lag) dimension. (See Nelson and Wedel (2017) for a similar kind of polarization for the voicing contrast with minimal pairs in conversational English.)

It is also worth noting that in the literature, word-initial voiced stops have generally been shown to be produced without significant prevoicing, which is consistent with a specification of {vl. unaspirated} (Docherty, 1989; Lisker & Abramson, 1964). However, word-initial stops often describe both stops produced in the utterance-initial position and stops produced in isolation, leaving questions open about how their phonetic implementation might differ across positions and degrees of prominence, as was also noted by Abramson and Whalen (2017). If the production of a voiced stop indeed makes reference to the phonetic feature {voiced} under prominence, its phonetic voicing is expected to be enhanced, showing a polarization effect similar to the Type 1 scheme in Fig. 1.

However, the feature specification can also be determined by the position in which a stop occurs. This possibility was discussed by Keating (1984): voiced stops can be specified as {voiced} in a medial position and as {vl. unaspirated} in an initial position, resulting in position-specific allophonic variations of phonologically voiced stops in English. Keating (1984), however, did not precisely define the term *initial* in the prosodic hierarchy, and she used the medial position to describe an intervocalic word-medial position flanked between vowels, which facilitates phonetic voicing. Given that phonetic voicing is also likely to be facilitated when a voiced stop is flanked by vowels phrase-medially, it is quite plausible to extend the medial position to include phrase-medial positions. As a reviewer pointed out, the specification of phonological features generally refers to what is specified in the lexicon or in a given lexical item. But the position-specific featural assignment hypothesis assumes that the phonological feature is phonetically implemented based on the phonetic substance of the phonetic feature, which could be assigned post-lexically. To the extent that this assumption holds, it is reasonable to hypothesize that phrase-initial stops are phonetically implemented based on {vl. unaspirated} and that phrase-medial stops are based on {voiced}, possibly showing a position-specific enhancement pattern of voicing contrast under prominence. Again, one way to test that hypothesis would be to use focus-driven prominence as a diagnostic for assessing which phonetic feature is involved in producing phonologically voiced stops. Phrase-medial voiced stops, if specified with {voiced}, would be produced with *more* phonetic voicing under prominence, which could be mapped as Type 1 of Fig. 1. On the other hand, phrase-initial voiced stops specified as {vl. unaspirated} would show *less* phonetic voicing in the same prominent condition, for possible mapping as Type 2 of Fig. 1.

One cannot, however, entirely rule out the possibility that prominence-induced strengthening could result in increasing voicelessness for both voiced and voiceless stops. Electromyographic studies by Ladefoged and his colleagues (e.g., Ladefoged, 1967; Ladefoged & Loeb, 2002) suggested that stress (broadly defined to include accentuation) might involve an increase in respiratory force. Such augmented respiratory power could facilitate glottal abduction, which could effectively increase voicelessness for both voiced and voiceless stops, schematized as Type 3 in Fig. 1.

Finally, in addition to the issues that have been discussed so far, we discuss the results of this present study in the framework of Articulatory Phonology, with a view to understanding the gestural underpinnings of voicing contrast in relation to its prosodic-structural modulation.

2. Methods

2.1. Participants

Eleven native speakers of American English (6 female, 5 male) were paid to participate in the recording. The participants, all in their 20 s and early 30 s, were from either the Midwest or the West Coast of the US[1] and were temporary residents in Korea as exchange students or English instructors at the time of recording.

[1] Jacewicz, Fox, and Lyle (2009) reported dialectal differences in the amount of prevoicing for initial voiced stops in American English. Speakers from North Carolina mostly produced fully voiced closures for the word-initial voiced stops at the juncture of two words (e.g., *small bids*), but stop closures produced by speakers from Wisconsin were not fully voiced. None of the participants in the current study is from the southern USA.

Table 1
List of target words.

	Trochaic words	Iambic words
Voiceless stop onset	*pánel, tánner*	*panáche, Tenı́se*
Voiced stop onset	*bánner, Dániel*	*banál, Denı́se*

2.2. Speech materials and recording procedure

The eight bisyllabic target words are listed in Table 1.[2] All the target words had the sequence of /CVNVC/. Half of the words were trochaic words, and the other half were iambic words. The word-initial stops were voiceless /p, t/ and voiced /b, d/. The initial stops were followed by the vowel /æ/ in trochaic words and by the unstressed vowel /ə/ in iambic words.

The target words were inserted in carrier sentences that consisted of a background and a test sentence, as shown in Table 2. The second sentence in the pair was always the target-bearing test sentence, whereas the first sentence was used to induce the intended prosodic conditions. To obtain tokens in the accented condition, the two sentences in each pair were constructed such that the target word in the second sentence (e.g., ***p**ánel*) received a focus in contrast with a corresponding word (e.g., ***b**ánner*) in the first sentence, as marked in bold upper case in Table 2a and c. For the unaccented condition, the contrastive focus fell somewhere else in the sentence, as in Table 2b and d. In addition, to induce an Intonational Phrase (IP) boundary, syntactically complex sentences were used, so that a prosodic boundary before the target word coincided with a major syntactic boundary between a subordinate clause and a main clause as in Table 2a and c (e.g., *But after **JOHN** says 'banana,' '**PANEL** again' will be the next phrase to say.*). For a Wd-boundary (IP-medial) condition, the two-word sequence formed part of a single object NP within the same syntactic phrase, as in Table 2c and d (e.g., *To say "banana panel again" with me...*), to increase the likelihood that the speaker would pronounce them phrase-internally as a chunk. The vowel of the word preceding the target word within a test sentence was controlled to be a schwa to prevent any potential confounding effect from a previous segment.

The recordings were made in a sound-attenuated booth at the Hanyang Phonetics and Psycholinguistics Lab. The acoustic data were collected at a sampling rate of 44 kHz using a SHURE KSN 44 dynamic microphone and a Tascam HD-P2 digital recorder. Sentences were presented on a computer screen in a randomized order and repeated four times across four blocks. Each block had a different randomization order. Participants were asked to read the carrier sentences with

Table 2
Sample carrier sentences with the target word *panel*. The accented words are marked in bold, and the target word is underlined

a. IP boundary, Accented
*After **I** say 'banana,' '**BANNER** again' will be the next phrase to say.*
*But after **JOHN** says 'banana,' '**<u>PANEL</u>** again' will be the next phrase to say.*
b. IP boundary, Unaccented
*After **I** say 'banana,' 'panel again' will be the **NEXT** phrase to say.*
*But after **JOHN** says 'banana,' '<u>panel</u> again' will be the **FINAL** phrase to say.*
c. Word boundary, Accented
*To say 'banana **BANNER** again' with me is going to be difficult.*
*But to say 'banana **<u>PANEL</u>** again' with me is going to be easy.*
d. Word boundary, Unaccented
*To say 'banana panel again' with **JOHN** is going to be difficult.*
*But to say 'banana <u>panel</u> again' with **ME** is going to be easy.*

the meaning contrast between the words marked in bold upper case in mind. The experimenter was a trained ToBI transcriber. During the recording session, when the experimenter thought that a speaker produced a sentence with hesitation or a rendition deviating from the intended one, the speaker was asked to read the sentence a few more times. Each recording comprised three sessions, and participants took a 5–10 min break between sessions. Thirty-two target sentences (2 boundary conditions × 2 accent conditions × 8 target words) with four repetitions resulted in 128 tokens per speaker, yielding a total of 1408 tokens (128 tokens × 11 speakers). After all the recordings were complete, two trained phoneticians cross-checked the data in terms of prosodic conditions. When either of the crosscheckers reported that the rendition of a token did not conform to the intended prosodic (accent/boundary) conditions, that token was excluded from further analyses. As a result of that crosschecking, 358 tokens were excluded, leaving 1022 tokens for the data analyses. The distribution of tokens across critical conditions is given in Appendix A (Table A1). Note that each speaker contributed an average of 5.3 tokens (ranging from 3 to 8) for each condition (Voice × Stress × Boundary × Accent) pooled across the two places of articulation.

2.3. Measurements

The following acoustic duration measures were taken from the initial syllable of each target word, using Praat (Boersma & Weenink, 2014).

***(Positive) VOT* (voicing lag):** Positive VOTs of both voiceless and voiced stops were measured from the stop release to the onset of voicing (the first regular waveform) for the following vowel, as guided by Davidson (2016) and Abramson and Whalen (2017). It should be noted that VOT included any observable voicing lag from the obvious release to the onset of voicing, even for the voiced stops, which were often produced with voicing during the closure. As Mikuteit and Reetz (2007) pointed out, this type of VOT may deviate from the classical definition of the positive VOT because voicing in this case is already initiated before the release. Mikuteit and Reetz therefore proposed a more neutral term, *After Closure Time* (ACT) that embraces both the positive VOT in a classical sense (i.e., without preceding voicing during closure) and the positive VOT preceded by voicing during the closure. Since the positive VOT in such a prevoiced token contains some degree of voicelessness (which Abramson and Whalen (2017) suggested to be included when examining the

[2] As pointed out by a reviewer, the target words were not controlled in terms of lexical statistics (e.g., word frequency and phonological neighborhood density). It was difficult to find (near) minimal pairs suitable for our purpose in this study and simultaneously control for lexical factors. Because the lexical factors are known to influence phonetic realization (e.g., Wright, 2004; Baese-Berk & Goldrick, 2009; Tomaschek et al., 2018), we acknowledge that the results we report here have some limitations. Nelson and Wedel (2017), however, suggested that hyperarticulation of stop voicing contrast in VOT is predicted primarily by a competitor that differs solely in word-initial voicing, while other lexical factors such as phonological neighborhood density is not a reliable predictor for contrastive hyperarticulation in VOT. So, even if other lexical factors may affect the degree of hyperarticulation in phonetic realization, we believe that similar conclusions would be drawn even when lexical factors were strictly controlled, albeit with a possible further fine-tuning of the phonetic realization of voicing.

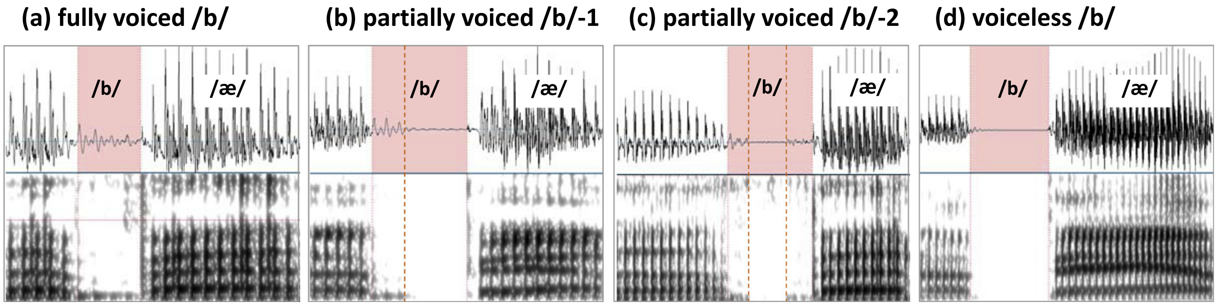

Fig. 2. Four typical types of voicing during closure for /b/ in banner: (a) fully voiced /b/, (b) partially voiced /b/-1 with continued voicing after the preceding vowel, (c) partially voiced /b/-2 with voicing at both edges of the closure; (d) voiceless /b/. The voiced intervals in (b) and (c) correspond to the 'bleed' and 'trough' patterns, respectively, described in Davidson (2016).

boundary-related strength), we assumed that this type of VOT would be useful in assessing the voicing of the token, thus included it in our analyses.

Voicing-in-Closure **(Voicing duration in closure):** The voiced interval during the stop closure was measured for both the voiced and voiceless stops, as indicated by voicing bars in the spectrogram in consultation with waveforms. (This measure can be taken to be negative VOT if it is defined to include any portion of prevoicing during the closure whether intermittent or continuous.) As exemplified in Fig. 2, the voiced interval included any continued voicing (with two or more clear voicing bars on the spectrogram) after the F2 offset of the preceding vowel and any voicing lead before the burst of the voiced consonants. As a result, Voicing-in-Closure included the voicing in fully voiced tokens (Fig. 2a) and partially voiced tokens, which corresponds to both the so-called *bleed* pattern (Fig. 2b) and the *trough* pattern (Fig. 2c), as discussed in Davidson (2016). It should be noted that Voicing-in-Closure was observed mostly for the voiced stops of IP-medial tokens, with only 5% of the IP-initial voiced stops produced with some voicing during closure.

Integrated Voicing Index **(IVI):** Whereas VOT and Voicing-in-Closure could each be used as a separate index of phonetic voicing, it was often the case, as mentioned above, that a stop was produced with both (positive) VOT and some voiced interval during the closure. Some of the voiced stops with voicing during the closure were indeed accompanied by a positive VOT (usually a short-lag VOT), and some of the voiceless stops, though less frequently, were also produced with some degree of voicing during the closure. Because of the complexity of the voicing realization, separate analyses of VOT and Voicing-in-Closure alone would make it difficult to assess the hypothesized polarization patterns of voicing contrast along a single phonetic dimension. We therefore devised the IVI, which we defined as a combined sum of VOT (as a positive value) and Voicing-in-Closure (as a negative value). This voicing index was negative when Voicing-in-Closure was longer than VOT and positive when the opposite was true. For the tokens in Fig. 2, for example, the IVI value was positive for the token in Fig. 2d, which has a short-lag VOT without voicing during the closure, and negative for the tokens in Fig. 2a–c, in which the voicing lag (positive VOT) is shorter than the voicing duration in closure (Voicing-in-Closure). The IVI thus weights the relative contribution of VOT and Voicing-in-Closure to voicing contrast, allowing us to assess the phonetic voicing of both voiceless and voiced stops along a single integrated dimension of phonetic voicing.

In the results section, we compare the observed effects on the IVI with those on VOT and Voicing-in-Closure to provide a complete picture of the prosodic modulation of voicing realization for the voiced and voiceless stops under investigation.

2.4. Statistical analyses

We conducted separate series of repeated measures (RM) ANOVAs on the **VOT**, **Voicing-in-Closure**, and **IVI** of trochaic and iambic words because the target words differed not only in their stress pattern but also in the following vowel context (i.e., the first vowel was always /æ/ in trochaic words but a schwa in iambic words). In the RM ANOVAs, the individual means per speaker contributed one averaged score for each condition. We considered three between-subject factors: Voice (voiced *vs*. voiceless), Boundary (IP-initial *vs*. IP-medial), Accent (Accented *vs*. Unaccented), and Place (alveolar *vs*. labial). We included the Place factor to control for the difference in the place of articulation, which varied across the target consonants. Although we report the results for the main effects, the interaction effects between Voice, on the one hand, and Boundary and Accent, on the other hand, are the most important because the primary focus of our study is investigating how Boundary and Accent influence voicing contrast and how differently or similarly the realization of phonetic voicing on voiced and voiceless stops is modulated by prosodic structural factors. When we observed a significant interaction effect between Voice and the crucial prosodic structural factors (Boundary and Accent), we performed separate one-way RM ANOVAs to examine where the interaction began. In all statistical analyses, *p*-values less than 0.05 were considered significant. (Note that in most cases, the statistical summaries of main effects and interactions are provided in the figures without repeating the statistical details in the text.)

3. Results

*3.1. Trochaic words (**p**ánel, **t**ánner, **b**ánner, **D**ániel)*

3.1.1. Main effects

The results of the RM ANOVAs on the main effects of the experimental factors are summarized in Fig. 3. There was a significant main effect of Voice, Boundary, Accent, and Place on the **IVI**. As shown in Fig. 3A.1, the IVI was shorter (and negative) for voiced stops than for voiceless stops (Voice effect). We attribute the negative IVI value for voiced stops to a substantial portion of **Voicing-in-Closure**, as shown in Fig. 3A.3,

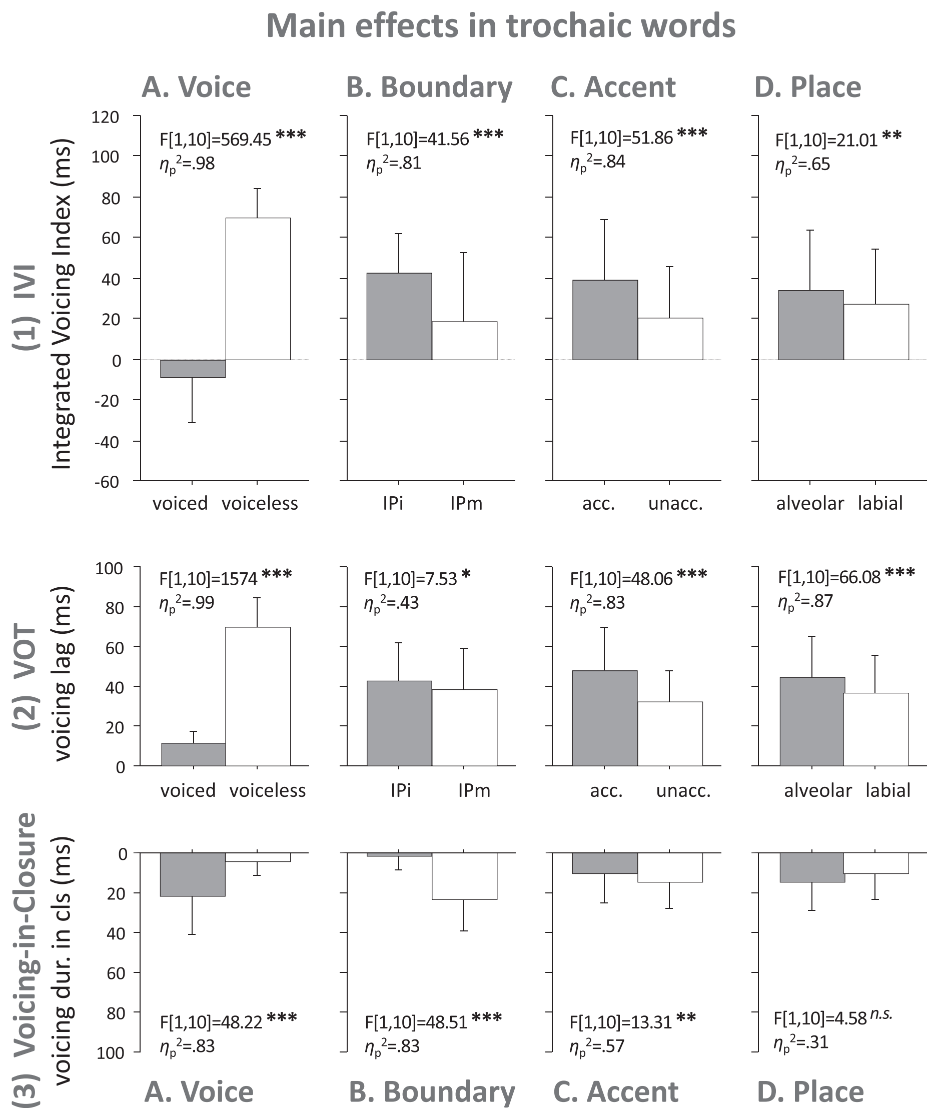

Fig. 3. Main effects of Voice, Boundary, Accent, and Place on (1) IVI (Integrated Voicing Index), (2) VOT (voicing lag), and (3) Voicing-in-Closure (voicing duration in closure) for trochaic words. Error bars refer to standard deviations; * refers to $p < 0.05$, ** to $p < 0.005$, and *** to $p < 0.001$.

although voiced stops were often accompanied by a short-lag VOT, as can be inferred from Fig. 3A.2. As for the boundary effect, IVI was longer in the IP-initial than in the IP-medial position (Fig. 3B.1), showing an overall DIS effect toward increased voicelessness. We attribute the boundary effect on the IVI to both VOT and Voicing-in-Closure—i.e., VOT was longer IP-initially (Fig. 3B.2), and Voicing-in-Closure was reduced IP-initially, centering near zero (Fig. 3B.3). The accent effect was similar to the boundary effect, showing an increase in the IVI under accent (Fig. 3C.1), but this time the accent effect on the IVI was attributable more to VOT (which was longer under accent) than to Voicing-in-Closure (which was reduced under accent), as can be inferred from the effect size difference shown in Fig. 3C.2–3. Finally, IVI was longer for the alveolar than for the labial stops, and this place effect occurred primarily because VOT was longer for alveolar stops than for labial stops, whereas Voicing-in-Closure showed no place effect.

As for interactions, Voice did not interact with Place in any of the three measures (IVI, VOT, Voicing-in-Closure), showing a general place-independent effect across the voicing conditions. Voice, however, did interact with Boundary and Accent, showing a further modulation of the voicing contrast as a function of prosodic factors, which will be explained in detail in the following subsections.

3.1.2. Voice × boundary in trochaic words

We found a significant Voice × Boundary interaction for the **IVI**. As shown in Fig. 4a, the interaction was caused in part by the fact that although the effect of Voice remained significant regardless of Boundary, its effect size was substantially *reduced* in the IP-initial position compared to the IP-medial position. In other words, the Boundary factor did not give rise to a polarization of the voicing contrast, but instead, the opposite occurred: stop voicing contrast was reduced in the DIS environment. This interaction effect was further evident in the **VOT** and **Voicing-in-Closure**. The reduced voicing contrast in the IVI resulted in part due from the voiced stops being produced with an increase in VOT but also from a decrease in the

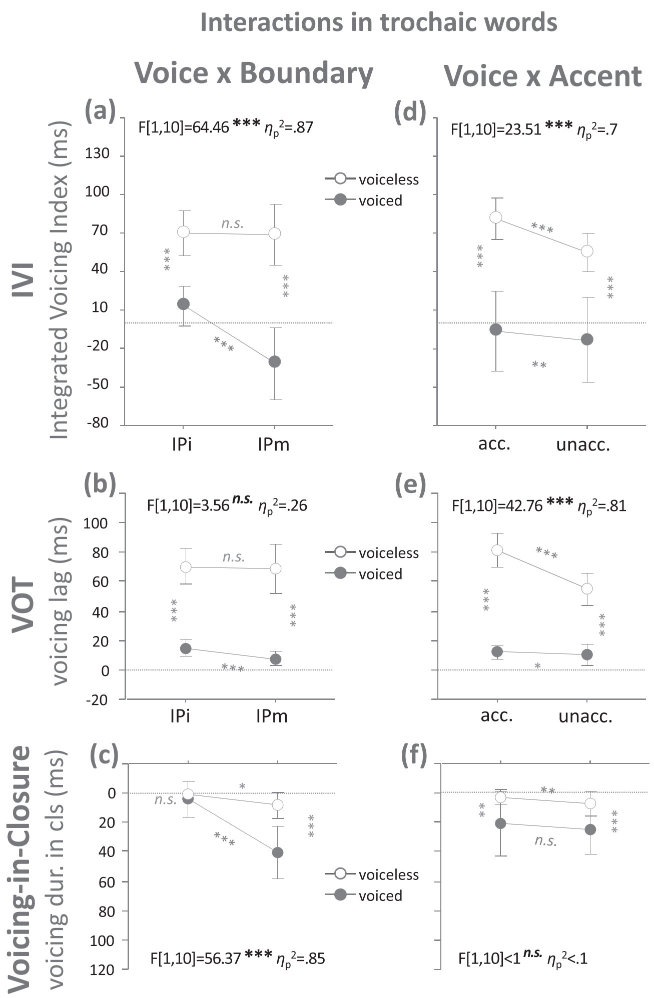

Fig. 4. Interactions of Voice × Boundary (a–c) and Voice × Accent (d–f) on the IVI (Integrated Voicing Index), VOT (voicing lag), and Voicing-in-Closure (voicing duration in closure) for trochaic words. Error bars refer to standard deviations. * refers to $p < 0.05$, ** to $p < 0.005$, and *** to $p < 0.001$ obtained from separate one-way RM ANOVAs.

Voicing-in-Closure in the IP-initial position (as indicated by the filled circles in Fig. 4b and c, respectively).

Furthermore, as can be seen in Fig. 4a, the boundary effect on voiced and voiceless stops was asymmetrical. On the one hand, the boundary effect (i.e., longer IVI in the IP-initial, IPi, condition, than in the IP-medial, IPm, condition) stemmed primarily from a reduction of phonetic voicing for voiced stops in the IP-initial position, which was reflected in both the VOT (which was longer IP-initially, as shown in Fig. 4b) and Voicing-in-Closure (which was reduced IP-initially, as shown in Fig. 4c). On the other hand, the voiceless stops showed no boundary effect on the IVI (as indicated by the empty circles in Fig. 4a), which was further confirmed by finding no boundary effect on the VOT of the voiceless stops (Fig. 4b). Note, however, that as reported below in the section about the Boundary × Accent interaction, further analyses indicated that the voiceless stops showed a boundary-induced lengthening of IVI in the *unaccented* condition.

3.1.3. Voice × accent in trochaic words

A significant Voice × Accent interaction occurred in the **IVI**. As shown in Fig. 4d, the IVI difference attributable to Voice remained significant in both the accented and unaccented conditions, but the difference between the voiced and voiceless stops was augmented under accent, showing an accent-induced enhancement of voicing contrast. The augmented voicing contrast was evident in the VOT, as can be seen in Fig. 4e, which also revealed a significant Voice × Accent interaction, but the **Voicing-in-Closure** did not contribute to the interaction, as shown in Fig. 4f (i.e., no significant Voice × Accent interaction).

Crucially, however, both the voiced and voiceless stops showed an increase in the IVI under accent, indicating that the augmented voicing contrast under accent did not occur because voiced stops were phonetically more voiced in a polarizing direction along the voicing continuum of the IVI. Instead, both the voiced and voiceless stops showed a unidirectional increase (both positive) under accent, which was further evident in both the VOT and Voicing-in-Closure, as can be seen in Fig. 4e and f, respectively. The observed interaction occurred because the unidirectional accent-induced increase in the IVI was much greater for voiceless stops than for voiced stops, which resulted in an enhancement of the voicing contrast. This pattern was also observed with VOT, as shown in Fig. 4e, indicating a significant Voice × Accent interaction. The Voicing-in-Closure did not appear to make a significant contribution to the interaction effect on the IVI, as evident by the lack of a significant interaction between Voice and Accent in Fig. 4f. However, post-hoc analyses revealed that the voiceless stops, as indicated by empty circles in Fig. 4f, were produced with a significant reduction in Voicing-in-Closure under accent, whereas the Voicing-in-Closure for the voiced stops remained largely unchanged. This asymmetrical accent effect on the Voicing-in-Closure appears to have contributed, though to a small degree, to the Voice × Accent interaction observed in the IVI.

3.1.4. Boundary × accent in trochaic words

We found a significant Boundary × Accent effect on the **IVI**, VOT, and **Voicing-in-Closure** (IVI, $F[1,10] = 24.04$, $p = .001$, $\eta^2 = 0.71$; VOT, $F[1,10] = 67.88$, $p < 0.001$, $\eta^2 = 0.87$; Voicing-in-Closure, $F[1,10] = 35.52$, $p < 0.001$, $\eta^2 = 0.77$). The interaction was consistently due to the fact that the boundary effect (generally showing an increase in voicelessness) was robust in the absence of prominence (i.e., in the *unaccented* condition). This interaction effect can be seen in Fig. 5, which plots the effect separately for voiced and voiceless stops. Note that we found no significant three-way Voice × Boundary × Accent interaction for the IVI or Voicing-in-Closure, indicating that the prominence-dependent boundary effect was consistent across the voiced and voiceless stops. But the VOT did show a significant three-way interaction ($F[1,10] = 12.49$, $p = .005$, $\eta^2 = 0.54$), which motivated the separate analyses of voiced and voiceless stops.

It is worth recalling that when the data were pooled across accent conditions (Fig. 4a), the voiceless stops showed no DIS effect on the IVI. But the RM ANOVAs run separately for the voiced and voiceless stops revealed some evidence of a boundary-induced DIS effect on voiceless stops. As shown in Fig. 5a, the voiceless stops showed a significant Boundary × Accent interaction for the IVI, which occurred because the boundary effect was significant in the less prominent,

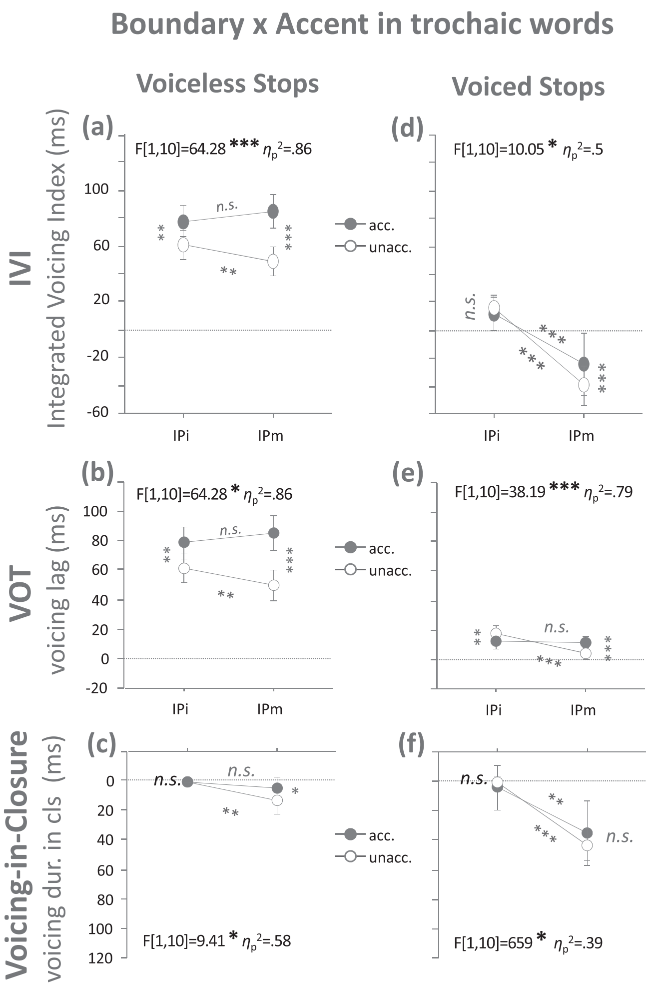

Fig. 5. Boundary × Accent interaction effects for voiceless stops (a–c) and voiced stops (d–f) on the IVI (Integrated Voicing Index), VOT (voicing lag), and Voicing-in-Closure (voicing duration in closure) in trochaic words. Error bars refer to standard deviations; * refers to $p < 0.05$; ** to $p < 0.005$, and *** to $p < 0.001$ obtained from separate one-way RM ANOVAs.

unaccented condition than in the accented condition. This prominence-dependent DIS effect was further evident in the VOT and Voicing-in-Closure, as shown in Fig. 5b–c.

On a related note, *voiced* stops also showed comparable effects. The voiced stops showed a significant Boundary × Accent interaction for all three measures, as shown in Fig. 5d–f, which we again attributed to a more robust boundary effect in the *unaccented* condition than in the accented condition. In other words, the boundary-induced *reduction* of voicing for the voiced stops was more robust in the *unaccented* condition than in the accented condition, which was evident in the IVI, as shown in Fig. 5d (accented, mean diff., 34.4 ms, $\eta^2 = 0.61$; unaccented, mean diff., 54.8 ms, $\eta^2 = 0.94$), and the Voicing-in-Closure, as shown in Fig. 5f (accented, mean diff., 18.8 ms, $\eta^2 = 0.65$; unaccented, mean diff., 28.3 ms, $\eta^2 = 0.91$). As shown in Fig. 5e, the VOT showed an even more extreme pattern; the boundary effect was significant only in the *unaccented* condition, and the effect disappeared in the accented condition.

Seen from a different angle, just as the boundary effect was *reduced* in the context of another strengthening effect (i.e., in the accented condition), so was the accent effect reduced in the context of DIS. The overall accent effect was reduced in the IP-initial position compared to the IP-medial position for both voiced and voiceless stops.

For the *voiceless* stops (seen in Fig. 5a–c), the accent effect was reduced from the IP-medial to the IP-initial position in all three measures (which was further confirmed by significant Boundary × Accent interactions for voiceless stops, as indicated by the statistical summaries in each panel of Fig. 5a–c). For the *voiced* stops, on the other hand, the directionality of the accent-induced modification of VOT was asymmetrical between the IP-initial and IP-medial positions (Fig. 5d–f). In the IP-*medial* position, the IVI of the voiced stops was significantly longer in the accented condition than in the unaccented condition (Fig. 5d), indicating an accent-induced modification of the IVI for *voiced* stops in a positive direction. This direction was further evidenced by a small but significantly longer VOT (Fig. 5e) under accent, as well as in the numerical direction of the Voicing-in-Closure (Fig. 5f), both of which indicate a phonetic reduction in voicing for voiced stops under accent. In the IP-*initial* position, however, the accent effect on the voicing contrast reflected in the IVI was observed only in the VOT, which turned out to be shorter under accent, whereas the Voicing-in-Closure in both the accented and unaccented conditions simply hovered near 0, showing no accent effect.

It is noteworthy that the voiced stops were generally produced with more variation than the voiceless stops, as shown by the IVI results and suggested by the error bars in Fig. 4a and d. The greater degree of variation in the voiced stops is largely attributable to the variation in the Voicing-in-Closure, as shown in Fig. 4c and f, whereas the opposite was true for VOT. This difference might result from physiological and aerodynamic constraints on maintaining voicing during closure, resulting in less consistent voicing realization during closure. This might be particularly true because English is not a true voicing language; phonetic implementation of the voiced stops is not based on the feature [voiced]. The observed variation might also be accounted for by the assumption that the laryngeal gesture for the voiced stops could be unspecified. In the discussion section, we provide a possible gestural account of the current findings.

*3.2. Iambic words (<u>**p**</u>anáche, <u>**T**</u>eníse, <u>**b**</u>anál, <u>**D**</u>eníse)*

3.2.1. Main effects

We found a significant main effect for Voice, Boundary, and Place on the **IVI** of stops in the unstressed syllable, but we did not find a main effect for Accent. As shown in Fig. 6A.1, the IVI was longer for voiceless stops than for voiced stops (Voice effect), which we attributed to both the **VOT** and **Voicing-in-Closure**. The VOT was shorter, but the Voicing-in-Closure was longer for voiced stops than for voiceless stops, as shown in Fig. 6A.2–3, respectively. The boundary effect, as can be seen in Fig. 6B.1, showed a DIS effect on the initial unstressed stops, with an increase in the IVI in the IP-initial position compared to the IP-medial position. This effect was also reflected in the VOT and Voicing-in-Closure, indicating an IP-initial lengthening of VOT (Fig. 6B.2) and an IP-initial shortening of Voicing-in-Closure (Fig. 6B.3). Once again, we found a place effect on the IVI, which was longer for alveolar stops than for

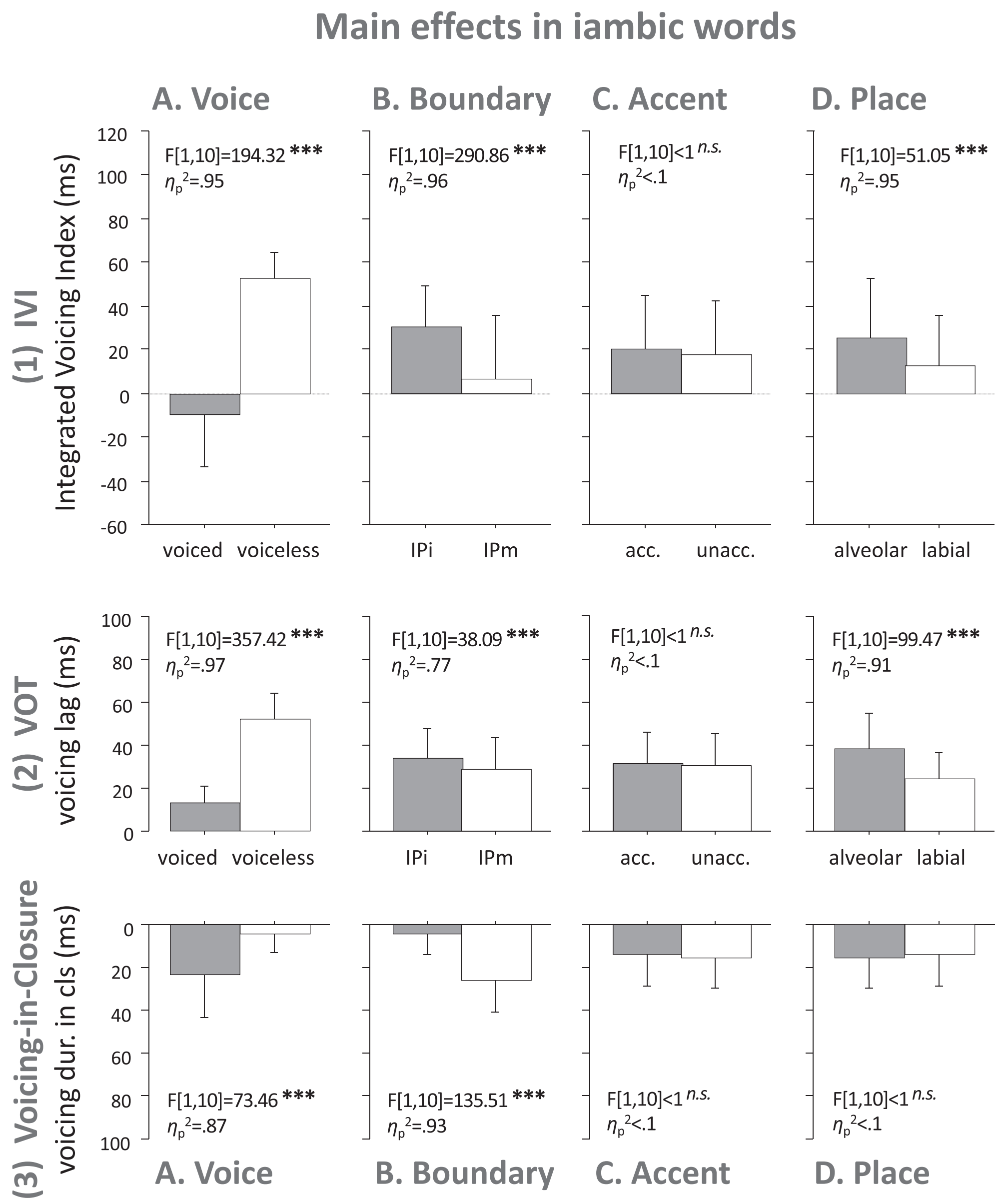

Fig. 6. Main effects of Voice, Boundary, Accent, and Place on the (1) IVI (Integrated Voicing Index), (2) VOT (voicing lag), and (3) Voicing-in-Closure (voicing duration in closure) for iambic words. Error bars refer to standard deviations; *** refers to $p < 0.001$.

labial stops (Fig. 6D.1). This was primarily attributable to the VOT, as can be inferred from Fig. 6D.2, whereas the Voicing-in-Closure remained unaffected by Place (Fig. 6D.3). Finally, we found no accent effect on any of the three measures, as shown in Fig. 6C.1–3, indicating that accentuating the following stressed syllable did not influence the VOT of the stops in the unstressed initial syllable, showing no leftward spreading of accentual lengthening.

As for interactions, we found no Voice × Accent interaction on any of the three measures (Fig. 7d–f), but we did find a significant Voice × Boundary interaction on all three measures, as we explain in detail in the following subsections. None of the measures showed a three-way interaction among Voice, Boundary, and Accent.

3.2.2. Voice × boundary in iambic words

We found a significant Voice × Boundary interaction on the **IVI**. As was the case with initial stops in the stressed syllable, the Voice × Boundary interaction on the IVI of unstressed initial stops occurred because the degree of voicing contrast was substantially reduced in the IP-initial vs. the IP-medial position. As shown in Fig. 7a, the effect size of Voice was smaller in the IP-initial position than in the IP-medial position, indicating a direction against the polarization of voicing contrast (Voice effect on the IVI: IPi, mean diff., 43.4 ms, $\eta^2 = 0.84$; IPm, mean diff., 83.4 ms, $\eta^2 = 0.97$). As in the case of an initial stress, the IP-initial reduction of the voicing contrast in the IVI was attributable in large part to a decrease in the phonetic voicing of voiced stops, as reflected in the **Voicing-in-Closure** (Fig. 7c) and also in part to an increase in the VOT of voiced stops in the IP-initial position, as shown in Fig. 7b. In other words, the voiced stops in the unstressed initial syllable were produced with a reduction in the IVI, showing a shift toward the positive dimension in the IP-initial vs. IP-medial position, whereas the IVI for voiceless stops remained unchanged.

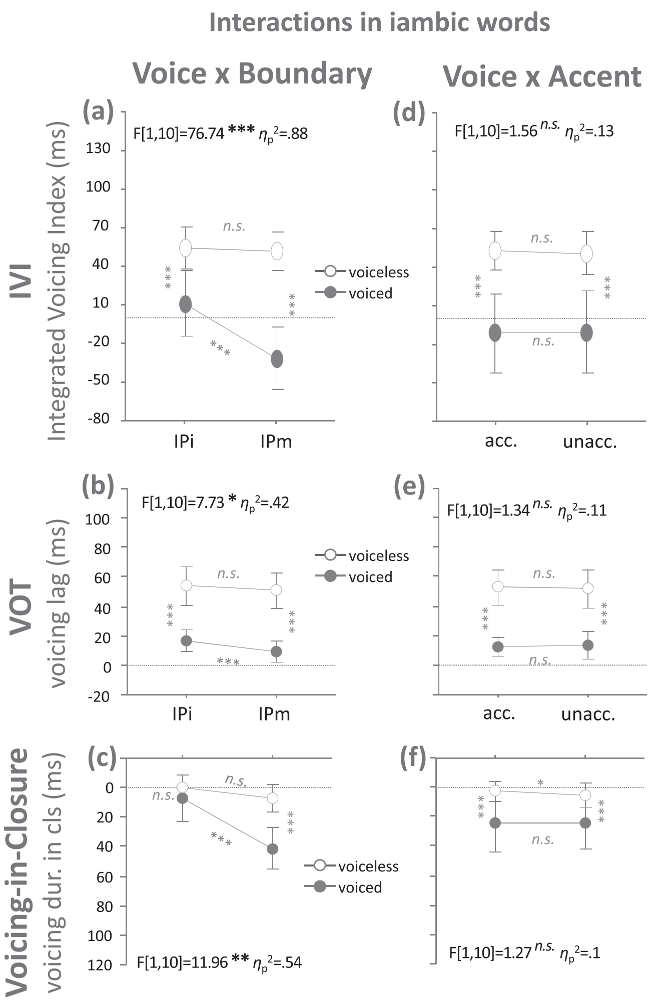

Fig. 7. Interactions of Voice × Boundary (a–c) and Voice × Accent (d–f) on the IVI (Integrated Voicing Index), VOT (voicing lag), and Voicing-in-Closure (voicing duration in closure) for iambic words. Error bars refer to standard deviations. * refers to $p < 0.05$, ** to $p < 0.005$, and *** to $p < 0.001$ obtained from separate one-way RM ANOVAs.

3.2.3. Boundary × accent in iambic words

We found a significant interaction between Boundary and Accent on the IVI, VOT, and Voicing-in-Closure, as illustrated in Fig. 8. As was generally observed with the stressed initial syllable, the interaction was due at least in part to a prominence-dependent DIS effect: The boundary-induced increase in the IVI was generally larger in the unaccented (less prominent) condition than in the accented condition (Fig. 8a), which was further evident in the VOT (Fig. 8b) and Voicing-in-Closure (Fig. 8c). There was no significant three-way Voice × Boundary × Accent interaction on any of the three measures, indicating that the DIS effect (i.e., the observed increase in voicelessness caused by the boundary) was consistent across voiced and voiceless stops, especially in the absence of prominence.

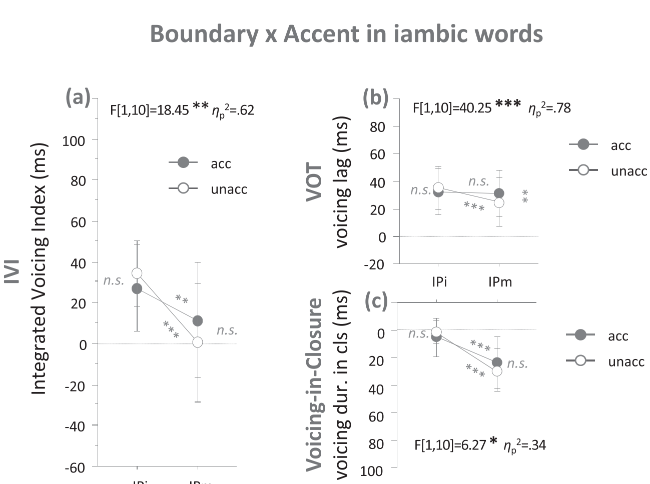

Fig. 8. Interactions of Boundary × Accent on the (a) IVI (Integrated Voicing Index), (b) VOT (voicing lag), and (c) Voicing-in-Closure (voicing duration in closure) for iambic words. Error bars refer to standard deviations. * refers to $p < 0.05$; ** to $p < 0.005$, and *** to $p < 0.001$ obtained from separate one-way RM ANOVAs.

4. General discussion

4.1. Effects of boundary-related domain-initial strengthening on voicing contrast

One of the primary questions that we aimed to answer in the present study was how **DIS** that might arise in association with boundary marking would modulate phonetic implementation of phonological voicing contrast between voiceless and voiced stops in American English. We considered three hypothetical polarizing patterns along the phonetic voicing continuum as possible types of voicing contrast as a function of prosodic strengthening, one of which was the DIS effect (Fig. 1).

First, consider DIS effects on the VOT of initial stops in trochaic words. Our results show that voiceless stops were not always produced with an increase in voicelessness under DIS, as reflected in the IVI, VOT, and Voicing-in-Closure. This is not fully consistent with what has been generally assumed about DIS—i.e., that domain-initial voiceless stops will be produced with articulatory strengthening of the laryngeal gesture (i.e., abduction gesture), resulting in a lengthening of VOT (Cooper, 1991; Pierrehumbert & Talkin, 1992). However, when we considered the data for the accented and unaccented conditions separately, we found that the voiceless stops were indeed produced with an increase in voicelessness in the IP-initial position compared with the IP-medial position in the unaccented condition, showing a prominence-dependent DIS effect. As discussed in the introduction, the prominence-dependent DIS effect is not new: VOTs for voiceless stops in English have often been found to be reliably longer under DIS when the initial syllable is not prominent, receiving no nuclear pitch accent (Cho & Keating, 2009; Cho et al., 2011, 2014; Cole et al., 2007). This has been discussed as a possible ceiling effect of prominence on VOT. As for voiced stops, on the other hand, our results show a consistent increase in voicelessness in a positive direction along the phonetic voicing continuum regardless of accent conditions. Moreover, driven primarily by voiced stops with an increased IVI (and a reduction in Voicing-in-Closure) in the positive dimension, the voicing contrast between voiced and voiceless stops turned out to be *reduced* rather than polarized in the IP-initial position. These results taken together reject the two hypothetical polarization effects schematized as Type 1 and Type 2 in Fig. 1; if anything, they appear to be better matched with Type 3, which shows no polarization of the voicing contrast.

A question that follows is then, what phonetic feature and featural enhancement can be involved in phonetic implementation of the voicing contrast under the influence of DIS. The increased IVI and VOT of voiceless stops under DIS, though limited to the *unaccented* condition, could well be interpreted as an enhancement of {vl. aspirated} (or {+spread glottis}), which has been assumed to be a phonetic feature in reference to which voiceless stops are phonetically realized in English (Keating, 1984, 1990). As for voiced stops, on the other hand, the boundary-induced VOT modulation could be seen as an enhancement of {vl. unaspirated} (or {-spread glottis}) rather than of {voiced} because the IVI for voiced stops changes from the negative to the positive dimension along the phonetic voicing continuum. However, if we follow the assumption that featural enhancements are driven by the principle of contrast maximization between phonological segments (i.e., *paradigmatic* contrast) to maximize lexical distinction (e.g., de Jong, 1995, 2004), it is not entirely clear whether featural enhancement is the actual driving force for the boundary-induced modification of VOT, which shows a reduction, rather than an enhancement, of voicing contrast.

Alternatively, when we consider the DIS effects on both voiced and voiceless stops, the results can be understood as driven by another kind of contrast enhancement—i.e., an enhancement of the structurally motivated *syntagmatic* contrast. The increased IVI and VOT can be seen as rendering the consonant more consonant-like (by increasing its voicelessness across the board), which maximizes its structural distinction from neighboring vowels, enhancing the CV contrast (e.g., Fougeron, 1999; Fougeron & Keating, 1997; Keating, Cho, Fougeron, & Hsu, 2003; Cho & Keating, 2009; see Cho, 2016, for a review).

Although we cannot tease these two accounts apart in full, we propose that the DIS effect on voicing contrast is more likely to be driven by a syntagmatic enhancement than a paradigmatic one for the following reasons. As just mentioned, if a paradigmatic (phonemic) enhancement drove the DIS effect, we should have observed some kind of polarization of voicing contrast. But we found the opposite, taking support away from the paradigmatic enhancement account.

Next, consider the extent to which DIS effects can be observed with initial stops in iambic words. Although we found a main effect of Boundary on the initial stops of iambic words (an overall increase in voicelessness under the influence of DIS), the Voice × Boundary interaction indicated that only the voiced stops showed a boundary-induced increase in voicelessness, as reflected in an increase in the IVI and a decrease the Voicing-in-Closure. Unlike the prominence (accent)-dependent effect seen on voiceless stops in trochaic words, voiceless stops in an unstressed (weak) syllable showed no boundary effect even in the unaccented condition. The results therefore suggest that the DIS effect might apply even to unstressed initial stops, but the effect is limited to *voiced* stops. The limited effect on the unstressed syllable, with no additional modulation as a function of accent, further implies that although the DIS effect might be prominence-dependent, it is not strictly in proportion to the degree of prominence—i.e., the least prominent condition (unstressed *and* unaccented) did not yield a particularly robust DIS effect.

With respect to the question of the type of enhancement that might underlie the observed DIS effect in an unstressed initial syllable, we found no evidence of a polarization of voicing contrast. Instead, similar to what we observed with trochaic words, the voicing contrast was substantially reduced in the IP-initial position compared with the IP-medial position, running counter to the polarization patterns schematized as Type 1 and Type 2 of Fig. 1. Because the voiceless stops in the unstressed initial syllable remained unchanged under the influence of DIS even in the *unaccented* condition, the observed pattern does not fit well with Type 3 either, at least not as closely as the voiceless stops in trochaic words, which showed an increase in voicelessness under DIS in the unaccented condition.

Our results therefore do not provide clear patterns with which to determine what kind of enhancement might underlie the observed DIS effects on the initial stops in iambic words. Increased voicelessness of voiced stops under the influence of DIS would have been more consistent with an enhancement of {vl. unaspirated} rather than {voiced}, if a phonetic featural enhancement underlay the DIS effect across the board. Likewise, the observed effect would have been more consistent with an enhancement of structurally motivated CV enhancement if voiceless stops were produced with an increase in voicelessness under DIS. Neither of those was true. Nevertheless, the DIS effect on the VOT of an unstressed-initial syllable is clearly not driven by the principle of maximizing phonological voicing contrast. We therefore propose that the DIS effect in the unstressed condition might also be better characterized as driven by a structurally motivated CV enhancement for the same reasons that we offered for the trochaic case: the phonetic voicing contrast was *reduced* under the influence of DIS, counter to polarization (see Kim et al., submitted, for related results in relational terms).

This interpretation leaves open the question of why the DIS effect was observed with voiced stops, but not with voiceless stops, in the *unstressed* syllable. One possible explanation that we can offer for now has to do with an overall phonetic weakening of segments in unstressed syllables. Segments in an unstressed syllable are generally reduced phonetically, so no compelling force (or linguistic motivation) appears to strengthen the degree of aspiration for voiceless stops insofar as the degree of aspiration is sufficiently maintained across positions. The VOT for voiceless stops in the unstressed syllable was on average well above 50 ms in both the IP-initial and IP-medial positions. The voiced stops, on the other hand, are likely to undergo a phonetic weakening in a prosodically weak position—i.e., the IP-medial position—making them more vulnerable to the coarticulatory force of voicing (or voicing overlap) from neighboring vowels, thus leading to some degree of phonetic voicing. But under the influence of DIS (i.e., in the IP-initial position), voiced stops appear to be strengthened in a way that increases their voicelessness, presumably to increase the distinction from neighboring vowels.

4.2. Effects of prominence on voicing contrast

Accent-driven prominence effects showed a strengthening pattern that differed from boundary-related DIS effects in terms of polarization type.

First, consider the accent effect on the IVI in trochaic words. The IVI increased for both voiced and voiceless stops under prominence (a nuclear pitch accent), driven by a contrastive focus. This pattern was further confirmed by an increase in the VOT and a decrease in the Voicing-in-Closure. That effect therefore rejects the maximal polarization of voicing contrast (Type 1 in Fig. 1), which would have been the case if voiced stops were produced with a more *negative* IVI in the accented condition. Nevertheless, a significant Voice × Accent interaction on the IVI indicated a kind of prominence-driven polarization of voicing contrast. The *voiceless* stops were produced with even more lengthening of the IVI (and VOT) than seen with the voiced stops, which augmented the voicing contrast between voiced and voiceless stops. This stands in sharp contrast with the *reduction* we saw in voicing contrast under the influence of DIS, which was caused primarily by the propensity for *voiced* stops to become more voiceless, narrowing the phonetic distance between voiced and voiceless stops along the phonetic voicing continuum.

As observed with voiced stops under the influence of DIS, voiced stops under accent were produced with a *lesser* degree of voicing (i.e., toward the positive dimension in the IVI and reduced Voicing-in-Closure) which reduced the voicing contrast, but the magnitude of the accent effect on voiceless stops was large enough to maximize voicing contrast. Furthermore, the IVI for voiced stops did not, on average, go substantially beyond '0' toward the positive dimension along the phonetic voicing continuum (as shown in Figs. 4d and 5d). Therefore, it did not encroach the territory that could be used by voiceless *aspirated* stops, contributing to the maximization of voicing contrast. The observed patterns thus largely match the Type 2 scheme in Fig. 1: voiceless stops primarily enhance the voicing contrast along the phonetic voicing continuum, whereas the IVI for voiced stops more or less centers on '0' under accent.

More specifically, even in the IP-medial position, in which, as discussed in the introduction, phonological voiced stops are often assumed to be specified with {voiced} (Keating, 1984), the IVI *increased* and the Voicing-in-Closure *decreased* under accent, showing no enhancement of {voiced}. In the IP-initial position, the IVI was, on average, positive, but it did not increase further under accent. If anything, the IVI tended to *decrease*, at least numerically. Therefore, these results do not lend strong support to the position-specific featural specification hypothesis (i.e., that voiced stops are specified with {vl. aspirated} in the initial position and with {voiced} in the medial position) suggested by Keating (1984). Instead, they are more consistent with the view that {vl. unaspirated} is involved in the phonetic implementation of phonologically voiced stops in English. (This is also consistent with the previously reported accent-induced shortening effect of VOT along the positive VOT dimension for phonological voiceless stops in Dutch (Cho & McQueen, 2005) and voiceless stops in /s/-stop clusters in English (Cho et al., 2014), which are assumed to be specified with {vl. unaspirated}, either underlyingly or allophonically.) As we proposed earlier, the phonetic voicing often observed in the IP-medial position is attributable to the coarticulatory propensity of voicing associated with voiced stops in a prosodically weak position.

The results taken together can then be interpreted to show an accent-induced enhancement of the phonetic feature {vl. unaspirated} for voiced stops and of {vl. aspirated} for voiceless stops, in line with the paradigmatic contrast enhancement assumed under focus-driven prominence (de Jong, 2004; see Cho, 2016 for related discussion).

Turning to iambic words, the initial stops in the unstressed syllable did not show robust prominence-related accent effects on voicing contrast. The voicing contrast was retained in the phonetic voicing dimension between initial voiced and voiceless stops in the unstressed syllable, but unlike the trochaic case, we found no main effect for Accent nor a significant Voice × Accent interaction. In other words, no further modification of phonetic realization occurred along the phonetic voicing continuum for the voicing contrast of initial stops when the following (stressed) syllable received a nuclear pitch accent driven by a contrastive focus. This suggests that the accent-driven prominence effect is localized to the stressed (second) syllable in iambic words, so that accentual lengthening does not extend to the preceding unstressed syllable (at least not as far as the VOT of the initial stop is concerned), which is largely consistent with previous observations (e.g., Turk & White, 1999; Dimitrova & Turk, 2012).

4.3. Possible gestural accounts

Finally, it is worth considering how our current findings can be understood in gestural terms in the framework of Articulatory Phonology (Browman & Goldstein, 1986, 1990, 1992). From an articulatory point of view at the laryngeal level, two different kinds of laryngeal gestures should be involved in contrasting voiceless and voiced stops: The glottal abduction (spreading) gesture is required for the former, and the glottal adduction (narrowing) gesture is required for the latter, all else being equal. On the one hand, the observed prosodic strengthening effects on voiceless (aspirated) stops in English can easily be translated in terms of articulatory strengthening of the glottal abduction gesture. An increase in the magnitude of the glottal abduction gesture would give rise to the observed increase in voicelessness (or VOT) for voiceless stops in prosodic strengthening environments.

For voiced stops, on the other hand, increasing the magnitude of the glottal adduction gesture for prosodic strengthening might constrict the vocal folds, and that could facilitate voicing of the vowel, shortening the VOT. The augmented glottal adduction gesture alone, however, cannot fully account for the variation in phonetic voicing of voiced stops, which straddles the boundary between the negative and positive dimensions of the phonetic voicing continuum. This is particularly true when following the general principle of Articulatory Phonology that the glottal adduction gesture affects the phonetic implementation of the voicing for voiced stops. We could still account for the phonetic variation by assuming that not only the magnitude of the glottal adduction gesture, but also the degree of glottal tension is modulated by prosodic strengthening factors. The increased glottal tension under prosodic strengthening might cause the vocal folds to become stiffer, which would suppress vocal-fold vibration during closure, as we observed in the Voicing-in-Closure values in the domain-initial position and accented condition of this study. In the same fashion, the phonetic voicing observed for voiced stops in prosodically weak positions could result from slack vocal folds in the absence of prosodic strengthening and thus be vulnerable to the coarticulatory influence of the flanking vowels.

As discussed above, the featural enhancement hypothesis would provide a unified account of the realization of phonetic voicing for voiced stops in an aspirated language such as English and the voiceless (unaspirated) stops in a true voicing language such as Dutch, as well as for the allophonically driven voiceless unaspirated variant of voiceless stops in /s/-stop clusters in English. But the featural account does not capture the possible gestural invariance underlying the surface phonetic realization. For example, in general terms, the increased VOT for voiceless (aspirated) stops and the reduced VOT of the stop in /s/-stop clusters in English can be taken to underlie the glottal *abduction* gesture (e.g., Goldstein, 1992; see Cho et al., 2014 for related discussion), whereas the featural account assumes separate phonetic features, {vl. aspirated} for the former and {vl. unaspirated} for the latter. On the other hand, it appears that the gestural account cannot easily explain the phonetically similar surface forms that could underlie different phonological categories. For example, at least in domain-initial positions, a phonetically voiceless unaspirated stop arises on the surface for both phonologically voiced stops in English and phonologically voiceless stops in Dutch, but they phonetically diverge in domain-medial positions such that the former becomes phonetically voiced (at least partially) and the latter remains voiceless and unaspirated. These language-specific modulations of voicing realization as a function of prosodic strengthening cannot be simply understood in terms of gestural dichotomy, whether the underlying gesture is an abduction or adduction gesture.

One way of resolving this issue in the frame of Articulatory Phonology might be to assume three types of gestural specifications. An abduction gesture, an adduction gesture, and underspecification for the laryngeal gesture. The idea of underspecification of the laryngeal gesture is indeed in line with an assumption made in Articulatory Phonology (e.g., Browman & Goldstein, 1986), which posits vocal fold vibration as a default mode in the absence of the laryngeal abduction gesture. It is then reasonable to assume that the voiced stops in English, unspecified for the laryngeal gesture, are phonetically realized as voiced in a context in which voicing is facilitated (such as being flanked by vowels in a phrase-medial position), whereas vocal fold vibration is impeded under prosodic strengthening, which could pose some aerodynamic constraints on voicing. (The observed variation in voiced stops in this study might also be attributable to the fact that the laryngeal gesture is not specified, allowing for some variation.) This possibility has also been mentioned in Davidson (2017), who discussed possible gestural representations for voiced and voiceless obstruents in English. Although gestural underspecification might thus explain phonetic behaviors for voiced stops in English, it is unclear what kind of gestural representation underlies the voiceless stops in a true voicing language (e.g., Dutch). It could be a glottal abduction gesture whose magnitude is smaller than that used in an aspiration language (e.g., English), resulting in voiceless unaspirated and aspirated forms. (See Davidson (2017) for a related discussion on how voicing ('phonation') during a stop closure can be actively maintained in true voicing languages.) Alternatively, as suggested by Cho and Ladefoged (1999) (also Ladefoged & Cho, 2001), speakers might control the timing of their voicing onset relative to the release (i.e., so-called *Articulatory VOT*) in a language-specific way, resulting in cross-linguistic differences.

The gestural account, as discussed so far, provides insights into the gestural underpinnings of surface acoustic variation as an alternative to the featural enhancement account. Recent years have seen some advancement in Articulatory Phonology regarding how articulatory realizations of gestures can be modulated by prosodic factors (see Krivokapić, in press for a review), but it remains to be seen how a single underlying laryngeal gesture can surface with different voicing patterns across languages under prosodic structural environments that are assumed to regulate the fine-tuning of the articulatory realization of gestures in language-specific ways.

5. Conclusion

The important findings of the present study are encapsulated as follows. The boundary-related DIS effect indicates that voicing contrast between voiced and voiceless stops is not polarized along the phonetic voicing continuum; instead, the degree of voicelessness increases for both voiced and voiceless stops. The increased voicelessness of voiced stops toward the positive dimension, however, results in a reduction of the phonetic distance between the voiceless and the voiced stops. We propose that the observed DIS effect is interpretable as an enhancement of a structurally motivated CV contrast. The DIS effect on the unstressed initial syllables of iambic words is, however, limited to initial *voiced* stops, whereas the effect on the stressed initial syllables of trochaic words can apply to both voiced and voiceless stops in a prominence-dependent way (being more robust in a *less* prominent, unaccented condition). The exact nature of the intricate interactions between DIS and prominence is a subject for further research. For now, the less clear DIS effect in unstressed syllables can be attributed to a general phonetic weakening of stops in the weak syllable—i.e., voiced stops could be more vulnerable than voiceless stops to coarticulatory voicing effects from neighboring vowels. Given that the intervocalic phonetic weakening effect disappears under the influence of DIS, which would effectively increase the CV contrast, the modification of phonetic voicing for voiced stops in an unstressed syllable can also be interpreted as a CV enhancement.

The accent-driven prominence effect shows quite a different enhancement pattern for voicing contrast. The voicing contrast is indeed maximized under accent. The voiced and voiceless stops, however, are not fully polarized in opposite directions along the phonetic voicing continuum. Instead, the degree of voicelessness for voiced stops increases from the phonetically voiced (negative) to the phonetically voiceless (positive) territory along the phonetic voicing continuum, eventually centering near '0' under accent. Thus, the phonetic implementation of voiced stops operates based on the phonetic feature {vl. unaspirated} rather than {voiced}, taking support away from the position-specific featural specification hypothesis. On the other hand, voiceless stops are produced with an even longer VOT under accent, operating on the phonetic feature {vl. aspirated}, which effectively maximizes voicing contrast. We also provided an alternative account in gestural terms to illuminate the gestural underpinnings of the variation that arises on the surface. These accounts spark questions about how the

Table A1
Distribution of tokens across conditions for each speaker. (For individual speakers' data, see Kim, Kim & Cho, submitted.)

Stress	Voice	Boundary	Accent	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	Total
str	vd	IPi	acc	6	6	6	5	6	8	6	8	8	6	7	72
str	vd	IPi	unacc	6	6	6	6	6	4	5	5	5	3	6	58
str	vd	IPm	acc	6	6	6	6	5	7	6	8	7	5	8	70
str	vd	IPm	unacc	6	6	6	6	6	6	6	5	6	5	6	64
str	vls	IPi	acc	6	6	6	6	6	8	6	8	6	8	9	75
str	vls	IPi	unacc	6	6	6	5	6	6	6	6	5	4	4	60
str	vls	IPm	acc	6	6	6	5	6	8	6	8	7	8	8	74
str	vls	IPm	unacc	6	6	6	6	6	5	6	4	6	3	6	60
unstr	vd	IPi	acc	6	6	6	6	6	7	6	6	8	7	9	73
unstr	vd	IPi	unacc	6	6	6	5	6	6	6	3	7	3	7	61
unstr	vd	IPm	acc	6	6	6	6	6	6	6	8	6	5	8	69
unstr	vd	IPm	unacc	6	6	4	6	6	6	6	5	4	5	5	59
unstr	vls	IPi	acc	6	6	6	6	6	6	6	6	7	8	5	68
unstr	vls	IPi	unacc	5	6	6	3	5	4	3	6	3	4	3	48
unstr	vls	IPm	acc	6	6	6	3	6	6	6	5	3	8	7	62
unstr	vls	IPm	unacc	6	5	6	4	6	5	3	3	4	4	3	49
Total				95	95	94	84	94	98	89	94	92	86	101	1022

voicing contrast in true voicing languages such as French, Russian, and Spanish could be phonetically implemented under prominence. If the phonetic implementation of voicing contrast indeed operates based on {voiced}, one could expect that the Voicing-in-Closure is actively controlled, presumably resulting in a pattern that would mirror what we found here in English. It would also remain to be seen how prominence-related strengthening would interact with boundary-related strengthening in those languages. Further research should explore the exact nature of the relationship between surface phonetic realization and the underlying representation from both the featural and gestural perspectives across languages.

All in all, these findings demonstrate that the phonetic implementation of voicing contrast along the VOT continuum in American English is modulated differentially by the delimitative (boundary marking) and culminative (prominence marking) functions of prosodic structure. Also, the observed effects can be seen as different types of prosodic strengthening driven by different linguistic contrast enhancements in reference to language-specific phonetic features. In other words, seemingly non-contrastive low-level phonetic variation in voicing contrast is in fact systematically modulated by prosodic structure in reference to phonetic representations that regulate the phonetic implementation of the phonological contrast in a given language. We hope that our study inspires further cross-linguistic research on the phonetics–prosody interface, including true voicing languages, to further illuminate the nature of voicing contrast across languages.

Acknowledgements

We thank the American English speakers who participated in this study. We also thank the reviewers and the Guest Editor (D. Whalen) for their constructive comments. Some of the data collected were used for the second author's MA thesis at Hanyang University. This work was supported by Global Research Network program through the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (Grant No. NRF-2016S1A2A2912410), and by 2018 Hongik University Research Fund.

Appendix A

See Table A1.

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.wocn.2018.07.004.

References

Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics, 63*, 75–86.
Baese-Berk, M., & Goldrick, M. (2009). Mechanisms of interaction in speech production. *Language and Cognitive Processes, 24*, 527–554.
Beckman, M. E. (1996). The parsing of prosody. *Language and Cognitive Processes, 11*, 17–67.
Beckman, M. E., & Edwards, J. (1994). Articulatory evidence for differentiating stress categories. In P. A. Keating (Ed.), *Papers in laboratory phonology III: Phonological structure and phonetic form* (pp. 7–33). Cambridge: Cambridge University Press.
Boersma, P., & Weenink, D. (2014). Praat: doing phonetics by computer [Computer program]. http://www.praat.org/.
Browman, C. P., & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*, 219–252.
Browman, C. P., & Goldstein, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. In J. Kingston & M. E. Beckman (Eds.), *Papers in laboratory phonology: Between the grammar and the physics of speech* (pp. 341–376). Cambridge: Cambridge University Press.
Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*, 155–180.
Byrd, D., Krivokapić, J., & Lee, S. (2006). How far, how long: On the temporal scope of prosodic boundary effects. *Journal of the Acoustical Society of America, 120*, 1589–1599.
Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /A, i/ in English. *Journal of the Acoustical Society of America, 33*, 121–157.
Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In L. Goldstein, D. H. Whalen, & C. T. Best (Eds.), *Laboratory phonology, Vol. 8: Varieties of phonological competence* (pp. 519–548). Berlin/New York: De Gruyter Mouton.
Cho, T. (2011). Laboratory phonology. In N. C. Kula, B. Botma, & K. Nasukawa (Eds.), *The continuum companion to phonology* (pp. 343–368). London & New York: Continuum.
Cho, T. (2016). Prosodic boundary strengthening in the phonetics-prosody interface. *Language and Linguistics Compass, 10*(3), 120–141.
Cho, T., & Jun, S.-A. (2000). Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society (CLS), 36*, 31–44.
Cho, T., Jun, S.-A., & Ladefoged, P. (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. *Journal of Phonetics, 30*, 193–228.
Cho, T., & Keating, P. A. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics, 37*, 466–485.
Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics, 27*, 207–229.
Cho, T., Lee, Y., & Kim, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics, 39*, 344–361.
Cho, T., Lee, Y., & Kim, S. (2014). Prosodic strengthening on the /s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics, 46*, 128–146.
Cho, T., & McQueen, J. M. (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics, 33*, 121–157.

Chodroff, E., & Wilson, C. (2017). Structure in talker-specific phonetic realization: Covariation of stop consonant VOT in American English. *Journal of Phonetics, 61*, 30–47.

Cole, J., Kim, H., Choi, H., & Hasegawa-Johnson, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News Speech. *Journal of Phonetics, 35*, 180–209.

Cooper, A. (1991). *Glottal Gestures and Aspiration in English* PhD Dissertation. Yale University.

Davidson, L. (2016). Variability in the implementation of voicing in American English obstruents. *Journal of Phonetics, 54*, 35–50.

Davidson, L. (2017). Phonation and laryngeal specification in American English voiceless obstruents. *Journal of the International Phonetic Association*. https://doi.org/10.1017/S0025100317000330.

de Jong, K. J. (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America, 97*, 491–504.

de Jong, K. J. (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics, 32*, 493–516.

de Jong, K. J., & Zawaydeh, B. A. (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics, 30*, 53–75.

Dimitrova, S., & Turk, A. E. (2012). Patterns of accentual lengthening in English four-syllable words. *Journal of Phonetics, 40*, 403–418.

Docherty, G. J. (1989). *An experimental phonetic study of the timing of voicing in English obstruents*. Edinburgh: University of Edinburgh Dissertation.

Fletcher, J. (2010). The prosody of speech: Timing and rhythm. In W. J. Hardcastle, J. Laver, & F. E. Gibbon (Eds.), *The handbook of phonetic sciences* (2nd ed., pp. 523–602). Oxford: Blackwell.

Fougeron (1999). Prosodically conditioned articulatory variations: a review. *UCLA Working Papers in Phonetics, 97*, 1–74.

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728–3740.

Goldstein, L. (1992). Comments on chapters 3 and 4. In G. Docherty & D. R. Ladd (Eds.), *Laboratory phonology, Vol. 2: Gesture, segment, prosody* (pp. 120–124). Cambridge: Cambridge University Press.

Jacewicz, E., Fox, R., & Lyle, S. (2009). Variation in stop consonant voicing in two in regional varieties of American English. *Journal of the International Phonetic Association, 39*, 313–334.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language, 60*, 286–319.

Keating, P. A. (1990). Phonetic representations in a generative grammar. *Journal of Phonetics, 18*, 321–334.

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. (2003). Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.), *Laboratory Phonology, Vol. 6: Phonetic Interpretation* (pp. 145–163). Cambridge: Cambridge University Press.

Keating, P. A., & Shattuck-Hufnagel, S. (2002). A prosodic view of word form encoding for speech production. *UCLA Working Papers in Phonetics, 101*, 112–156.

Kim, S., Kim, J., & Cho, T. (Submitted). Stop voicing contrast in American English: Data of individual speakers in trochaic and iambic words in different prosodic structural contexts.

Kingston, J., & Diehl, R. L. (1994). Phonetic knowledge. *Language, 70*, 419–454.

Krivokapić, J. (in press). Prosody in articulatory phonology. In S. Shattuck-Hufnagel & J. Barnes (Eds.), Prosodic Theory and Practice. MIT Press.

Ladefoged, P. (1967). *Three areas of experimental phonetics*. London: Oxford University Press.

Ladefoged, P. & Cho T. (2001) Linking linguistic contrasts to reality: The case of VOT. In N. Gronnum & J. Rischel (Eds.), Travaux Du Cercle Linguistique De Copenhague, vol. XXXI. (To Honour Eli Fischer-Forgensen.), C.A. Reitzel; Copenhagen, 212–223.

Ladefoged, P., & Loeb, G. (2002). Preliminary studies on respiratory activity in speech. *UCLA Working Papers in Phonetics, 101*, 50–60.

Lindblom, B. (1990). Explaining phonetic variation: A sketch of the H & H theory. *Speech Production and Speech Modeling, 55*, 403–439.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word, 20*, 384–422.

Mikuteit, S., & Reetz, H. (2007). Caught in the ACT: The timing of aspiration and voicing in East Begnali. *Language and Speech, 50*(2), 247–277.

Nelson, N. R., & Wedel, A. (2017). The phonetic specificity of competition: Contrastive hyperarticulation of voice onset time in conversational English. *Journal of Phonetics, 64*, 51–70.

Pierrehumbert, J., & Talkin, D. (1992). Lenition of /h/ and glottal stop. In G. Docherty & D. R. Ladd (Eds.), *Laboratory Phonology, Vol. 2: Gesture, Segment, Prosody* (pp. 90–117). Cambridge: Cambridge University Press.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*, 193–247.

Smiljanic, R., & Bradlow, A. R. (2008). Stability of temporal contrasts across speaking styles in English and Croatian. *Journal of Phonetics, 36*, 91–113.

Tomaschek, F., Arnold, D., Bröker, F., & Baayen, H. (2018). Lexical frequency co-determines the speed-curvature relation in articulation. *Journal of Phonetics, 68*, 103–116. B.

Turk, A. E., & White, L. (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics, 27*, 171–206.

Wright, R. (2004). Facotrs of lexical competition in vowel articulation. In J. Local, R. Ogden, & R. Temple (Eds.), *Papers in laboratory phonology VI* (pp. 26–50). Cambridge: Cambridge University Press.